



# **AFFECTIVE COMPUTING AND REGULATION IN BRAIN COMPUTER INTERFACE**

EDITED BY: Zehong Jimmy Cao, Hiroshi Higashi and Jane Zhen Liang

PUBLISHED IN: Frontiers in Neuroscience and

Frontiers in Computational Neuroscience



# frontiers

## Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence.

The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714

ISBN 978-2-88976-551-5

DOI 10.3389/978-2-88976-551-5

## About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

## Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

## Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

## What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: [frontiersin.org/about/contact](https://frontiersin.org/about/contact)

# AFFECTIVE COMPUTING AND REGULATION IN BRAIN COMPUTER INTERFACE

Topic Editors:

**Zehong Jimmy Cao**, University of South Australia, Australia

**Hiroshi Higashi**, Kyoto University, Japan

**Jane Zhen Liang**, Shenzhen University, China

**Citation:** Cao, Z. J., Higashi, H., Liang, J. Z., eds. (2022). Affective Computing and Regulation in Brain Computer Interface. Lausanne: Frontiers Media SA.  
doi: 10.3389/978-2-88976-551-5

# Table of Contents

- 04** *Inferior Frontal Gyrus-Based Resting-State Functional Connectivity and Medium Dispositional Use of Reappraisal Strategy*  
Wenjuan Li, Ke Xie, Ronald K. Ngetich, Junjun Zhang, Zhenlan Jin and Ling Li
- 16** *Multimodal Neuroimaging Predictors of Learning Performance of Sensorimotor Rhythm Up-Regulation Neurofeedback*  
Linling Li, Yinxue Wang, Yixuan Zeng, Shaohui Hou, Gan Huang, Li Zhang, Nan Yan, Lijie Ren and Zhiguo Zhang
- 28** *An Empirical Comparative Study on the Two Methods of Eliciting Singers' Emotions in Singing: Self-Imagination and VR Training*  
Jin Zhang, Ziming Xu, Yueying Zhou, Pengpai Wang, Ping Fu, Xijia Xu and Daoqiang Zhang
- 40** *Identification of Emotion Using Electroencephalogram by Tunable Q-Factor Wavelet Transform and Binary Gray Wolf Optimization*  
Siyu Li, Xiaotong Lyu, Lei Zhao, Zhuangfei Chen, Anmin Gong and Yunfa Fu
- 52** *Relaxation Degree Analysis Using Frontal Electroencephalogram Under Virtual Reality Relaxation Scenes*  
Yue Zhang, Lulu Zhang, Haoqiang Hua, Jianxiu Jin, Lingqing Zhu, Lin Shu, Xiangmin Xu, Feng Kuang and Yunhe Liu
- 65** *Monitoring and Evaluation of Emotion Regulation by Aerobic Exercise and Motor Imagery Based on Functional Near-Infrared Spectroscopy*  
Peng Ding, Fawang Wang, Siyu Li, Wei Zhang, Hongquan Li, Zhuangfei Chen, Lei Zhao, Anmin Gong and Yunfa Fu
- 78** *Efficacy Evaluation of Neurofeedback-Based Anxiety Relief*  
Chao Chen, Xiaolin Xiao, Abdelkader Nasreddine Belkacem, Lin Lu, Xin Wang, Weibo Yi, Penghai Li, Changming Wang, Sha Sha, Xixi Zhao and Dong Ming
- 89** *Hierarchical Spatiotemporal Electroencephalogram Feature Learning and Emotion Recognition With Attention-Based Antagonism Neural Network*  
Pengwei Zhang, Chongdan Min, Kangjia Zhang, Wen Xue and Jingxia Chen
- 106** *Measuring and Modeling the Effect of Audio on Human Focus in Everyday Environments Using Brain-Computer Interface Technology*  
Aia Haruvi, Ronen Kopito, Noa Brande-Eilat, Shai Kalev, Eitan Kay and Daniel Furman
- 123** *Is Mate Preference Recognizable Based on Electroencephalogram Signals? Machine Learning Applied to Initial Romantic Attraction*  
Guangjie Yuan, Wenguang He and Guangyuan Liu
- 134** *Closed-Loop Tracking and Regulation of Emotional Valence State From Facial Electromyogram Measurements*  
Luciano R. F. Branco, Arian Ehteshami, Hamid Fekri Azgomi and Rose T. Faghih





# Inferior Frontal Gyrus-Based Resting-State Functional Connectivity and Medium Dispositional Use of Reappraisal Strategy

Wenjuan Li, Ke Xie, Ronald K. Ngetich, Junjun Zhang, Zhenlan Jin\* and Ling Li\*

MOE Key Laboratory for Neuroinformation, High-Field Magnetic Resonance Brain Imaging Key Laboratory of Sichuan Province, Center for Psychiatry and Psychology, School of Life Sciences and Technology, University of Electronic Science and Technology of China, Chengdu, China

## OPEN ACCESS

### Edited by:

Jane Zhen Liang,  
Shenzhen University, China

### Reviewed by:

Linling Li,  
Shenzhen University, China  
Binlong Zhang,  
Guang'anmen Hospital, China  
Academy of Chinese Medical  
Sciences, China

### \*Correspondence:

Zhenlan Jin  
jinzi@uestc.edu.cn  
Ling Li  
liling@uestc.edu.cn

### Specialty section:

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

**Received:** 17 March 2021

**Accepted:** 11 May 2021

**Published:** 17 June 2021

### Citation:

Li W, Xie K, Ngetich RK, Zhang J,  
Jin Z and Li L (2021) Inferior Frontal  
Gyrus-Based Resting-State  
Functional Connectivity and Medium  
Dispositional Use of Reappraisal  
Strategy. *Front. Neurosci.* 15:681859.  
doi: 10.3389/fnins.2021.681859

The previous neuroimaging functional connectivity analyses have indicated that the association between the inferior frontal gyrus (IFG) and other brain regions results in better emotion regulation in reappraisal tasks. However, no study has explored the relationship between IFG-based resting-state functional connectivity (rsFC) and the dispositional use of reappraisal strategy. Therefore, the present study examined the potential associations between rsFC patterns of both left and right IFG and dispositional reappraisal use. One hundred healthy participants completed the Emotion Regulation Questionnaire (ERQ) and underwent a resting-state functional magnetic resonance imaging (fMRI) acquisition. An approach of the seed-based rsFC analysis was recruited to estimate the functional connectivity maps of bilateral IFG with other brain regions, and the reappraisal scores from the ERQ were then correlated with the functional maps. Our findings showed that IFG-based rsFC was positively correlated with dispositional reappraisal only in the range of 4 to 5.5 points [medium reappraisal group (MRG)]. Specifically, medium dispositional reappraisal was positively correlated with rsFC between left/right IFG and bilateral temporal gyrus. Besides, medium dispositional reappraisal was positively correlated with rsFC between left IFG and bilateral superior parietal lobe (SPL), middle cingulate cortex (MCC), and right insula, as well as between right IFG and dorsomedial prefrontal cortex (DMPFC) and anterior cingulate cortex (ACC). In conclusion, these results indicate that bilateral IFG plays an important role in the medium use of the reappraisal strategy.

**Keywords:** emotion regulation, inferior frontal gyrus, prediction, resting-state functional connectivity, medium reappraisal

## INTRODUCTION

Effective emotion regulation is necessary for our daily social life. Essentially, various strategies can be employed to achieve successful emotion regulation, e.g., distraction, cognitive reappraisal, and expressive suppression (Webb et al., 2012; Morawetz et al., 2017b). Among these strategies, reappraisal, which entails the changing of the emotional value

of stimuli that evokes emotions (Kanske et al., 2011; Webb et al., 2012), is the most frequently applied and studied strategy of emotion regulation (Wager et al., 2008; Kalisch, 2009; Ochsner et al., 2012; Buhle et al., 2014). Moreover, Gross and John (2003) developed a self-report Emotion Regulation Questionnaire (ERQ) to measure the dispositional use of two strategies, reappraisal (center on reinterpretation) and suppression. Assessment of this kind of personality habitude can reflect the utilization frequency of strategy, which may finally implicate the individual differences in abilities of emotion regulation. Also, more frequent use of reappraisal strategy has been demonstrated to be associated with better regulation of emotions, social interactions, and mental and physical health (Gross and John, 2003; Meyer et al., 2012; Hu et al., 2014; Picó-Pérez et al., 2018; Zaehring et al., 2020).

The neural underpinnings related to reappraisal strategy have usually been evaluated by measurement of functional activation during experimental reappraisal tasks with functional magnetic resonance imaging (fMRI). The previous meta-analytic studies have shown that reappraisal recruits a widespread network that includes dorsomedial prefrontal cortex (DMPFC), dorsolateral prefrontal cortex (DLPFC)/superior frontal gyrus (SFG), ventrolateral prefrontal cortex (VLPFC)/inferior frontal gyrus (IFG), parietal lobes, temporal gyrus, and cingulate cortex (Phillips et al., 2008; Buhle et al., 2014; Kohn et al., 2014; Morawetz et al., 2017b). Importantly, the IFG/VLPFC is well known as a critical region for processes of selection and inhibition (Schulz et al., 2009; Kohn et al., 2014; Morawetz et al., 2017b), language (Ochsner et al., 2012; Messina et al., 2015), and social cognition (Kohn et al., 2014; Hartwigsen et al., 2019) in emotion regulation. In particular, the IFG has been observed with increased activation when multiple appropriate reinterpretations emerge and a choice must be made to achieve goal-directed behavior, as well as when required to inhibit goal-inappropriate reinterpretations (Morawetz et al., 2016; Braunstein et al., 2017).

Moreover, methods of IFG-based (with IFG serving as seed regions) functional connectivity have also been used to explore the neural correlations of reappraisal strategy in experimental settings. Morawetz et al. (2017a) defined left IFG as a seed region and examined effective connectivity between the seed region and the remaining brain regions, with reappraisal success scores as covariate. They have observed positive effective coupling between the left IFG and DLPFC, DMPFC, right middle temporal gyrus (MTG), and superior temporal gyrus (STG) during downregulation of emotion. Furthermore, in another study, Morawetz et al. (2016) found that the inhibitory effect on connectivity from IFG to DLPFC could facilitate successful reappraisal, deducing that the IFG may choose one from the many feasible goal-relevant reinterpretations actively maintained in the working memory (associated with DLPFC's increased activity), and suppress the DLPFC as soon as the selection process is finished. On the other hand, Wager et al. (2008) found that the right IFG could effectively predict reappraisal success with some cortical and subcortical regions as mediators, such as DMPFC, SFG, inferior temporal gyrus (ITG), and subgenual anterior cingulate cortex (ACC).

Recent evidence has shown that the formation of intrinsic resting-state functional architecture is influenced by repeated task-based co-activation within a network (Mackey et al., 2013; Guerra-Carrillo et al., 2014; Uchida et al., 2014), suggesting a close correspondence between task-specific brain activation and intrinsic brain connectivity, which is reflected by resting-state functional connectivity (rsFC). In essence, Smith et al. (2009) compared task-based activation networks derived from a large database of functional imaging studies with the covarying networks from 36 subjects' resting fMRI data, and found that these task-related networks closely matched the networks when at resting state. Intriguingly, another study using a sample of 4- to 18-year-old healthy participants found that task-related functional connectivity could even predict rsFC of up to 2 years after the initial experimental task (Gabard-Durnam et al., 2016). Thus, it seems possible that recurring activation caused by a specific task may share an association with resting-state connectivity pattern, and this may also apply to the emotion regulation domain with IFG-based functional connectivity during the reappraisal task. Besides, one study demonstrated that activation of IFG in reappraisal task is positively correlated with the frequency of dispositional reappraisal in daily life (Grecucci et al., 2013). Consequently, it can be assumed that the reappraisal task and dispositional reappraisal may share a similar IFG-based functional connectivity pattern. However, since no study has examined the association between dispositional use of reappraisal strategy and rsFC with IFG seed regions, it is uncertain whether IFG-based rsFC could also facilitate habitual reappraisal, thus resulting in better emotion regulation, although previous evidence indicates that both left and right IFG show associations with DMPFC and temporal gyrus during the performance of a reappraisal task. Nevertheless, it remains unknown whether the functional connectivity pattern of the left IFG seed concerning habitual reappraisal is the same as that of the right IFG seed.

Moreover, neural efficiency supposes that more adept individuals optimally use the functional connectivity to undertake minute neural processing and, hence, display diminished neural activity alongside the performance facilitation (Neubauer and Fink, 2009; Di Domenico et al., 2015; Curtin et al., 2019). It is anticipated that the higher the scores of dispositional reappraisal, the more frequent the use of reappraisal, and the better the ability of emotion regulation. Presumably, the neural efficiency may also be suitable for dispositional reappraisal, a daily used specific strategy of emotion regulation, with a changeable connectivity pattern along with a variation of reappraisal scores. However, there is no evidence supporting how the individual difference in frequency of reappraisal use may affect functional connectivity during the resting state. Considering the aforementioned close association between rsFC and task-related functional connectivity, we asked another question: Could it be possible that the IFG-based rsFC pattern vary with the level of frequency of dispositional reappraisal?

Therefore, in an endeavor to answer the questions raised, the present study applied seed-based rsFC analysis and prediction analysis with an aim of (1) examining whether IFG-based rsFC is related to individual dispositional use of reappraisal; (2) investigating whether habitual reappraisal related rsFC pattern of

the left IFG is the same as that of the right IFG; (3) exploring whether IFG-based rsFC is specific to the frequency level of the use of dispositional reappraisal. Based on the limited evidence mentioned above, we hypothesized that dispositional use of reappraisal would be positively correlated with IFG-based rsFC and that the rsFC pattern of the left IFG and that of the right IFG would be very similar, with both showing associations with DMPFC and temporal gyrus. Moreover, according to the theory of neural efficiency (Neubauer and Fink, 2009; Di Domenico et al., 2015; Curtin et al., 2019), the individual differences in dispositional reappraisal may be associated with the difference in the IFG-based rsFC pattern, with a higher level of dispositional reappraisal corresponding to a less IFG-based rsFC.

## MATERIALS AND METHODS

### Participants

One hundred and seven ( $N = 107$ ) healthy, right-handed adults (59 females, 17 to 26 years old, mean age  $21.36 \pm 2.052$  years) participated in the experiment after giving their written informed consent. All participants reported no history of mental disorders, head injury, or cardiovascular diseases. The study protocol was approved by the Local Committee for the Protection of Human Subjects of the University of Electronic Science and Technology of China and was conducted according to the declaration of Helsinki.

### Behavioral Assessment

All participants completed the ERQ (Gross and John, 2003) before scanning. The ERQ consisted of two subscales, cognitive reappraisal and expressive suppression. Ten items, with six for reappraisal and four for suppression, are included in the scale, whose choices ranged from strongly disagree (1) to strongly agree (7). In the current study, Cronbach's  $\alpha$  coefficient of reappraisal subscale was 0.828 and that of suppression subscale was 0.601. According to the scale used, four is the median score. In our interpretation, the individuals who scored less than four points do not frequently use either the reappraisal or suppression strategy. On the other hand, a score of more than four implies a frequent use of the aforementioned strategies. In our study, a very small number of participants ( $n = 7$ ) recorded less than four points on the reappraisal subscale. Therefore, we found it appropriate to examine the intrinsic neural mechanisms of

reappraisal strategy only among the individuals who frequently apply this strategy in their daily life. Hence, we excluded the data of  $n = 7$  participants with low frequent use of reappraisal strategy ( $< 4$  points). Subsequently, we categorized the remaining  $n = 100$  participants into two groups, according to their reappraisal scores. Those who had a score of between 4 (median) and 5.5 were classified as moderate users of reappraisal strategy and thus put into the medium reappraisal group (MRG). Similarly, those with a score of between 5.5 and 7 were considered as high-frequency reappraisal strategy users and, therefore, categorized into high reappraisal group (HRG). Ultimately, all the remaining participants ( $n = 100$ ) were assigned into either MRG ( $n = 80$ ) or HRG ( $n = 20$ ). In MRG, neither scores of reappraisal and suppression nor age showed significant gender differences (all  $p > 0.08$ ). This was the same case with HRG (all  $p > 0.4$ ) (Table 1). On the other hand, the reappraisal score was higher than the suppression score in both groups [MRG:  $t_{(79)} = 11.989$ ,  $p < 0.001$ ; HRG:  $t_{(19)} = 11.995$ ,  $p < 0.001$ ]. Finally, to compare MRG with HRG, a subsample of 20 participants from MRG (named as sMRG) were selected, with similar age, gender, and suppression scores as HRG (all  $p > 0.4$ ) (Table 2).

### Image Acquisition and Data Analysis

A 3.0-T GE Sigma scanner was used to collect resting-state fMRI images with a gradient echo planar imaging (EPI) sequence (TR, 2,000 ms; TE, 30 ms; FA,  $90^\circ$ ; FOV,  $240 \text{ mm} \times 240 \text{ mm}$ ; matrix size,  $64 \times 64$ ; voxel size,  $3.75 \text{ mm} \times 3.75 \text{ mm} \times 3 \text{ mm}$ ; slices, 43). The T1-weighted structural image was acquired with a high-resolution T1-weighted scan (TR, 5.96 ms; TE, 1.96 ms; FA,  $9^\circ$ ; FOV,  $256 \text{ mm} \times 256 \text{ mm}$ ; matrix size,  $256 \times 256$ ; voxel size,  $1 \text{ mm} \times 1 \text{ mm} \times 1 \text{ mm}$ ; number of slices, 176). Participants were instructed to rest with their eyes closed but not to fall asleep during the scan.

Resting-state fMRI data analysis was conducted using the data processing assistant for resting-state fMRI toolbox (DPARF<sup>1</sup>) and statistical parametric mapping software (SPM12<sup>2</sup>). To keep magnetic field stabilization, the first five EPI volumes of the fMRI images were removed. Preprocessing consisted of the following steps: Slice timing correction, 3D motion correction, nuisance covariates regression (Friston-24 motion parameters; white matter, cerebrospinal fluid, and global signals), spatial

<sup>1</sup><http://rfmri.org/DPARF>

<sup>2</sup><http://www.fil.ion.ucl.ac.uk/spm>

**TABLE 1 |** Demographic characteristics and behavioral assessment.

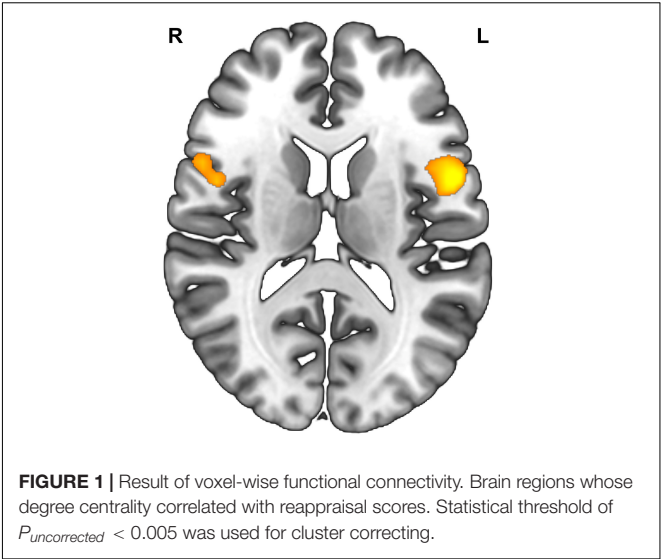
	MRG				HRG				
	<i>n</i>	Age	Reappraisal	Suppression	<i>n</i>	Age	Reappraisal	Suppression	
		M (SD)	M (SD)	M (SD)		M (SD)	M (SD)	M (SD)	M (SD)
Female	44	21.20 (2.11)	4.93 (0.35)	3.40 (0.91)	12	21.25 (1.82)	6.04 (0.44)	3.52 (0.89)	
Male	36	21.58 (2.09)	4.92 (0.33)	3.76 (0.94)	8	21.25 (2.05)	6.21 (0.54)	3.50 (1.16)	
<i>p</i>		0.425	0.891	0.088		1.000	0.460	0.964	

MRG, medium reappraisal group; HRG, high reappraisal group.

normalization to the Montreal Neurological Institute (MNI) template and resampling to 3 mm × 3 mm × 3 mm, removing the linear trends, temporal band-pass filtering (0.01–0.1 Hz), and spatial smoothing with a Gaussian kernel of full-width half-maximum 6 mm. The head motion exclusion was applied with translation not exceeding 3 mm and rotation not exceeding 3°, and mean FD\_Jenkinson not exceeding 0.2 (Power et al., 2012; van Dijk et al., 2012). According to this threshold, no participant was excluded.

Seed Definition

Degree centrality refers to the number of brain connections from a voxel to others across the whole brain (Zuo et al., 2012; Yan et al., 2017; Lv et al., 2018). The measure of degree centrality has been widely performed to examine node characteristics of intrinsic connectivity networks, especially for the identification of functional hubs in functional connectivity analysis (Zuo et al., 2012). In the present study, a voxel-wise functional connectivity analysis was performed using DPARSFA, and the examination of degree centrality was recruited to identify functional hubs that were related to individual reappraisal scores. The correlation threshold was set at  $r > 0.25$  for the degree centrality calculation. Then, the resulting degree centrality was used to conduct a multiple regression analysis, with gender, age and suppression scores controlled as covariates of noninterest and reappraisal scores as a predictor of interest. For the exploratory purpose, we lowered the statistical threshold to  $P_{uncorrected} < 0.005$  with a cluster size  $> 50$  voxels. At this threshold, only two clusters emerged, the left and right IFG, whose degree centrality showed a positive correlation with reappraisal scores (Table 3 and Figure 1). Therefore, based on the results of the above voxel-wise



functional connectivity analysis, left IFG (−48 9 6) and right IFG (54 12 18) were defined as seed regions for further seed-based rsFC analysis. The two seed regions were separately built as a 6-mm radius sphere centered around the peak activation using the Marsbar toolbox<sup>3</sup>.

Seed-Based Voxel-Wise rsFC Analysis

After seeds extraction, seed-based voxel-wise rsFC analyses were performed to explore brain regions connected with left/right IFG, and the connectivity correlating with individual reappraisal scores. Firstly, time series of all voxels located within these two seed regions were abstracted and averaged, respectively. Secondly, a Pearson correlation was conducted between each seed region’s time series and those of all other brain voxels of each participant. Thereafter, the resulting correlation coefficients were transformed into Fisher’s  $z$  scores, representing the rsFC for each connection of each participant. Subsequently, multiple regression models were performed with reappraisal scores as a predictor of interest, and the effect of gender, age, and suppression scores simultaneously eliminated. All activations were applied at the whole-brain level with a statistical significance of false discovery rate  $P_{FDR} < 0.05$  and a cluster extent  $> 50$  voxels. Besides, multiple regression models were also performed with suppression scores as a predictor of interest, suggested by Picó-Pérez et al. (2018) in their similar study on the association between dispositional use of emotional regulation strategies and rsFC, but with the amygdala as seed regions. However, we did not observe any significant activation when we used the suppression scores as a predictor of interest at the same threshold of  $P_{FDR} < 0.05$ . Consequently, the suppression strategy was not included in the result.

In addition, we further extracted the rsFC strength value of each region of interest (ROI), which amounted to a sphere of 6-mm radius centered around the peak of activation using Marsbar toolbox (see text footnote 3). Then, the partial

TABLE 2 | Comparison of behavioral assessment between HRG and sMRG.

	Gender	Age	Reappraisal	Suppression
	Female/Male	M (SD)	M (SD)	M (SD)
HRG	12/8	21.25 (1.86)	6.11 (0.48)	3.51 (0.98)
sMRG	12/8	21.70 (1.95)	4.90 (0.29)	3.43 (0.85)
$p$	1.000	0.460	$<0.001$	0.764

HRG, high reappraisal group; sMRG, subsample of medium reappraisal group.

TABLE 3 | Results of voxel-wise functional connectivity.

Region	H	K	T	MNI coordinates		
				x	y	z
Inferior frontal gyrus	L	106	4.33	−54	6	15
			3.85	−45	9	18
			3.27	−51	12	3
Inferior frontal gyrus	R	69	3.85	45	12	9
			3.46	48	3	15
			3.42	54	12	15

H, hemisphere; L, left; R, right; K, cluster size in number of activated voxels; T,  $T$  value; MNI, the Montreal Neurological Institute coordinates. Statistical threshold of  $P_{uncorrected} < 0.005$  was used for cluster correcting.

<sup>3</sup><http://marsbar.sourceforge.net>



correlation analyses were performed between connectivity strength and reappraisal scores after controlling for gender, age, and suppression scores.

Initially, we tried to investigate the potential association between IFG-based rsFC and reappraisal scores across the entire range observed in the sample of  $n = 100$ . Unfortunately, we did not find any significant activation at the threshold of  $P_{FDR} < 0.05$ . The previous literature, especially the theory of neural efficiency, may provide a reasonable conjecture that individuals with a higher level of frequency of dispositional reappraisal use may show less connectivity between IFG and the remaining regions. Hence, we considered that, probably, the high-frequency level of dispositional use of reappraisal affects the association between IFG-based rsFC and reappraisal scores. Therefore, as reported before, the 100 participants were allocated into two groups (MRG and HRG) according to the range of reappraisal scores.

### Prediction Analysis Using Cross-Validation

To test whether the observed functional brain features in MRG could reliably predict reappraisal scores of new individuals, internal cross-validation analyses were performed using the Pattern Recognition for Neuroimaging Toolbox (PRoNTo v2.1<sup>4</sup>). The input vectors were mean-centered using the training data (Xie et al., 2020), while the effect of covariates of noninterest (gender, age, and suppression scores) was regressed out. The predictive power was assessed by calculating Pearson's correlation coefficients between the predicted and actual reappraisal scores. Additionally, the statistical significance of the correlation was determined by 5,000 times of permutation testing without replacement.

In order to examine whether the reappraisal related IFG-based rsFC pattern in MRG is the same as that in HRG, internal cross-validation analyses were conducted using the reappraisal related rsFC of sMRG to predict the reappraisal scores of HRG.

## RESULTS

### Functional Connectivity Analysis

In MRG, with the left IFG as a seed region, reappraisal scores were positively correlated with rsFC between left IFG and most bilateral regions consisting of STG/MTG/ITG, superior parietal lobe (SPL), middle cingulate cortex (MCC), postcentral/precentral gyrus, rolandic operculum, cerebellum, and fusiform gyrus; and between left IFG and some left regions including inferior parietal lobe (IPL), supplementary motor area (SMA), precuneus, and occipital gyrus, as well as between left IFG and right insula (Table 4 and Figure 2A). However, there was no significant activation at the same threshold level in HRG.

Functional connectivity analysis with right IFG as a seed region in MRG showed that reappraisal scores were positively correlated with rsFC between right IFG and brain areas such as the bilateral medial SFG, bilateral STG/MTG/ITG, bilateral precuneus, bilateral postcentral, bilateral occipital gyrus, bilateral fusiform, left ACC, left SFG, left insula, and right supramarginal

gyrus (Table 4 and Figure 3A). Although in HRG, there were still no significant activation yielded at the same threshold level.

### Association of rsFC Strength With Reappraisal

When considering the left IFG as a seed region in MRG, the following ROIs showed significant correlation between strength values and reappraisal scores: left SPL ( $-21 -57 48$ ),  $r = 0.432$ ,  $p < 0.001$ ; right SPL ( $24 -60 51$ ),  $r = 0.388$ ,  $p < 0.001$ ; right insula ( $42 -12 6$ ),  $r = 0.431$ ,  $p < 0.001$ ; left STG ( $-57 -27 12$ ),  $r = 0.476$ ,  $p < 0.001$ ; right STG ( $54 -21 6$ ),  $r = 0.450$ ,  $p < 0.001$ ; right MCC ( $9 -15 42$ ),  $r = 0.378$ ,  $p = 0.001$  (Figure 2B).

On the other hand, when the right IFG is considered as a seed region in MRG, the following ROIs showed significant correlation between strength and reappraisal scores: DMPFC (rostral cluster) ( $-6 51 18$ ),  $r = 0.506$ ,  $p < 0.001$ ; DMPFC (caudal cluster) ( $-6 36 51$ ),  $r = 0.442$ ,  $p < 0.001$ ; left ACC ( $-6 48 12$ ),  $r = 0.442$ ,  $p < 0.001$ ; left MTG ( $-48 -57 6$ ),  $r = 0.395$ ,  $p < 0.001$ ; right MTG ( $42 -60 3$ ),  $r = 0.428$ ,  $p < 0.001$ ; right STG ( $57 -42 21$ ),  $r = 0.378$ ,  $p = 0.001$  (Figure 3B).

### Prediction Analysis

In MRG, the left IFG seed-region-related rsFC could effectively predict for individual reappraisal scores ( $r = 0.370$ ,  $p = 0.002$ ) (Figure 4A), and this was also true for the right IFG seed region ( $r = 0.330$ ,  $p = 0.004$ ) (Figure 4B). However, there was no significant predictive power for individual reappraisal scores of HRG with left IFG or right IFG seed-region-related rsFC in sMRG (left IFG:  $r = -0.380$ ,  $p = 0.636$ , Figure 5A; right IFG:  $r = -0.590$ ,  $p = 0.856$ , Figure 5B).

## DISCUSSION

Previous studies have emphasized the key functional role of IFG in collaboration with other regions in emotional regulation task-based functional connectivity, while the present study constitutes the first investigation into the associations between rsFC patterns of bilateral IFG and the dispositional use of reappraisal. Interestingly, we observed that medium dispositional use of reappraisal was positively related to IFG-based rsFC. Specifically, the medium habitual use of reappraisal was associated with a significant positive coupling between: (1) bilateral IFG and temporal gyrus; (2) left IFG and bilateral SPL, MCC, left IPL, and right insula; and (3) right IFG and DMPFC/ACC. However, no significant correlation emerged between left or right IFG-related rsFC and the high dispositional reappraisal use. The predictive analyses also showed that both left and right IFG-related rsFC could effectively and separately predict for individual reappraisal scores in MRG. However, IFG-related rsFC of sMRG had no significant predictive power for reappraisal scores of HRG.

In line with our hypotheses, the associations of both left and right IFG with temporal gyrus were linked to the habitual use of reappraisal. Particularly, using the reappraisal strategy altered the emotional value of stimuli by manipulating the conceptual knowledge and creating opposite interpretations. This suggests that the involvement of the semantic process is a core part

<sup>4</sup>www.ml.nl.cs.ucl.ac.uk/pronto/

**TABLE 4 |** Functional connectivity results.

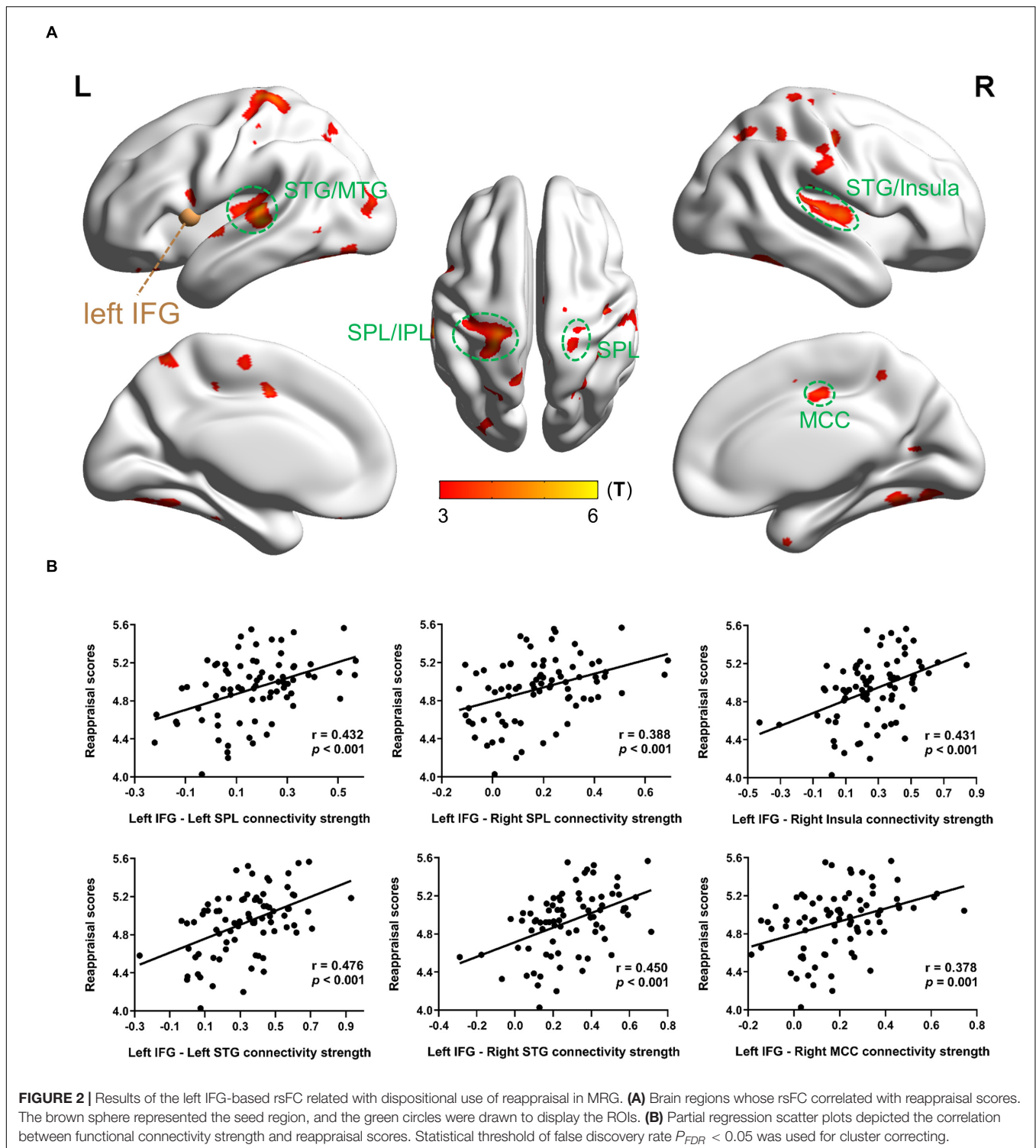
Region	H	K	T	MNI coordinates		
				x	y	z
Left IFG as a seed						
Superior temporal gyrus/middle temporal gyrus/rolandic operculum	L	320	5.50	−57	−30	3
			5.26	−57	−27	12
Postcentral gyrus/precentral gyrus/superior parietal lobe/inferior parietal lobe/middle cingulate cortex/supplementary motor area/precuneus	L	949	5.09	−21	−36	66
			4.64	−21	−51	51
Cerebellum/fusiform/inferior temporal gyrus	R	314	5.01	24	−54	−18
			4.59	39	−54	−15
			4.46	48	−48	−21
Superior temporal gyrus/middle temporal gyrus/rolandic operculum/insula	R	562	4.81	72	−30	0
			4.57	39	−12	6
			4.49	63	3	6
Postcentral gyrus/precentral gyrus/superior parietal lobe	R	320	4.61	21	−36	66
Cerebellum/fusiform/inferior temporal gyrus	L	385	4.68	−42	−57	−24
			4.42	−39	−66	−18
Middle occipital gyrus	L	128	4.65	−30	−78	12
Right IFG as a seed						
Superior medial frontal gyrus/anterior cingulate cortex	L	294	5.45	−6	51	18
Cerebellum/fusiform/superior occipital gyrus/hippocampus/middle temporal gyrus/inferior temporal gyrus	L	1560	5.00	−36	−33	−27
			4.71	−30	−9	−21
Superior temporal gyrus/middle temporal gyrus	R	136	4.19	57	−42	21
Superior temporal gyrus/middle temporal gyrus	L	299	4.43	−45	−45	12
			3.88	−48	−57	6
Precuneus	L	111	4.40	−9	−54	45
Superior medial frontal gyrus/superior frontal gyrus	L	158	4.35	−6	36	51
			3.87	3	39	48
Superior occipital gyrus	R	152	4.09	15	−90	21
Middle temporal gyrus/insula	L	114	4.11	−60	−15	−9
			4.07	−45	−9	3
Postcentral gyrus/Precuneus	R	101	3.92	27	−39	51
			3.47	12	−54	45
Supramarginal gyrus/postcentral gyrus	L	144	3.66	−54	−27	30
			3.61	−54	−12	24

H, hemisphere; L, left; R, right; K, cluster size in number of activated voxels; T, T value; MNI, the Montreal Neurological Institute coordinates. Statistical threshold of false discovery rate  $P_{FDR} < 0.05$  was used for cluster correcting.

of emotion regulation. Evidently, the semantic system plays a critical role in the storage and controlled retrieval of conceptual knowledge (Binder et al., 2009), contributing to representations of relevant emotional information from emotional experiences (Neumann and Lozo, 2012). Specifically, the temporal gyrus, which has often been observed with enhanced activation in reappraisal task (Goldin et al., 2008; Kanske et al., 2011; Dörfel et al., 2014), is usually considered as part of semantic system (Binder et al., 2009) and plays a role in both the storage and the strategic retrieval of semantic knowledge (Davey et al., 2016). Recent evidence on functional connectivity between IFG and MTG at both task (Zhang et al., 2019) and resting-state contexts (Kohn et al., 2014; Davey et al., 2016) indicates that the cooperation between IFG and MTG makes strategic access of semantic information possible. Similarly, the present study also revealed strong functional connectivity between IFG and

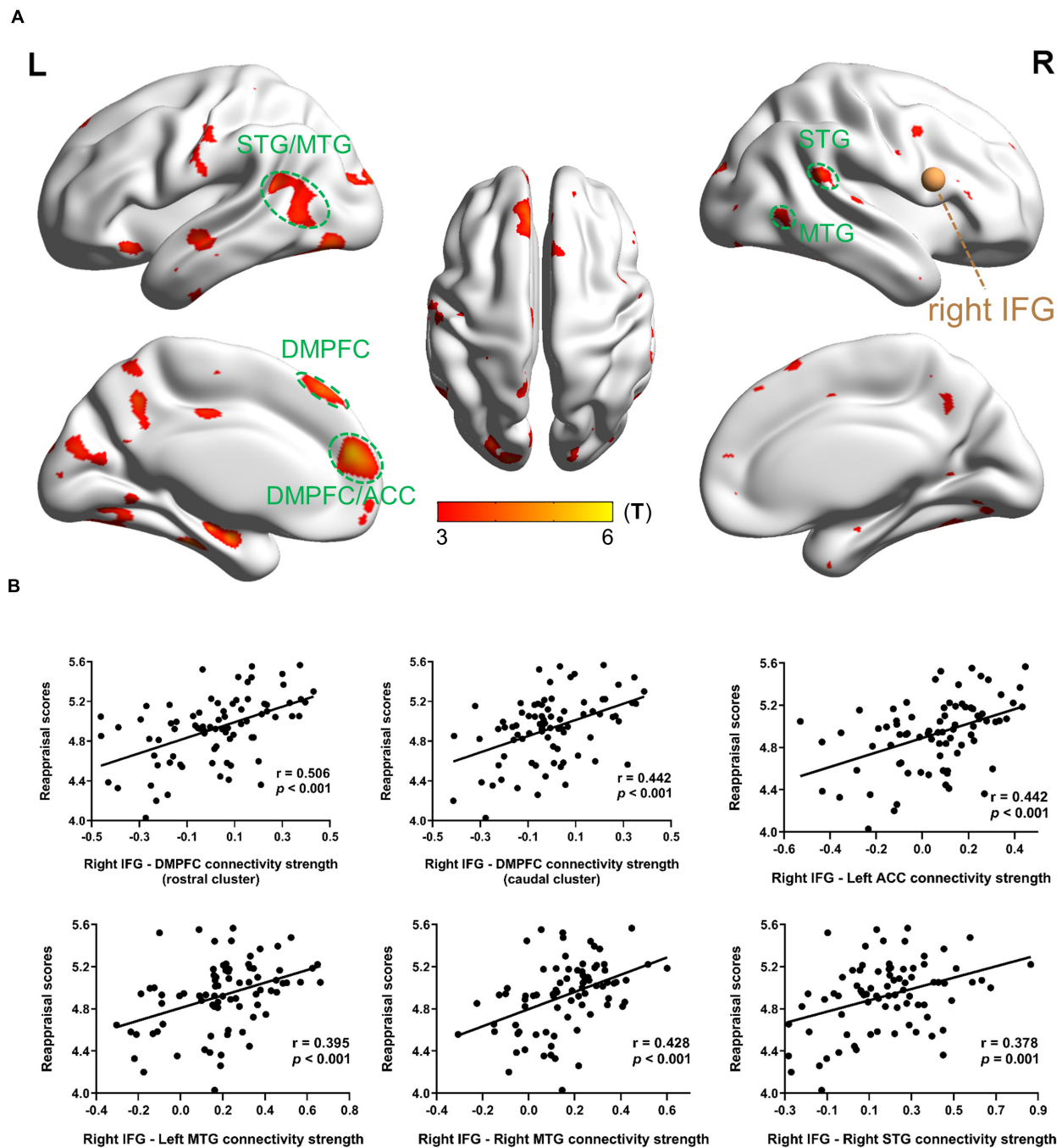
temporal gyrus, alongside a positive correlation with medium habitual use of reappraisal, perhaps supporting the capacity to potentially engage and sustain semantic retrieval, in line with goal-driven control of subjective emotional feelings.

Inconsistent with our preliminary expectation, only the right IFG displayed an association with DMPFC. The DMPFC has generally been proved to be involved in semantic and self-reflective processes (Olsson and Ochsner, 2008; Binder et al., 2009), and has repeatedly been observed to be significantly activated in the reappraisal tasks (Kanske et al., 2011; Buhle et al., 2014; Morawetz et al., 2017a, 2016). In particular, DMPFC is implicated with the elaboration of the affective meaning of stimuli and representation of value information concerning mental states (Ochsner et al., 2012; Dixon et al., 2017). Thus, the right IFG, with a close connection with DMPFC, may facilitate the evaluation of the changing mental states, in relation to outcomes



of appropriate or inappropriate interpretations of emotional stimuli. Besides, with a correlation with medium habitual reappraisal use, right IFG-based rsFC also showed a strong link with ACC. Recent evidence suggests that ACC constitutes a core part of the neural circuitry of valuation (Amemori and Graybiel, 2012; Bartra et al., 2013; Clithero and Rangel, 2013),

playing an important role in evaluating interoceptive signals based on self-referential and conceptual emotion knowledge (Dixon et al., 2017). The evaluation role of ACC may thus facilitate a better understanding of subjective emotional feelings, by assigning conceptual meaning to these bodily sensations. Notably, through interaction with ACC, DMPFC contributes to



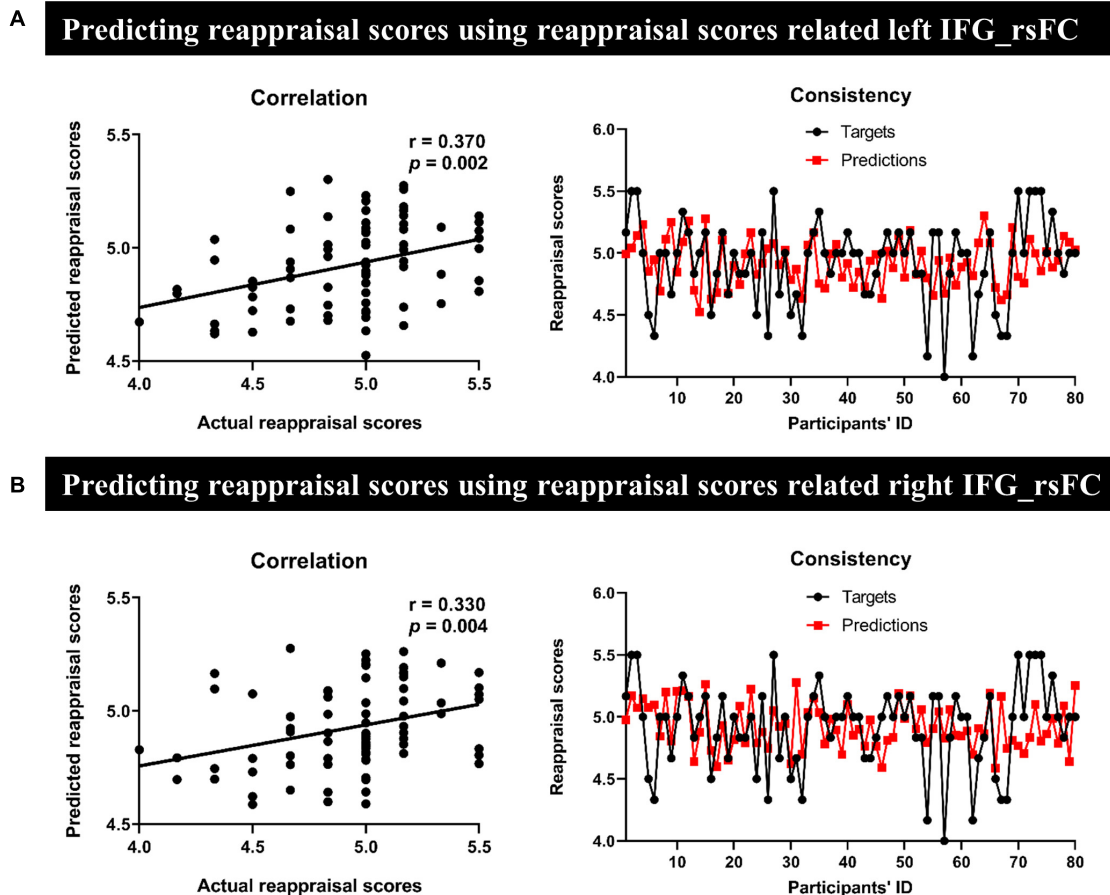
**FIGURE 3 |** Results of the right IFG-based rsFC related with dispositional use of reappraisal in MRG. **(A)** Brain regions whose rsFC correlated with reappraisal scores. The brown sphere represented the seed region, and the green circles were drawn to display the ROIs. **(B)** Partial regression scatter plots depicted the correlation between functional connectivity strength and reappraisal scores. Statistical threshold of false discovery rate  $P_{FDR} < 0.05$  was used for cluster correcting.

the maintenance of mental representations of an individual's feelings active in affective working memory (Lane et al., 2015) and may subsequently transfer these internal state information to IFG *via* a feed-forward mechanism (Phillips et al., 2008). Therefore, it is highly plausible that the right IFG (roles in selecting appropriate or inhibiting inappropriate interpretations

from semantic memory) strongly connects with DMPFC and ACC (roles in perceiving and evaluating subjective emotional feelings), exhibiting correspondingly more frequent reappraisal skill, to achieve goal-directed emotional states.

Furthermore, as indicated above, we also observed a positive correlation between medium dispositional use of reappraisal



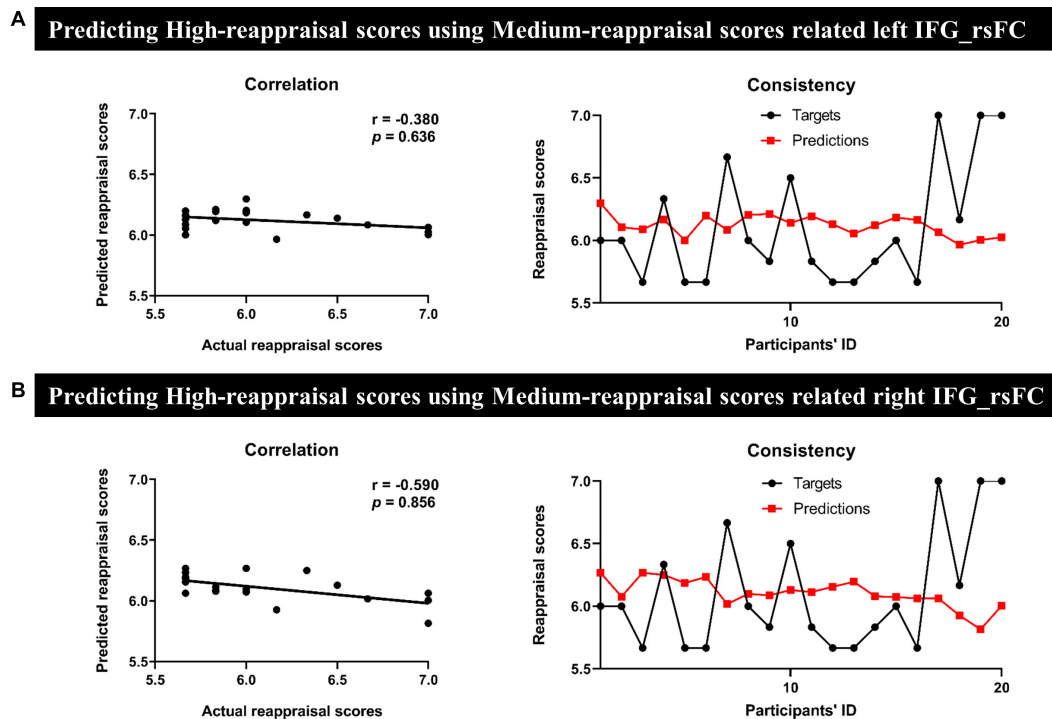


**FIGURE 4 |** Results of prediction in MRG. Severally using (A) the left and (B) the right IFG-based rsFC to predict reappraisal scores. The scatter plots and line charts [in both panels (A,B)] described a significant correlation and consistency between actual and predicted reappraisal scores, respectively.

and connectivity of left IFG with bilateral SPL, MCC, left IPL, and right insula during resting state. Consistent with the previous findings on the activation of parietal lobes in reappraisal task (Buhle et al., 2014; Morawetz et al., 2017b), our study also found an association between habitual use of reappraisal and co-activation of left IFG and parietal lobes (including SPL and IPL), which usually engage in the attention control process (Luks et al., 2007; Hutchinson et al., 2014; Salo et al., 2017; Sang et al., 2017; Wang et al., 2020). More importantly, it is documented that the parietal lobes and prefrontal gyrus engage in cognitive control by exerting influence on the temporal regions to change the semantic and perceptual representations, so as to facilitate the selection of appropriate behaviors, and the inhibition of maladaptive habitual actions (Buhle et al., 2014; Dixon, 2015). Similarly, MCC has also been found to be strongly involved in the allocation of attention to emotional information and action monitoring (McRae et al., 2008; Kragel et al., 2018). Indeed, the cognitive function of MCC in performance monitoring may help guide the changing emotional responses through reappraisal strategy in an intended way (Ochsner et al., 2012). Moreover, MCC also combines with the insula, which contains bodily information and sensations

(including interoceptive representation of emotions) (Craig, 2009; Lane et al., 2015), to project the affective information to the neighboring IFG/VLPFC (Craig, 2009; Kohn et al., 2014), indicating a motivation to IFG to select an appropriate response in the final stage. During this process, IFG may work in concert with the parietal lobes and temporal gyrus to focus attention on the subjective feelings and select appropriate interpretation to obtain a desirable emotional state.

As initially anticipated, we did not observe a significant correlation between left or right IFG-related rsFC and the high dispositional reappraisal use. More so, the IFG-related rsFC pattern of MRG could not effectively predict for reappraisal scores of HRG. This supports the neural efficiency view, that the more adept the skill, the lesser the neural connectivity, but the more enhanced the performance becomes (Neubauer and Fink, 2009; Di Domenico et al., 2015; Curtin et al., 2019). However, there is a conspicuous lack of evidence on trait emotion regulation with reappraisal disposition to support our findings. Nevertheless, some promising evidence from cognitive training provides a potential explanation. For example, Vartanian et al. (2013) proved that working memory training can augment performance on divergent thinking task



**FIGURE 5 |** Results of prediction between HRG and sMRG. We separately used (A) the left and (B) the right IFG-based rsFC of sMRG to predict the reappraisal scores of HRG. Scatter plots and line charts indicated non-significant correlation and consistency between actual and predicted reappraisal scores, respectively.

and lead to lower activation in the prefrontal gyrus (VLPFC and DLPFC) in adult participants, while Motes et al. (2018) found that cognitive training results in a faster processing speed along with reduced activation in the prefrontal gyrus in elderly participants. Therefore, it is possible that individuals with high-frequency daily use of reappraisal strategy, possess a more adept skill of emotion regulation, and recruit less functional connectivity between IFG and other regions in the resting state. On the other hand, the limited sample size in HRG may also be a potential reason to explain the non-significant results of both the association between rsFC and reappraisal scores, and the correlation in the internal cross-validation analysis. If so, future study should bring into consideration the sample size factor to ascertain whether the IFG-based rsFC pattern is linked to the variation of frequency level of dispositional use of reappraisal.

Overall, previous studies have emphasized the important role of the IFG in cooperation with other brain regions in task-based functional connectivity, in the selection and inhibition processes of emotional regulation. The present study expands on these findings by explicitly investigating how patterns of functional connectivity between IFG and other brain regions change during resting state, and how these changes may be linked to individuals' habitual reappraisal use. Specifically, besides coupling with temporal gyrus in the function of general semantic control, we found that the left IFG, along with SPL, MCC, left IPL and right insula, predominantly engages in monitoring the emotional performance (cognitive control of emotion), while the right IFG,

coupling with DMPFC and ACC, predominantly engages in representation of mental states (evaluation of emotion).

Despite the important contributions of our study, there are several limitations needed to be noted. Firstly, as mentioned above, the sample size of HRG is relatively smaller than that of the MRG. Possibly, there may be some other potential intrinsic functional connectivity patterns recruited by individuals with higher emotion regulation capacity. To fully understand the neural substrates of emotion regulation, more participants with a higher ability of emotion regulation need to be included in future studies. Secondly, the present study only recruited healthy participants. Although the neural association observed in these participants may provide potential neural evidence for clinical practice, the dispositional reappraisal use may show a different association with the resting-state networks in the context of emotional dysregulation. Therefore, future studies should consider a comparative analysis of the relevance of habitual strategy use and regulation networks, between patients with emotional disorders and healthy populations, to improve the clinical understanding and intervention. Finally, although we have not observed a significant association between IFG-based rsFC and suppression scores, which is also a personality trait of emotion regulation, it may exist in other potential functional hubs and neural substrates relevant to the dispositional use of the suppression strategy. Thus, future studies should explore the potential neural mechanism underlying trait suppression in both healthy and clinical populations.

In conclusion, the present study investigated the intrinsic neural underpinnings of dispositional reappraisal employing the IFG-based function connectivity approach during resting state. Our findings demonstrate that the medium dispositional reappraisal use relies on the cooperation of the functional hubs of the bilateral IFG and other regions within the emotion regulation cortex. These findings may explain how individuals cope with emotional events in daily life, as well as applied in clinical intervention for emotion-regulation-related disorders.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding authors.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Local Committee for the Protection of Human

Subjects of the University of Electronic Science and Technology of China. Written informed consent to participate in this study was provided by the participants.

## AUTHOR CONTRIBUTIONS

WL, JZ, ZJ, and LL conceived and designed the experiments. WL performed the experiments. WL and KX analyzed the data. WL wrote the main manuscript text. RN revised the manuscript. All authors reviewed the manuscript.

## FUNDING

This work was supported by grants from NSFC (61773092, 61673087, and 61773096), by the Sichuan Province Science and Technology Support Program (2020YFS0230), by the Higher Education Discipline Innovation Project (111 project, grant number B12027), and by the Fundamental Research Funds for the Central Universities.

## REFERENCES

- Amemori, K. I., and Graybiel, A. M. (2012). Localized microstimulation of primate pregenual cingulate cortex induces negative decision-making. *Nat. Neurosci.* 15, 776–785. doi: 10.1038/nn.3088
- Bartra, O., McGuire, J. T., and Kable, J. W. (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* 76, 412–427. doi: 10.1016/j.neuroimage.2013.02.063
- Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796. doi: 10.1093/cercor/bhp055
- Braunstein, L. M., Gross, J. J., and Ochsner, K. N. (2017). Explicit and implicit emotion regulation: a multi-level framework. *Soc. Cogn. Affect. Neurosci.* 12, 1545–1557. doi: 10.1093/scan/nsx096
- Buhle, J. T., Silvers, J. A., Wage, T. D., Lopez, R., Onyemkwu, C., Kober, H., et al. (2014). Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. *Cereb. Cortex* 24, 2981–2990. doi: 10.1093/cercor/bht154
- Cliethero, J. A., and Rangel, A. (2013). Informatic parcellation of the network involved in the computation of subjective value. *Soc. Cogn. Affect. Neurosci.* 9, 1289–1302. doi: 10.1093/scan/nst106
- Craig, A. D. (2009). How do you feel - now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi: 10.1038/nrn2555
- Curtin, A., Ayaz, H., Tang, Y., Sun, J., Wang, J., and Tong, S. (2019). Enhancing neural efficiency of cognitive processing speed via training and neurostimulation: an fNIRS and TMS study. *Neuroimage* 198, 73–82. doi: 10.1016/j.neuroimage.2019.05.020
- Davey, J., Thompson, H. E., Hallam, G., Karapanagiotidis, T., Murphy, C., De Caso, I., et al. (2016). Exploring the role of the posterior middle temporal gyrus in semantic cognition: integration of anterior temporal lobe with executive processes. *Neuroimage* 137, 165–177. doi: 10.1016/j.neuroimage.2016.05.051
- Di Domenico, S. I., Rodrigo, A. H., Ayaz, H., Fournier, M. A., and Ruocco, A. C. (2015). Decision-making conflict and the neural efficiency hypothesis of intelligence: a functional near-infrared spectroscopy investigation. *Neuroimage* 109, 307–317. doi: 10.1016/j.neuroimage.2015.01.039
- Dixon, M. L. (2015). Cognitive control, emotional value, and the lateral prefrontal cortex. *Front. Psychol.* 6:758. doi: 10.3389/fpsyg.2015.00758
- Dixon, M. L., Thiruchselvam, R., Todd, R., and Christoff, K. (2017). Emotion and the prefrontal cortex: an integrative review. *Psychol. Bull.* 143, 1033–1081. doi: 10.1037/bul0000096
- Dörfel, D., Lamke, J. P., Hummel, F., Wagner, U., Erk, S., and Walter, H. (2014). Common and differential neural networks of emotion regulation by detachment, reinterpretation, distraction, and expressive suppression: a comparative fMRI investigation. *Neuroimage* 101, 298–309. doi: 10.1016/j.neuroimage.2014.06.051
- Gabard-Durnam, L. J., Gee, D. G., Goff, B., Flannery, J., Telzer, E., Humphreys, K. L., et al. (2016). Stimulus-elicited connectivity influences resting-state connectivity years later in human development: a prospective study. *J. Neurosci.* 36, 4771–4784. doi: 10.1523/JNEUROSCI.0598-16.2016
- Goldin, P. R., McRae, K., Ramel, W., and Gross, J. J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biol. Psychiatry* 63, 577–586. doi: 10.1016/j.biopsych.2007.05.031
- Greccucci, A., Giorgetta, C., Bonini, N., and Sanfey, A. G. (2013). Reappraising social emotions: the role of inferior frontal gyrus, temporo-parietal junction and insula in interpersonal emotion regulation. *Front. Hum. Neurosci.* 7:523. doi: 10.3389/fnhum.2013.00523
- Gross, J. J., and John, O. P. (2003). Individual differences in two emotion regulation processes: implications for affect, relationships, and well-being. *J. Pers. Soc. Psychol.* 85, 348–362. doi: 10.1037/0022-3514.85.2.348
- Guerra-Carrillo, B., MacKey, A. P., and Bunge, S. A. (2014). Resting-state fMRI: a window into human brain plasticity. *Neuroscientist* 20, 522–533. doi: 10.1177/1073858414524442
- Hartwigsen, G., Neef, N. E., Camilleri, J. A., Margulies, D. S., and Eickhoff, S. B. (2019). Functional segregation of the right inferior frontal gyrus: evidence from coactivation-based parcellation. *Cereb. Cortex* 29, 1532–1546. doi: 10.1093/cercor/bhy049
- Hu, T., Zhang, D., Wang, J., Mistry, R., Ran, G., and Wang, X. (2014). Relation between emotion regulation and mental health: a meta-analysis review. *Psychol. Rep.* 114, 341–362. doi: 10.2466/03.20.PR0.114k22w4
- Hutchinson, J. B., Uncapher, M. R., Weiner, K. S., Bressler, D. W., Silver, M. A., Preston, A. R., et al. (2014). Functional heterogeneity in posterior parietal cortex across attention and episodic memory retrieval. *Cereb. Cortex* 24, 49–66. doi: 10.1093/cercor/bhs278
- Kalisch, R. (2009). The functional neuroanatomy of reappraisal: time matters. *Neurosci. Biobehav. Rev.* 33, 1215–1226. doi: 10.1016/j.neubiorev.2009.06.003
- Kanske, P., Heissler, J., Schönfelder, S., Bongers, A., and Wessa, M. (2011). How to regulate emotion? Neural networks for reappraisal and distraction. *Cereb. Cortex* 21, 1379–1388. doi: 10.1093/cercor/bhq216
- Kohn, N., Eickhoff, S. B., Scheller, M., Laird, A. R., Fox, P. T., and Habel, U. (2014). Neural network of cognitive emotion regulation – an ALE meta-analysis

- and MACM analysis. *Neuroimage* 87, 345–355. doi: 10.1016/j.neuroimage.2013.11.001
- Kragel, P. A., Kano, M., Van Oudenhove, L., Ly, H. G., Dupont, P., Rubio, A., et al. (2018). Generalizable representations of pain, cognitive control, and negative emotion in medial frontal cortex. *Nat. Neurosci.* 21, 283–289. doi: 10.1038/s41593-017-0051-7
- Lane, R. D., Weihs, K. L., Herring, A., Hishaw, A., and Smith, R. (2015). Affective agnosia: expansion of the alexithymia construct and a new opportunity to integrate and extend Freud's legacy. *Neurosci. Biobehav. Rev.* 55, 594–611. doi: 10.1016/j.neubiorev.2015.06.007
- Luks, T. L., Simpson, G. V., Dale, C. L., and Hough, M. G. (2007). Preparatory allocation of attention and adjustments in conflict processing. *Neuroimage* 35, 949–958. doi: 10.1016/j.neuroimage.2006.11.041
- Lv, H., Wang, Z., Tong, E., Williams, L. M., Zaharchuk, G., Zeineh, M., et al. (2018). Resting-state functional MRI: everything that nonexperts have always wanted to know. *Am. J. Neuroradiol.* 39, 1390–1399. doi: 10.3174/ajnr.A5527
- Mackey, A. P., Singley, A. T. M., and Bunge, S. A. (2013). Intensive reasoning training alters patterns of brain connectivity at rest. *J. Neurosci.* 33, 4796–4803. doi: 10.1523/JNEUROSCI.4141-12.2013
- McRae, K., Reiman, E. M., Fort, C. L., Chen, K., and Lane, R. D. (2008). Association between trait emotional awareness and dorsal anterior cingulate activity during emotion is arousal-dependent. *Neuroimage* 41, 648–655. doi: 10.1016/j.neuroimage.2008.02.030
- Messina, I., Bianco, S., Sambin, M., and Viviani, R. (2015). Executive and semantic processes in reappraisal of negative stimuli: insights from a meta-analysis of neuroimaging studies. *Front. Psychol.* 6:956. doi: 10.3389/fpsyg.2015.00956
- Meyer, T., Smeets, T., Giesbrecht, T., and Merckelbach, H. (2012). The efficiency of reappraisal and expressive suppression in regulating everyday affective experiences. *Psychiatry Res.* 200, 964–969. doi: 10.1016/j.psychres.2012.05.034
- Morawetz, C., Bode, S., Baudewig, J., and Heekeren, H. R. (2017a). Effective amygdala-prefrontal connectivity predicts individual differences in successful emotion regulation. *Soc. Cogn. Affect. Neurosci.* 12, 569–585. doi: 10.1093/scan/nsw169
- Morawetz, C., Bode, S., Baudewig, J., Kirilina, E., and Heekeren, H. R. (2016). Changes in effective connectivity between dorsal and ventral prefrontal regions moderate emotion regulation. *Cereb. Cortex* 26, 1923–1937. doi: 10.1093/cercor/bhv005
- Morawetz, C., Bode, S., Derntl, B., and Heekeren, H. R. (2017b). The effect of strategies, goals and stimulus material on the neural mechanisms of emotion regulation: a meta-analysis of fMRI studies. *Neurosci. Biobehav. Rev.* 72, 111–128. doi: 10.1016/j.neubiorev.2016.11.014
- Motes, M. A., Yezhuvath, U. S., Aslan, S., Spence, J. S., Rypma, B., and Chapman, S. B. (2018). Higher-order cognitive training effects on processing speed-related neural activity: a randomized trial. *Neurobiol. Aging* 62, 72–81. doi: 10.1016/j.neurobiolaging.2017.10.003
- Neubauer, A. C., and Fink, A. (2009). Intelligence and neural efficiency: measures of brain activation versus measures of functional connectivity in the brain. *Intelligence* 37, 223–229. doi: 10.1016/j.intell.2008.10.008
- Neumann, R., and Lozo, L. (2012). Priming the activation of fear and disgust: evidence for semantic processing. *Emotion* 12, 223–228. doi: 10.1037/a0026500
- Ochsner, K. N., Silvers, J. A., and Buhle, J. T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Ann. N. Y. Acad. Sci.* 1251, E1–E24. doi: 10.1111/j.1749-6632.2012.06751.x
- Olsson, A., and Ochsner, K. N. (2008). The role of social cognition in emotion. *Trends Cogn. Sci.* 12, 65–71. doi: 10.1016/j.tics.2007.11.010
- Phillips, M. L., Ladouceur, C. D., and Drevets, W. C. (2008). A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Mol. Psychiatry* 13, 833–857. doi: 10.1038/mp.2008.65
- Picó-Pérez, M., Alonso, P., Contreras-Rodríguez, O., Martínez-Zalacáín, I., López-Solà, C., Jiménez-Murcia, S., et al. (2018). Dispositional use of emotion regulation strategies and resting-state cortico-limbic functional connectivity. *Brain Imaging Behav.* 12, 1022–1031. doi: 10.1007/s11682-017-9762-3
- Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E. (2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018
- Salo, E., Salmela, V., Salmi, J., Numminen, J., and Alho, K. (2017). Brain activity associated with selective attention, divided attention and distraction. *Brain Res.* 1664, 25–36. doi: 10.1016/j.brainres.2017.03.021
- Sang, N., Zhang, L., Hao, L., Wang, Y., Wang, X., Zhang, F., et al. (2017). Human sensory cortex structure and top-down controlling brain network determine individual differences in perceptual alternations. *Neurosci. Lett.* 636, 113–119. doi: 10.1016/j.neulet.2016.10.048
- Schulz, K. P., Clerkin, S. M., Halperin, J. M., Newcorn, J. H., Tang, C. Y., and Fan, J. (2009). Dissociable neural effects of stimulus valence and preceding context during the inhibition of responses to emotional faces. *Hum. Brain Mapp.* 30, 2821–2833. doi: 10.1002/hbm.20706
- Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., et al. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13040–13045. doi: 10.1073/pnas.0905267106
- Uchida, M., Biederman, J., Gabrieli, J. D. E., Micco, J., De Los Angeles, C., Brown, A., et al. (2014). Emotion regulation ability varies in relation to intrinsic functional brain architecture. *Soc. Cogn. Affect. Neurosci.* 10, 1738–1748. doi: 10.1093/scan/nsv059
- van Dijk, K. R. A., Sabuncu, M. R., and Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* 59, 431–438. doi: 10.1016/j.neuroimage.2011.07.044
- Vartanian, O., Jobidon, M. E., Bouak, F., Nakashima, A., Smith, I., Lam, Q., et al. (2013). Working memory training is associated with lower prefrontal cortex activation in a divergent thinking task. *Neuroscience* 236, 186–194. doi: 10.1016/j.neuroscience.2012.12.060
- Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A., and Ochsner, K. N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron* 59, 1037–1050. doi: 10.1016/j.neuron.2008.09.006
- Wang, M., Yu, B., Luo, C., Fogelson, N., Zhang, J., Jin, Z., et al. (2020). Evaluating the causal contribution of fronto-parietal cortices to the control of the bottom-up and top-down visual attention using fMRI-guided TMS. *Cortex* 126, 200–212. doi: 10.1016/j.cortex.2020.01.005
- Webb, T. L., Miles, E., and Sheeran, P. (2012). Dealing with feeling: a meta-analysis of the effectiveness of strategies derived from the process model of emotion regulation. *Psychol. Bull.* 138, 775–808. doi: 10.1037/a0027600
- Xie, K., Jin, Z., Ni, X., Zhang, J., and Li, L. (2020). Distinct neural substrates underlying target facilitation and distractor suppression: a combined voxel-based morphometry and resting-state functional connectivity study. *Neuroimage* 221:117149. doi: 10.1016/j.neuroimage.2020.117149
- Yan, C. G., Yang, Z., Colcombe, S. J., Zuo, X. N., and Milham, M. P. (2017). Concordance among indices of intrinsic brain function: insights from inter-individual variation and temporal dynamics. *Sci. Bull.* 62, 1572–1584. doi: 10.1016/j.scib.2017.09.015
- Zaehring, J., Jennen-Steinmetz, C., Schmahl, C., Ende, G., and Paret, C. (2020). Psychophysiological effects of downregulating negative emotions: insights from a meta-analysis of healthy adults. *Front. Psychol.* 11:470. doi: 10.3389/fpsyg.2020.00470
- Zhang, Q., Wang, H., Luo, C., Zhang, J., Jin, Z., and Li, L. (2019). The neural basis of semantic cognition in Mandarin Chinese: a combined fMRI and TMS study. *Hum. Brain Mapp.* 40, 5412–5423. doi: 10.1002/hbm.24781
- Zuo, X. N., Ehmke, R., Mennes, M., Imperati, D., Castellanos, F. X., Sporns, O., et al. (2012). Network centrality in the human functional connectome. *Cereb. Cortex* 22, 1862–1875. doi: 10.1093/cercor/bhr269

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Li, Xie, Ngetich, Zhang, Jin and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.





# Multimodal Neuroimaging Predictors of Learning Performance of Sensorimotor Rhythm Up-Regulation Neurofeedback

Linling Li<sup>1,2,3†</sup>, Yinxue Wang<sup>1,2,3†</sup>, Yixuan Zeng<sup>4</sup>, Shaohui Hou<sup>1,2,3</sup>, Gan Huang<sup>1,2,3</sup>, Li Zhang<sup>1,2,3</sup>, Nan Yan<sup>5</sup>, Lijie Ren<sup>4\*</sup> and Zhiguo Zhang<sup>1,2,3,6\*</sup>

<sup>1</sup> School of Biomedical Engineering, Health Science Center, Shenzhen University, Shenzhen, China, <sup>2</sup> Guangdong Provincial Key Laboratory of Biomedical Measurements and Ultrasound Imaging, Shenzhen, China, <sup>3</sup> Marshall Laboratory of Biomedical Engineering, Shenzhen University, Shenzhen, China, <sup>4</sup> Department of Neurology, Shenzhen Second People's Hospital, The First Affiliated Hospital of Shenzhen University, Shenzhen, China, <sup>5</sup> CAS Key Laboratory of Human-Machine Intelligence-Synergy Systems, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China, <sup>6</sup> Peng Cheng Laboratory, Shenzhen, China

## OPEN ACCESS

### Edited by:

Zehong Jimmy Cao,  
University of Tasmania, Australia

### Reviewed by:

Minghao Dong,  
Xidian University, China  
Feng Wan,  
University of Macau, China

### \*Correspondence:

Zhiguo Zhang  
zgzhang@szu.edu.cn  
Lijie Ren  
13631605966@126.com

<sup>†</sup> These authors have contributed  
equally to this work

### Specialty section:

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

**Received:** 25 April 2021

**Accepted:** 25 June 2021

**Published:** 20 July 2021

### Citation:

Li L, Wang Y, Zeng Y, Hou S,  
Huang G, Zhang L, Yan N, Ren L and  
Zhang Z (2021) Multimodal  
Neuroimaging Predictors of Learning  
Performance of Sensorimotor Rhythm  
Up-Regulation Neurofeedback.  
Front. Neurosci. 15:699999.  
doi: 10.3389/fnins.2021.699999

Electroencephalographic (EEG) neurofeedback (NFB) is a popular neuromodulation method to help one selectively enhance or inhibit his/her brain activities by means of real-time visual or auditory feedback of EEG signals. Sensory motor rhythm (SMR) NFB protocol has been applied to improve cognitive performance, but a large proportion of participants failed to self-regulate their brain activities and could not benefit from NFB training. Therefore, it is important to identify the neural predictors of SMR up-regulation NFB training performance for a better understanding the mechanisms of individual difference in SMR NFB. Twenty-seven healthy participants (12 males, age:  $23.1 \pm 2.36$ ) were enrolled to complete three sessions of SMR up-regulation NFB training and collection of multimodal neuroimaging data [resting-state EEG, structural magnetic resonance imaging (MRI), and resting-state functional MRI (fMRI)]. Correlation analyses were performed between within-session NFB learning index and anatomical and functional brain features extracted from multimodal neuroimaging data, in order to identify the neuroanatomical and neurophysiological predictors for NFB learning performance. Lastly, machine learning models were trained to predict NFB learning performance using features from each modality as well as multimodal features. According to our results, most participants were able to successfully increase the SMR power and the NFB learning performance was significantly correlated with a set of neuroimaging features, including resting-state EEG powers, gray/white matter volumes from MRI, regional and functional connectivity (FC) of resting-state fMRI. Importantly, results of prediction analysis indicate that NFB learning index can be better predicted using multimodal features compared with features of single modality. In conclusion, this study highlights the importance of multimodal neuroimaging technique as a tool to explain the individual difference in within-session NFB learning performance, and could provide a theoretical framework for early identification of individuals who cannot benefit from NFB training.

**Keywords:** neurofeedback (NFB), multimodal neuroimaging, sensorimotor rhythm (SMR), learning, functional connectivity (FC)

## INTRODUCTION

Electroencephalographic (EEG) neurofeedback (NFB) training is a popular neuromodulation method to train brain functions. Through EEG NFB, one can learn to selectively enhance or inhibit his/her brain activities by means of real-time visual or auditory feedback of EEG signals (Sitaram et al., 2017). Since its first attempts in the 1960s, EEG NFB has rapidly received much attention because of its numerous potential applications for healthy participants and patients (Gruzelier, 2014a,b,c; Sitaram et al., 2017; Omejc et al., 2019). Among the diversity of NFB training protocols, one popular protocol is to up-regulate the amplitude of the sensory motor rhythm (SMR, 12–15 Hz), which is associated with a mental state of “relaxed alertness” (Witte et al., 2013). SMR NFB protocol has been applied to improve cognitive performance, such as sustained attention and visuo-motor skills of healthy participants (Gruzelier, 2014a; Kober et al., 2020). In clinical situation, SMR NFB training could improve the impaired cognitive functions in post-stroke patients (Kober et al., 2015). However, the efficacy of EEG NFB has been questioned recently because many sham-controlled studies have shown that a large proportion of users (16% to 57%) failed to self-regulate their brain activities and could not benefit from EEG NFB training (Gruzelier, 2014c; Thibault and Raz, 2016; Alkoby et al., 2018; Weber et al., 2020; Zhang et al., 2020a). It is of great importance to investigate the neural mechanisms of the huge individual differences in EEG NFB learning performance, because the successful learning during EEG NFB training can directly contribute to the improvement of disease symptoms in clinical patients (Gruzelier, 2014c). To identify in advance those participants who are likely to benefit from EEG NFB is a crucial step toward individualized and precise neuromodulation. Therefore, it is highly desirable to discover predictors of EEG NFB learning success and to establish a machine learning model to predict the learning performance based on identified predictors.

Some recent reviews concerned with the inefficiency problem of EEG NFB training have summarized various types of predictors of training performance (Alkoby et al., 2018; Weber et al., 2020). However, according to these reviews, only a very limited number of related studies have been carried out on and their findings are not convergent and complete. Standardized questionnaires or behavioral tasks are naturally considered as candidate predictors, but existing evidence showed that these questionnaires or tasks can predict NFB learning performance to a limited extent (Kleih et al., 2010; Nan et al., 2012; Witte et al., 2013; Alkoby et al., 2018). A batch of studies were focused on neurophysiological signals, mainly EEG recorded prior to training, for the prediction of performance during NFB training. In a resting-state EEG study, Nan et al. found that the amplitude of low beta (12–15 Hz) EEG rhythm measured before training could predict the NFB learning ability of low beta (15–18 Hz)/theta (4–7 Hz) ratio training (Nan et al., 2015). This research group later proposed that, eye-closed resting-state EEG activities in broad frequency bands, including lower alpha and theta, measured before training could distinguish learners/non-learners of alpha down-regulating NFB (Nan et al., 2018).

Similarly, resting-state SMR power before training was related to the NFB training target at SMR activities (Reichert et al., 2015).

Besides above EEG predictors of NFB learning performance, magnetic resonance imaging (MRI) has also been more and more popularly used to investigate the problem of NFB inefficacy, because multimodal MRI can provide various types of predictors from brain structure, function, and connectivity with high spatial resolution. For example, structural MRI (sMRI) studies found that, the gray/white matter volume (GMV/WMV) could be the predictors of NFB performance and they were related to the neuroanatomical basis of the ability to learn to self-regulate one's own brain activity (Weber et al., 2020). Resting-state functional MRI (rsfMRI) have also been applied to study neural mechanisms of individual differences in self-regulation of many other behaviors and brain functions (Kelley et al., 2015). Resting-state functional MRI also offers many metrics of local brain activity, such as regional homogeneity (ReHo) (Zang et al., 2004) and the amplitude of low-frequency fluctuations (ALFF) or fractional ALFF (fALFF) (Zou et al., 2008). All of this metrics can be used to investigated correlation between baseline brain activities and cognitive performance (Dong et al., 2015; Xiang et al., 2016; Xie et al., 2021). Particularly, because the interactions between brain regions are crucial for supporting cognitive functions, recent studies suggested that functional connectivity (FC) might be more promising for predicting complex high-order cognitive processes than those measured based on local brain regions (Qi et al., 2019; Horien et al., 2020). An fMRI-based NFB study suggested that rsfMRI FC can be used to identify individuals who are likely to benefit from fMRI NFB training to control anxiety symptom (Scheinost et al., 2014). However, rsfMRI (no matter which type of metrics) is still seldom used to investigate the inefficiency problem of SMR NFB.

In summary, the predictors of SMR up-regulation in NFB training has not been well understood yet. Multimodal brain imaging techniques could provide complementary and/or novel information about brain function, structure, and connectivity, so it is promising that more predictors could be discovered from multimodal brain imaging data and the accuracy to predict NFB learning performance can be improved. Thus, in the present study, we collected multimodal neuroimaging data (resting-state EEG, sMRI, and rsfMRI) before SMR-NFB experiments, with aims to identify the multimodal neuroanatomical and neurophysiological predictors for within-session learning performance of SMR up-regulation. The work could obtain an early identification of individuals who would not benefit from NFB training, and could increase the efficacy and cost-effectiveness of NFB technique in practical uses.

## MATERIALS AND METHODS

### Participants

This study was approved by the Medical Ethics Committee of the Health Science Center of Shenzhen University. Thirty-four healthy participants (17 males, mean  $\pm$  SD age:  $22.8 \pm 2.23$ ) were recruited from Shenzhen University. The participants had no history of major medical illness, psychiatric or

neurological disorder and had normal corrected-to-normal vision. All participants gave their written informed consent and the experimental procedure was approved by the local ethics committee.

## Data Acquisition

Electroencephalographic (EEG) signal during the NFB training was recorded with a 32-bit OpenBCI Board (Cyton Biosensing) connected to a lithium-ion polymer rechargeable battery. The OpenBCI board was connected to Ag/AgCl wet electrodes secured within an elastomeric EEG head-cap (Easy-Cap; Brain Products GmbH, Munich, Germany). Electrodes were placed at C1, C2, Cz, CPz, TP9, and TP10 positions according to the 10–20 electrode montage system. The reference electrode was located at FCz, and the ground electrode was located at the forehead.

Resting-state whole-brain EEG measurements (3-min eyes-open and eyes-closed) were also collected for each participant in a separate session before the first session of NFB training. EEG recordings were obtained with 64 Ag/AgCl electrodes placed on the EasyCap (Brain Products GmbH, Munich, Germany) with a reference electrode positioned at FCz. Vertical electrooculogram was recorded with the electrode placed on the superior to the nasion of the right eye. Input impedances were kept below 10 k $\Omega$  and the records were taken simultaneously at a sampling frequency of 1000 Hz.

Structural and resting-state functional MRI Images were collected using a 3.0 Tesla Siemens Trio scanner (Siemens Medical, Erlangen, Germany) at Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. A standard 12-channel birdcage head coil was used and each participant's head was fixed by foam pads in order to reduce head movement. Functional images were acquired with echo planar imaging sequence with the following parameters: 31 contiguous slices with a slice thickness of 4 mm; TR/TE = 2000/30 ms, 90° flip angle; field of view (FOV) 192  $\times$  192 mm<sup>2</sup>; 64  $\times$  64 data matrix. During the scanning of resting-state fMRI data, participants were asked to remain motionless, keep their eyes open, stay awake, and stare at the “+” sign. High resolution T1-weighted images were collected with a volumetric three-dimensional spoiled gradient recall sequence with the parameters: TR/TE = 2000/30 ms, FOV = 240  $\times$  240 mm<sup>2</sup>, matrix size = 256  $\times$  256, flip angle = 90°, slice number = 176, voxel size = 0.9  $\times$  0.9  $\times$  1 mm<sup>3</sup>.

## NFB Training

The SMR (12–15 Hz) up-regulating EEG NFB protocol was employed in this study, and each participant had to complete three sessions within one week and usually one day apart for continuous sessions. Each session consisted of 10 NFB training runs (3 min each). The first run is the baseline run during which the participants saw moving feedback bar but were instructed to relax themselves without trying to control the feedback bar voluntarily. The EEG signal recorded in the NFB baseline run was used to calculate initial individual threshold for SMR feedback and the threshold of SMR power was adapted after each run using the median value of SMR power in previous run.

During NFB training, the EEG signal was sampled at 250 Hz and band-pass filtered between 0.5 Hz and 45 Hz. The SMR

power was calculated by fast Fourier-transforms (FFT) every 100 ms with a 10-s data window. The real-time SMR power (12–15 Hz) was presented on the screen in front of the participants as one vertical feedback bar. We only gave the participants a minimum degree of guidance, telling them to relax and concentrate physically during NFB training to increase the SMR power. When the SMR power exceeded the threshold, the color of the feedback bar changed from red to blue. Furthermore, when participants were able to move SMR feedback bars above the threshold and keeping for more than 1 s, they were rewarded with one more point displayed on the top of the screen.

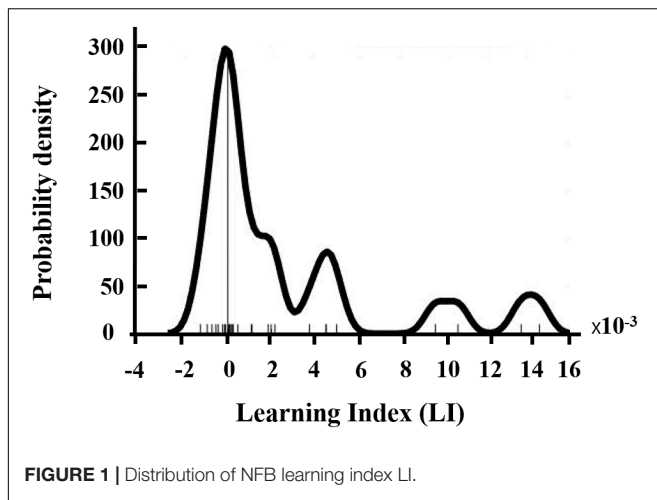
The within-session NFB learning performance was assessed using a learning index (LI), which is calculated from the SMR powers of all NFB training runs and has been successfully applied in previous studies (Wan et al., 2014). Specifically, for each session, the median values of SMR power were calculated for all 12 training runs and then a linear regression was performed on these median values. Then LI was calculated as the average of the regression slopes across three sessions. The learning performance (LI) was checked for normality using Shapiro–Wilk test. It should be noted that the NFB learning index LI was calculated using SMR power values obtained from online EEG recorded by the OpenBCI cap during NFB training, because these on-line EEG results could better reflect the training performance.

## Analysis of Resting-State EEG

Resting-state EEG data recorded by the 64-channel EasyCap were analyzed by the Letswave toolbox<sup>1</sup> and self-written MATLAB scripts. Continuous data were bandpass filtered between 1 Hz and 100 Hz. After visual inspection, bad channels were interpolated with adjacent channels. Then the signal was split into 1-s epochs and epochs with artefacts were rejected after visual inspection. These remaining epochs were submitted to an informax algorithm to decompose into their independent components (Makeig et al., 1997; Jung et al., 2001; Olbrich et al., 2011). The components related to eye blinking or movement were removed from the original data. Finally, all EEG signals were re-referenced to a common reference.

After preprocessing, data were transformed into the frequency domain by the Welch's method (with 1000-point FFT, 50% overlapping, and Hamming windows). Then we extracted the power values at four frequency bands: theta (4–8 Hz), alpha (8–12 Hz), SMR (12–15 Hz) and beta (12.5–30 Hz). Because EEG showed great inter-individual variabilities in the power, relative SMR power was calculated as the ratio of the power within a specific frequency band and the power and the total power in the full frequency band. For each NFB run, the relative power values were calculated for each epoch and then standardized across all epochs. Finally, the median value of relative powers of all epochs was taken as the relative power for the entire run. Furthermore, to assess the topological distribution of relative powers, nine scalp regions of interest (ROI) were defined (as shown in **Figure 1**) and they are: left frontal (F5, F3, FC5, FC3), middle frontal (F1, Fz, F2, FC1, FCz, FC2), right frontal (F4, F6, FC4, FC6), left central (C5, C3, CP5, CP3), middle central (C1, Cz, C2, CP1, CPz, CP2),

<sup>1</sup><https://github.com/NOCIONS/letswave6>



**FIGURE 1 |** Distribution of NFB learning index LI.

right central (C4, C6, CP4, CP6), left parietal (P5, P3, PO7, PO3), middle parietal (P1, Pz, P2, POz), right parietal (P4, P6, PO4, PO8) (Reichert et al., 2015). The relative power of each ROI was calculated as the average across all the electrodes within this ROI.

Because the learning index (LI) of NFB training is not normally distributed ( $P = 0.0002$ ), Spearman's correlation analysis was performed to evaluate the correlation between resting-state EEG power and the learning performance (LI) of NFB training for each ROI at each frequency band. False discovery rate (FDR) correction was performed to address the multiple comparison problem (Benjamini and Hochberg, 1995).

## Analysis of Structural MRI Data

Voxel based morphometry analysis of sMRI images were performed using Computational Anatomy Toolbox<sup>2</sup>, which is an extension toolbox of Statistical Parametric Mapping (SPM12<sup>3</sup>). The sMRI images were segmented into gray matter, white matter and cerebrospinal fluid areas using the unified standard segmentation option in SPM12. The individual structural images were then normalized into standard Montreal Neurological Institute (MNI) template. Spatial normalization into the MNI standard space was done by the high-dimensional DARTEL (Diffeomorphic Anatomical Registration Through Exponentiated Lie Algebra) approach. Then the generated gray matter and white matter images were smoothed with an 8-mm full-width at half-maximum (FWHM) Gaussian kernel.

The voxel-wise statistical analyses were performed using statistical non-parametric mapping (SnPM) toolbox<sup>4</sup>. The non-parametric permutation approach was applied because it did not require any assumption on data normality (Nichols and Holmes, 2002). The standard general linear model (GLM) design setup was used by creating design matrices for multiple regression analysis of GMV or WMV, with individual NFB performance (LI), age and gender as regressors. General linear model was used to construct pseudo t-statistic images, which were then assessed

for significance using a standard non-parametric multiple comparisons procedure based on randomization/permutation testing ( $N = 10000$ ). Significant clusters were extracted with voxel-level  $P > 0.005$  and cluster size  $> 50$ .

## Analysis of Resting-State fMRI

The rsfMRI data were processed by using the Data Processing Assistant for Resting State fMRI (DPARSF<sup>5</sup>), which is based on SPM12 and the toolbox for Data Processing & Analysis of Brain Imaging (DPABI<sup>6</sup>). The preprocessing procedure was as follows. The first 10 volumes were removed to avoid T1 equilibration effects. Slice timing used the middle slice as the reference. Then the time series of rsfMRI images for each participant were realigned using a six-parameter (rigid body) linear transformation. The T1 images were co-registered to the mean functional image and then segmented into gray matter, white matter and cerebrospinal fluid. The Friston 24-parameter model was used to remove the linear trend and other nuisance signals (including Friston's 24 motion parameters, cerebrospinal fluid, white matter). The rsfMRI data were then normalized to the MNI space and re-sampled to 3-mm isotropic voxels. A 6 mm FWHM Gaussian kernel were applied to smooth the rsfMRI data. Finally, a bandpass filter (0.01–0.1 Hz) was then performed on the rsfMRI data.

Three rsfMRI regional features, including ALFF, fALFF, and ReHo, were calculated. Amplitude of low-frequency fluctuations is the mean of amplitudes within a specific frequency domain of a voxel's time course, and fALFF represents the relative contribution of specific oscillations to the whole detectable frequency range. ReHo is a rank-based Kendall's coefficient of concordance (KCC) which shows the synchronization among a given voxel and its nearest neighbors (26 voxels) time courses. ALFF/fALFF was calculated using smoothed (unfiltered) rsfMRI timeseries, and ReHo was calculated using unsmoothed time series. The metric maps for ALFF/fALFF and ReHo were z-standardized (subtracting the mean value for the entire brain from each voxel, and dividing by the corresponding standard deviation). Normalized ReHo maps were then smoothed using a 6-mm FWHM.

The relationship between the learning index LI and three regional features were explored using SnPM toolbox. Standard GLM design setup was used by creating design matrices for multiple regression analysis of gray matter intensity, with individual NFB performance (learning index LI), age and gender as regressors. General linear model was used to construct pseudo t-statistic images, which were then assessed for significance using a standard non-parametric multiple comparisons procedure based on randomization/permutation testing ( $N = 10000$ ). Significant clusters were extracted with voxel level  $P > 0.005$  and cluster size  $> 50$ .

Besides three regional rsfMRI features mentioned above, we also calculated the whole-brain FC using 164 ROIs (160 ROIs from the Dosenbach atlas and four emotional related ROIs) (Dosenbach et al., 2010). The 164 ROIs were assigned into

<sup>2</sup><http://www.neuro.uni-jena.de/cat/>

<sup>3</sup><https://www.fil.ion.ucl.ac.uk/spm/software/spm12/>

<sup>4</sup><http://warwick.ac.uk/snpm>

<sup>5</sup><http://rfmri.org/DPARSF>

<sup>6</sup><http://rfmri.org/dpabi>



7 intrinsic connectivity networks: (1) cerebellar, (2) cingulo-opercular, (3) default mode, (4) frontoparietal, (5) occipital, (6) sensorimotor, and (7) emotional. Each of the 164 ROIs was defined as a sphere with an 8 mm radius and the mean time series from all of the voxels within the ROI was extracted from preprocessed rsfMRI timeseries. We then calculated Pearson's correlations between all pairs of ROIs for each participant to generate a  $164 \times 164$  correlation matrix. The obtained correlation matrix for each participant was normalized using Fisher's z-transformation.

Subsequently, we calculated the correlation between the learning index LI and rsfMRI FC strength between each ROI pairs, using the Spearman partial correlation analysis with age and gender as nuisance regressors. Significant resting-state FC were extracted with  $P < 0.005$ .

### Prediction of NFB Learning Performance

Lastly, machine learning models were trained to predict NFB learning performance (LI) from multimodal neuroimaging predictors (correlates), which were identified based on above mentioned analyses of multimodal neuroimaging data (Sections "Analysis of resting-state EEG," "Analysis of structural MRI data," and "Analysis of resting-state fMRI"). According to previous correlation analysis results between the NFB learning index and features of each imaging modality, four sets of features were employed in the prediction analysis: (1) band-limited relative power of resting-state EEG, (2) GMV/WMV of sMRI, (3) three regional features (ALFF, fALFF, ReHo) of rsfMRI, (4) FC of rsfMRI. To investigate whether the combination of multimodal features can improve the performance in predication of NFB learning index, we further compare the prediction performance using features from each modality as well as multimodal features. The set of multimodal features were constructed by concatenating all four types of feature vectors. Across-individual normalization was performed for each feature as well as the learning index LI before training the machine learning models.

Four machine learning techniques, namely, linear support vector regression (SVR), linear regression (LR), Bayesian automatic relevance determination regression (ARDR), and random forest regression (RFR), were used here to quantitatively predict the NFB learning index from multimodal neuroimaging features (Pereira et al., 2009). Note that, because there is only one EEG correlate (SMR power of the left central region) found to be significantly correlated with LI (see Section "EEG correlates of NFB learning performance"), a linear regression model was used instead of SVR to predict LI from this EEG predictor. A leave-one-out-cross-validation (LOOCV) strategy was adopted to evaluate the performance of the prediction models. For each iteration in LOOCV, one participant was selected as the test sample and fed to the linear regression model (for EEG) or the SVR model (for MRI and combined multimodal features) trained with remaining samples, and the iterations were repeated for each participant. To quantify the prediction performance, mean absolute error (MAE), which was calculated as the average of the absolute difference between actual and predicted values, as well as Pearson correlation between actual and predicted values were used. Furthermore, we compared the Pearson correlation

coefficient obtained using multimodal features and the Pearson correlation coefficient obtained using each set of features from single modality. The calculation and comparison of Pearson correlations in this part were carried out using SPSS (SPSS Statistics, version 22, IBM, Armonk, NY).

## RESULTS

### NFB Learning Performance

Seven participants were excluded from analysis due to incomplete training or excessive EEG artifacts. Therefore, data from 27 participants (12 males, age:  $23.1 \pm 2.36$ ) were available for subsequent analysis. Most of the participants (19 of 27) were able to increase their SMR power within-session as suggested by their positive NFB learning index LI (as shown in **Figure 1**).

### EEG correlates of NFB Learning Performance

We performed correlation analysis between the NFB learning index LI and EEG powers of nine scalp ROIs. As shown in **Figure 2**, LI had significant positive correlations with SMR power of left central ROI (including EEG channels C5, C3, CP5, and CP3) during the eyes-open resting-state conditions ( $P < 0.05$ , FDR-corrected).

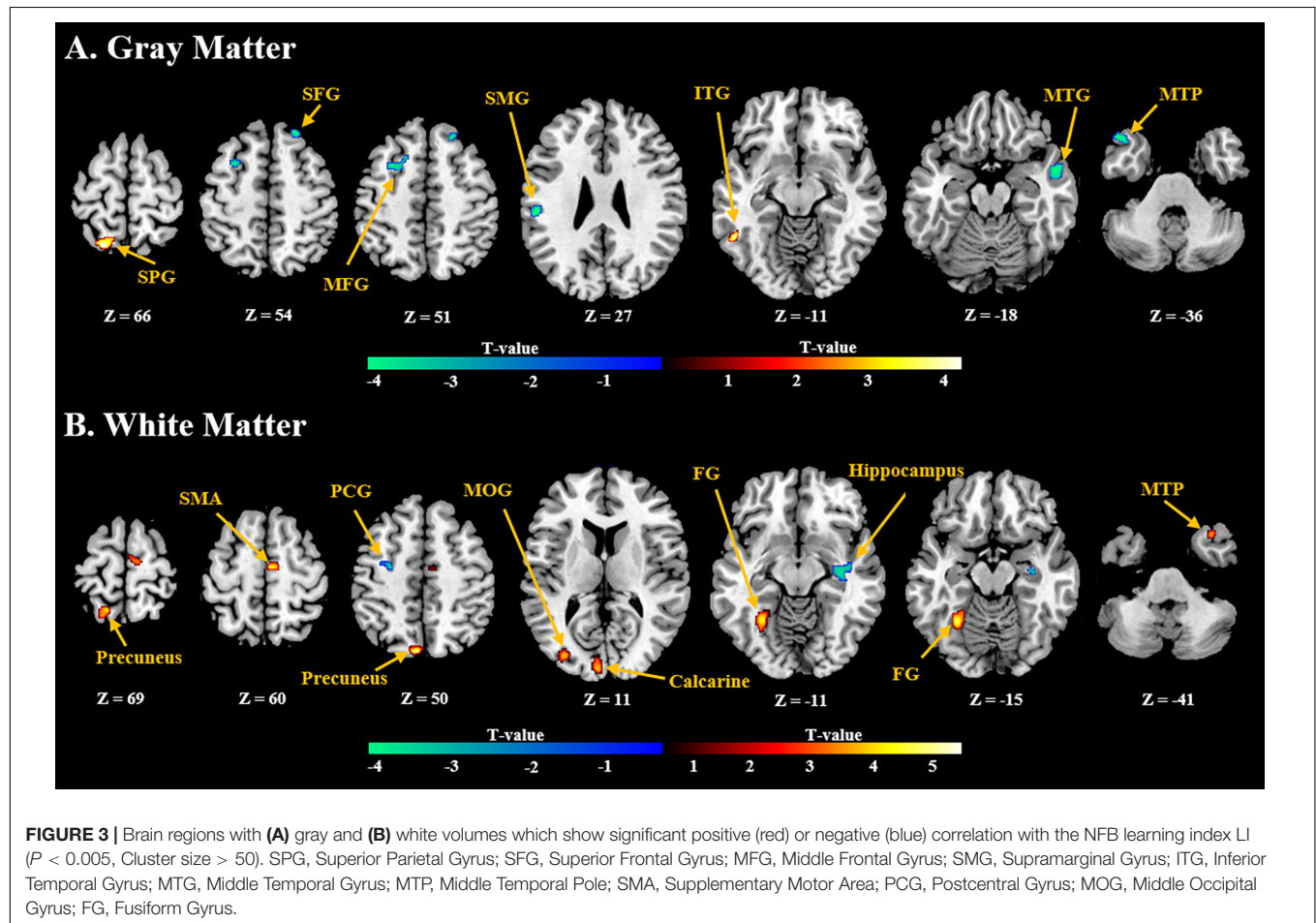
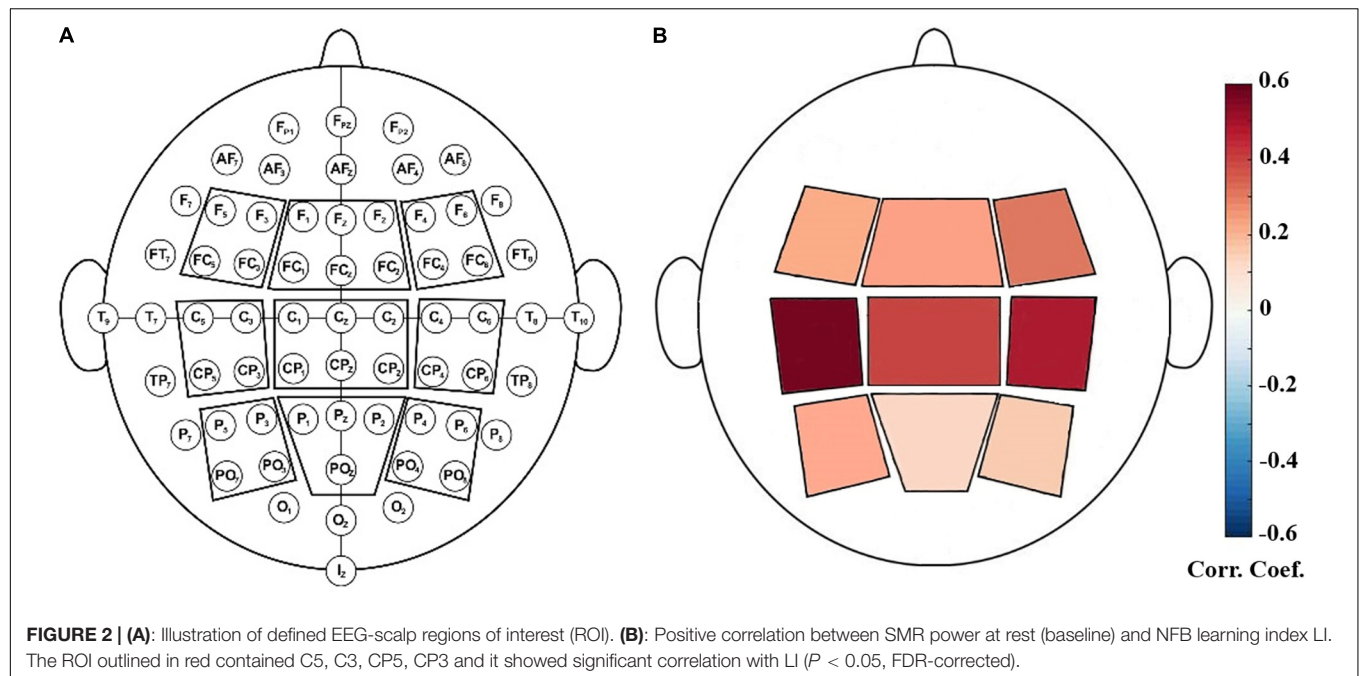
### sMRI Correlates of NFB Learning Performance

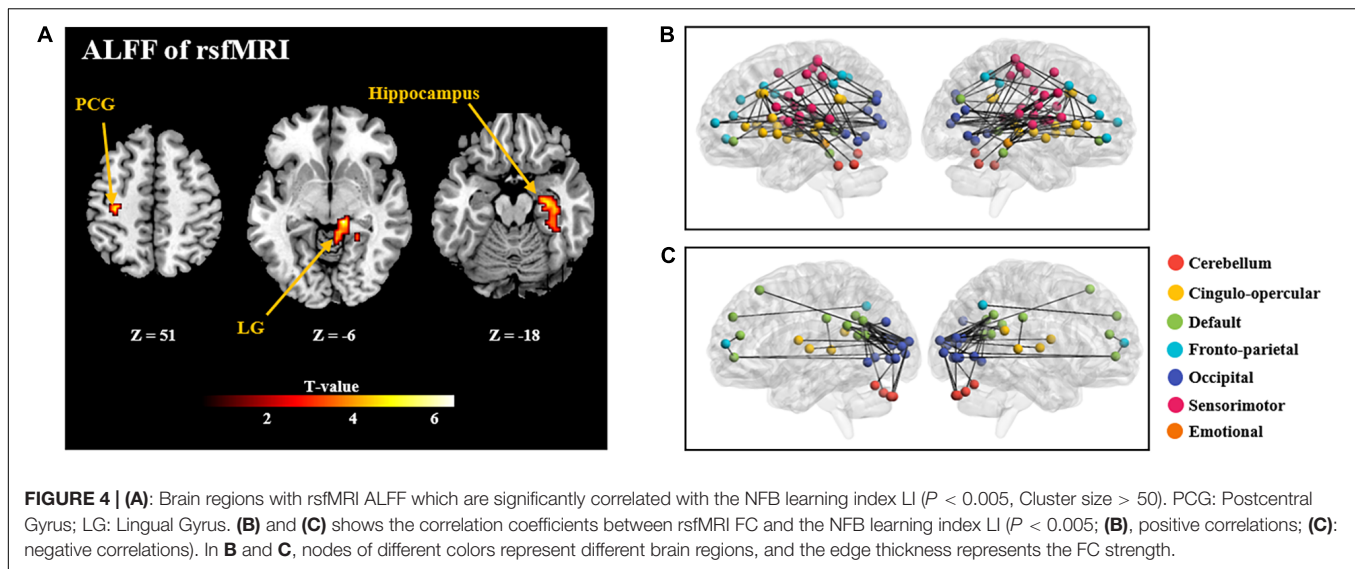
General linear model (GLM) analysis revealed that the NFB learning index LI was positively associated with GMV localized in the inferior temporal gyrus and superior parietal gyrus, and negatively associated with GMV localized in the frontal gyrus, middle temporal gyrus, and supramarginal gyrus (**Figure 3A**). For WMV, positive correlation was observed mainly in the supplementary motor area, precuneus, and medial occipitotemporal gyrus, and negative correlation was observed in the precentral gyrus and hippocampus (**Figure 3B**). Detailed information of these clusters with significant correlation can be found in **Supplementary Table 1**.

### rsfMRI Correlates of NFB Learning Performance

GLM analysis found positive correlation between ALFF and NFB learning index LI in the postcentral gyrus, lingual gyrus, and hippocampus (as shown in **Figure 4** and **Supplementary Table 2**). However, other two types of regional rsfMRI features (fALFF and ReHo) did not show any significant correlation results with LI. As a result, only the identified ALFF features were used to build a model for prediction of the NFB learning index.

Further, significant positive correlations between strength of rsfMRI FC and NFB learning index LI were mainly observed within the cingulo-opercular network, between the sensorimotor network and occipital network ( $P < 0.005$ ; **Figure 4B**); and significant negative correlation was observed between the default mode network and occipital network ( $P < 0.005$ ; **Figure 4C**).





## Prediction of NFB Learning Performance

Prediction analysis of the NFB learning index LI was performed using features selected from each modality, as well as from multimodal features. According to the results using SVR model, the NFB learning index LI can be better predicted using multimodal neuroimaging features, as compared with using features of single modality (as shown in **Figures 5**). Specifically, the MAE of multimodal features is 0.4341, which is smaller than those of single modality (0.6136 for EEG features, 0.6565 for sMRI features, 0.6297 for rsfMRI ALFF features, 0.7994 for rsfMRI FC features) (**Figures 5A–D**). Moreover, the correlation coefficient between actual and predicted values using multimodal features are significantly higher than the correlation coefficients derived using features of single modality (**Figure 5F**). Prediction of NFB learning index LI using other models can be found in **Supplementary Table 3** and **Supplementary Figure 1**, and generally better results could be obtained using multimodal neuroimaging features.

## DISCUSSION

Investigating the neural correlates of an individual's learning ability during NFB training is important since the learning ability has a crucial mediation link with the behavioral or clinical outcome after NFB training (Gruzelier, 2014a). Prediction of NFB learning performance would help prevent unnecessary time and resources used on participants who cannot learn to modulate their brain rhythms, and could make these participants choose other treatment means earlier. However, reliable predictors of NFB learning performance remain elusive. Thus, this study was aimed to investigate the association between learning performance during the SMR up-regulating NFB-training and multimodal neuroimaging data (resting-state EEG, sMRI, and rsfMRI), and then to assess whether NFB learning performance can be better predicted using multimodal features. According to our results, most participants were able to successfully increase

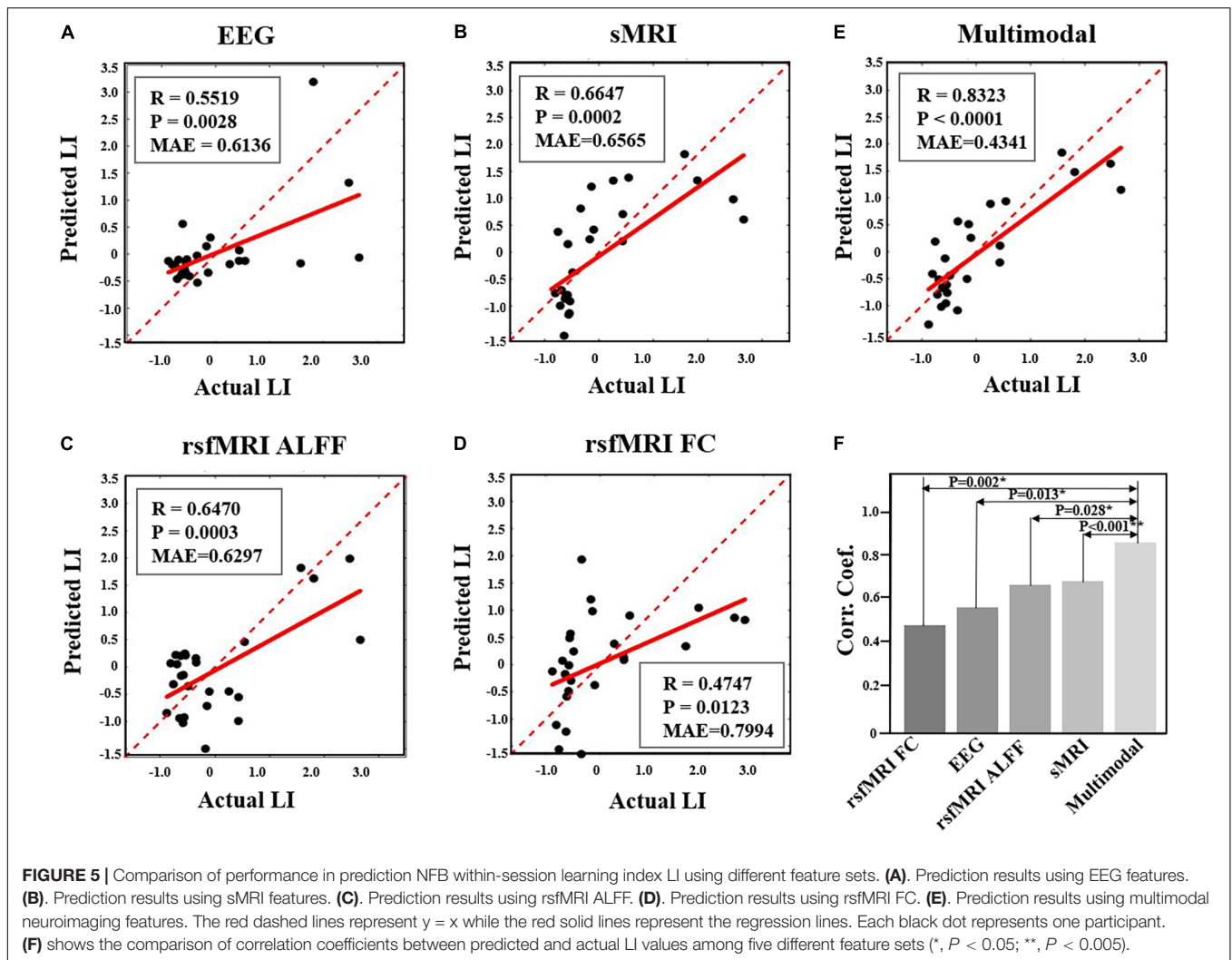
the SMR power and the learning performance was significantly correlated with a set of EEG or MRI features. Importantly, results of prediction analysis indicate that NFB learning index can be better predicted using multimodal features compared with features of single modality. These results will be discussed in detail below.

## Advantages of Using Multimodal Data

Neurofeedback (NFB) training is generally based on real-time feedback of voluntarily induced changes of brain activities, and it is a process of operant conditioning which leads to self-regulation of brain activity. Successful self-regulation of brain activity during NFB training can be considered as a personal skill, and a relatively high proportion of participants cannot achieve stable self-regulation of the target brain activity (Alkoby et al., 2018). In line with the trend to study the relationship between an individual's brain structure or function and individual differences in behavior, many NFB studies explained the individual differences of NFB learning performance based on neural recordings and brain imaging data, such as resting-state EEG features and neuroanatomical features (Alkoby et al., 2018; Weber et al., 2020). A number of EEG NFB studies suggested that participants can be grouped as "learners" or "non-learners," based on their brain ability or inability to regulate their brain activity during the NFB training. Then, machine learning models can be trained to predict NFB learning performance using single-modality neuroimaging predictors (Wan et al., 2014; Nan et al., 2015; Reichert et al., 2015; Nan et al., 2018).

More and more studies used multimodal neuroimaging in human brain researches because it overcomes the limitations of individual modalities. Different neuroimaging techniques have different biochemical/biophysical mechanisms, which lead to different capabilities in probing the human brain's structure and function (Zhang et al., 2020b). Specifically, EEG passively records electric changes induced by extra- and intra-cellular electric currents associated with neuronal brain activity





(Michel and Murray, 2012). Electroencephalographic has a high temporal resolution and it is safe, cheap, and easy to operate. On the other hand, MRI has a good spatial resolution and can provide a more complete as well as more detailed picture of the brain. Structural MRI can provide static anatomical information (Symms et al., 2004) and fMRI depicts brain activity by detecting the changes in brain hemodynamics (Howseman and Bowtell, 1999). Multimodal neuroimaging data analysis could take the advantages from multiple imaging techniques, such as improving both spatial and temporal resolution and illustrating the anatomical basis for functional activities (Zhang et al., 2020b). Numerous neuroimaging studies have shown that, multiple neuroimaging modalities may provide a more comprehensive understanding of the complex interplay between the brain (including structure, function, and networks) and behavior. For example, integrating both functional and structural features could improve prediction accuracy of intelligence of healthy subjects (Jiang et al., 2020). Here in the current study, we identified a set of NFB learning-performance-related multimodal neuroimaging features, which greatly broaden

our knowledge about the neural mechanisms of NFB learning effects. The results suggested that, in view of the underspecified and complex character of NFB training task, the individual difference in NFB learning performance is not attributed to single modality, the individual difference in SMR NFB performance is not attributed to one single factor, but modulated by the brain's baseline neural activity, structure, function, and functional connections.

## Multimodal Predictors of SMR NFB Learning Index

### EEG Predictors

According to existing findings, a well-known predictor of NFB learning index should be the baseline EEG activity. A higher baseline level of the training parameter (brain activity of the training target) is advantageous for better NFB training performance (Wan et al., 2014; Nan et al., 2015; Reichert et al., 2015). In consistence with previous studies (Reichert et al., 2015; Nan et al., 2018), we found that the

eyes-open resting-state SMR power showed positive correlation with NFB learning index. Generally, the most relevant resting-state SMR power located in central regions, especially the left central regions. It has been proposed that a certain baseline level of SMR power is associated with a higher neural adaptability which may allow a better modulation of this EEG rhythm (Reichert et al., 2015).

### fMRI Predictors

Similar to rest EEG, rsfMRI data are also easy to collect and have gained widespread applications. Two types of features, regional activities (ALFF/fALFF/ReHo) and FC of intrinsic brain networks, can be calculated from rsfMRI. However, there is no study to examine rsfMRI-based predictors for EEG NFB training performance. We found significant positive correlations between the NFB learning performance and the strength of a set of FC features, which were mainly observed within the cingulo-opercular network, between the sensorimotor network and occipital network. Similarly, FC having significant negative correlations with the learning performance were observed between the default mode network and occipital network.

Considering the underspecified content of NFB training task, successful learning performance requires participants to show self-initiated multitasking, such as to focus on inner mental thoughts and external stimuli at the same time. Therefore, while concentrating on real-time feedback, participants have to redirect attention away from irrelevant thoughts and toward goal-related thoughts (Kober et al., 2017). Hence, positive correlation between resting-state within-network FC of cingulo-opercular network might be explained by its positive effects for redirection of attention and maintenance of tonic alertness (Sadaghiani and D'Esposito, 2015). A previous study showed that the intention to control the moving bar of sham feedback is sufficient to engage a broad cingulate-opercular network related to cognitive control (Ninaus et al., 2013).

The default mode network is implicated in self-referential and integrative processes and it generally has negative FC with other task-positive networks during resting-state (Fox et al., 2005). The anticorrelation between the default mode network and task-positive visual networks may reflect the dichotomy between NFB training requiring introspectively oriented and extrospectively oriented attentional modes (Fransson, 2005). The strength of this negative correlation could be considered as an index of the degree of regulation in the default mode and task-positive networks, and showed positive correlation with behavioral performance (Kelly et al., 2008). Consistently, better NFB training performance showed negative correlation with resting-state FC between the default mode and occipital networks, i.e., the higher the negative correlation, the better the NFB training performance. Besides, performance during NFB training is also positively correlated with resting-state FC between occipital and sensorimotor networks. Both occipital and sensorimotor networks are known to support more specialized and mostly externally driven functions, and the FC between these two networks is normally low during resting-state (Doucet et al., 2011; Gu et al., 2015; Lee and Frangou, 2017). In addition to FC, we also observed positive correlation between ALFF and NFB

learning performance in brain regions within the sensorimotor and occipital networks, such as the postcentral gyrus and lingual gyrus. Considering increased FC between the networks responsible for processing different types of sensory information, it can be inferred that better NFB learning performance might be related to multisensory integration.

### sMRI Predictors

VBM analysis of sMRI data measures the neuroanatomy characteristic of the human brain, and it has been used in previous studies to investigate the link between brain anatomical properties and the individual difference in cognitive function (Kanai and Rees, 2011). In order to reveal the neuroanatomical basis of the ability to achieve self-regulation of one's own brain activity, two studies investigated structural predictors of learning performance during up-regulation of SMR power (Ninaus et al., 2015; Kober et al., 2017). Compared with these two studies, the current study observed relevant findings on GMV and WMV of more widely distributed brain regions. Specifically, for the default mode network, a positive relationship between GMV and NFB performance was observed in the superior parietal gyrus, precuneus, and inferior temporal gyrus, and a negative relationship was observed in the frontal gyrus and temporal pole. Besides, a number of brain regions within the occipital network, including the middle occipital gyrus, calcarine sulcus, lingual gyrus, and fusiform gyrus, showed positive correlation between WMV and NFB performance. As mentioned before, NFB performance showed significant correlation with the rsfMRI FC strength between the default mode network and occipital network. These results underscore the importance of these two functional networks to the capability of learning self-regulation of brain activity from the neuroanatomical perspective. So far, a very limited number of studies investigated structural predictors of NFB performance, and their results also explained the neuroanatomical basis for learning self-regulation of brain activity (Enriquez-Geppert et al., 2013; Ninaus et al., 2013, 2015; Kober et al., 2017). There are some differences between the results obtained in this study and those reported before, and the differences can be attributed to a number of reasons, such as different NFB protocols, training characteristics, and measurements of training performance. Future study should attempt to reveal a more general "NFB network" in the brain, regarding overlapping neuroanatomical correlates for different NFB protocols (Weber et al., 2020).

### Limitations

This present study has some limitations. First, the present study was constrained in terms of the sample size, which could limit and weaken the result of this study. Second, a corrected threshold was only used for extraction of EEG correlates. The results are sufficient to support our conclusion, but studies with more strict feature selection approach are needed to verify the results obtained here. Third, the multimodal prediction was primarily achieved by simple concatenating brain features from different modalities horizontally into a single, combined feature space, thus not allowing for a full use of the joint information among modalities (Sui et al., 2020). Fourth, only within-session learning

effect was evaluated in this study. There is no generally accepted best measure for assessing NFB learning success so far. One might speculate that observed predictors for NFB performance might be different if we had used another measure of learning performance. But here in this study, our main purpose is to validate the necessity of utilizing multimodal data when investigating predictors of NFB learning success, therefore we used a learning index (LI), which has been successfully applied in previous studies. Last but not least, we haven't collected any behavioral data to characterize subject's psychological state during NFB training, such as mental strategy. This should be considered in future study together with pre-NFB baseline factors to make a more comprehensive investigation of NFB learning performance predictors.

## CONCLUSION

In conclusion, inter-individual differences concerning the ability to regulate one's brain activity are in the focus of current NFB research. Existing studies proposed that the individual differences in NFB learning performance can be attributed to electrophysiological and anatomical baseline characteristics. Our results support and extend these findings, since we found reliable predictors of within-session NFB learning performance from multimodal neuroimaging data. The results of this study highlight the importance of multimodal neuroimaging technique as a tool to explain the individual difference in learning performance during NFB training, and could provide a basic theoretical framework for development of individualization of NFB protocols.

## DATA AVAILABILITY STATEMENT

The data supporting the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES

- Alkoby, O., Aburmileh, A., Shriki, O., and Todder, D. (2018). Can we predict who will respond to neurofeedback? A review of the inefficacy problem and existing predictors for successful EEG neurofeedback learning. *Neuroscience* 15, 155–164. doi: 10.1016/j.neuroscience.2016.12.050
- Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* 57, 289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x
- Dong, M., Li, J., Shi, X., Gao, S., Fu, S., Liu, Z., et al. (2015). Altered baseline brain activity in experts measured by amplitude of low frequency fluctuations (ALFF): a resting state fMRI study using expertise model of acupuncturists. *Front. Hum. Neurosci.* 9:99. doi: 10.3389/fnhum.2015.00099
- Dosenbach, N. U., Nardos, B., Cohen, A. L., Fair, D. A., Power, J. D., Church, J. A., et al. (2010). Prediction of individual brain maturity using fMRI. *Science* 329, 1358–1361. doi: 10.1126/science.1194144
- Doucet, G., Naveau, M., Petit, L., Delcroix, N., Zago, L., Crivello, F., et al. (2011). Brain activity at rest: a multiscale hierarchical functional organization. *J. Neurophysiol.* 105, 2753–2763. doi: 10.1152/jn.00895.2010

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of Shenzhen University. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

ZZ, LR, and NY conceived and designed the experiments. LL, YW, YZ, and SH performed the experiments. LL and YW analyzed the data and wrote the main manuscript. GH and LZ verified the analytical methods. All authors reviewed the manuscript.

## FUNDING

This research was supported by grants from the Science, Technology, and Innovation Commission of Shenzhen Municipality Technology Fund (Nos. JCYJ20170818093322718, JCYJ20190808173819182, and 2021SHIBS0003) and National Natural Science Foundation of China (Nos. 81901831, 61771461, and U1736202).

## ACKNOWLEDGMENTS

We gratefully acknowledge all the participants.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnhum.2021.699999/full#supplementary-material>

- Enriquez-Geppert, S., Huster, R. J., Scharfenort, R., Mokom, Z. N., Vosskuhl, J., Figge, C., et al. (2013). The morphology of midcingulate cortex predicts frontal-midline theta neurofeedback success. *Front. Hum. Neurosci.* 7:453. doi: 10.3389/fnhum.2013.00453
- Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–9678. doi: 10.1073/pnas.0504136102
- Fransson, P. (2005). Spontaneous low-frequency BOLD signal fluctuations: an fMRI investigation of the resting-state default mode of brain function hypothesis. *Hum. Brain Mapp.* 26, 15–29. doi: 10.1002/hbm.20113
- Gruzelier, J. H. (2014a). EEG-neurofeedback for optimising performance. I: a review of cognitive and affective outcome in healthy participants. *Neurosci. Biobehav. Rev.* 44, 124–141. doi: 10.1016/j.neubiorev.2013.09.015
- Gruzelier, J. H. (2014b). EEG-neurofeedback for optimising performance. II: creativity, the performing arts and ecological validity. *Neurosci. Biobehav. Rev.* 44, 142–158. doi: 10.1016/j.neubiorev.2013.11.004
- Gruzelier, J. H. (2014c). EEG-neurofeedback for optimising performance. III: a review of methodological and theoretical considerations. *Neurosci. Biobehav. Rev.* 44, 159–182. doi: 10.1016/j.neubiorev.2014.03.015

- Gu, S., Satterthwaite, T. D., Medaglia, J. D., Yang, M., Gur, R. E., Gur, R. C., et al. (2015). Emergence of system roles in normative neurodevelopment. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13681–13686. doi: 10.1073/pnas.1502829112
- Horien, C., Greene, A. S., Constable, R. T., and Scheinost, D. (2020). Regions and connections: complementary approaches to characterize brain organization and function. *Neuroscientist* 26, 117–133. doi: 10.1177/1073858419860115
- Howseman, A. M., and Bowtell, R. W. (1999). Functional magnetic resonance imaging: imaging techniques and contrast mechanisms. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 354, 1179–1194. doi: 10.1098/rstb.1999.0473
- Jiang, R. T., Calhoun, V. D., Cui, Y., Qi, S. L., Zhuo, C. J., Li, J., et al. (2020). Multimodal data revealed different neurobiological correlates of intelligence between males and females. *Brain Imaging Behav.* 14, 1979–1993. doi: 10.1007/s11682-019-00146-z
- Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., and Sejnowski, T. J. (2001). Analysis and visualization of single-trial event-related potentials. *Hum. Brain Mapp.* 14, 166–185. doi: 10.1002/hbm.1050
- Kanai, R., and Rees, G. (2011). The structural basis of inter-individual differences in human behaviour and cognition. *Nat. Rev. Neurosci.* 12, 231–242. doi: 10.1038/nrn3000
- Kelley, W. M., Wagner, D. D., and Heatherton, T. F. (2015). In search of a human self-regulation system. *Annu. Rev. Neurosci.* 38, 389–411. doi: 10.1146/annurev-neuro-071013-014243
- Kelly, A. M., Uddin, L. Q., Biswal, B. B., Castellanos, F. X., and Milham, M. P. (2008). Competition between functional brain networks mediates behavioral variability. *Neuroimage* 39, 527–537. doi: 10.1016/j.neuroimage.2007.08.008
- Kleih, S. C., Nijboer, F., Halder, S., and Kübler, A. (2010). Motivation modulates the P300 amplitude during brain–computer interface use. *Clin. Neurophysiol.* 121, 1023–1031. doi: 10.1016/j.clinph.2010.01.034
- Kober, S. E., Neuper, C., and Wood, G. (2020). Differential effects of up- and down-regulation of SMR coherence on EEG activity and memory performance: A neurofeedback training study. *Front. Hum. Neurosci.* 14:606684. doi: 10.3389/fnhum.2020.606684
- Kober, S. E., Schweiger, D., Witte, M., Reichert, J. L., Grieshofer, P., Neuper, C., et al. (2015). Specific effects of EEG based neurofeedback training on memory functions in post-stroke victims. *J. Neuroeng. Rehabil.* 12:107. doi: 10.1186/s12984-015-0105-6
- Kober, S. E., Witte, M., Ninaus, M., Koschutnig, K., Wiesen, D., Zaiser, G., et al. (2017). Ability to gain control over one's own brain activity and its relation to spiritual practice: A multimodal imaging study. *Front. Hum. Neurosci.* 11:271. doi: 10.3389/fnhum.2017.00271
- Lee, W. H., and Frangou, S. (2017). Linking functional connectivity and dynamic properties of resting-state networks. *Sci. Rep.* 7:16610. doi: 10.1038/s41598-017-16789-1
- Makeig, S., Jung, T. P., Bell, A. J., Ghahremani, D., and Sejnowski, T. J. (1997). Blind separation of auditory event-related brain responses into independent components. *Proc. Natl. Acad. Sci. U.S.A.* 94, 10979–10984. doi: 10.1073/pnas.94.20.10979
- Michel, C. M., and Murray, M. M. (2012). Towards the utilization of EEG as a brain imaging tool. *NeuroImage* 61, 371–385. doi: 10.1016/j.neuroimage.2011.12.039
- Nan, W. Y., Rodrigues, J. O. P., Ma, J. L., Qu, X. T., Wan, F., Mak, P. L., et al. (2012). Individual alpha neurofeedback training effect on short term memory. *Int. J. Psychophysiol.* 86, 83–87. doi: 10.1016/j.ijpsycho.2012.07.182
- Nan, W. Y., Wan, F., Tang, Q., Wong, C. M., Wang, B. Y., and Rosa, A. (2018). Eyes-closed resting EEG predicts the learning of alpha down-regulation in neurofeedback training. *Front. Psychol.* 9:1607. doi: 10.3389/fpsyg.2018.01607
- Nan, W. Y., Wan, F., Vai, M. I., and Da Rosa, A. C. (2015). Resting and initial beta amplitudes predict learning ability in beta/theta ratio neurofeedback training in healthy young adults. *Front. Hum. Neurosci.* 9:677. doi: 10.3389/fnhum.2015.00677
- Nichols, T. E., and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25. doi: 10.1002/hbm.1058
- Ninaus, M., Kober, S. E., Witte, M., Koschutnig, K., Neuper, C., and Wood, G. (2015). Brain volumetry and self-regulation of brain activity relevant for neurofeedback. *Biol. Psychol.* 110, 126–133. doi: 10.1016/j.biopsycho.2015.07.009
- Ninaus, M., Kober, S. E., Witte, M., Koschutnig, K., Stangl, M., Neuper, C., et al. (2013). Neural substrates of cognitive control under the belief of getting neurofeedback training. *Front. Hum. Neurosci.* 7:914. doi: 10.3389/fnhum.2013.00914
- Olbrich, S., Jödicke, J., Sander, C., Himmerich, H., and Hegerl, U. (2011). ICA-based muscle artefact correction of EEG data: what is muscle and what is brain? Comment on McMenamin et al. *Neuroimage* 54, 1–3; discussion4–9. doi: 10.1016/j.neuroimage.2010.04.256
- Omejc, N., Rojc, B., Battaglini, P. P., and Marusic, U. (2019). Review of the therapeutic neurofeedback method using electroencephalography: EEG Neurofeedback. *Bosn. J. Basic. Med. Sci.* 19, 213–220. doi: 10.17305/bjbm.2018.3785
- Pereira, F., Mitchell, T., and Botvinick, M. (2009). Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45, S199–S209. doi: 10.1016/j.neuroimage.2008.11.007
- Qi, P., Ru, H., Gao, L. Y., Zhang, X. B., Zhou, T. S., Tian, Y., et al. (2019). Neural mechanisms of mental fatigue revisited: new insights from the brain connectome. *Engineering* 5, 276–286. doi: 10.1016/j.eng.2018.11.025
- Reichert, J. L., Kober, S. E., Neuper, C., and Wood, G. (2015). Resting-state sensorimotor rhythm (SMR) power predicts the ability to up-regulate SMR in an EEG-instrumental conditioning paradigm. *Clin. Neurophysiol.* 126, 2068–2077. doi: 10.1016/j.clinph.2014.09.032
- Sadaghiani, S., and D'Esposito, M. (2015). Functional characterization of the cingulo-opercular network in the maintenance of tonic alertness. *Cereb. Cortex* 25, 2763–2773. doi: 10.1093/cercor/bhu072
- Scheinost, D., Stoica, T., Wasyluk, S., Gruner, P., Saksa, J., Pittenger, C., et al. (2014). Resting state functional connectivity predicts neurofeedback response. *Front. Behav. Neurosci.* 8:338. doi: 10.3389/fnbeh.2014.00338
- Sitaram, R., Ros, T., Stoeckel, L., Haller, S., Scharnowski, F., Lewis-Peacock, J., et al. (2017). Closed-loop brain training: the science of neurofeedback. *Nat. Rev. Neurosci.* 18, 86–100. doi: 10.1038/nrn.2016.164
- Sui, J., Jiang, R. T., Bustillo, J., and Calhoun, V. (2020). Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: Methods and Promises. *Biol. Psychol.* 88, 818–828. doi: 10.1016/j.biopsycho.2020.02.016
- Symms, M., Jäger, H. R., Schmierer, K., and Yousry, T. A. (2004). A review of structural magnetic resonance neuroimaging. *J. Neurol. Neurosurg. Psychiatry* 75, 1235–1244. doi: 10.1136/jnnp.2003.032714
- Thibault, R. T., and Raz, A. (2016). When can neurofeedback join the clinical armamentarium? *Lancet Psychiatry* 3, 497–498. doi: 10.1016/s2215-0366(16)30040-2
- Wan, F., Nan, W. Y., Vai, M. I., and Rosa, A. (2014). Resting alpha activity predicts learning ability in alpha neurofeedback. *Front. Hum. Neurosci.* 8:500. doi: 10.3389/fnhum.2014.00500
- Weber, L. A., Ethofer, T., and Ehls, A. C. (2020). Predictors of neurofeedback training outcome: a systematic review. *NeuroImage Clin.* 27:102301. doi: 10.1016/j.nicl.2020.102301
- Witte, M., Kober, S. E., Ninaus, M., Neuper, C., and Wood, G. (2013). Control beliefs can predict the ability to up-regulate sensorimotor rhythm during neurofeedback training. *Front. Hum. Neurosci.* 7:478. doi: 10.3389/fnhum.2013.00478
- Xiang, Y., Kong, F., Wen, X., Wu, Q., and Mo, L. (2016). Neural correlates of envy: regional homogeneity of resting-state brain activity predicts dispositional envy. *Neuroimage* 142, 225–230. doi: 10.1016/j.neuroimage.2016.08.003
- Xie, Y., Li, Y., Guan, M., Duan, H., Xu, X., and Fang, P. (2021). Audiovisual working memory and association with resting-state regional homogeneity. *Behav. Brain Res.* 411, 113382. doi: 10.1016/j.bbr.2021.113382
- Zang, Y., Jiang, T., Lu, Y., He, Y., and Tian, L. (2004). Regional homogeneity approach to fMRI data analysis. *Neuroimage* 22, 394–400. doi: 10.1016/j.neuroimage.2003.12.030
- Zhang, R., Li, F. L., Zhang, T., Yao, D. Z., and Xu, P. (2020a). Subject inefficiency phenomenon of motor imagery brain-computer interface: Influence factors

- and potential solutions. *Brain Sci. Adv.* 6, 224–241. doi: 10.26599/BSA.2020.9050021
- Zhang, Y. D., Dong, Z. C., Wang, S. H., Yu, X., Yao, X. J., Zhou, Q. H., et al. (2020b). Advances in multimodal data fusion in neuroimaging: Overview, challenges, and novel orientation. *Inform. Fusion* 64, 149–187. doi: 10.1016/j.inffus.2020.07.006
- Zou, Q. H., Zhu, C. Z., Yang, Y., Zuo, X. N., Long, X. Y., Cao, Q. J., et al. (2008). An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. *J. Neurosci. Methods* 172, 137–141. doi: 10.1016/j.jneumeth.2008.04.012

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Li, Wang, Zeng, Hou, Huang, Zhang, Yan, Ren and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.





# An Empirical Comparative Study on the Two Methods of Eliciting Singers' Emotions in Singing: Self-Imagination and VR Training

Jin Zhang<sup>1</sup>, Ziming Xu<sup>2</sup>, Yueying Zhou<sup>2</sup>, Pengpai Wang<sup>2</sup>, Ping Fu<sup>3</sup>, Xijia Xu<sup>4</sup> and Daoqiang Zhang<sup>2\*</sup>

<sup>1</sup> College of Arts, Nanjing University of Aeronautics and Astronautics, Nanjing, China, <sup>2</sup> MIT Key Laboratory of Pattern Analysis and Machine Intelligence, College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China, <sup>3</sup> Department of Library Services, Central Washington University, Ellensburg, WA, United States, <sup>4</sup> Department of Psychiatry, Affiliated Nanjing Brain Hospital, Nanjing Medical University, Nanjing, China

## OPEN ACCESS

### Edited by:

Jane Zhen Liang,  
Shenzhen University, China

### Reviewed by:

Panagiotis Kourtesis,  
Inria Rennes - Bretagne Atlantique  
Research Centre, France  
Yunfa Fu,  
Kunming University of Science  
and Technology, China

### \*Correspondence:

Daoqiang Zhang  
dqzhang@nuaa.edu.cn

### Specialty section:

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

**Received:** 11 April 2021

**Accepted:** 27 July 2021

**Published:** 12 August 2021

### Citation:

Zhang J, Xu Z, Zhou Y, Wang P,  
Fu P, Xu X and Zhang D (2021) An  
Empirical Comparative Study on  
the Two Methods of Eliciting Singers'  
Emotions in Singing: Self-Imagination  
and VR Training.  
*Front. Neurosci.* 15:693468.  
doi: 10.3389/fnins.2021.693468

Emotional singing can affect vocal performance and the audience's engagement. Chinese universities use traditional training techniques for teaching theoretical and applied knowledge. Self-imagination is the predominant training method for emotional singing. Recently, virtual reality (VR) technologies have been applied in several fields for training purposes. In this empirical comparative study, a VR training task was implemented to elicit emotions from singers and further assist them with improving their emotional singing performance. The VR training method was compared against the traditional self-imagination method. By conducting a two-stage experiment, the two methods were compared in terms of emotions' elicitation and emotional singing performance. In the first stage, electroencephalographic (EEG) data were collected from the subjects. In the second stage, self-rating reports and third-party teachers' evaluations were collected. The EEG data were analyzed by adopting the max-relevance and min-redundancy algorithm for feature selection and the support vector machine (SVM) for emotion recognition. Based on the results of EEG emotion classification and subjective scale, VR can better elicit the positive, neutral, and negative emotional states from the singers than not using this technology (i.e., self-imagination). Furthermore, due to the improvement of emotional activation, VR brings the improvement of singing performance. The VR hence appears to be an effective approach that may improve and complement the available vocal music teaching methods.

**Keywords:** vocal music teaching, singing emotion, self-imagination, virtual reality, electroencephalogram, emotion classification

## INTRODUCTION

The generation of human emotion has a certain regularity that reflects the degree of experience and cognitive relationship between objectivity and subjectivity (Xi, 2010). Singing is a performing art that uses sound as a tool to awaken the human spirit, which fully embodies the self-creation and pursuit of human emotional expression (Xu, 2008). The art of singing is the art of human mood.

The singing activities of singers are the expression of their inner psychology. These activities are inseparable from important psychological factors such as feeling, perception, consciousness, will, memory, imagination, emotion, and thinking (Mu, 2011). Singers need to pay attention to their emotional expression, truly integrate their thoughts and feelings with the song, and give life and soul to the song (Shi, 2002). Therefore, in the process of learning vocal music, students not only need to master the vocalization skills, but also need to actively invest in emotions, effectively express the inner emotions of musical works, and give them vitality and appeal.

Current vocal singing learning in Chinese universities is mainly based on the teaching of vocal theories and knowledge. The traditional self-imagination training for vocal learners does not efficiently engage and attracts students' attention. We have observed that vocal music singers in our institution have insufficient perception and lack the ability to sing emotionally. In the process of song interpretation, the emotional expression of the singers is not ideal, and the emotional state is not actively mobilized. Thus it is difficult to achieve the requirements of the singing state with a strong voice.

Immersive virtual reality (VR) technology has existed for 50 years (Slater, 2009) and has been regarded as a promising and clinical tool (Kourtesis et al., 2019a). It enables researchers to collect advanced cognitive and behavioral data through dynamic stimuli and interactions in an ecologically valid environment, and it can be combined with non-invasive brain imaging techniques (Kourtesis et al., 2020). Virtual environments are sophisticated human-computer interfaces that can be used for a wide variety of applications (Skarbez et al., 2017). In recent years, VR technology has greatly met the needs of art education due to its multi-sensing, immersive, interactive, and imaginative characteristics. Though it has been widely used in design (Huang and Chen, 2017; Pei, 2017) and art education (Gao, 2013; Zhang, 2018), VR training has scarcely been used in the field of vocal teaching performance. Vocal music requires a comprehensive response of multiple psychological factors such as perception, thinking, imagination, and movement.

An emotion is a subjective state of being that we often describe as our feelings that impact human behavior and mental health (Zhang G. H. et al., 2019). The neuroscience research shows that the trigger of emotion is closely related to physiological activities, especially brain activities (LeDoux, 2000), which provides a theoretical basis for identifying emotional states by analyzing brain activities. Electroencephalogram (EEG) signal has the advantages of high time resolution, portability, and non-invasiveness (Baig and Kavakli, 2018). Emotion recognition based on EEG has received widespread attention (Alarcao and Fonseca, 2017). Generally, researchers collect EEG data during the elicitation of emotions by presenting videos, pictures, and other emotionally stimulating means. Then they extract relevant EEG features to explore the correlation between EEG features and different emotion categories (Hajcak et al., 2010). Some studies use machine learning algorithms to predict emotion based on EEG features (Wang et al., 2014; Li et al., 2019). Today EEG-based emotion recognition has been utilized in the rehabilitation treatment of patients with impaired consciousness

(Huang et al., 2019), soldier mental state assessment (Lin et al., 2017), driving status monitoring (Halim and Rehan, 2020), but has not yet used in the area of eliciting singers' emotion in singing.

In recent years, VR combined with EEG has been adopted in rehabilitation (Calabrò et al., 2017), stress relief (Tarrant et al., 2018), and teaching (Sood and Singh, 2018). Also, immersive VR and brainwave technologies have been adopted across education and training fields (Yang and Ren, 2019).

Inspired by these researches, an empirical study of combining VR training and EEG to investigate the effect of the method in eliciting emotions was designed. It is important to emphasize that combined VR training and EEG has not yet been found in the literature for vocal-music teaching and performance evaluation. In this study, 16 music students were invited to participate in a 2-stage experiment. The two-stage experiment was designed to collect data for the two teaching methods: self-imagination and VR training. The two methods were compared by analyzing the three types of data: self-rating reports, third-party evaluation, and EEG emotion classification data. An empirical judgment of the study is, as long as we elicit the emotional state corresponding to the song emotion (such as the positive song corresponds to the positive emotion), the singing effect will be improved; the better the effect of the emotional elicitation, the more obvious the improvement of the singing effect. Therefore, this study aimed to verify the hypothesis that VR can better elicit the positive, neutral, and negative emotions states than self-imagination, and further improve the singing performance due to the improvement of emotional activation.

## METHODS

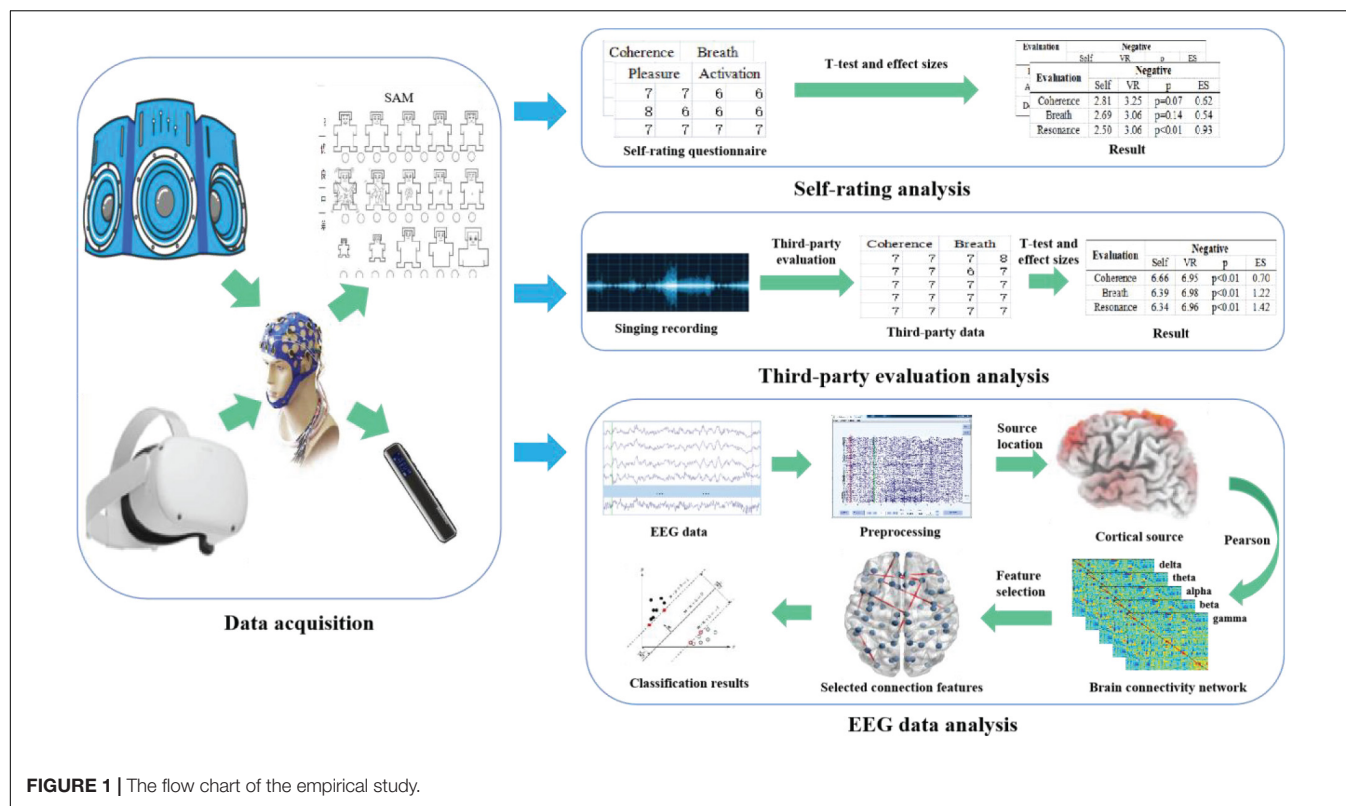
**Figure 1** shows a flowchart of the proposed method, mainly including data collection and data analysis. The data collection part collected the EEG data, self-rating reports, and third-party evaluation data. The data analysis part conducted a statistical investigation on self-rating reports, third-party evaluation data, and emotion recognition based on functional connectivity features of EEG data.

### Experiment and Data Collection Participants

In this empirical study, 16 college students (8 male students and 8 female students, mean age of  $19.5 \pm 1.54$ ) from the Art School of Nanjing University of Aeronautics and Astronautics were invited to participate in this experiment. All participants were right-handed and have normal hearing, normal vision, or vision correction, and no brain or mental illness. They were asked to maintain adequate sleep time before the experiment. They agreed and signed a written informed consent form. The experiment and data in this empirical study complied with all relevant ethical laws and regulations. Ethical guidelines were followed for conducting experiments with human participants.

### The Description of the Designed Experiment

Six songs in three emotion categories (positive, neutral, and negative) were selected as emotional stimulation materials. Each



**FIGURE 1 |** The flow chart of the empirical study.

emotion category has been assigned by two songs. Each song was edited into a 3-min core segment, the lyrics of each song were removed, and only the background music was retained. The emotional category tags and song names of the selected songs are shown in **Table 1**.

In the experiment, each subject was asked to participate in a two-stage experiment on the emotion elicited by the self-imagination and VR training. The complete experimental paradigm is shown in **Figure 2**. In the self-imagination stage, the vocal music teaching scene was simulated, the edited songs were played, and the subjects were allowed to imagine themselves according to the songs they heard, and mobilized the scenes, and elicited emotions required for singing. In the VR training

stage, participants were asked to wear VR glasses (Quest2, Oculus, United States) and watched VR videos made according to the emotional background of the song. To better simulate the real vocal music teaching scene, in the self-imagining stage, a teaching guide commentary was added before the background music of the song was played, which introduced the relevant background of the song and the emotion required for the singing scene of the song. However, the background music of the VR training video was only the background music that the edited song played in the VR training.

For each emotion category, each stage of the experiment was divided into three groups, negative, neutral, and positive. And each group included an experiment of two songs in the same emotion category. The experimental process of each song included self-imagination and watching a VR video for 3 min to elicit emotion in singing. The subject sang for 1 min and the singing was recorded for evaluating the subject's singing performance. Then the subject was asked to rest for 1 min and continued the experiment of the next song. After completing the experiment of two songs in each group, the subject was asked to fill out two self-rating reports: Self-Assessment Manikin (SAM) emotion self-rating form (Bradley and Lang, 1994), and vocal performance self-rating form. The two songs of each group were played in the order of negative, neutral, and positive, with a 3-min rest between each group. Participants rested for 10 min after the end of the three groups in each stage. In the experiment, the subjects were asked to stay as still as possible to prevent the influence of artifacts on the EEG. After all the experiments of a subject were completed, third-party evaluators evaluated

**TABLE 1 |** Emotional stimulation materials.

Order	Emotion category	Translated name of song	Original name of song
1	Negative	Mama in the candlelight	烛光里的妈妈
2	Negative	Where does the time go	时间都去哪了
3	Neutral	Pastoral song	牧歌
4	Neutral	By Lake Baikal	贝加尔湖畔
5	Positive	My motherland and me	我和我的祖国
6	Positive	In the field of hope	在希望的田野上

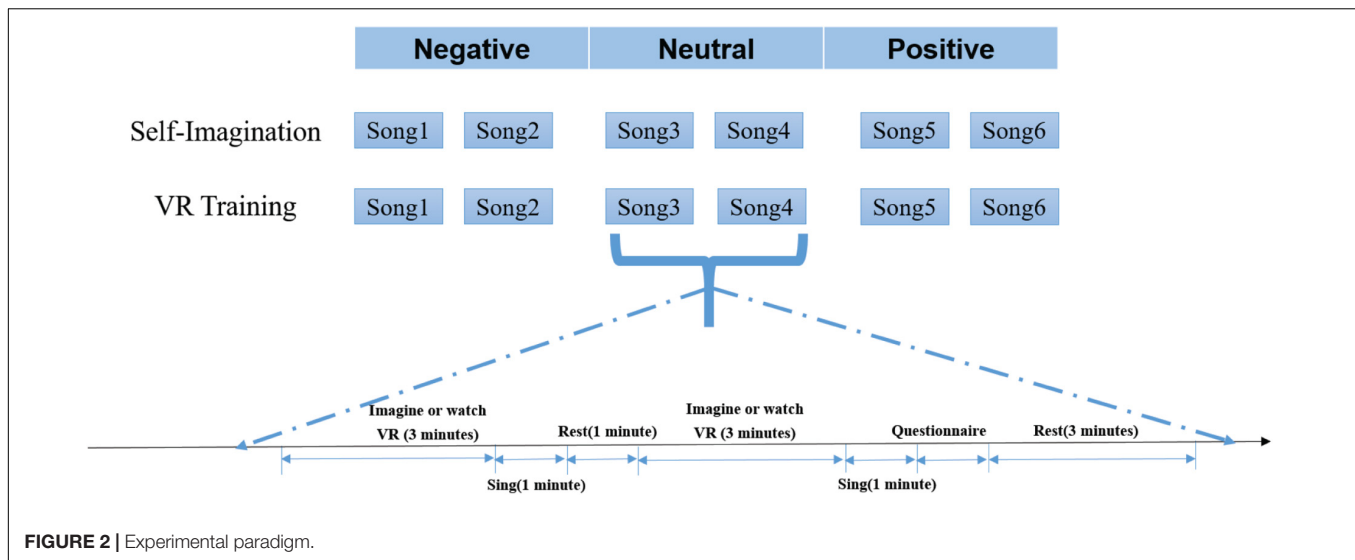


FIGURE 2 | Experimental paradigm.

the 12 songs according to the different dimensions of vocal singing (2 scenes  $\times$  6 songs) and gave a score based on the subject's singing performance. The evaluators were five selected professional vocal teachers.

The SAM self-rating scale was set from one to nine points and used to measure emotional pleasure, activation, and dominance. The higher the score, the stronger the emotion indicating higher pleasure, higher arousal, and higher dominance. The vocal self-rating form provides self-rating scores on singing performance from seven aspects: coherence of song singing, use of breath, use of resonance, intonation, rhythm, language use, musicality, and emotional expression. A rubric with four levels for the assessment is used: excellent, good, medium, and poor. The corresponding numbers were 4, 3, 2, and 1. The higher the score, the better the singing performance. As for the third-party evaluation, the same seven aspects used for the vocal self-rating form were evaluated except the scale was set to from 1 to 10 points. The higher the score, the better the singing performance.

### The Description of the EEG Recording

A 64-channel portable wireless EEG system (NeuSen.W64, Neuracle, China) was used for EEG data collection, and the sampling rate was set to 1,000 Hz. According to the international standard 10–20 system, the EEG data were recorded of 59 electrodes, AF3/4, AF7/8, Fp1/2, Fz, F1/2, F3/4, F5/6, F7/8, FC1/2, FC3/4, FC5/6, FCz, Cz, C1/2, C3/4, C5/6, FT7/8, T7/8, TP7/8, CP1/2, CP3/4, CP5/6, FPz, Pz, P3/4, P5/6, P7/8, PO3/4, PO5/6, PO7/8, POz, Oz, O1/2, with CPz as the reference electrode and AFz as the ground electrode. Throughout the experiment, the impedance of all electrodes was kept below 5 k $\Omega$ .

## Experimental Method

### EEG Signal Preprocessing

In this study, the EEGLAB (Delorme and Makeig, 2004) toolbox was used to preprocess the EEG data. After channel location, the original EEG data was band-pass filtered at 1–45 Hz, the whole brain was averaged re-referenced, and down-sampled to 250 Hz.

Then, independent component analysis was used to decompose the EEG signal into independent components, and the ICLabel (Pion-Tonachini et al., 2019) was used to remove artifacts such as eye movement and muscle movement. After preprocessing, the last 150 s EEG signal of each song was divided into 1 s period EEG signal as a sample to obtain the maximum emotional response (Li et al., 2020).

### Source Location and Functional Connectivity Estimation

Due to the volume conduction effect, the brain activation area derived from the scalp EEG is not accurate, and the electrical signals from the brain can be recorded from multiple nearby sensors (Dimitrakopoulos et al., 2017). Therefore, a brain network calculation method was used similar to some literature described (Dimitrakopoulos et al., 2017; Wang et al., 2020). First, the source localization brain activation of each frequency band was calculated, and then the cortical functional network was calculated. The Brainstorm software (Tadel et al., 2011) for source location was used. The ICBM152 template was used to construct the head model and source space, and map the scalp electrodes to the ICBM152 (T1) head model. The head model used in this study was divided into three tissue types: scalp, brain, and skull. Their default conductivity was 1, 0.0125, and 1 s/m, respectively. The direction of each source was restricted to be perpendicular to the surface of the cortex, and the Boundary Element Method (BEM) (Mosher and Leahy, 1999) was used to calculate the guiding field matrix L. The cerebral cortex source signal was reconstructed by standardized low-resolution electromagnetic tomography (sLORETA) (Pascual-Marqui, 2002), the direction was fixed and perpendicular to the cortex, and all the parameters of sLORETA uses the default settings.

According to the Desikan-Killiany map (Desikan et al., 2006), the cerebral cortex gray matter was divided into 68 regions. The Pearson correlation coefficient was used to construct functional connectivity, which has been widely used in functional magnetic resonance imaging research (Dimitrakopoulos et al., 2017). The



source signal was split into five frequency bands (delta: 1–4 HZ, theta: 4–8 HZ, alpha: 8–13 HZ, beta: 13–30 HZ, and gamma: 30–45 HZ), for the source of each frequency band in each area. The signals were averaged to obtain the source signal of the brain region. The functional network was calculated by the Pearson correlation coefficient between the source signals of each brain region pair. Each sample got 11,390 functional connectivity features for classification, that was,  $68 \times (68-1)/2 = 2,278$  features for each frequency band.

### Feature Selection and Emotion Classification

In this empirical study, emotions were divided into negative, neutral, and positive. For each sample, the source functional connectivity features were extracted separately, and different tags were marked according to different emotional types of songs.

The Max-relevance and min-redundancy (mRMR) algorithm (Peng et al., 2005) was used for feature selection and select the most important 10 features for emotion classification. The algorithm uses mutual information to measure the relationship between features and categories. It maximized the correlation between features and categorical variables and minimized the redundancy between features. In other words, the algorithm found  $m$  features from the feature space that has the greatest correlation with the target category and the least redundancy with other features. The definition of maximum correlation is shown in Eq. 1, and the definition of minimum redundancy is shown in Eq. 2:

$$\max D(S, c) = \frac{1}{|S|} \sum_{f_i \in S} I(f_i; c) \quad (1)$$

$$\min R(S) = \frac{1}{|S|^2} \sum_{f_i, f_j \in S} I(f_i; f_j) \quad (2)$$

where  $S$  represents the feature set;  $c$  represents the target category;  $I(f_i; c)$  represents the mutual information between the feature  $i$  and the target category  $c$ ;  $I(f_i; f_j)$  is the mutual information between feature  $i$  and feature  $j$ .

The support vector machine (SVM) with a radial basis function (RBF) kernel was used for emotion recognition. As a widely used classifier, SVM has been proved by a large number of studies to be a practical EEG classification method (Bashivan et al., 2015). The 10-fold cross-validation was used to calculate the accuracy of the classification. The SVM kernel function parameters  $g$  (0.5–4, step size 0.5) and penalty coefficient  $c$  ( $10^{-2}$  to  $10^2$ , step size coefficient is 10) were traversed to get the optimal parameters. To explore the influence of functional connectivity features of different frequency bands on the results, the functional connectivity features of 5 frequency bands in the 2 scenarios were, respectively, cross-verified 100 times. The five frequency band features were classified separately and the five frequency band features were spliced into a feature vector for classification.

### Statistical Method

A statistical analysis of three rating scales was performed, including SAM self-rating report, vocal performance self-rating report, and third-party evaluation scores. Given different emotion categories, the three rating scales of emotion data in both

self-imagination and VR training methods were gone through the test by  $t$ -test and effect sizes (Lakens, 2013). The detailed results are described in the following section.

## RESULTS

### SAM Self-Rating

The SAM self-rating data were collected from the 16 subjects, from the perspective of emotional pleasure, activation and dominance. The average score and statistical test results of SAM are shown in **Table 2**. In the case of negative emotion, participants reported lower evaluations of VR training ( $M = 3.00$ ) than self-imagination ( $M = 4.69$ ),  $p < 0.01$ , 95%CI [1.02, 2.35], Hedges's  $g = 1.27$  for pleasure; higher evaluations of VR training ( $M = 7.00$ ) than self-imagination ( $M = 5.56$ ),  $p < 0.01$ , 95%CI [0.85, 2.02], Hedges's  $g = 1.03$  for activation; lower evaluations of VR training ( $M = 4.63$ ) than self-imagination ( $M = 6.06$ ),  $p < 0.01$ , 95%CI [0.61, 2.26], Hedges's  $g = 0.89$  for dominance. Under neutral emotion, participants reported higher evaluations of VR training ( $M = 6.19$ ) than self-imagination ( $M = 5.94$ ),  $p = 0.43$ , 95%CI [-0.41, 0.91], Hedges's  $g = 0.18$  for pleasure; higher evaluations of VR training ( $M = 6.13$ ) than self-imagination ( $M = 5.69$ ),  $p = 0.34$ , 95%CI [-0.52, 1.39], Hedges's  $g = 0.29$  for activation; higher evaluations of VR training ( $M = 6.69$ ) than self-imagination ( $M = 6.50$ ),  $p = 0.59$ , 95%CI [-0.55, 0.92], Hedges's  $g = 0.14$  for dominance. Under positive emotion, participants reported higher evaluations of VR training ( $M = 8.06$ ) than self-imagination ( $M = 7.00$ ),  $p < 0.01$ , 95%CI [0.57, 1.56], Hedges's  $g = 1.13$  for pleasure; higher evaluations of VR training ( $M = 7.56$ ) than self-imagination ( $M = 6.31$ ),  $p < 0.01$ , 95%CI [0.89, 1.61], Hedges's  $g = 1.42$  for activation; lower evaluations of VR training ( $M = 6.06$ ) than self-imagination ( $M = 6.44$ ),  $p = 0.33$ , 95%CI [-0.42, 1.17], Hedges's  $g = 0.22$  for dominance.

### Self-Rating on Vocal Performance

The vocal self-rating data were collected from the 16 subjects, and the score range was set to 1–4. The average scores and statistical test results of the scale are shown in **Table 3**. In the self-rating scale of negative emotion, participants reported higher evaluations of VR training ( $M = 3.25$ ) than self-imagination ( $M = 2.81$ ),  $p = 0.07$ , 95%CI [-0.04, 0.91], Hedges's  $g = 0.62$  for coherence; higher evaluations of VR training ( $M = 3.06$ ) than self-imagination ( $M = 2.69$ ),  $p = 0.14$ , 95%CI [-0.14, 0.89], Hedges's  $g = 0.54$  for breath; higher evaluations of VR training ( $M = 3.06$ ) than self-imagination ( $M = 2.50$ ),  $p < 0.01$ , 95%CI [0.23, 0.90], Hedges's  $g = 0.93$  for resonance; higher evaluations of VR training ( $M = 3.06$ ) than self-imagination ( $M = 2.94$ ),  $p = 0.54$ , 95%CI [-0.30, 0.55], Hedges's  $g = 0.20$  for intonation; higher evaluations of VR training ( $M = 3.06$ ) than self-imagination ( $M = 2.94$ ),  $p = 0.43$ , 95%CI [-0.20, 0.45], Hedges's  $g = 0.25$  for language; higher evaluations of VR training ( $M = 3.44$ ) than self-imagination ( $M = 3.16$ ),  $p = 0.06$ , 95%CI [-0.01, 0.63], Hedges's  $g = 0.50$  for musicality; higher evaluations of VR training ( $M = 3.31$ ) than self-imagination ( $M = 2.75$ ),  $p < 0.01$ , 95%CI [0.23, 0.90], Hedges's  $g = 0.88$  for expression. In the self-rating scale for neutral emotion, participants reported higher

**TABLE 2 |** Average score and statistical test results of Self-Assessment Manikin (SAM).

Evaluation	Negative				Neutral				Positive			
	Self	VR	<i>p</i>	ES	Self	VR	<i>p</i>	ES	Self	VR	<i>p</i>	ES
Pleasure	4.69	3.00	$p < 0.01$	1.27	5.94	6.19	$p = 0.43$	0.18	7.00	8.06	$p < 0.01$	1.13
Activation	5.56	7.00	$p < 0.01$	1.03	5.69	6.13	$p = 0.34$	0.29	6.31	7.56	$p < 0.01$	1.42
Dominance	6.06	4.63	$p < 0.01$	0.89	6.50	6.69	$p = 0.59$	0.14	6.44	6.06	$p = 0.33$	0.22

evaluations of VR training ( $M = 3.38$ ) than self-imagination ( $M = 3.25$ ),  $p = 0.43$ , 95%CI  $[-0.20, 0.45]$ , Hedges's  $g = 0.16$  for coherence; higher evaluations of VR training ( $M = 3.38$ ) than self-imagination ( $M = 3.19$ ),  $p = 0.19$ , 95%CI  $[-0.10, 0.48]$ , Hedges's  $g = 0.29$  for breath; higher evaluations of VR training ( $M = 3.25$ ) than self-imagination ( $M = 2.88$ ),  $p = 0.05$ , 95%CI  $[-0.01, 0.76]$ , Hedges's  $g = 0.63$  for resonance; higher evaluations of VR training ( $M = 3.31$ ) than self-imagination ( $M = 3.13$ ),  $p = 0.27$ , 95%CI  $[-0.16, 0.54]$ , Hedges's  $g = 0.26$  for intonation; higher evaluations of VR training ( $M = 3.38$ ) than self-imagination ( $M = 3.06$ ),  $p = 0.06$ , 95%CI  $[-0.01, 0.63]$ , Hedges's  $g = 0.59$  for language; higher evaluations of VR training ( $M = 3.56$ ) than self-imagination ( $M = 3.44$ ),  $p = 0.33$ , 95%CI  $[-0.14, 0.39]$ , Hedges's  $g = 0.22$  for musicality; higher evaluations of VR training ( $M = 3.44$ ) than self-imagination ( $M = 3.00$ ),  $p < 0.01$ , 95%CI  $[0.10, 0.77]$ , Hedges's  $g = 0.76$  for expression. In the self-rating scale of positive emotion, participants reported higher evaluations of VR training ( $M = 3.44$ ) than self-imagination ( $M = 3.31$ ),  $p = 0.34$ , 95%CI  $[-0.14, 0.39]$ , Hedges's  $g = 0.23$  for coherence; higher evaluations of VR training ( $M = 3.31$ ) than self-imagination ( $M = 3.13$ ),  $p = 0.19$ , 95%CI  $[-0.10, 0.48]$ , Hedges's  $g = 0.31$  for breath; higher evaluations of VR training ( $M = 3.06$ ) than self-imagination ( $M = 3.19$ ),  $p = 0.5$ , 95%CI  $[-0.26, 0.51]$ , Hedges's  $g = 0.22$  for resonance; higher evaluations of VR training ( $M = 3.00$ ) than self-imagination ( $M = 2.69$ ),  $p = 0.14$ , 95%CI  $[-0.11, 0.74]$ , Hedges's  $g = 0.44$  for intonation; higher evaluations of VR training ( $M = 3.31$ ) than self-imagination ( $M = 3.00$ ),  $p < 0.05$ , 95%CI  $[0.06, 0.57]$ , Hedges's  $g = 0.92$  for language; lower evaluations of VR training ( $M = 3.31$ ) than self-imagination ( $M = 3.38$ ),  $p = 0.75$ , 95%CI  $[-0.35, 0.47]$ , Hedges's  $g = 0.09$  for musicality; higher evaluations of VR training ( $M = 3.63$ ) than self-imagination ( $M = 3.25$ ),  $p < 0.05$ , 95%CI  $[0.05, 0.70]$ , Hedges's  $g = 0.63$  for expression.

### Third-Party Evaluation

The third-party evaluation data on singing performance were collected from the five professional vocal teachers. The average score of the subjects and the results of statistical test are shown in **Table 4**. Under negative emotion, participants reported higher evaluations of VR training ( $M = 6.95$ ) than self-imagination ( $M = 6.66$ ),  $p < 0.01$ , 95%CI  $[0.07, 0.50]$ , Hedges's  $g = 0.70$  for coherence; higher evaluations of VR training ( $M = 6.98$ ) than self-imagination ( $M = 6.39$ ),  $p < 0.01$ , 95%CI  $[0.38, 0.80]$ , Hedges's  $g = 1.22$  for breath; higher evaluations of VR training ( $M = 6.96$ ) than self-imagination ( $M = 6.34$ ),  $p < 0.01$ , 95%CI  $[0.33, 0.92]$ , Hedges's  $g = 1.42$  for resonance; higher evaluations of VR training ( $M = 6.26$ ) than self-imagination ( $M = 6.18$ ),

$p = 0.18$ , 95%CI  $[-0.11, 0.29]$ , Hedges's  $g = 0.16$  for intonation; lower evaluations of VR training ( $M = 6.58$ ) than self-imagination ( $M = 6.66$ ),  $p = 0.07$ , 95%CI  $[-0.04, 0.21]$ , Hedges's  $g = 0.17$  for language; higher evaluations of VR training ( $M = 7.34$ ) than self-imagination ( $M = 6.58$ ),  $p < 0.01$ , 95%CI  $[0.50, 1.02]$ , Hedges's  $g = 1.52$  for musicality; higher evaluations of VR training ( $M = 7.81$ ) than self-imagination ( $M = 6.61$ ),  $p < 0.01$ , 95%CI  $[0.95, 1.45]$ , Hedges's  $g = 2.21$  for expression. Under neutral emotion, participants reported higher evaluations of VR training ( $M = 7.00$ ) than self-imagination ( $M = 6.73$ ),  $p < 0.01$ , 95%CI  $[0.08, 0.47]$ , Hedges's  $g = 0.83$  for coherence; higher evaluations of VR training ( $M = 7.13$ ) than self-imagination ( $M = 6.65$ ),  $p < 0.01$ , 95%CI  $[0.23, 0.72]$ , Hedges's  $g = 1.09$  for breath; higher evaluations of VR training ( $M = 7.38$ ) than self-imagination ( $M = 6.58$ ),  $p < 0.01$ , 95%CI  $[0.46, 1.14]$ , Hedges's  $g = 1.42$  for resonance; higher evaluations of VR training ( $M = 6.36$ ) than self-imagination ( $M = 6.29$ ),  $p = 0.22$ , 95%CI  $[-0.13, 0.28]$ , Hedges's  $g = 0.14$  for intonation; lower evaluations of VR training ( $M = 6.63$ ) than self-imagination ( $M = 6.64$ ),  $p = 0.43$ , 95%CI  $[-0.15, 0.17]$ , Hedges's  $g = 0.03$  for language; higher evaluations of VR training ( $M = 7.35$ ) than self-imagination ( $M = 6.89$ ),  $p < 0.01$ , 95%CI  $[0.19, 0.73]$ , Hedges's  $g = 1.02$  for musicality; higher evaluations of VR training ( $M = 7.73$ ) than self-imagination ( $M = 6.98$ ),  $p < 0.01$ , 95%CI  $[0.54, 0.96]$ , Hedges's  $g = 1.49$  for expression. Under positive emotion, participants reported higher evaluations of VR training ( $M = 7.10$ ) than self-imagination ( $M = 6.80$ ),  $p < 0.01$ , 95%CI  $[0.12, 0.48]$ , Hedges's  $g = 0.74$  for coherence; higher evaluations of VR training ( $M = 7.08$ ) than self-imagination ( $M = 6.60$ ),  $p < 0.01$ , 95%CI  $[0.27, 0.68]$ , Hedges's  $g = 0.90$  for breath; higher evaluations of VR training ( $M = 7.39$ ) than self-imagination ( $M = 6.56$ ),  $p < 0.01$ , 95%CI  $[0.58, 1.07]$ , Hedges's  $g = 1.32$  for resonance; higher evaluations of VR training ( $M = 6.45$ ) than self-imagination ( $M = 6.38$ ),  $p = 0.20$ , 95%CI  $[-0.11, 0.26]$ , Hedges's  $g = 0.13$  for intonation; lower evaluations of VR training ( $M = 6.68$ ) than self-imagination ( $M = 6.69$ ),  $p = 0.42$ , 95%CI  $[-0.12, 0.14]$ , Hedges's  $g = 0.03$  for language; higher evaluations of VR training ( $M = 7.41$ ) than self-imagination ( $M = 6.70$ ),  $p < 0.01$ , 95%CI  $[0.51, 0.91]$ , Hedges's  $g = 1.38$  for musicality; higher evaluations of VR training ( $M = 7.75$ ) than self-imagination ( $M = 6.71$ ),  $p < 0.01$ , 95%CI  $[0.78, 1.30]$ , Hedges's  $g = 2.02$  for expression.

### Emotion Classification

In the 2 scenarios of self-imagination and VR training, based on the top 10 features selected by the mRMR algorithm, **Table 5** shows the emotion classification accuracy using the SVM under the optimal parameters (the average of 1,000

**TABLE 3 |** Average scores and statistical test results of the self-rating.

Evaluation	Negative				Neutral				Positive			
	Self	VR	<i>p</i>	ES	Self	VR	<i>p</i>	ES	Self	VR	<i>p</i>	ES
Coherence	2.81	3.25	<i>p</i> = 0.07	0.62	3.25	3.38	<i>p</i> = 0.43	0.16	3.31	3.44	<i>p</i> = 0.34	0.23
Breath	2.69	3.06	<i>p</i> = 0.14	0.54	3.19	3.38	<i>p</i> = 0.19	0.29	3.13	3.31	<i>p</i> = 0.19	0.31
Resonance	2.50	3.06	<i>p</i> < 0.01	0.93	2.88	3.25	<i>P</i> = 0.05	0.63	3.06	3.19	<i>p</i> = 0.50	0.22
Intonation	2.94	3.06	<i>p</i> = 0.54	0.20	3.13	3.31	<i>p</i> = 0.27	0.26	2.69	3.00	<i>p</i> = 0.14	0.44
Language	2.94	3.06	<i>p</i> = 0.43	0.25	3.06	3.38	<i>p</i> = 0.06	0.59	3.00	3.31	<i>p</i> < 0.05	0.92
Musicality	3.16	3.44	<i>p</i> = 0.06	0.50	3.44	3.56	<i>p</i> = 0.33	0.22	3.38	3.31	<i>p</i> = 0.75	0.09
Expression	2.75	3.31	<i>p</i> < 0.01	0.88	3.00	3.44	<i>p</i> < 0.01	0.76	3.25	3.63	<i>p</i> < 0.05	0.63

We rename the performance from seven aspects as coherence, breath, resonance, intonation, language, musicality, and expression, respectively.

**TABLE 4 |** Average score of the subjects and the statistical test results of the third-party evaluation.

Evaluation	Negative				Neutral				Positive			
	Self	VR	<i>p</i>	ES	Self	VR	<i>p</i>	ES	Self	VR	<i>p</i>	ES
Coherence	6.66	6.95	<i>p</i> < 0.01	0.70	6.73	7.00	<i>p</i> < 0.01	0.83	6.80	7.10	<i>p</i> < 0.01	0.74
Breath	6.39	6.98	<i>p</i> < 0.01	1.22	6.65	7.13	<i>p</i> < 0.01	1.09	6.60	7.08	<i>p</i> < 0.01	0.90
Resonance	6.34	6.96	<i>p</i> < 0.01	1.42	6.58	7.38	<i>p</i> < 0.01	1.42	6.56	7.39	<i>p</i> < 0.01	1.32
Intonation	6.18	6.26	<i>p</i> = 0.18	0.16	6.29	6.36	<i>p</i> = 0.22	0.14	6.38	6.45	<i>p</i> = 0.20	0.13
Language	6.66	6.58	<i>p</i> = 0.07	0.17	6.64	6.63	<i>p</i> = 0.43	0.03	6.69	6.68	<i>p</i> = 0.42	0.03
Musicality	6.58	7.34	<i>p</i> < 0.01	1.52	6.89	7.35	<i>p</i> < 0.01	1.02	6.70	7.41	<i>p</i> < 0.01	1.38
Expression	6.61	7.81	<i>p</i> < 0.01	2.21	6.98	7.73	<i>p</i> < 0.01	1.49	6.71	7.75	<i>p</i> < 0.01	2.02

classification accuracy). All bands mean directly splicing all the features of five frequency bands together. In the results of six cases, the emotion classification accuracy of VR training was greater than the accuracy of self-imagining. In the self-imagination scenario, the highest classification accuracy was 82.82% based on all frequency bands. In a VR training scenario, the highest emotion classification accuracy was 85.84% acquired with the gamma frequency band. The best results were bold in Table 5.

## Selected Connectivity Features

This section showed the top 10 most important functional connectivity features selected by the mRMR algorithm in the 2 scenarios and displays them in the form of charts. The specific functional connectivity features and corresponding brain areas are shown in Figure 3 and Table 6. The abbreviations of these brain areas were adopted from Román et al. (2017) study. It can be seen from Table 6 that the functional connectivity features were only distributed in the beta and gamma frequency bands. In both scenarios, three functional connectivity features were located in the beta frequency band and the rest seven were in the gamma frequency band. Also, the asymmetry between the cerebral hemispheres was observed. In the self-imagination scenario, there were 10 brain areas in the right hemisphere and 7 brain areas in the left hemisphere; in the VR training scenario, there were 7 brain areas in the right hemisphere and 10 brain areas in the left hemisphere. Besides, the connectivity features between the two hemispheres were strong.

In both scenarios, there were five connectivity features between the hemispheres.

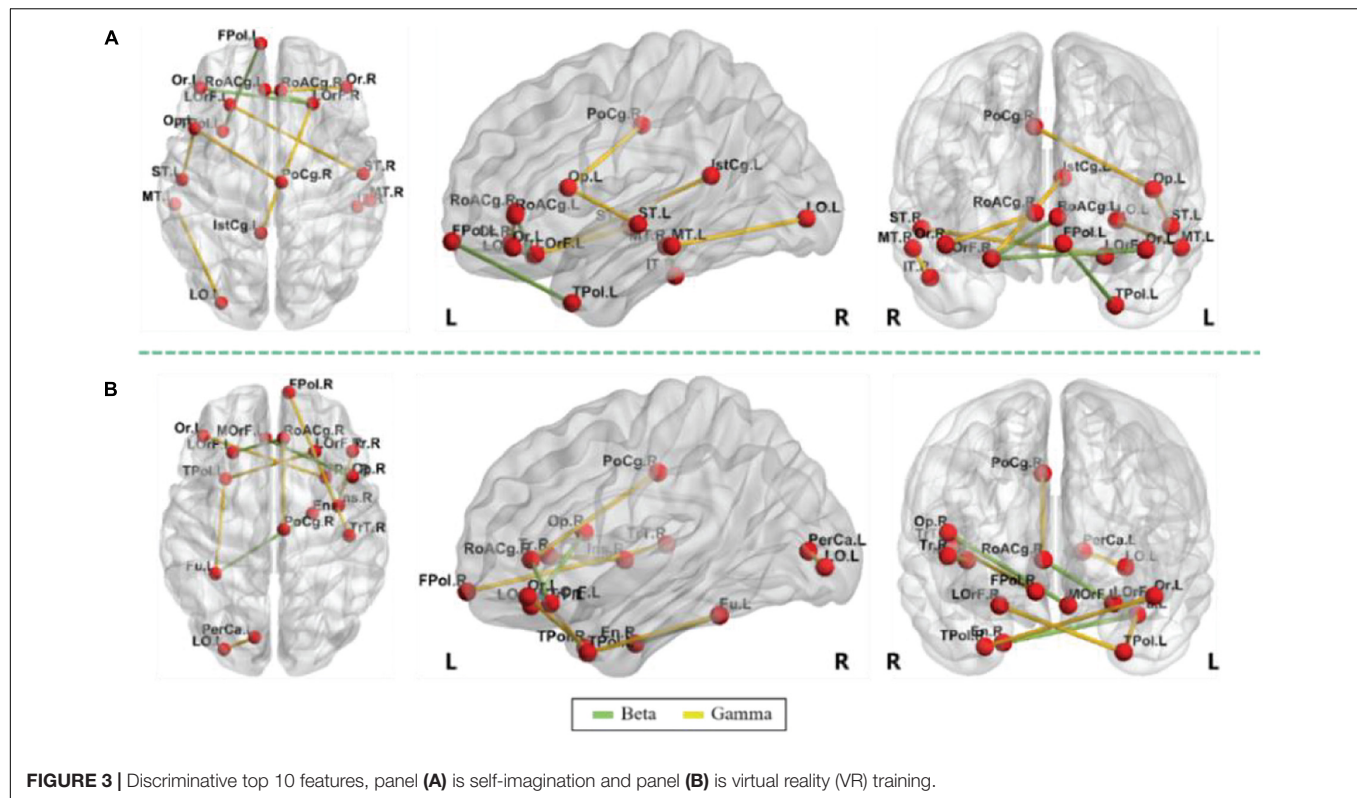
## DISCUSSION

This empirical study conducted a statistical test on self-rating report data and third-party teacher evaluation data for the two methods: self-imagination and VR training. Emotion classification based on the EEG functional connectivity features was also conducted. From the perspective of self-rating reports and third-party evaluation, the differences in an emotional state and singing performance and the differences in functional connectivity were compared and revealed. There were significant differences between the two methods to some extent. The VR training got higher evaluation scores than self-imagination. The emotion classification accuracy in the VR training was higher than that of the self-imagination. These results supported this empirical study's initial hypothesis that compared with the traditional self-imagination method, VR training can better elicit the emotional state of the singers, further improve their singing performance, and provide a new teaching aid method for vocal music teaching.

**TABLE 5 |** Classification accuracy of emotion recognition.

Cases	Delta	Theta	Alpha	Beta	Gamma	All bands
Self-imagination	43.00	47.75	53.00	79.05	80.47	<b>82.82</b>
VR training	45.68	51.80	55.55	80.55	<b>85.84</b>	84.69





**FIGURE 3 |** Discriminative top 10 features, panel (A) is self-imagination and panel (B) is virtual reality (VR) training.

## Self-Rating Scale

In this empirical study, the SAM scale and the vocal self-rating scale were used to rate the emotional state and singing performance. As shown in **Table 2**, regarding the self-rating scales under the negative emotion and the positive emotion, there were significant differences between the self-imagination and the VR training in terms of pleasure and activation. The average scores were in line with the expectations, that is, the VR training under the negative emotion reduced pleasure, while under the positive emotion, it increased pleasure. And, the VR training under both negative and positive increased activation. However, in terms of emotional dominance, there were significant differences between the two methods under the negative emotion, and the VR training reduced dominance, while under the positive emotion, there was no significant difference between the two methods ( $p = 0.33$ ), this may be because VR training under the negative emotion was easier to achieve good emotion-elicited effects (Liao et al., 2019). Since the valence and arousal in the neutral emotion were generally medium, there was no significant difference in terms of pleasure, arousal, and dominance for the two methods.

It can be seen from **Table 3**, regarding the self-rating scale for a vocal performance, there were significant differences in the scores of emotional expression between the two methods. And the average score of the VR training was higher than the average score of self-imagination, indicating that the improvement of emotional activation by VR training could be applied well in the singers' singing emotion.

Regarding the evaluation of third-party professional vocal teachers on singing performance, **Table 4** showed that in terms

of the singing coherence, breath use, resonance use, musical sense, and emotional expression, the singing performance score after VR training was significantly higher than that of the self-imagination. This could be due to that the VR training transmitted emotional information to the subjects in an audio-visual way and enhanced the subjects' emotional state. The emotional information in a relaxed and natural state improved the subjects' singing level. In terms of pronouncing words and intonation rhythm, there was no significant difference between the two methods. The correct pronouncing of words is the premise of language purity and could have an impact on the performance of vocal art (Li, 2016). Singing pronouncing words vary from person to person. The participants in the experiment were all students who have received systematic vocal training. They have undergone long-term professional training in pronouncing words and intonation rhythm. The singing songs had been practiced before the subjects participated in the experiment, and the correct intonation rhythm had also been mastered. It would not be significantly changed in a short time through third-party evaluation data on singing performance.

## Emotion-Related Functional Connectivity

This empirical study classified emotion from the two methods based on the functional connectivity features after source location. The accuracy of emotion classification under the VR training was greater than the self-imagination. Among the five frequency bands, the beta and gamma frequency bands had high classification accuracy, and the top 10 most important functional connectivity features selected were all located in the



**TABLE 6 |** Discriminative top 10 functional connectivity features.

Cases	Selected features
Self-imagination	<p><b>Beta:</b> frontal pole L–temporal pole L (FPol L–TPol L), lateral orbitofrontal R–pars orbitalis L (LORF R–Or L), lateral orbitofrontal R–rostral anterior cingulate L (LORF R–RoACg L).</p> <p><b>Gamma:</b> inferior temporal R–middle temporal R (IT R–MT R), isthmus cingulate L–lateral orbitofrontal R (IstCg L–LORF R), lateral occipital L–middle temporal L (LO L–MT L), lateral orbitofrontal L–superior temporal R (LORF L–ST R), pars opercularis L–posterior cingulate R (Op L–PoCg R), pars opercularis L–superior temporal L (Op L–ST L), pars orbitalis R–rostral anterior cingulate R (Or R–RoACg R).</p>
VR training	<p><b>Beta:</b> entorhinal R–fusiform L (En R–Fu L), lateral orbitofrontal L–rostral anterior cingulate R (LORF L–RoACg R), medial orbitofrontal L–pars opercularis R (MORF L–Op R).</p> <p><b>Gamma:</b> frontal pole R–transverse temporal R (FPol R–TrT R), fusiform L–temporal pole L (Fu L–TPol L), insula R–pars triangularis R (Ins R–Tr R), lateral occipital L–pericalcarine L (LO L–PerCa L), lateral orbitofrontal R–temporal pole L (LORF R–TPol L), pars orbitalis L–temporal pole R (Or L–TPol R), posterior cingulate R–rostral anterior cingulate R (PoCg R–RoACg R).</p>

beta and gamma frequency bands. This was consistent with previous studies, indicating that the beta and gamma bands of brain activity are more related to emotional processing than other frequency bands (Zheng and Lu, 2015; Zheng et al., 2017; Li et al., 2018).

Among the functional connectivity features for the self-imagination method, the right lateral orbitofrontal was related to both beta and gamma bands, which was consistent with the findings of previous studies. The lateral prefrontal cortex controls the experience and expression of emotion, especially playing an important role in the processing of negative emotion (Hooker and Knight, 2006). There were also four brain regions in the beta frequency band, including the left frontal pole, left temporal pole, left pars orbitalis, and left rostral anterior cingulate. There were 13 brain regions in the gamma band, right inferior temporal, left and right middle temporal, left isthmus cingulate, left lateral occipital, left lateral orbitofrontal, left and right superior temporal, left pars opercularis, right posterior cingulate, right pars orbitalis, right rostral anterior cingulate. The prefrontal cortex played an important role in emotional processing (Esslen et al., 2004). The temporal cortex was involved in the processing of humor, laughter, and smiles (Wild et al., 2006). The rostral anterior cingulate had a regulatory effect on the brain regions that produce emotional responses and can be enhanced by emotional valence (Chiu et al., 2008). The pars opercularis played a role in regulating the perception of emotional rhythm (Patel et al., 2018). Emotional stimulation activated the posterior cingulate cortex (Maddock et al., 2003) and the pars orbitalis (Krautheim et al., 2020).

Among the functional connectivity features for the VR training method, the left fusiform and the right rostral anterior cingulate were related to both beta and gamma bands. This could

be due to the fusiform provided perception for participants in VR scenes to detect the difference between the virtual environment and the real physical world (Garcia et al., 2012), and the right rostral anterior cingulate regulated emotion (Chiu et al., 2008; Alvarez et al., 2011). There were also five brain regions in the beta band, including the right entorhinal, left lateral orbitofrontal, right rostral anterior cingulate, and left medial orbitofrontal, and right pars opercularis. There were 13 brain regions in the gamma band, including the left frontal pole, right transverse temporal, left temporal pole, right insula, and right inferior frontal pole, right pars triangularis, left lateral occipital, left pericalcarine, right lateral orbitofrontal, left pars orbitalis, right temporal pole, and right posterior cingulate. The entorhinal cortex was a key neural structure for spatial navigation (Jacobs et al., 2010). Many studies have studied the activation and changes of the entorhinal cortex in the VR environment (Jacobs et al., 2010; Howett et al., 2019). The medial frontal cortex (including the medial orbitofrontal cortex) played a crucial role in the processing of negative emotion (Sitaram et al., 2011). Insula activation involves emotional processing, and studies have shown that the posterior and anterior sub-regions of the human insula showed obvious response characteristics to auditory emotional stimuli (Zhang Y. et al., 2019).

Seven brain regions existed in the features selected for the two methods, including the right rostral anterior cingulate, left lateral occipital, right lateral orbitofrontal, left pars orbitalis, right posterior cingulate, left lateral orbitofrontal, and left temporal pole, were all related to the activation and processing of emotion (Maddock et al., 2003; Esslen et al., 2004; Wild et al., 2006; Güntekin and Basar, 2007; Chiu et al., 2008; Krautheim et al., 2020). The brain regions and functional connectivity features involved in this research provided references for research on brain emotion activation mechanisms, the impact of VR on the brain, and EEG-based emotion recognition.

## Limitations

One limitation of the empirical study is that the common template (ICBM152) was not used to obtain the head model and source space representation of all participants. Studies have shown that using an individualized template to construct a head model and source space for each participant can obtain more accurate results in the source location process (Lei et al., 2011; Hassan and Wendling, 2018). The individualized template helps further improve the performance of emotion recognition and discover more physiologically relevant features. Other network analysis methods, such as Phase Lock Value (PLV) (Li et al., 2019), Phase Slope Index (PSI) (Basti et al., 2018), Partially Directed Coherence (PDC) (Wang et al., 2020) can also be used for functional network analysis after source location. Therefore, individualized templates will be used in future work to establish a head model in source space, and other functional connectivity construction methods will be used to improve EEG-based emotion recognition.

This study demonstrated the utility of VR for eliciting emotions to trainees-singers and further improving their performance. However, the cost of VR equipment may still be

unaffordable for some institutions with a limited budget. Another limitation is that EEG signals may be easily disturbed by singing and/or moving actions. Finally, the sample size of this empirical study was relatively small. Future attempts should strive to address these limitations and collect data from a larger sample.

Furthermore, the adverse symptoms and effects (i.e., cybersickness) during VR training may undermine the health and safety standards, and compromise the reliability of the experimental results (Kourtesis et al., 2019a). The current study did not examine the symptoms of cybersickness, while they may negatively affect cognitive and behavioral performance, as well as EEG data (Kourtesis et al., 2019a; Weech et al., 2019). Also, this empirical study did not examine the immersive user experience. These factors play a central role in the efficiency of the VR experience (Slater, 2009; Skarbez et al., 2017) and the incidence of cybersickness (Kourtesis et al., 2019b; Weech et al., 2019). Future attempts should consider the administration of recently published cybersickness questionnaires (Somrak et al., 2021).

## CONCLUSION

In this study, a VR training task was implemented to elicit emotions from singers and assist them with further improving their emotional singing performance. The VR training method was compared against the traditional self-imagination method. By conducting a two-stage experiment, the two methods were compared in terms of emotions' elicitation and emotional singing performance. In the first stage, electroencephalographic (EEG) data were collected from the subjects. In the second stage, self-rating reports and third-party teachers' evaluations were collected. The EEG data were analyzed by adopting the max-relevance and min-redundancy algorithm for feature selection and the SVM for emotion recognition. The experimental results have validated that VR can better elicit the positive, neutral, and negative emotional states of singers than self-imagination, and further improve the singing performance due to the

improvement of emotional activation. As such, we argue the VR training method can be seen as an effective approach that will improve and complement the available vocal music teaching methods.

## 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

The studies involving human participants were reviewed and approved by the Ethics Committee of the Affiliated Nanjing Brain Hospital, Nanjing Medical University Nanjing. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

JZ and DZ proposed the research topic and outline. ZX, YZ, PW, and XX designed the procedures of experiments, collected, and analyzed the data. ZX, YZ, and PF contributed to the structuring and writing of the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the National Key Research and Development Program of China (Nos. 2018YFC2001600 and 2018YFC2001602), the National Natural Science Foundation of China (Nos. 61876082 and 61861130366), and the Fundamental Research Funds for the Central Universities (No. ND2021009).

## 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
- Alvarez, R. P., Chen, G., Bodurka, J., Kaplan, R., and Grillon, C. (2011). Phasic and sustained fear in humans elicits distinct patterns of brain activity. *Neuroimage* 55, 389–400. doi: 10.1016/j.neuroimage.2010.11.057
- Baig, M. Z., and Kavakli, M. (2018). "Analyzing novice and expert user's cognitive load in using a multi-modal interface system," in *Proceedings of the 2018 26th International Conference on Systems Engineering (ICSEng, Sydney*, 1–7.
- Bashivan, P., Yeasin, M., and Bidelman, G. M. (2015). "Single trial prediction of normal and excessive cognitive load through EEG feature fusion," in *Proceedings 2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, (Piscataway, NJ: IEEE), 1–5.
- Basti, A., Pizzella, V., Chella, F., Romani, G. L., Nolte, G., and Marzetti, L. (2018). Disclosing large-scale directed functional connections in MEG with the multivariate phase slope index. *Neuroimage* 175, 161–175. doi: 10.1016/j.neuroimage.2018.03.004
- Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exper. Psychiatry* 25, 49–59. doi: 10.1016/0005-7916(94)90063-9
- Calabrò, R. S., Naro, A., Russo, M., and Bramanti, P. (2017). The role of virtual reality in improving motor performance as revealed by EEG: a randomized clinical trial. *J. Neuroeng. Rehabil.* 14, 1–16. doi: 10.1186/s12984-017-0268-4
- Chiu, P. H., Holmes, A. J., and Pizzagalli, D. A. (2008). Dissociable recruitment of rostral anterior cingulate and inferior frontal cortex in emotional response inhibition. *Neuroimage* 42, 988–997. doi: 10.1016/j.neuroimage.2008.04.248
- Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Desikan, R. S., Ségonne, F., Fischl, B., and Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* 31, 968–980. doi: 10.1016/j.neuroimage.2006.01.021
- Dimitrakopoulos, G. N., Kakkos, I., Dai, Z., Lim, J., deSouza, J. J., Bezerianos, A., et al. (2017). Task-independent mental workload classification based upon common multiband EEG cortical connectivity. *IEEE Trans. Neural Syst. Rehabil. Eng.* 25, 1940–1949. doi: 10.1109/TNSRE.2017.2701002

- Esslen, M., Pascual-Marqui, R. D., Hell, D., Kochi, K., and Lehmann, D. (2004). Brain areas and time course of emotional processing. *Neuroimage* 21, 1189–1203. doi: 10.1016/j.neuroimage.2003.10.001
- Gao, Y. J. (2013). Virtual technology in relation to teaching and research of art. *J. Hebei Norm. Univ.* 15, 95–96.
- Garcia, L., Kartolo, A., and Méthot-Curtis, E. (2012). “A discussion of the use of virtual reality in dementia,” in *Virtual Reality in Psychological, Medical and Pedagogical Applications*, ed. E. Christiane (Rijeka: IntechOpen), 123–136.
- Güntekin, B., and Başar, E. (2007). Emotional face expressions are differentiated with brain oscillations. *Intern. J. Psychophysiol.* 64, 91–100. doi: 10.1016/j.ijpsycho.2006.07.003
- Hajcak, G., MacNamara, A., and Olvet, D. M. (2010). Event-related potentials, emotion, and emotion regulation: an integrative review. *Dev. Neuropsychol.* 35, 129–155. doi: 10.1080/87565640903526504
- Halim, Z., and Rehan, M. (2020). On identification of driving-induced stress using electroencephalogram signals: a framework based on wearable safety-critical scheme and machine learning. *Inform. Fusion* 53, 66–79. doi: 10.1016/j.inffus.2019.06.006
- Hassan, M., and Wendling, F. (2018). Electroencephalography source connectivity: aiming for high resolution of brain networks in time and space. *IEEE Signal Process. Magaz.* 35, 81–96. doi: 10.1109/MSP.2017.2777518
- Hooker, C. I., and Knight, R. T. (2006). “The role of lateral Orbitofrontal cortex in the inhibitory control of emotion,” in *The Orbitofrontal Cortex*, eds. D. Zald, and S. Rauch (Oxford: Oxford University Press), 307–324.
- Howett, D., Castegnaro, A., Krzywicka, K., and Chan, D. (2019). Differentiation of mild cognitive impairment using an entorhinal cortex-based test of virtual reality navigation. *Brain* 142, 1751–1766. doi: 10.1093/brain/awz116
- Huang, H., Xie, Q., Pan, J., He, Y., Wen, Z., Yu, R., et al. (2019). An EEG-based brain computer interface for emotion recognition and its application in patients with disorder of consciousness. *IEEE Trans. Affect. Comput.* 1–1. doi: 10.1109/TAFFC.2019.2901456
- Huang, X. Y., and Chen, B. J. (2017). Virtual reality art design strategy based on immersion communication. *Modern Commun.* 39, 85–89. doi: 10.3969/j.issn.1007-8770.2017.01.015
- Jacobs, J., Kahana, M. J., Ekstrom, A. D., Mollison, M. V., and Fried, I. (2010). A sense of direction in human entorhinal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 107, 6487–6492. doi: 10.1073/pnas.0911213107
- Kourtesis, P., Collina, S., Doumas, L. A., and MacPherson, S. E. (2019a). Technological competence is a pre-condition for effective implementation of virtual reality head mounted displays in human neuroscience: a technological review and meta-analysis. *Front. Hum. Neurosci.* 13:342. doi: 10.3389/fnhum.2019.00342
- Kourtesis, P., Collina, S., Doumas, L. A., and MacPherson, S. E. (2019b). Validation of the virtual reality neuroscience questionnaire: maximum duration of immersive virtual reality sessions without the presence of pertinent adverse symptomatology. *Front. Hum. Neurosci.* 13:417. doi: 10.3389/fnhum.2019.00417
- Kourtesis, P., Korre, D., Collina, S., Doumas, L. A., and MacPherson, S. E. (2020). Guidelines for the development of immersive virtual reality software for cognitive neuroscience and neuropsychology: the development of virtual reality everyday assessment lab (VR-EAL), a neuropsychological test battery in immersive virtual reality. *Front. Comput. Sci.* 1:12. doi: 10.3389/fcomp.2019.00012
- Krauthaim, J. T., Steines, M., Dannlowski, U., and Kircher, T. (2020). Emotion specific neural activation for the production and perception of facial expressions. *Cortex* 127, 17–28. doi: 10.1016/j.cortex.2020.01.026
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* 4:863. doi: 10.3389/fpsyg.2013.00863
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annu. Rev. Neurosci.* 23, 155–184. doi: 10.1146/annurev.neuro.23.1.155
- Lei, X., Xu, P., Luo, C., Zhao, J., Zhou, D., and Yao, D. (2011). fMRI functional networks for EEG source imaging. *Hum. Brain Mapp.* 32, 1141–1160. doi: 10.1002/hbm.21098
- Li, J., Zhang, Z., and He, H. (2018). Hierarchical convolutional neural networks for EEG-based emotion recognition. *Cogn. Comput.* 10, 368–380. doi: 10.1007/s12559-017-9533-x
- Li, P., Liu, H., Si, Y., and Xu, P. (2019). EEG based emotion recognition by combining functional connectivity network and local activations. *IEEE Trans. Biomed. Eng.* 66, 2869–2881. doi: 10.1109/TBME.2019.2897651
- Li, W., Hu, X., Long, X., Tang, L., Chen, J., Wang, F., et al. (2020). EEG responses to emotional videos can quantitatively predict big-five personality traits. *Neurocomputing* 415, 368–381. doi: 10.1016/j.neucom.2020.07.123
- Li, Y. (2016). Some views on articulation in vocal music performance. *N. Music* 36, 54–55.
- Liao, D., Shu, L., Liang, G., Li, Y., Zhang, Y., Zhang, W., et al. (2019). Design and evaluation of affective virtual reality system based on multimodal physiological signals and self-assessment manikin. *IEEE J. Electromagnet. RF Microwav. Med. Biol.* 4, 216–224. doi: 10.1109/JERM.2019.2948767
- Lin, T., Gilam, G., Raz, G., Or-Borichev, A., Bar-Haim, Y., Fruchter, E., et al. (2017). Accessible neurobehavioral anger-related markers for vulnerability to post-traumatic stress symptoms in a population of male soldiers. *Front. Behav. Neurosci.* 11:38. doi: 10.3389/fnbeh.2017.00038
- Maddock, R. J., Garrett, A. S., and Buonocore, M. H. (2003). Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. *Hum. Brain Mapp.* 18, 30–41. doi: 10.1002/hbm.10075
- Mosher, J. C., and Leahy, R. M. (1999). Source localization using recursively applied and projected (RAP) MUSIC. *IEEE Trans. Signal Process.* 47, 332–340. doi: 10.1109/78.740118
- Mu, J. (2011). Analysis of psychological factors affecting singing. *Literat. Life* 8, 115–116.
- Pascual-Marqui, R. D. (2002). Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Methods Find Exp. Clin. Pharmacol.* 24, 5–12.
- Patel, S., Oishi, K., Wright, A., Sutherland-Foggio, H., Saxena, S., Sheppard, S. M., et al. (2018). Right hemisphere regions critical for expression of emotion through prosody. *Front. Neurol.* 9:224. doi: 10.3389/fneur.2018.00224
- Pei, X. Y. (2017). Requirements and application of virtual reality technology in the modern environmental art design. *Automat. Instrument.* 6, 216–217. doi: 10.14016/j.cnki.1001-9227.2017.06.216
- Peng, H., Long, F., and Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Patt. Analys. Mach. Intellig.* 27, 1226–1238. doi: 10.1109/TPAMI.2005.159
- Pion-Tonachini, L., Kreutz-Delgado, K., and Makeig, S. (2019). ICLabel: an automated electroencephalographic independent component classifier, dataset, and website. *Neuroimage* 198, 181–197. doi: 10.1016/j.neuroimage.2019.05.026
- Román, C., Guevara, M., Valenzuela, R., and Guevara, P. (2017). Clustering of whole-brain white matter short association bundles using HARDI data. *Front. Neuroinform.* 11:73. doi: 10.3389/fninf.2017.00073
- Shi, W. Z. (2002). *Fundamentals of Vocal Music*. Beijing: People's Music Publishing House.
- Sitaram, R., Lee, S., Ruiz, S., Rana, M., Veit, R., and Birbaumer, N. (2011). Real-time support vector classification and feedback of multiple emotional brain states. *Neuroimage* 56, 753–765. doi: 10.1016/j.neuroimage.2010.08.007
- Skarbez, R., Brooks, F. P. Jr., and Whitton, M. C. (2017). A survey of presence and related concepts. *ACM Comput. Surv.* 50, 1–39. doi: 10.1145/3134301
- Slater, M. (2009). Place illusion and plausibility can lead to realistic behaviour in immersive virtual environments. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 3549–3557. doi: 10.1098/rstb.2009.0138
- Somrak, A., Pogačnik, M., and Guna, J. (2021). Suitability and comparison of questionnaires assessing virtual reality-induced symptoms and effects and user experience in virtual environments. *Sensors* 21:1185. doi: 10.3390/s21041185
- Sood, S. K., and Singh, K. D. (2018). An optical-Fog assisted EEG-based virtual reality framework for enhancing E-learning through educational games. *Comput. Appl. Eng. Educ.* 26, 1565–1576. doi: 10.1002/cae.21965
- Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., and Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intellig. Neurosci.* 2011:879716. doi: 10.1155/2011/879716
- Tarrant, J., Viczko, J., and Cope, H. (2018). Virtual reality for anxiety reduction demonstrated by quantitative EEG: a pilot study. *Front. Psychol.* 9:1280. doi: 10.3389/fpsyg.2018.01280
- Wang, H., Wu, X., and Yao, L. (2020). Identifying cortical brain directed connectivity networks from high-density EEG for emotion recognition. *IEEE Trans. Affect. Comput.* 1–1. doi: 10.1109/TAFFC.2020.3006847

- Wang, X. W., Nie, D., and Lu, B. L. (2014). Emotional state classification from EEG data using machine learning approach. *Neurocomputing* 129, 94–106. doi: 10.1016/j.neucom.2013.06.046
- Weech, S., Kenny, S., and Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: a review. *Front. Psychol.* 10:158. doi: 10.3389/fpsyg.2019.00158
- Wild, B., Rodden, F. A., Rapp, A., Erb, M., Grodd, W., and Ruch, W. (2006). Humor and smiling: cortical regions selective for cognitive, affective, and volitional components. *Neurology* 66, 887–893. doi: 10.1212/01.wnl.0000203123.68747.02
- Xi, X. (2010). Two aspects of emotional expression in vocal art. *Home Drama* 7:21.
- Xu, N. (2008). On the harmonious unity of aesthetic and psychological pleasure in vocal music teaching. *Sci. Educ. Article Collects* 8:185. doi: 10.3969/j.issn.1672-7894.2008.24.158
- Yang, X. Z., and Ren, Y. Q. (2019). Development of virtual reality and EEG linkage system and exploration of its educational research function. *J. Dist. Educ.* 37, 47–54.
- Zhang, G. H., Yu, M. J., Chen, G., and Liu, Y. J. (2019). A review of EEG features for emotion recognition. *Sci. China* 49, 1097–1118. doi: 10.1360/N112018-00337
- Zhang, Y., Zhou, W., Wang, S., and Wang, X. (2019). The roles of subdivisions of human insula in emotion perception and auditory processing. *Cereb. Cortex* 29, 517–528. doi: 10.1093/cercor/bhx334
- Zhang, Z. D. (2018). Research on the application of virtual reality technology in art education in digital context. *Sichuan Theatre* 12, 186–188.
- 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
- Zheng, W. L., Zhu, J. Y., and Lu, B. L. (2017). Identifying stable patterns over time for emotion recognition from EEG. *IEEE Trans. Affect. Comput.* 10, 417–429. doi: 10.1109/TAFFC.2017.2712143

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships 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.

Copyright © 2021 Zhang, Xu, Zhou, Wang, Fu, Xu and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.





# Identification of Emotion Using Electroencephalogram by Tunable Q-Factor Wavelet Transform and Binary Gray Wolf Optimization

Siyu Li<sup>1,2</sup>, Xiaotong Lyu<sup>1,2</sup>, Lei Zhao<sup>2,3</sup>, Zhuangfei Chen<sup>2,4</sup>, Anmin Gong<sup>5\*</sup> and Yunfa Fu<sup>1,2,4,6\*</sup>

<sup>1</sup> School of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, China,

<sup>2</sup> Brain Cognition and Brain-Computer Intelligence Integration Group, Kunming University of Science and Technology, Kunming, China, <sup>3</sup> Faculty of Science, Kunming University of Science and Technology, Kunming, China, <sup>4</sup> School

of Medicine, Center for Brain Science and Visual Cognition, Kunming University of Science and Technology, Kunming, China,

<sup>5</sup> College of Information Engineering, Engineering University of PAP, Xi'an, China, <sup>6</sup> Computer Technology Application Key Lab of Yunnan Province, Kunming University of Science and Technology, Kunming, China

## OPEN ACCESS

### Edited by:

Jane Zhen Liang,  
Shenzhen University, China

### Reviewed by:

Dong Wen,  
Yanshan University, China  
Peng Xu,  
University of Electronic Science  
and Technology of China, China

### \*Correspondence:

Yunfa Fu  
fyf@ynu.edu.cn  
Anmin Gong  
gonganmincapf@163.com

**Received:** 01 July 2021

**Accepted:** 16 August 2021

**Published:** 08 September 2021

### Citation:

Li S, Lyu X, Zhao L, Chen Z, Gong A and Fu Y (2021) Identification of Emotion Using Electroencephalogram by Tunable Q-Factor Wavelet Transform and Binary Gray Wolf Optimization. *Front. Comput. Neurosci.* 15:732763. doi: 10.3389/fncom.2021.732763

Emotional brain-computer interface based on electroencephalogram (EEG) is a hot issue in the field of human-computer interaction, and is also an important part of the field of emotional computing. Among them, the recognition of EEG induced by emotion is a key problem. Firstly, the preprocessed EEG is decomposed by tunable-Q wavelet transform. Secondly, the sample entropy, second-order differential mean, normalized second-order differential mean, and Hjorth parameter (mobility and complexity) of each sub-band are extracted. Then, the binary gray wolf optimization algorithm is used to optimize the feature matrix. Finally, support vector machine is used to train the classifier. The five types of emotion signal samples of 32 subjects in the database for emotion analysis using physiological signal dataset is identified by the proposed algorithm. After 6-fold cross-validation, the maximum recognition accuracy is 90.48%, the sensitivity is 70.25%, the specificity is 82.01%, and the Kappa coefficient is 0.603. The results show that the proposed method has good performance indicators in the recognition of multiple types of EEG emotion signals, and has a better performance improvement compared with the traditional methods.

**Keywords:** emotion recognition, emotional brain-computer interface, tunable-Q wavelet transform, binary gray wolf optimization algorithm, EEG

## INTRODUCTION

Emotion is a psychological phenomenon mediated by the subject's needs and desires. It has three components: physiological arousal, subjective experience, and external manifestation (Peng, 2004). Emotions have an important impact on people's production and life, physical and mental health, and interpersonal relationships. For example, for patients with depression or schizophrenia, abnormal emotions are the main clinical manifestations. If negative emotions can be identified before the onset, medical staff can intervene and treat in time. For the field of human-computer interaction, computer recognition can be realized, understand and adapt to human

emotions, the human-computer interaction environment is more natural (Nie et al., 2012). Therefore, the decoding and recognition of emotions is an important research goal in the field of emotion computing.

Common emotion recognition methods are mainly divided into two categories: recognition based on non-physiological signals and recognition based on physiological signals. Recognition based on non-physiological signals mainly includes expression recognition and speech recognition, but these two methods have the risk of artificial disguise. In contrast, physiological signals can objectively reflect the true emotional state of a person. Physiological signals caused by emotions include heart rate, respiration, skin temperature, electromyography, electroencephalogram (EEG), and so on. Among them, EEG is not easy to be disguised, and the recognition rate is higher than other physiological signal recognition methods, so it is increasingly used in emotion recognition research (Nie et al., 2012).

Brain-computer interface (BCI) directly connects the brain and external devices, and realizes the information exchange between the brain and the device by decoding EEG (Wolpaw et al., 2000). With the rapid development of BCI and emotional computing, emotional BCI (e-BCI) that automatically recognize emotions have received extensive attention from all walks of life (Fattouh et al., 2013). Among them, decoding the individual's emotional state from EEG information is the core content and key technology of the e-BCI (Molina et al., 2009).

So far, there are many EEG-based emotion recognition methods, and wavelet transform is one of the widely used ones. For example (Asghar et al., 2020) used the wavelet transform method to represent the EEG as a two-dimensional time-frequency distribution image, and then used a neural network method based on deep feature clustering (DFC) to evaluate the emotional state of the subjects, and achieved the recognition accuracy of 81.3% for four types of emotional states. On the basis of wavelet transform (Zhou et al., 2020), extracted Mel-frequency cepstral coefficient (MFCC) features, fused EEG features, and used deep residual network (Resnet18) to recognize two kinds of emotions in wake-up and price effect dimensions, with recognition accuracy of 86.01 and 85.46%. Luo et al. (2020) studied three algorithms of discrete wavelet transform (DWT), variance and fast fourier transform (FFT) to extract features of EEG signals, and spike neural network (SNN) to further classify the emotion signal, the two types of recognition accuracy of valence, arousal, dominance, and liking dimensions are 74, 78, 80, and 86.27%, respectively. Mohammadpour et al. (2017) used DWT to extract features, and then used artificial neural networks (ANN) performs emotion classification and achieves a recognition accuracy of 55.58% for six types of emotional states. Wei et al. (2020) used dual tree-complex wavelet transform (DT-CWT) to decompose and reconstruct EEG, and then extract features from time domain, frequency domain and non-linear analysis and use different integration strategies to obtain the recognition accuracy of the three types of emotions is 83.13%.

Although the wavelet transform can perform positioning in the time domain and the frequency domain at the same time, it is very convenient to perform the round-trip transform

between the time domain and the frequency domain for time-varying signals, but a single wavelet basis function of the wavelet transform is difficult to accurately represent the local characteristics of the signal. It is easy to lose the original time domain characteristics when reconstructing the signal. Therefore, a new tunable Q-factor wavelet transform (TQWT) has been proposed in recent years (Selesnick, 2011). Compared with traditional wavelet transform, TQWT is more flexible and can better reflect complex oscillation signals including EEG by adjusting parameters, so it has quickly attracted the attention of scholars in related fields.

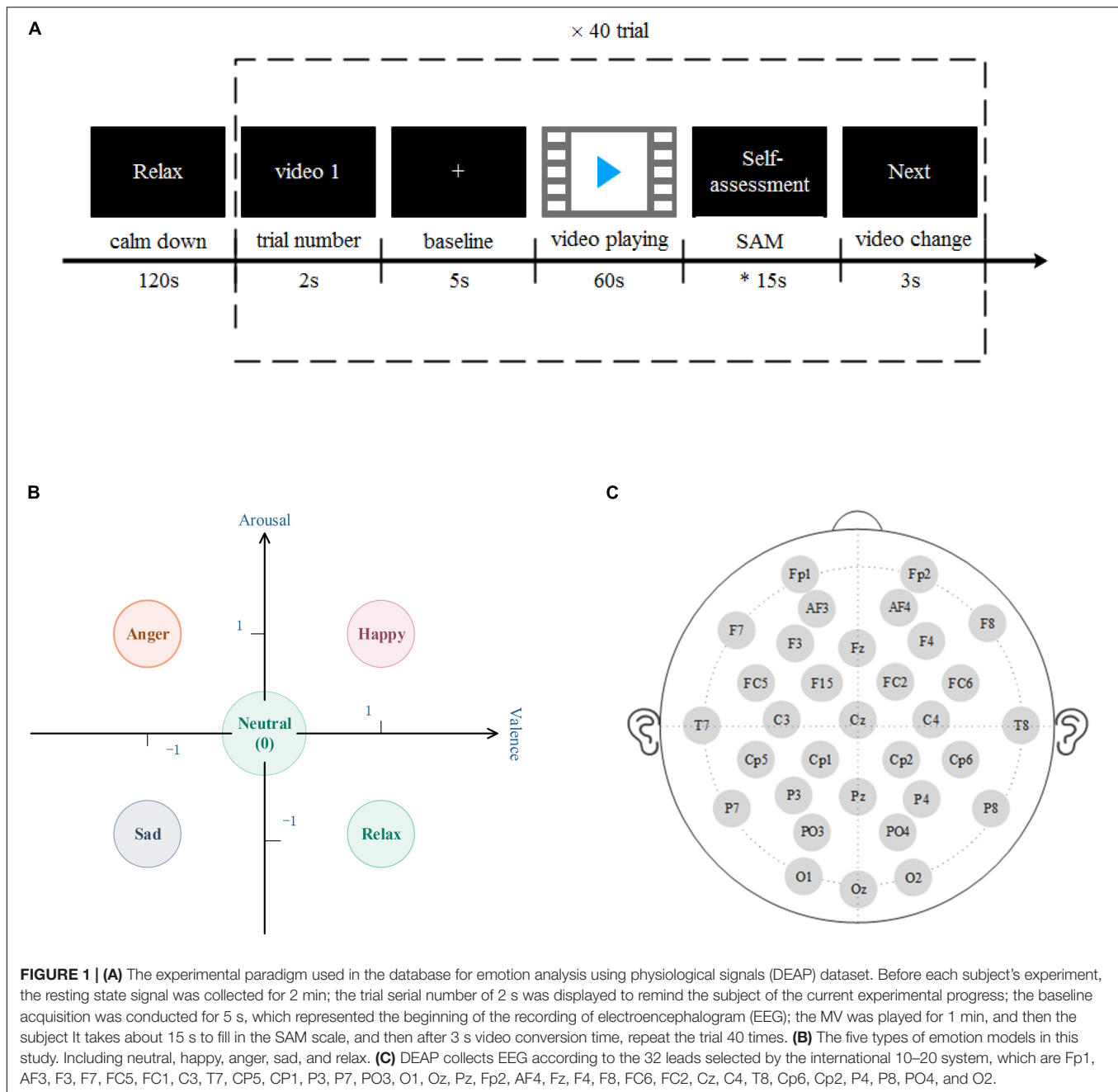
However, the decomposition of the signal also increases the amount of data to identify the features, which affects the performance of the system. This study introduces a feature selection algorithm to solve this problem. Traditional feature selection methods include principal component analysis (PCA), least absolute shrinkage and selection operator (LASSO), and recursive feature elimination (RFE), etc. (Mao et al., 2007). Among them, the binary gray wolf optimization (BGWO; Too et al., 2018) is an improved version of the gray wolf optimization (GWO; Sm et al., 2014), which was also inspired by the prey hunting activities of the gray wolf. An optimized search method of, it has the characteristics of strong convergence performance, few parameters, and easy implementation. It has been used in many fields by many researchers.

Therefore, this manuscript proposes a joint EEG recognition algorithm based on TQWT and BGWO. The algorithm first decomposes the sub-band from the original emotional EEG, and then extracts the signal sample entropy, Hjorth parameter, second-order difference mean and normalized second-order difference mean as features, and then optimizes the feature set through BGWO, and finally input to support vector machine (SVM) for classification. The follow-up structure of this article is as follows: First, the experimental materials and methods are described, including; the relevant description of the experimental data, the basic process of the TQWT algorithm, the feature extraction index, the basic process of the BGWO algorithm, and the classifier and algorithm evaluation index. The result part shows the classification effect of the algorithm on the data set, the analysis of the influence of different decomposition sub-bands and experimental parameters on the experimental results, and the comparative analysis with the classification effect of the classic algorithm. Finally, the experimental results are summarized and discussed.

## MATERIALS AND METHODS

### Experimental Data and Preprocessing

This research uses a database for emotion analysis using physiological signals (DEAP; Koelstra, 2012), and its experimental paradigm is shown in **Figure 1A**. The DEAP data set includes the multi-modal physiological signals induced by 32 subjects watching 40–60-s music video materials and the subjects' ratings of the video's valence, arousal, dominance, and liking. Among them, the physiological signals include: 32 channels of EEG, eight channels of peripheral physiological



**FIGURE 1 | (A)** The experimental paradigm used in the database for emotion analysis using physiological signals (DEAP) dataset. Before each subject's experiment, the resting state signal was collected for 2 min; the trial serial number of 2 s was displayed to remind the subject of the current experimental progress; the baseline acquisition was conducted for 5 s, which represented the beginning of the recording of electroencephalogram (EEG); the MV was played for 1 min, and then the subject It takes about 15 s to fill in the SAM scale, and then after 3 s video conversion time, repeat the trial 40 times. **(B)** The five types of emotion models in this study. Including neutral, happy, anger, sad, and relax. **(C)** DEAP collects EEG according to the 32 leads selected by the international 10–20 system, which are Fp1, AF3, F7, FC5, FC1, C3, T7, CP5, CP1, P3, P7, PO3, O1, Oz, Pz, Fp2, AF4, Fz, F4, F8, FC6, FC2, Cz, C4, T8, Cp6, Cp2, P4, P8, PO4, and O2.

signals: ① current skin response, ② skin temperature, ③ blood volume pulse, ④ respiration, EMG ⑤ main muscles and ⑥ trapezius, ⑦ horizontal, and ⑧ vertical electrooculograms (EOGs). In terms of subjective evaluation, the experiment used self-assessment manikin (SAM; Morris, 1995) with a scale of 1–9 to quantify the participants' ratings of the value, arousal, advantage, and liking of video-induced emotions.

In this study, we set the threshold to 3, and divide each emotion sample into three levels according to the 9 scales of valence and arousal, 1–3, 4–6, and 7–9, respectively, mapped to “–1,” “0,” and “1” on the rectangular coordinate system, five types of emotion recognition are performed in two dimensions

(as shown in **Figure 1B**; Fang et al., 2021), each type. The emotion setting rules are as follows:

- Happy (label 1): valence = 1 and arousal = 1
- Anger (label 2): valence = 1 and arousal = –1
- Sad (label 3): valence = –1 and arousal = –1
- Relax (label 4): valence = –1 and arousal = 1
- Neutral (label 5): valence = 0 or arousal = 0

In this study, a 32-channel EEG in the data set was selected for emotion recognition. The position of the EEG channel is shown in **Figure 1C**. Downsample the

EEG data to 128 HZ, remove the EOGs artifacts, filter the signal to 4–45 HZ through a band-pass filter, and perform a whole-brain average reference. Each piece of data includes 60 s video-induced EEG data and 3 s video conversion.

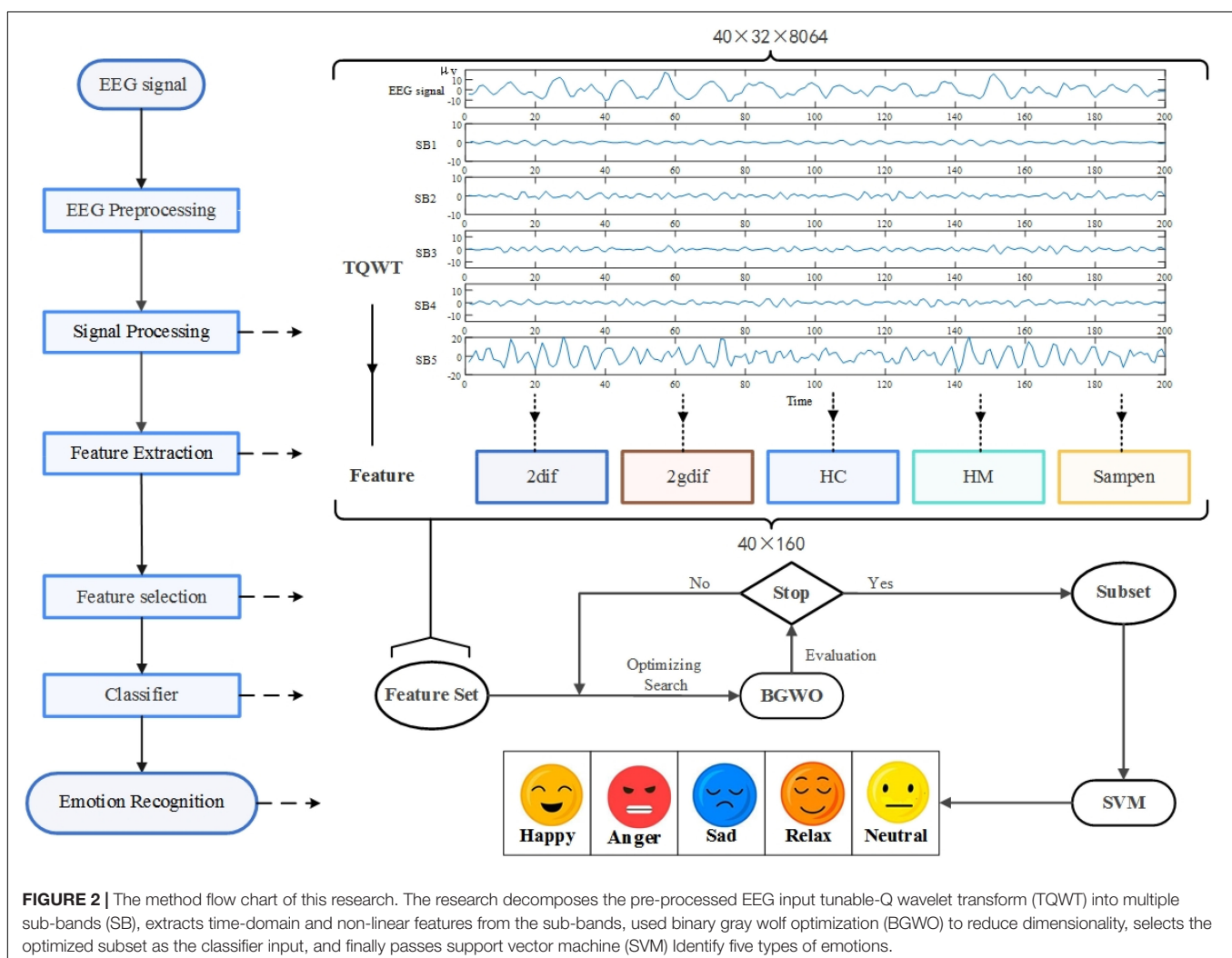
The shape of the preprocessed EEG data of the 32 subjects is  $\text{trial} \times \text{channel} \times \text{data}$ , which is  $40 \times 32 \times 8,064$ ; the shape of the label data is  $\text{trial} \times \text{label}$  (1–5), which is  $40 \times 1$ .

## Method

The algorithm flow is shown in **Figure 2**. In this study, the original EEG was preprocessed and decomposed into multiple sub-bands through TWQT, and then five features of sample entropy, second-order difference mean, normalized second-order difference mean, mobility and complexity were extracted from each sub-band. And then use BGWO to reduce the dimensionality of the feature set, and finally identify the five types of emotions: neutral, happy, anger, sad and relax through SVM.

## Tunable-Q Wavelet Transform

Tunable-Q wavelet transform is a flexible DWT, a lifting algorithm based on wavelet transform, which can analyze complex oscillation signals more effectively (Selesnick, 2011), and has been used for the decomposition of EEG (Hassan et al., 2016). Its parameters are adjustable, so the transformation can be tuned and applied according to the oscillation behavior of the signal. The main parameters of TQWT are quality factor  $Q$ , total oversampling rate  $r$  and number of stages  $J$ . The degree to which  $Q$  affects the duration of wavelet oscillation is the ratio of its center frequency to its bandwidth.  $r$  is the total oversampling rate (redundancy) when calculating TQWT when  $J \geq 1$ , that is, the total sampling rate coefficient of all sub-bands, which controls the excessive ringing of the system by affecting the scaling factor ( $l, h$ ; Krishna et al., 2019).  $J$  represents the number of stages of the wavelet transform, which consists of a sequence of two-channel filter banks, and the low-pass output of each filter bank is used as the input of the continuous filter bank. The sub-bands ( $J + 1$ ) obtained by signal decomposition are composed of the output signal of the high-pass filter of each filter bank



**FIGURE 2 |** The method flow chart of this research. The research decomposes the pre-processed EEG input tunable-Q wavelet transform (TQWT) into multiple sub-bands (SB), extracts time-domain and non-linear features from the sub-bands, used binary gray wolf optimization (BGWO) to reduce dimensionality, selects the optimized subset as the classifier input, and finally passes support vector machine (SVM) Identify five types of emotions.



and the output signal of the low-pass filter of the final filter bank (Selesnick, 2011).

The low-pass filter frequency response  $H_0^l(\omega)$  and the high-pass filter frequency response  $H_1^l(\omega)$  after level  $J$  should be defined as:

$$H_0^l(\omega) = \begin{cases} \prod_{m=0}^{J-1} H_0\left(\frac{\omega}{l^m}\right), & |\omega| \leq l^J \pi \\ 0, & l^J \pi < |\omega| \leq \pi \end{cases} \quad (1)$$

$$H_1^l(\omega) = \begin{cases} H_1\left(\frac{\omega}{l^{J-1}}\right) \prod_{m=0}^{J-2} H_0\left(\frac{\omega}{l^m}\right), & (1-h)l^{J-1}\pi \leq |\omega| \leq l^{J-1}\pi \\ 0, & \omega \in [-\pi, \pi] \end{cases} \quad (2)$$

where low-pass scaling factor ( $l$ ) and the high-pass scaling factor ( $h$ ) are defined as:

$$l = 1 - \frac{h}{r} \quad (3)$$

$$h = \frac{2}{Q+1} \quad (4)$$

In this study, the EEG is decomposed into five sub-bands with a  $Q$  factor of 3 and an oversampling rate ( $r$ ) of 3 by TQWT, and feature extraction from the sub-bands ( $Q = 3$ ,  $r = 3$ , and  $J = 4$ ). **Figure 3** is a time-frequency diagram of TQWT decomposing the Fp1 channel EEG into five sub-bands.

### Feature Extraction

Extract five time-domain non-linear features for each sub-band signal decomposed by TQWT, namely sample entropy, two differential features and two Hjorth parameters as classification features:

#### (1) Sample entropy

Sample Entropy (SampEn; Richman and Moorman, 2000) measures the complexity of time series by measuring the probability of generating a new pattern in the signal. It is similar to approximate entropy (AE) but is more consistent. Define the sample entropy of a finite array as:

$$\text{SampEn} = -\ln \left[ \frac{A^m(r)}{B^m(r)} \right] \quad (5)$$

where  $\ln$  represents the natural logarithm;  $B^m(u)$  is defined as:

$$B^m(u) = \frac{1}{N-m} \sum_{i=1}^{N-m} B_i^m(u) \quad (6)$$

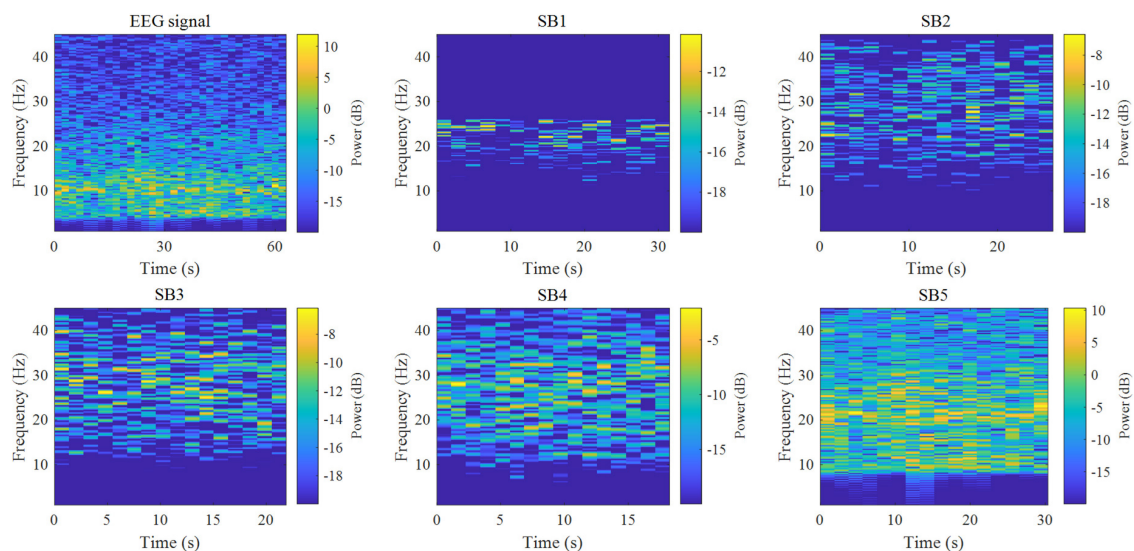
$$B_i^m(u) = \frac{1}{N-m-1} B_i \quad (7)$$

$A^m(u)$  is defined as:

$$A^m(u) = \frac{1}{N-m} \sum_{i=1}^{N-m} A_i^m(u) \quad (8)$$

$$A_i^m(u) = \frac{1}{N-m-1} A_i \quad (9)$$

where  $m$  represents the dimension of the vector, generally 1 or 2;  $N$  represents the length of the sequence;  $u$  represents the measure of “similarity,” generally choose  $u = 0.1 \times \text{std} - 0.25 \times \text{std}$ , where  $\text{std}$  represents the standard deviation of the original time series. In this study,  $m = 2$ ,  $u = 0.2 \times \text{std}$ .



**FIGURE 3 |** Time-frequency diagrams of five sub-bands obtained from single-channel EEG and TQWT decomposed signal. TQWT decomposes the EEG data of a single channel into five sub-bands ( $Q = 3$ ,  $r = 3$ , and  $J = 4$ ), among which sub-band 1 has energy fluctuations at 15–25 HZ, and both sub-bands 2 and 3 are at 10–45 HZ. There are energy fluctuations from time to time. Sub-band 4 has energy fluctuations in the time-frequency axis within 4–45 HZ just like the original data, and the energy fluctuation of sub-band 5 is 8–45 HZ.

## (2) Second-order difference mean (2dif)

$$2dif = \frac{1}{N-2} \sum_{n=1}^{N-2} (x(n+2) - x(n)) \quad (10)$$

where  $x(n)$  represents the time series vector.

## (3) Normalized second-order difference mean (2ndif)

$$2ndif = \frac{2dif}{\sigma_x} \quad (11)$$

where  $\sigma_x$  is the standard deviation.

## (4) Hjorth parameter: mobility

The Hjorth parameter was proposed by Hjorth (1970). Among them, Hjorth-Mobility (HM) is a parameter to estimate the mean frequency, which measures the mobility of EEG:

$$HM = \sqrt{\frac{\text{var}\left(\frac{dx(n)}{dn}\right)}{\text{var}(x(n))}} \quad (12)$$

where var represents the variance.

## (5) Hjorth parameter: complexity

Hjorth-Complexity (HC) is often used to estimate the bandwidth of the signal and measure the complexity of the EEG:

$$HC = \frac{\text{Mobility}\left(\frac{dx(n)}{dn}\right)}{\text{Mobility}(x(n))} \quad (13)$$

In this study, five types of features are extracted from the five sub-bands decomposed by TQWT, and the data shape of each sub-band is a trial feature, namely  $40 \times 32$ . A total of 32 feature matrices of  $40 \times 160$  are obtained for subsequent feature selection.

## Feature Selection

The feature matrix extracted from the TQWT sub-band is selected by the binary gray wolf optimization algorithm (BGWO; Too et al., 2018). The GWO algorithm is an optimized search method developed by simulating the hierarchy and hunting process of the wolf pack. The  $\alpha$ ,  $\beta$ ,  $\delta$ , and  $\omega$  wolves in the wolf pack represent different social classes, respectively. This algorithm has been used by many researchers in the research fields of feature selection, parameter optimization and motor control because of its considerable optimization performance and simplicity and ease of implementation (Wei et al., 2017).

Emary et al. (2016) proposes two BGWO algorithms (BGWO1 and BGWO2) are proposed for feature selection. Among them, BGWO1 uses a crossover operator to update the wolf's position, while BGWO2 uses a crossover operator to update the wolf's position, while BGWO2 updates the wolf by converting the

position into a binary vectors position. In this study, the BGWO2 method will be selected to optimize the feature set by dimensionality reduction, and the formula is as follows:

$$Y^n(t+1) = \begin{cases} 1, & \text{if } S\left(\frac{Y_1^n + Y_2^n + Y_3^n}{3}\right) \geq r_0 \\ 0, & \text{else} \end{cases} \quad (14)$$

where  $r_0$  is a random number in  $[0,1]$ ;  $t$  is the number of iterations;  $n$  is the dimension of the search space;  $Y_1$ ,  $Y_2$ , and  $Y_3$  are defined as binary steps affected by  $\alpha$ ,  $\beta$ , and  $\delta$  wolves, respectively;  $Y^n(t+1)$  is iteration the updated binary position in dimension  $n$  at time  $t$ .  $S(a)$  is defined as:

$$S(a) = \frac{1}{1 + e^{(-10(a-0.5))}} \quad (15)$$

This article discusses the optimization effect of BGWO in three situations: (1) Fusion of the sub-band data of 32 subjects, and optimization of the five feature sets through BGWO, and the data is reduced from  $40 \times 160$  to  $40 \times 57-92$ ; (2) Fusion All the test data are optimized for the feature sets of the five sub-bands, and the data is reduced from  $1,280 \times 32$  to  $1,280 \times 7-17$ ; (3) Fusion of all test and sub-band data, the optimized data length is reduced from 1,280,160 to  $1,280 \times 43-67$  not waiting.

## Classifier and Evaluation Index

This research uses a SVM classifier. The basic idea of SVM is to solve the separation hyperplane that can correctly divide the training data set and have the largest geometric interval (Hsu and Lin, 2002). Originally to solve the two-classification problem, it is now widely used in the recognition of multiple types of emotional EEG (Kawintiranon et al., 2016; Samara et al., 2017).

In order to evaluate the effectiveness of the method proposed in this manuscript, four indicators of accuracy (Acc), sensitivity (Sen), specificity (Spe) and Kappa coefficient (Chu et al., 2021) are calculated through 6-fold cross-validation. The calculation formula of each indicator is as follows:

$$\text{Acc} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (16)$$

where TP refers to true positive, TN is true negative, FP is false positive, and FN is false negative.

$$\text{Sen} = \frac{TP}{TP + FN} \times 100\% \quad (17)$$

$$\text{Spe} = \frac{TN}{FP + TN} \times 100\% \quad (18)$$

$$\text{Kappa} = \frac{\text{Acc} - p_e}{1 - p_e} \quad (19)$$

where  $p_e$  is the completely random classification accuracy. For the five classification problems in this manuscript,  $p_e = 0.2$ .

Accuracy is our most common evaluation index. Generally speaking, the higher the accuracy, the better the classifier. Sensitivity represents the proportion of all positive examples that are matched and measures the classifier's ability to recognize positive examples. Specificity represents the proportion of all negative cases that are matched and measures the ability of the classifier to recognize negative cases. The Kappa coefficient is usually used for consistency testing. It can be used as an index to measure the accuracy of classification, and it can also be used as a normalized index to measure the accuracy of different classification numbers.

## Experimental Results

The experiment is carried out on MATLAB R2019b platform under Windows 8.1 64 bit operating system. The system CPU is AMD Radeon R5 and the memory is 8 GB. This study uses the DEAP data set to verify the effectiveness of the algorithm for emotion recognition from five aspects: (1) The data of each subject is decomposed by TQWT, and the features are extracted after fusing the sub-bands. The feature sets are classified by SVM before and after BGWO. In order to explore the classification performance of the algorithm to individuals, and the improvement effect of BGWO on the algorithm. (2) Extract features from each sub-band decomposed by TQWT, and merge the feature sets of all subjects into SVM classification before and after BGWO to explore the classification performance of the algorithm for different sub-bands of TQWT. (3) Fuse the data of the subject and the sub-bands, and the total feature set obtained after BGWO optimization is used as the classification feature to explore the overall recognition performance of the algorithm. (4) On the basis of experiment (1), the influence of key parameters of the algorithm on the accuracy of individual recognition is explored. (5) Compare and analyze other EEG emotion recognition methods of the same data set, in order to explore the effectiveness of this method for multi-type emotion recognition.

### Accuracy of Individual Recognition

The EEG of 32 subjects were decomposed into five sub-bands by TQWT, the sub-bands were fused and features were extracted and then classified by SVM, the emotion recognition accuracy rate of each subject was obtained as shown in **Figure 4**. Among them, "Before" indicates that the feature set is not optimized by BGWO, and "After" indicates the accuracy information after feature selection by BGWO. Judging from the recognition accuracy of the five categories in the figure, the average recognition accuracy of the two differences, two Hjorth parameters and sample entropy as the classification features is 53.37%; the maximum recognition accuracy of the individual is 87.7%, appearing in 20th subject. After the feature set is optimized by BGWO, the average recognition accuracy of the five types of features is 60.44%; the maximum recognition accuracy of the individual is 88.1%, which appears in the 18th subject. The accuracy of each participant increased by 7.07% on average.

In order to show the time complexity of the algorithm, **Table 1** counts the time consumption information of the 63 s emotion recognition process of a single trial.

### The Recognition Accuracy of Different TQWT Sub-Bands

The feature sets of 32 subjects were fused, and the feature matrixes of five sub-bands were respectively, passed through BGWO, and the classification accuracy before and after optimization of each feature was obtained as shown in **Table 2**. It can be seen from the table that the classification accuracy of each feature when the five sub-bands are not optimized are  $57.168 \pm 1.34$ ,  $58.36 \pm 2.08$ ,  $58.28 \pm 1.34$ ,  $58.28 \pm 1.22$ , and  $57.578 \pm 1.34\%$ , respectively. The recognition accuracy of each sub-band after BGWO was  $63.36 \pm 2.64$ ,  $62.89 \pm 1.32$ ,  $62.764 \pm 2.07$ ,  $62.768 \pm 0.95$ , and  $62.734 \pm 1.88\%$ , and the Acc of each sub-band increased by  $4.97 \pm 0.28\%$  on average.

### Recognition Results of Fusion of Subjects and Sub-Band Data

**Table 3** counts the accuracy (All Acc)%, sensitivity (Sen)%, specificity (Spe)%, and Kappa coefficient information of the five types of emotion recognition overall data (fusion of subjects and sub-band data) with five types of features. The overall recognition accuracy of the algorithm in this manuscript is 62.34%, the average sensitivity is 65.22%, the average specificity is 78.13%, and the Kappa coefficient is 0.53. It can be seen that the classification performance of time-domain non-linear features is similar in accuracy and Kappa coefficient; the optimal performance of sensitivity and specificity are both differential features.

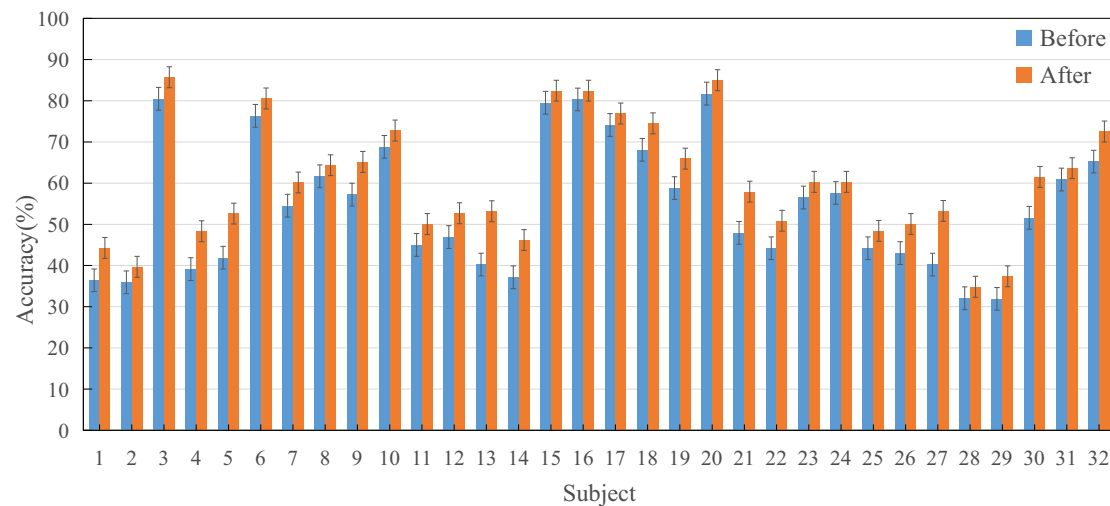
### The Impact of Key Parameters on Recognition Accuracy

Tunable-Q wavelet transform can adjust three parameters to apply to different individuals and achieve the best classification effect. Selesnick (2011) provides suggestions for the selection of TQWT parameters, that is,  $Q \geq 1$ ,  $r$  value of 3.0 or 4.0, and  $J \geq 1$ . On this basis, this article will specifically explore the influence of parameters on the accuracy of different individual emotion recognition. Through repeated trials, the optimal recognition accuracy of 32 subjects and their corresponding parameter combinations are shown in **Table 4**. It can be seen from the table that the average recognition accuracy rate of the subjects obtained after the personalized parameters is 65.2%, and the accuracy rate increases to 68.24% after passing BGWO.

**Figure 5** shows the Kappa coefficient of the best combination of parameters for each subject. It can be seen from the figure that the maximum Kappa coefficient is 0.88, and the average Kappa coefficient is 0.603. The Kappa coefficients of different subjects are quite different.

### Comparative Analysis of Related Research

**Table 5** compares some EEG emotion recognition research methods based on the DEAP data set, and normalizes the classification accuracy of different categories into the Kappa coefficient. Among them, the creator of this database (Koelstra, 2012) used the correlation coefficient to do a 2-classification study, and the Kappa coefficients of the arousal, valence and dominance dimensions were 0.24, 0.15, and 0.11. Yin et al. (2020)



**FIGURE 4 |** The accuracy information of the feature set before and after BGWO. Where before refers to the average recognition accuracy  $\pm$  standard error of the five types of feature sets without BGWO optimization; After refers to the average recognition accuracy  $\pm$  standard error of the feature set after BGWO optimization.

**TABLE 1 |** Time consumption statistics of this research method(s).

Subject	TQWT	Feature extraction	BGWO	SVM	Total time
1	0.0052	0.0797	1.0827	0.7377	1.9850
2	0.0046	0.0789	1.1811	0.5733	1.9168
3	0.0055	0.0830	1.1099	0.2136	1.4950
4	0.0049	0.0815	1.1091	0.4496	1.7266
5	0.0054	0.0812	1.1513	0.5508	1.8699
6	0.0055	0.0820	1.1019	0.2911	1.5625
7	0.0047	0.0825	1.0904	0.3029	1.5630
8	0.0053	0.0808	1.1171	0.5324	1.8164
9	0.0048	0.0848	1.1645	0.3213	1.6602
10	0.0045	0.0800	1.1011	0.5356	1.8012
11	0.0047	0.0822	1.1166	0.6990	1.9847
12	0.0059	0.0796	1.1249	0.3097	1.5997
13	0.0048	0.0791	1.1123	0.4574	1.7327
14	0.0045	0.0807	1.1144	0.6253	1.9056
15	0.0052	0.0825	1.1876	0.4849	1.8427
16	0.0046	0.0801	1.1382	0.4296	1.7326
17	0.0063	0.0829	1.1028	0.2028	1.4777
18	0.0062	0.0792	1.2041	0.1932	1.5619
19	0.0050	0.0803	1.3384	0.5900	2.0940
20	0.0060	0.0787	1.4798	0.2691	1.9123
21	0.0050	0.0845	1.1740	0.2876	1.6356
22	0.0047	0.0812	1.2576	0.4934	1.9181
23	0.0059	0.0804	1.3163	0.4459	1.9289
24	0.0051	0.0810	1.2210	0.2974	1.6855
25	0.0056	0.0869	1.1947	0.2966	1.6707
26	0.0054	0.0884	1.2080	0.5649	1.9551
27	0.0044	0.0802	1.3141	0.3100	1.7889
28	0.0056	0.0814	1.1724	0.6294	1.9702
29	0.0054	0.0838	1.2144	0.6628	2.0502
30	0.0048	0.0828	1.2136	0.3896	1.7736
31	0.0056	0.0803	1.2812	0.4850	1.9324
32	0.0048	0.0820	1.1622	0.2938	1.6248



**TABLE 2 |** Statistics of accuracy (%) of various features of each sub-band before and after binary gray wolf optimization (BGWO).

BGWO	Feature	Sub-band				
		SB1	SB2	SB3	SB4	SB5
Before	2dif	57.81	58.2	58.59	57.81	<b>59.38</b>
	2ndif	58.98	58.59	60.16	57.42	56.64
	HC	55.86	61.33	58.59	59.37	58.59
	HM	55.86	55.47	56.64	57.03	56.25
	Sampen	57.42	58.21	57.42	59.77	57.03
	Average	57.186	58.36	58.28	58.28	57.578
	Std	1.34	2.08	1.34	1.22	1.34
After	2dif	63.67	63.28	61.72	61.36	62.11
	2ndif	<b>67.58</b>	<b>64.45</b>	<b>66.41</b>	63.28	62.11
	HC	60.94	63.67	61.33	<b>63.67</b>	<b>65.62</b>
	HM	61.33	61.72	62.25	62.25	63.28
	Sampen	63.28	61.33	62.11	63.28	60.55
	Average	63.36	62.89	62.764	62.768	62.734
	Std	2.64	1.32	2.07	0.95	1.88

*Bold values means the maximum accuracy of each sub-band.*

**TABLE 3 |** The overall accuracy (All Acc)%, sensitivity (Sen)%, specificity (Spe)%, and Kappa coefficient of five types of emotion recognition based on tunable-Q wavelet transform (TQWT) and BGWO.

	All Acc	Sen	Spe	Kappa
2dif	62.1	<b>70.25</b>	81.33	0.5263
2ndif	62.5	64.19	<b>82.01</b>	0.5313
HC	<b>62.6</b>	62.12	79.29	<b>0.5325</b>
HM	62.4	65.13	68.7	0.53
Sampen	62.1	64.4	79.3	0.5263
Average	62.34	65.22	78.13	0.53
Std	0.23	3.03	5.41	0.003

*All means fusion of all subjects and sub-band data. Bold value means the optimal value of each indicator.*

used locally-robust feature selection (LRFS) method is used to conduct a 2-classification study, and the Kappa coefficients of the arousal and valence dimensions are 0.3 and 0.36. Gupta et al. (2016) used the graph-theoretic feature extraction method for three classification studies, the Kappa coefficients in four dimensions were 0.54, 0.51, 0.48, and 0.48. Tao and Dan (2021) proposed a multi-source co-adaptation framework for mining diverse correlation information (MACI), the Kappa coefficient of the three categories was 0.45. Zhang et al. (2016) used the ReliefF feature selection method to conduct the four-category study, and the Kappa coefficient of the category was 0.45. Gupta et al. (2019) used flexible analytic wavelet transform (FAWT) Extract features for four classification studies, the Kappa coefficient was 0.45. Atkinson and Campos (2016) proposed an emotion recognition model combining the feature selection method based on minimum-Redundancy-Maximum-Relevance (mRMR) and kernel classifier. The Kappa coefficients of the two categories were 0.46 and 0.46, the kappa coefficients of the three categories were 0.43 and 0.41, and the Kappa coefficients of the five categories were 0.33 and 0.32, respectively. Generally speaking, the higher

**TABLE 4 |** The optimal recognition accuracy of different individuals and their corresponding TQWT parameter combinations.

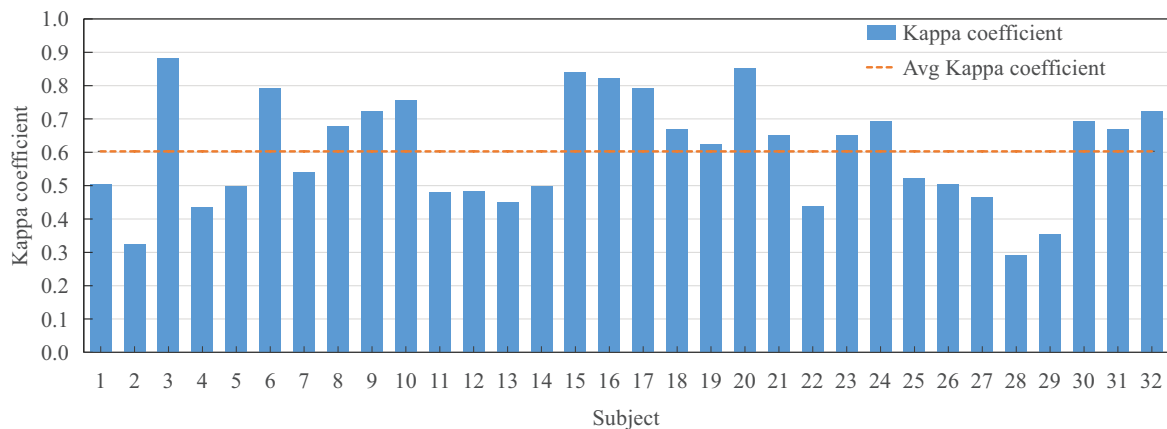
Subject	TQWT parameter			Accuracy (%)	
	Q	r	J	Before BGWO	After BGWO
1	1	3	4	52.38	60.32
2	3	3	2	42.86	46.03
3	5	3	3	<b>89.68</b>	<b>90.48</b>
4	1	3	4	52.78	54.76
5	1	3	4	59.52	59.92
6	1	3	2	82.14	83.33
7	1	3	1	59.52	63.1
8	4	3	1	67.86	74.21
9	1	3	4	77.38	77.78
10	3	3	1	79.76	80.56
11	1	3	3	57.94	58.33
12	4	3	3	57.54	58.73
13	4	3	1	52.38	55.95
14	1	3	5	57.54	59.92
15	2	3	3	85.71	87.3
16	5	3	3	85.32	85.71
17	3	3	3	82.94	83.33
18	3	3	1	70.24	73.41
19	2	3	6	68.25	69.84
20	3	3	1	88.10	88.10
21	4	3	4	62.70	72.22
22	4	3	1	51.59	55.16
23	1	3	3	67.86	72.22
24	1	3	3	70.24	75.40
25	5	3	6	54.37	61.9
26	1	3	6	51.98	60.32
27	1	3	1	55.16	57.14
28	3	3	1	42.46	43.25
29	1	3	4	45.63	48.41
30	1	3	1	67.06	75.4
31	3	3	1	73.02	73.41
32	3	3	3	74.60	77.78
Average				65.20	68.24
Std				13.72	12.99

*Bold values indicate the subject maximum accuracy.*

the number of machine learning classifications, the lower the classification accuracy (Kong et al., 2021). It can be seen from the table that the classification performance of this method has reached a 5-class Kappa coefficient of 0.603. Compared with the classification performance of the above methods, this research has achieved outstanding classification effects.

## DISCUSSION

In this study, for the EEG-based e-BCI, the method of TQWT and BGWO was used to identify the five types of emotions in the DEAP dataset: neutral, happy, sad, relax, and anger. First of all, in terms of the number of identifications, this research has improved compared with previous traditional studies, and



**FIGURE 5 |** The Kappa coefficient of the optimal parameter combination of each subject and the average Kappa coefficient of all subjects.

**TABLE 5 |** Comparison of classification performance between the method in this manuscript and other studies on the database for emotion analysis using physiological signals (DEAP) dataset.

Authors	Method	Accuracy (%)				Kappa
		2 Classes	3 Classes	4 Classes	5 Classes	
Koelstra, 2012	Correlation coefficient + NB	62, 57.6, 55.4	–	–	–	0.24, 0.15, 0.11
Yin et al., 2020	LRFS + LSSVM/NB	65, 68	–	–	–	0.30, 0.36
Gupta et al., 2016	Graph-theoretic + SVM/RVM	–	69, 67, 65, 65	–	–	0.54, 0.51, 0.48, 0.48
Tao and Dan, 2021	MACI	–	63.31	–	–	0.45
Zhang et al., 2016	ReliefF + SVM	–	–	58.75	–	0.45
Gupta et al., 2019	FAWT + SVM	–	–	59.06	–	0.45
Atkinson and Campos, 2016	mRMR + SVM	73.14, 73.06	62.33, 60.7	–	46.69, 45.32	0.46, 0.46/0.43, 0.41/0.33, 0.32
This work	TQWT-BGWO + SVM	–	–	–	68.24	0.603

increased the types of signal identification. Secondly, for EEG-based emotion recognition, the current more innovative TQWT algorithm is selected to analyze the signal. In addition, for the EEG feature selection method, BGWO is used for the first time to optimize the EEG emotional features, and its optimization effect on the emotion recognition task is verified.

It can be seen from **Figure 4** that it is feasible to extract the time domain and non-linear dynamic characteristics from TQWT and use SVM to identify five types of emotions. The classification accuracy of different subjects for the same trial is significantly different, indicating that the same emotion-inducing material has different emotion-inducing effects for different subjects. Excluding subjects or trials with poor emotion-inducing effects may improve the overall recognition accuracy. Therefore, designing an emotion-induced paradigm suitable for different subjects is still a prominent problem of e-BCI. Secondly, **Figure 4** shows that after using BGWO, the accuracy of a single subject has been enhanced, showing better applicability. In addition, it can be seen from part Feature Selection that BGWO can effectively reduce the data size of the feature set. It shows that BGWO is also an effective optimization method for EEG emotion recognition tasks. It can be seen from **Table 1** that the time consumption of each stage of the method proposed in this manuscript basically meets the online BCI system.

It can be seen from **Table 2** that the optimal classification accuracy can be obtained by taking the TQWT sub-band as the classification axis. In addition, the standard deviation  $Std \leq 2.64$  of the classification accuracy of the same classification feature in different sub-bands indicates that the stability of the recognition accuracy of the TQWT decomposition signal is good. **Table 3** shows the accuracy, sensitivity, specificity, and Kappa coefficient of the five features under this research method. Among them, the average accuracy rate is 62.34%, which exceeds the probability of random guessing (above-chance level) by 42.34%. For sensitivity and specificity, this method has a 65.22% ability to recognize positive cases and 78.13% on negative cases. If understood from a medical point of view, sensitivity and specificity measure the missed diagnosis rate and the misdiagnosis rate, respectively. The Kappa coefficient is 0.53, which represents the ratio of the error reduction of the classification and the chance level.

**Table 4** statistics the optimal TQWT parameter combination of each subject and the recognition accuracy information before and after the obtained BGWO. Compared with **Figure 4**, tuning the TQWT parameters for different subjects can achieve better recognition performance, achieving a maximum individual recognition accuracy of 90.48% and an average recognition accuracy of 68.24%. **Figure 5** shows the Kappa coefficient information of the subjects. It can be seen that the average Kappa

coefficient of all subjects is 0.603, and the Kappa coefficients of all subjects are linearly related. In addition, for all subjects, the optimal value of  $Q$  factor is between 1 and 5, the optimal value of  $r$  is both 3, and the optimal value of  $J$  is between 1 and 6. Individual differences are not only manifested in the inducing effect of emotions, but also in system parameters. Therefore, TQWT with adjustable parameters is an effective method to overcome individual differences. It is worth noting that the EEG emotion recognition methods based on TQWT and BGWO use simple and common features and classifiers. If try other advanced or improved features and classifiers, can achieve good classification results, or you can switch the emotion category It is a control instruction for BCI equipment, which will be more conducive to the development of e-BCI.

## CONCLUSION

In this study, the TQWT-BGWO method was used to recognize five types of emotions from EEG. TQWT decomposes the EEG into sub-bands, extracts features from the sub-bands, and used the SVM classifier to classify after BGWO optimization to realize the recognition of five types of emotion signals: neutral, happy, sad, relaxed, and anger. The parameterized TQWT signal decomposition can overcome individual differences to a certain extent, and combined with the BGWO feature selection method with fast convergence speed and good optimization performance, it can effectively improve the recognition accuracy of the system.

## REFERENCES

- Asghar, M.A., Khan, M.J., Rizwan, M., Mehmood, R.M., and Kim, S.H. (2020). An innovative multi-model neural network approach for feature selection in emotion recognition using deep feature clustering. *Sensors* 20:3765. doi: 10.3390/s20133765
- Atkinson, J., and Campos, D. (2016). Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers. *Expert Syst. Appl.* 47, 35–41. doi: 10.1016/j.eswa.2015.10.049
- Hsu, C.W., and Lin, C.J. (2002). A comparison of methods for multiclass support vector machines. *IEEE Trans. Neural Networks* 13, 415–425. doi: 10.1109/72.991427
- Chu, Y., Zhu, B., Zhao, X., and Zhao, Y. (2021). Convolutional neural network based on temporal-spatial feature learning for motor imagery electroencephalogram signal decoding. *J. Biomed. Engin.* 38, 1–9. doi: 10.7507/1001-5515.202007006
- Emary, E., Zawba, H.M., and Hassanien, A.E. (2016). Binary grey wolf optimization approaches for feature selection. *Neurocomputing* 172, 371–381. doi: 10.1016/j.neucom.2015.06.083
- Fang, Y., Yang, H., Zhang, X., Liu, H., and Tao, B. (2021). Multi-feature input deep forest for EEG-based emotion recognition. *Front. Neurobot.* 14:617531. doi: 10.3389/fnbot.2020.617531
- Fattouh, A., Horn, O., and Bourhis, G. (2013). Emotional BCI control of a smart wheelchair. *Int. J. Comp. Sci. Issues* 3, 32–36.
- Gupta, R., Laghari, K., and Falk, T.H. (2016). Relevance vector classifier decision fusion and EEG graph-theoretic features for automatic affective state characterization. *Neurocomputing* 174, 875–884. doi: 10.1016/j.neucom.2015.09.085
- Gupta, V., Chopda, M., and Pachori, R.B. (2019). Cross-subject emotion recognition using flexible analytic wavelet transform from EEG signals. *IEEE Sensors J.* 19, 2266–2274. doi: 10.1109/JSEN.2018.2883497
- Through the DEAP data set, the effectiveness of the proposed algorithm is verified. The experimental results show that the research method in this manuscript has an average recognition accuracy of 68.24%, a sensitivity of 65.22%, a specificity of 78.13% and a Kappa coefficient of 0.603 for the five types of emotions. The proposed algorithm can effectively identify multiple types of emotional states, and provides new ideas for emotional BCIs.
- 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 authors.
- AUTHOR CONTRIBUTIONS**
- SL and XL conceived the experiment. SL carried out the experiment and data analysis. LZ, ZC, and AG provided the suggestions. AG and YF revised the manuscript. All authors contributed to the article and approved the submitted version.
- FUNDING**
- Fund projects: The National Natural Science Foundation of China (81771926, 61763022, 81470084, 61463024, and 62006246).
- Hassan, A.R., Siuly, S., and Zhang, Y.J.C.M. (2016). Epileptic seizure detection in EEG signals using tunable-Q factor wavelet transform and bootstrap aggregating. *Comput. Methods Programs Biomed.* 137, 247–259. doi: 10.1016/j.cmpb.2016.09.008
- Hjorth, B. (1970). EEG analysis based on time domain properties. *Electroencephalogr. Clin. Neurophysiol.* 29, 306–310. doi: 10.1016/0013-4694(70)90143-4
- Kawintiranon, K., Buatong, Y., and Vateekul, P. (2016). Online music emotion prediction on multiple sessions of EEG data using SVM, in: *Joint Conference on Computer Science and Software Engineering (JCSSE)*. New York: IEEE.
- Koelstra, S. (2012). DEAP: A database for emotion analysis; using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15
- Kong, W., Song, X., and Sun, J. (2021). Emotion recognition based on sparse representation of phase synchronization features. *Multimedia Tools Appl.* 80, 21203–21217. doi: 10.1007/s11042-021-10716-3
- Krishna, A.H., Sri, A.B., Priyanka, K., Taran, S., and Bajaj, V. (2019). Emotion classification using EEG signals based on tunable-Q wavelet transform. *IET Sci. Measur. Technol.* 13, 375–380. doi: 10.1049/iet-smt.2018.5237
- Luo, Y., Fu, Q., Xie, J., Qin, Y., and Ding, X. (2020). EEG-based emotion classification using spiking neural networks. *IEEE Access* 8, 46007–46016. doi: 10.1109/ACCESS.2020.2978163
- Mao, Y., Zhou, X., Xia, Z., Yin, Z., and Sun, Y. (2007). A survey for study of feature selection algorithms. *Pattern Recog. Artif. Intel.* 20, 211–218.
- Mohammadpour, M., Hashemi, S., and Houshmand, N. (2017). *Classification of EEG-based emotion for BCI applications, in: Intelligence and Robotics (IRANOPEN)*. New York: IEEE.
- Molina, G.G., Tsoneva, T., and Nijholt, A. (2009). Emotional brain-computer interfaces. *Int. J. Autonomous Adapt. Commun. Syst.* 6, 9–25.
- Morris, J.D. (1995). Observations: SAM: The selfassessment manikin; an efficient cross-cultural measurement of emotional response. *J. Adv. Res.* 23, 63–68.

- Nie, D., Wang, X., Duan, R., and Lu, B. (2012). A survey on EEG based emotion recognition. *Chinese J. Biomed. Engin.* 31, 595–606. doi: 10.3969/j.issn.0258-8021.2012.04.018
- Peng, H. (2004). *General Psychology*. China: Beijing Normal University Press.
- Richman, J.S., and Moorman J.R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Heart Circ. Physiol.* 278, H2039–H2049. doi: 10.1152/ajpheart.2000.278.6.H2039
- Samara, A., Menezes, M., and Galway, L. (2017). *Feature Extraction for Emotion Recognition and Modelling Using Neurophysiological Data*, in: 2016 15th International Conference on Ubiquitous Computing and Communications and 2016 International Symposium on Cyberspace and Security (IUCC-CSS). New York: IEEE
- Selesnick, I.W. (2011). Wavelet transform with tunable Q-Factor. *IEEE Trans. Signal Proces.* 59, 3560–3575.
- Sm, A., Smm, B., and Al, A. (2014). Grey wolf optimizer. *Adv. Engin. Software* 69, 46–61. doi: 10.1016/j.advengsoft.2013.12.007
- Tao, J.W., and Dan, Y.F. (2021). Multi-Source Co-adaptation for EEG-based emotion recognition by mining correlation information. *Front. Neurosci.* 15:677106. doi: 10.3389/fnins.2021.677106
- Too, J., Abdullah, A., Saad, N.M.O., Ali, N.M.O., and Tee, W. (2018). A new competitive binary grey wolf optimizer to solve the feature selection problem in EMG signals classification. *J. Comp.* 7:58. doi: 10.3390/computers7040058
- Wei, C., Chen, L.L., Song, Z.Z., Lou, X., and Li, D.D. (2020). EEG-based emotion recognition using simple recurrent units network and ensemble learning. *Biomed. Signal Proces. Control* 58:101756.
- Wei, Y., Ni, N., Liu, D., Chen, H., Wang, M., Li, Q., et al. (2017). An improved grey wolf optimization strategy enhanced SVM and its application in predicting the second major. *Math. Problems Engin.* 2017, 1–12. doi: 10.1155/2017/9316713
- Wolpaw, J.R., Birbaumer, N., Heetderks, W.J., Mcfarland, D.J., and Peckham, P.H. (2000). Brain-computer interface technology: a review of the first international meeting. *IEEE Trans. Rehabil. Eng.* 8, 164–173. doi: 10.1109/TRE.2000.847807
- Yin, Z., Liu, L., Chen, J., Zhao, B., and Wang, Y. (2020). Locally robust EEG feature selection for individual-independent emotion recognition. *Expert Syst. Appl.* 162:113768. doi: 10.1016/j.eswa.2020.113768
- Zhang, J., Ming, C., Zhao, S., Hu, S., Shi, Z., and Yu, C. (2016). ReliefF-Based EEG sensor selection methods for emotion recognition. *Sensors* 16:1558. doi: 10.3390/s16101558
- Zhou, Y., Li, D., Wang, Z., and Gao, D. (2020). Emotion classification of EEG based on cepstrum features. *Comp. Engin. Appl.* 56, 170–175.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships 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.

Copyright © 2021 Li, Lyu, Zhao, Chen, Gong and Fu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.





# Relaxation Degree Analysis Using Frontal Electroencephalogram Under Virtual Reality Relaxation Scenes

Yue Zhang<sup>1†</sup>, Lulu Zhang<sup>2†</sup>, Haoqiang Hua<sup>1</sup>, Jianxiu Jin<sup>1</sup>, Lingqing Zhu<sup>1</sup>, Lin Shu<sup>1\*</sup>, Xiangmin Xu<sup>1,3</sup>, Feng Kuang<sup>1</sup> and Yunhe Liu<sup>1</sup>

<sup>1</sup> School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China, <sup>2</sup> Department of Psychiatry, Guangzhou First People's Hospital, The Second Affiliated Hospital, South China University of Technology, Guangzhou, China, <sup>3</sup> Zhongshan Institute of Modern Industrial Technology of South China University of Technology, Zhongshan, China

## OPEN ACCESS

### Edited by:

Jane Zhen Liang,  
Shenzhen University, China

### Reviewed by:

Xiujuan Zheng,  
Sichuan University, China  
Chao Chen,  
Tianjin University of Technology, China

### \*Correspondence:

Lin Shu  
shul@scut.edu.cn

<sup>†</sup>These authors share first authorship

### Specialty section:

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

**Received:** 03 June 2021

**Accepted:** 13 August 2021

**Published:** 24 September 2021

### Citation:

Zhang Y, Zhang L, Hua H, Jin J, Zhu L, Shu L, Xu X, Kuang F and Liu Y (2021) Relaxation Degree Analysis Using Frontal Electroencephalogram Under Virtual Reality Relaxation Scenes. *Front. Neurosci.* 15:719869. doi: 10.3389/fnins.2021.719869

Increasing social pressure enhances the psychological burden on individuals, and the severity of depression can no longer be ignored. The characteristics of high immersion and interactivity enhance virtual reality (VR) application in psychological therapy. Many studies have verified the effectiveness of VR relaxation therapy, although a few have performed a quantitative study on relaxation state (R-state). To confirm the effectiveness of VR relaxation and quantitatively assess relaxation, this study confirmed the effectiveness of the VR sightseeing relaxation scenes using subjective emotion scale and objective electroencephalogram (EEG) data from college students. Moreover, some EEG features with significant consistent differences after they watched the VR scenes were detected including the energy ratio of the alpha wave, gamma wave, and differential asymmetry. An R-state regression model was then built using the model stacking method for optimization, of which random forest regression, AdaBoost, gradient boosting (GB), and light GB were adopted as the first level, while linear regression and support vector machine were applied at the second level. The leave-one-subject-out method for cross-validation was used to evaluate the results, where the mean accuracy of the framework achieved 81.46%. The significantly changed features and the R-state model with over 80% accuracy have laid a foundation for further research on relaxation interaction systems. Moreover, the VR relaxation therapy was applied to the clinical treatment of patients with depression and achieved preliminary good results, which might provide a possible method for non-drug treatment of patients with depression.

**Keywords:** VR, EEG, relaxation state, regression model, machine learning, depression therapy

## INTRODUCTION

The rapidly developing society enhances pressure on individuals, while mental health problems are getting increasingly critical. By the end of 2020, mental health problems became the second most critical disease worldwide (Randy et al., 2019). Statistics from the World Health Organization in 2020 showed that more than 300 million individuals were suffering from depression, with over 80% not receiving appropriate treatment (WHO, 2020). Patients with mental disorders have a profound

negative impact on their personal development, bringing burdens to their families and society (Sidra et al., 2013). Since many studies have confirmed that relaxation could relieve depression (DeBerry, 1982; Lolak et al., 2008), it is highly important to effectively reduce stress and relax.

The current primary methods to relax include deep respiration (Ilse et al., 2014), muscular relaxation (Wesley and Douglas, 1977), music relaxation (Claas et al., 2008; Olga et al., 2011), meditation (Lazar et al., 2000; Narendra et al., 2017), and autogenic training (Ernst and Kanji, 2000). These methods were easily subject to the environment and devices.

The advantages of free space, high immersion, and interactivity have enhanced virtual reality (VR) application in psychological therapy with its rapid development, thus achieving good results (Imran et al., 2014; Allison et al., 2017). Nevertheless, most of these VR scenes are static in nature without scene transitions (Andersen et al., 2017; Kiefl et al., 2018), which might cause boredom and affect the relaxation effect. Moreover, current studies have only made a subjective and qualitative evaluation (Freeman et al., 2017; Linda et al., 2020).

Subjective scales and physiological parameters are generally used to assess the relaxation degree, where the former refers to Perceived Stress Scale (Cohen et al., 1983) and State-trait Anxiety Inventory (Theresa and Hilary, 1992), and the latter includes electroencephalogram (EEG) (Knott et al., 1997; Xing et al., 2019), heart rate variability (Patil and Shirley, 2006; Shu et al., 2020), galvanic skin response (Alexandros et al., 2015), and respiration (Joseph et al., 2016). EEG can relevantly reflect people's emotional state more accurately among the physiological parameters (Soraia and Manuel, 2019) since emotion is a natural product of neural activity in the brain. Consequently, EEG would be an ideal parameter for measuring relaxation state (R-state), and frontal EEG is considered as the first choice considering its simple operation.

Relevant studies have proved that different frequencies of brain electricity reflect different brain states (Hou et al., 2020), in which alpha, theta, and gamma waves show stronger relevance with R-state. Cahn and Delorme found that the long-term training of Vipassana meditation could increase gamma power (Baruch et al., 2016). Du and Lee observed that low-frequency alpha waves in the left frontal lobe while high-frequency alpha waves in the right frontal lobe increased significantly during positive emotional audio stimulation, where the experimental materials were from the standard International Affective Digital Sounds dataset (Du and Lee, 2015). However, there is a lack of EEG-based relaxation regression models under a VR environment as well as an effective VR relaxation system.

To explore the relaxation effect of VR scenes and the correlation between R-state and frontal EEG, VR relaxation scenes were used as emotion-evoked materials in this study, and a relaxation rating model was established based on EEG data. The VR relaxation scenes and R-state model were then used on patients with depression to explore the possibility of VR relaxation therapy for depression. The study is organized as follows. Section "Materials and Methods" introduces the methods of relaxation VR scene design, data collecting, and model building. Section "Results" shows the results of the analysis

of EEG data and the effect of the relaxation model. Section "Application" introduces the application of the research. Section "Discussion" and section "Conclusion" present the discussion and conclusion, respectively.

## MATERIALS AND METHODS

### Design of Virtual Reality Relaxation Scenes

Four sightseeing-relaxation VR scenes were selected as experimental materials including National Park, Snow Mountain, the Great Wall, and Yunnan. Western classical and new age music were chosen, as background music since O'Sullivan's research had proved that relaxation music was mainly soft music composed of slow rhythm, low pitch, low volume, beautiful melody, and orchestral instruments (O'Sullivan, 1991), which were copyright free and had been evaluated (Zhu et al., 2019). The selected background music was absolute music without lyrics to avoid cognitive and cultural differences. The description of the scenes and music is shown in **Table 1**.

The design process of VR relaxation scenes is illustrated in **Figure 1**. Appropriate VR scenes and background music were chosen and combined to get visual and auditory fusion materials. EEG was collected during the whole period of watching the VR scene, after which a subjective scale was completed. A designed relaxation VR scene was officially completed when it was verified to achieve an ideal relaxation effect through subjective scale and EEG data evaluation.

### Methods of R-State Evaluation





The forehead prefrontal EEG electrodes of FP1, FP2, and FPZ were chosen to acquire EEG signals for analyzing the relaxation degree of the participants, since the forehead region of the brain was found to be most associated with emotions (Suranjita and Rajesh, 2019).

Furthermore, subjective emotion scale Self-Assessment Manikin (SAM) and R-state were used as subjective emotion labels. SAM was based on the valence-arousal-dominance emotion model, which assessed emotion state through three indices. Each score of the three indices ranged from 1 to 9. A higher score indicated a more intense emotion state (higher valence, arousal, and dominance) (Bradley and Lang, 1994; Shu et al., 2018). R-state was based on the R-state pyramid theory proposed by Smith (2005). To keep the grading uniform, the value of the R-state also ranged from 1 to 9. A value of R-state greater than five indicated relaxation, and the numbers 5–9 corresponded to the five levels of R-state, as shown in **Figure 2** (value 5 of R-state corresponded to level 1 of stress relief). The higher the score, the more relaxed the participant was.

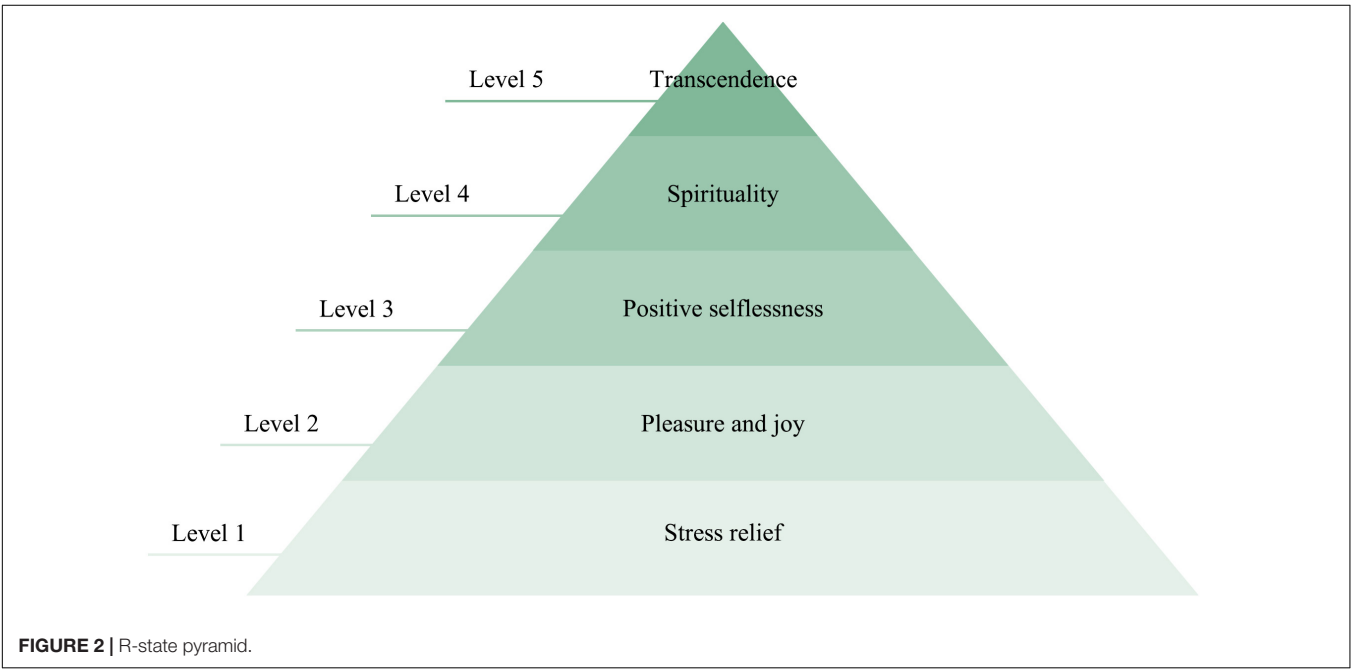
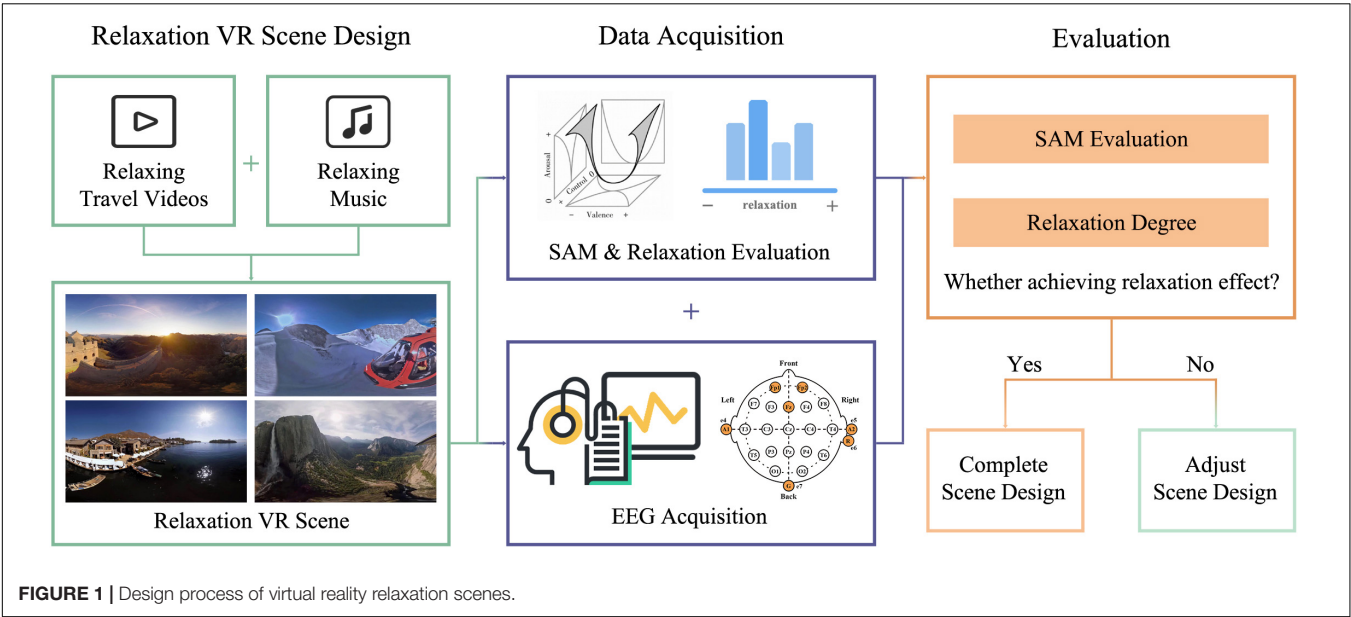
### Participants and Experimental Procedure

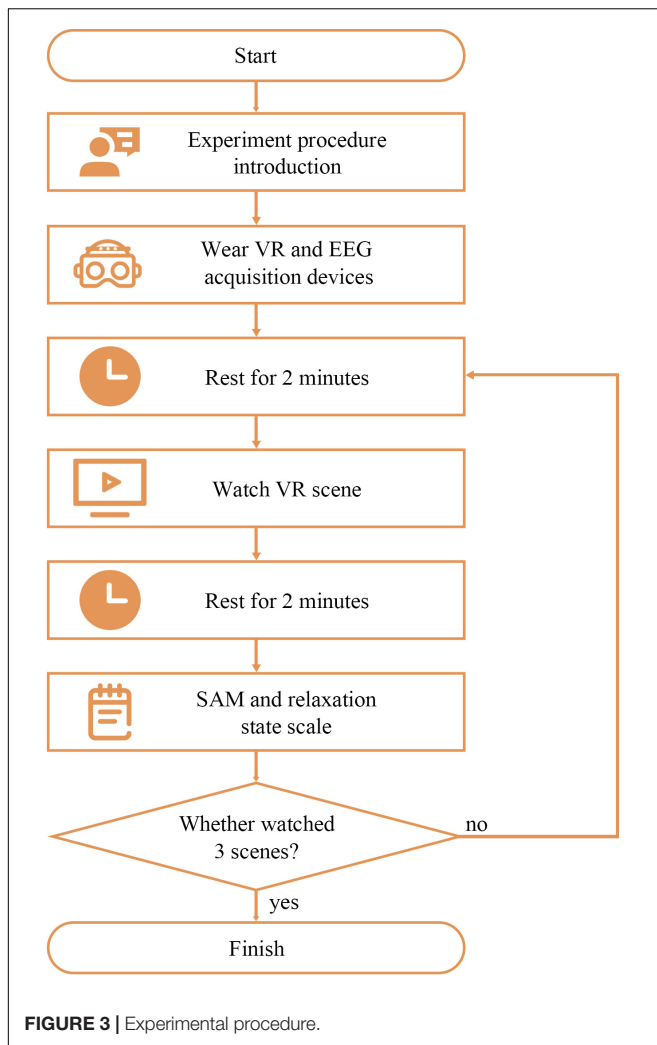
Thirty-three healthy college students (age ranging from 20 to 26 years) including 19 men and 14 women participated in the experiment, with data of only 30 participants valid (16 men and 14 women) for the reason that there were three participants

TABLE 1 | Selected VR relaxation scenes.

Sample figure of VR scenes					
	Scene name	National Park	Snow Mountain	The Great Wall	Yunnan
	Scene length	199 s	156 s	90 s	144 s
	Background music	Pastoral Symphony	Wight Light	Sonata for Spring	The Reiki Gold

VR, virtual reality.





whose EEG data were not fully collected due to the instability of electrode–scalp interface of the EEG collection device. The experiment was conducted in a 30-dB closed soundproof room (Hengqi, Dongguan, China), with experimental equipment, two comfortable chairs, and a table. A pre-training was conducted to make the participants familiar with the experimental steps and SAM scale evaluation method. The procedure of the experiment is shown in **Figure 3**. After offering personal information and wearing VR glasses together with the EEG acquisition device, participants needed to rest for 2 min with a black screen insight, before and after watching each relaxation scene, the duration of which lasted 90–199 s. Each participant was asked to randomly watch three of the four scenes. The participants were asked to keep their eyes open during the whole experiment to control the variables. The experiment procedure was based on other relevant studies (Zhu et al., 2019). The experimental procedures were approved by the Guangzhou First People's Hospital (202002030262, on April 1, 2020).

The VR scenes were watched by HTC Vive, and the EEG acquisition device was a Mangold-10 multichannel physiological instrument with an acquisition frequency of 256 Hz. **Figure 4**

shows the experimental equipment and the data collection settings for the participants. Three flexible EEG electrodes were embedded in the sponge of the VR device to collect forehead EEG data. The subjective scale was finished after each scene.

## Electroencephalogram Data Processing

Considering the EEG data collection from watching one section of the VR scene as one segment of data, 80 segments of effective data of normal participants were collected for further analysis (20 segments for each scene). To explore the relaxing effect of the VR scenes, data sections in each segment of EEG data before and after watching scenes for the 30 s (noted as pre EEG and post EEG) and 30-s data during the period of watch VR scenes (noted as begin EEG and end EEG) were selected (Nitin et al., 2016), as shown in **Figure 5**.

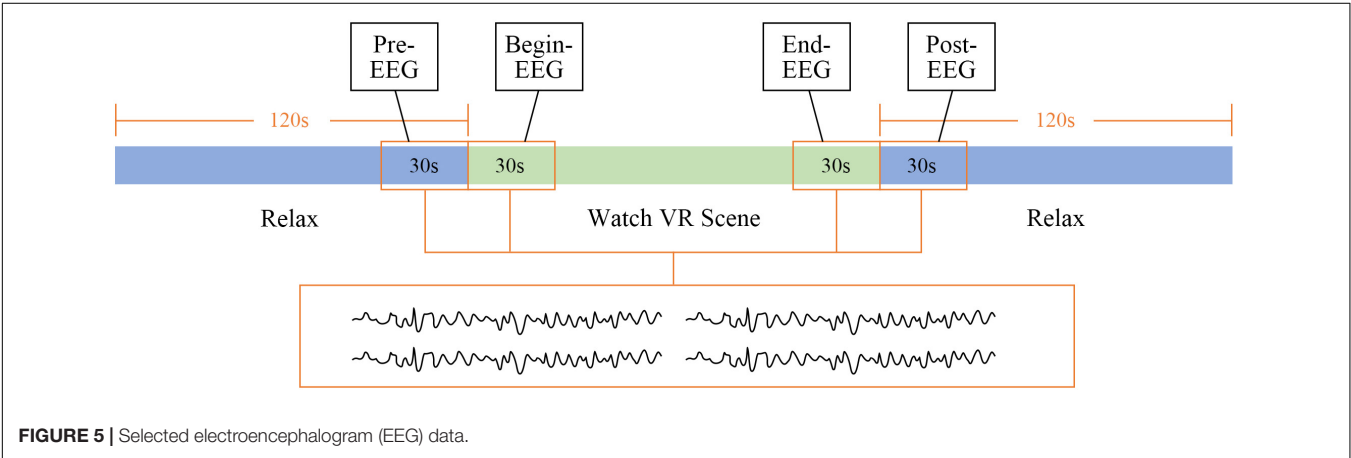
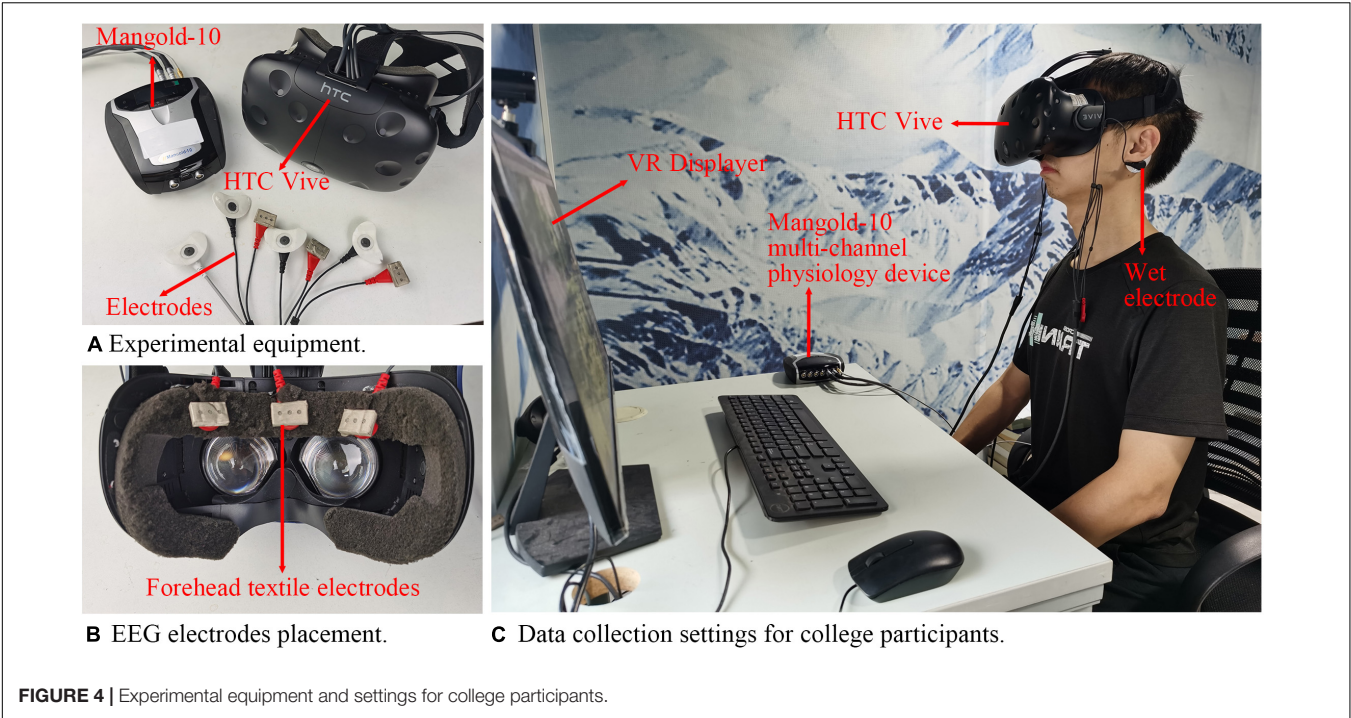
Selected EEG data were filtered using a 1- to 45-Hz Butterworth filter to eliminate the power interference and the baseline drift, after which each segment of de-noised EEG data was decomposed into seven frequency bands according to the different frequencies including delta ( $\delta$ , 1–3 Hz), theta ( $\theta$ , 4–7 Hz), low alpha ( $\alpha_l$ , 8–10 Hz), high alpha ( $\alpha_h$ , 10–12 Hz) (Dombrowe and Hilgetag, 2014), low beta ( $\beta_l$ , 12–20 Hz), high beta ( $\beta_h$ , 20–30 Hz) (Kwon et al., 2011), and gamma waves ( $\gamma$ , 31–50 Hz). Then, EEG features of each band were extracted including the EEG energy value (E), energy ratio (ER), energy entropy (EE), differential entropy (DE), power spectral density, the asymmetry (ASM), rational asymmetry (RASM), differential asymmetry (DASM) of the alpha wave, and the asymmetry of EEG energy ratio of each frequency band, as shown in **Table 2** (Jenke et al., 2014; Tortella-Feliu et al., 2014; Zheng and Lu, 2015; Deng et al., 2021).

The feature changes of the four data sections extracted from each segment of EEG data including pre–post EEG, pre–end EEG, and begin–end EEG were tested using *t*-test except begin–post EEG for the reason that the mood swings were evident by the VR scene and the goal emotion had not been fully aroused with huge mood swings at the beginning. Moreover, the relaxing emotion would be somewhat diminished during the period of post EEG. Consequently, comparing the begin–post EEG involves multiple variables that cannot be controlled. The features with significant variance ( $p < 0.05$ ) after watching VR scenes were selected for further study. Since participants were exposed to visual and auditory stimuli during the begin EEG period of time, this EEG might be different from pre EEG collected in the resting state. Therefore, both pre–end EEG and begin–end EEG are worth analyzing.

## R-State Model

Since four of the 30 normal participants did not finish the subjective emotion scale, which meant that some of their EEG data lacked an R-state label, and seven segments of the EEG data had obvious noises due to the large body or eye movements during the experiments, only 71 segments of EEG data of 26 participants were used for relaxation model building. After being preprocessed, 147 EEG features were selected to train the R-state model.





**TABLE 2 |** Selected features.

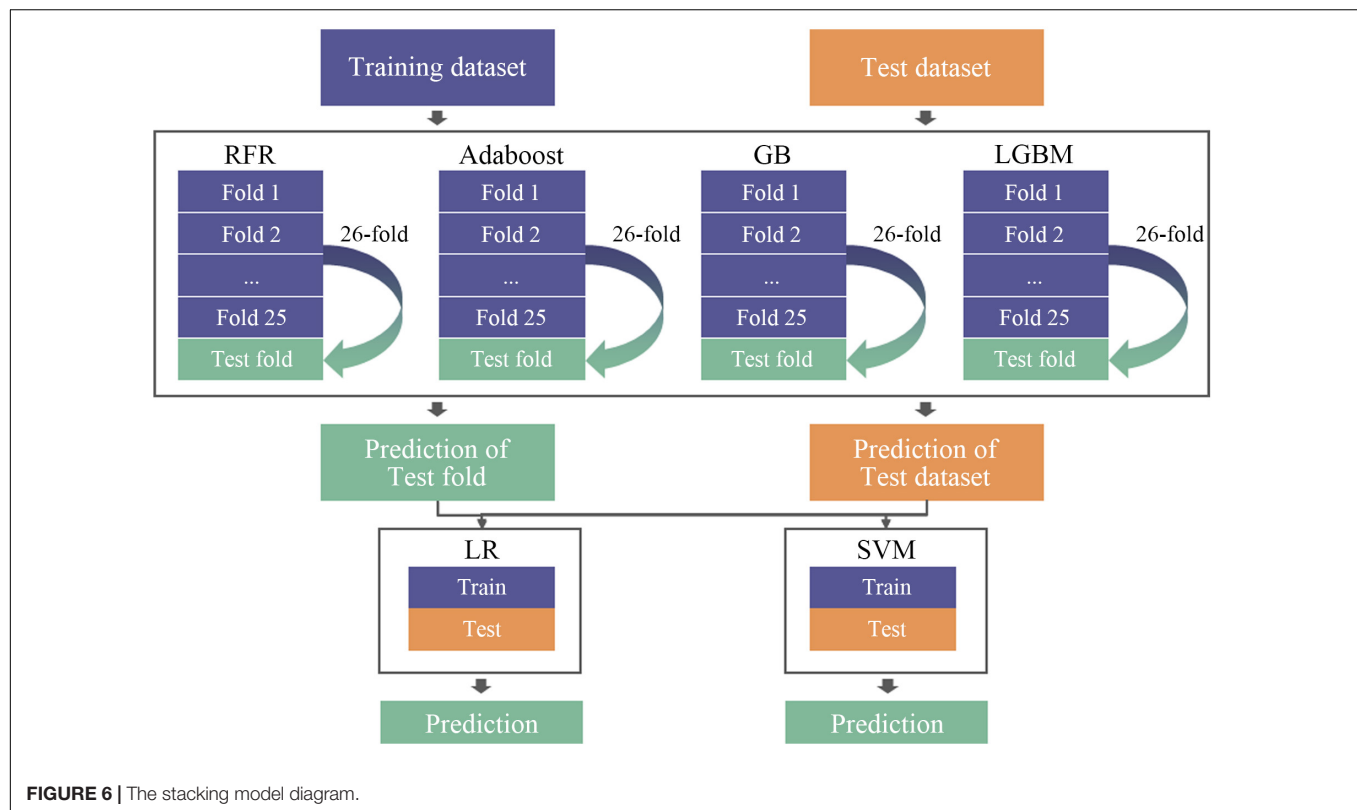
Features	Selected band	Description
E (energy)	All	Transformed and calculated by Fourier transform
ER (energy ratio)	All	Ratio of EEG energy in different frequency bands $ER(ab) = E(a)/E(b)$
EE (energy entropy)	All	$EE = -\sum_i \log(P_i^2)$
DE (differential entropy)	All	$DE(X) = \frac{1}{2} \log(2\pi e \sigma^2)$
PSD (power spectral density)	All	$PSD(f) = \frac{\hat{p}}{ 1 + \sum_{k=1}^p \hat{a}_p(k)e^{-j2\pi k} }$
(ASM) Energy asymmetry	Alpha	$ASM = E(right) - E(left)$
ERASM (energy ratio asymmetry)	All	$ERASM = ER(right) - ER(left)$
DASM	Alpha	$DASM = DE(Xleft) - DE(Xright)$
RASM	Alpha	$RASM = DE(Xleft)/DE(Xright)$

EEG, electroencephalogram; DASM, differential asymmetry; RASM, rational asymmetry.

**TABLE 3** | Input data groups.

Training group	EEG length (s)	Window size (s)	Step length (s)	Input data number
Group 1	30	2	1	2,059
Group 2	30	4	2	994
Group 3	30	6	3	639
Group 4	60	2	1	4,189
Group 5	60	4	2	2,059
Group 6	60	6	3	1,349
Group 7	60	8	4	994

EEG, electroencephalogram.

**FIGURE 6** | The stacking model diagram.

First, different lengths of EEG were selected for training. The last 30 and 60 s of EEG data while watching the scenes were chosen to extract selected features for EEG regression model training. Cross-subject research was adopted to make the model more generalized. The leave-one-subject-out (LOSO) method for cross-validation was used to evaluate the accuracy (Tommaso et al., 2006). The LOSO would be performed with  $n$  iterations when given a dataset from  $n$  participants. The classifier would be trained with EEG data of  $n - 1$  participant and tested on the remaining single subject in each iteration. In this study, the whole segments of EEG data of one participant were considered as one subject data.

### Data Enhancement

To increase the number of the existing dataset to increase training accuracy, each segment of EEG data was divided into several fragments, and all the fragments in one segment were tagged with

the same label. Window sizes that were tried included 2, 4, 6, and 8 s, whereas the overlapping remaining 50%, which meant that the 2-s window corresponded to 1-s step and the 8-s window corresponded to 4-s step. The input data groups are shown in Table 3.

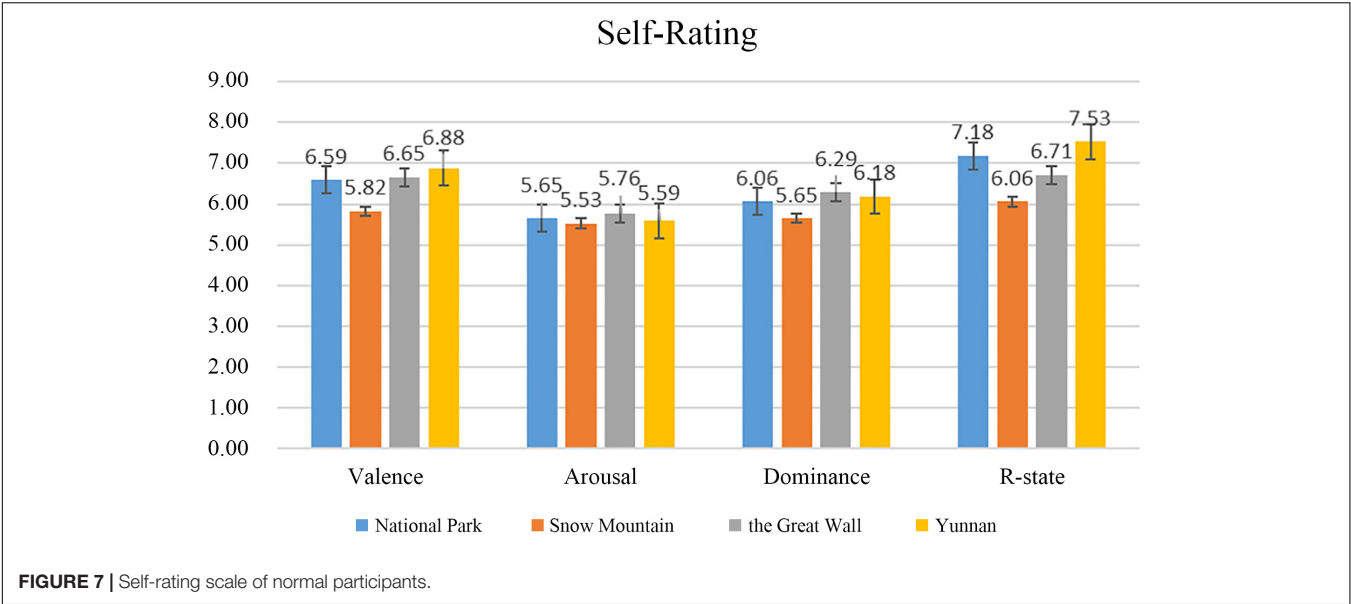
The amount of data per segment of EEG noted as  $N$  was calculated by the equation below.

$$N = \frac{L - W}{\text{step}} + 1$$

where  $L$  indicates the length of one segment of the EEG,  $W$  represents window size, and step is the overlapping length.

### Regression Model

After all the 147 features from every second of data were extracted and the mean value of each fragment was calculated, the results were then put into different regression models. Eight models were



**TABLE 4 |** EEG features with significant variance.

Begin-end		Pre-post		Pre-end			
Feature	p	Feature	p	Feature	p	Feature	p
dsp_alpha_l2	0.03240	theta1/alpha_l1	0.00450	SE_theta3	0.00008	EE_theta1	0.01359
gamma3	0.03790	EE_delta2	0.01011	EE_theta2	0.00013	alpha1_h/gamma1	0.01521
gamma2	0.03956	delta1/alpha_l1	0.01133	EE_theta3	0.00016	alpha3_h/gamma3	0.02441
alpha_h3	0.03968	delta1/theta1	0.01970	SE_theta2	0.00021	delta2/theta2	0.02842
dsp_beta_h3	0.04018	EE_alpha_l1	0.02283	SE_alpha_h2	0.00079	alpha2_h/gamma2	0.03056
delta3	0.04132	delta2/alpha_l2	0.02962	SE_alpha_h1	0.00080	alpha_h1/beta_h1	0.03092
beta_l3	0.04230	EE_beta_h3	0.03087	SE_alpha_h3	0.00083	EE_alpha_h1	0.03184
		SE_beta_l1	0.03722	alpha_h3/beta_l3	0.00152	alpha_h3/beta_h3	0.04503
		SE_beta_h3	0.03838	alpha_h1/beta_l1	0.00170	delta1/beta_l1	0.04796
		EE_delta3	0.03839	alpha_h2/beta_l2	0.00286	EE_alpha_h3	0.04892
		EE_beta_h2	0.03896	SE_delta3	0.00916		
		SE_delta2	0.04095	SE_theta1	0.00949		

EEG, electroencephalogram.

used for a comparison including linear regression (LR), support vector machine (SVM), random forest regression (RFR), adaptive boosting (AdaBoost), BootStrap aggregation (Bagging), gradient boosting (GB), eXtreme GB (XGB), and light GB (LGBM). Stacking regression was then used, which was first proposed by Leo (1996). It was a method that could integrate the outputs of multiple models to produce a new model to improve prediction accuracy. The stacking model generally consists of two levels. Several different high-prediction models with complementary advantages and disadvantages were often used at the first level; and at the second level, one simple model would be used. In this study, four different types of base regression models including RFR, AdaBoost, GB, and LGBM were used at the first level to train the original dataset referring to the results of the eight models training and previous studies (Kim et al., 2020). RFR and boosting models are the most commonly used models at the first level of stacking because these two models belong to the parallel

model and the serial model, which are quite different and have generalization to the results. And at the second level, a simple model such as LR or SVM will be used to integrate the results of the models used at the first level to prevent overfitting (David, 1992; Bohdan, 2020). LOSO method was also used in the first level so that 26-fold cross-validation would be done by each model to get predicted labels. The predictions of each test fold were then put into the second level as the training dataset, and the average of the 26-fold predictions would be taken as test datasets in the second level. Two simple models LR and SVM were tried at the second level to make a comparison. The diagram of model stacking in the regression work is shown in **Figure 6**.

Evaluation Index

Mean absolute error (MAE) and mean relative accuracy (ACC) (Li et al., 2019) were used as indices to evaluate the results of different model training. MAE calculated the absolute error

Feature	Pre → End	Pre → Post	Begin → End
E-delta/alpha_l		↑	
E-alpha_h/gamma	↓	↓	
E-alpha_l/beta_h_ASM		↑	
EE-beta_l		↓	
EE-beta_h		↓	
EE-gamma		↓	
EE-theta			↓
SE-alpha_h	↓		
SE-theta	↑		
DASM_alpha_l		↓	↓

**FIGURE 8 |** Summary of changes in electroencephalogram (EEG) characteristics of normal participants after watching the relaxation scenes.

between the predicted value and the true value. The formula is illustrated below, where  $n$  indicates the number of the EEG data,  $y_i$  is the true value, and  $\hat{y}_i$  is the predicted value. The lower the value, the better is the training model.

$$MAE = \frac{1}{n} \sum_{k=0}^n |y_i - \hat{y}_i|$$

The calculation formula of ACC is shown below with an index ranging in value from 0 to 1. The closer the value is to 1, the better is the training model. ACC reflected the relative error between the predicted and true values, which would be more comparable than MAE.

$$ACC = 1 - \frac{1}{n} \sum_{k=0}^n \frac{|y_i - \hat{y}_i|}{y_i}$$

## RESULTS

### Subjective Emotion Scale Result

The result of the subjective scale is shown in **Figure 7**. Since the relaxation degree score greater than five indicated that the scene had a relaxing effect, all the four relaxation scenes were effective (National Park 7.18, Snow Mountain 6.06, the Great wall 6.71, and Yunnan 7.53), in which Yunnan was the most relaxing VR scene. Furthermore, the results also showed that with increase in the relaxation degree, the value of valence also increased.

### Electroencephalogram Feature Analysis Result

The typically changed EEG features of the participants are shown in **Table 4** including energy features, energy ratio features, and SE features of each band.

The results of pre EEG to end EEG, pre EEG to post EEG, and begin EEG to end EEG were used as comparison groups. Ten features were found from significantly changed features, as was shown in **Figure 8**, to have consistency differences after watching the VR relaxation scenes, which meant the EEG feature values of all the 26 participants showed an increasing or a decreasing trend after watching the VR scene for each participant. The significantly changed features included E-delta/alpha\_l, E-alpha\_h/gamma, E-alpha\_l/beta\_h-ASM, EE-beta\_l, EE-beta\_h, EE-gamma, EE-theta, SE-alpha\_h, SE-theta, and DASM-alpha\_l. Moreover, features E-delta/alpha\_l, E-alpha\_l/beta\_h-ASM, and SE-theta showed an increasing trend after watching the relaxation VR scene. Features E-alpha\_h/gamma, EE-beta\_l, EE-beta\_h, EE-gamma, SE-alpha\_h, SE-theta, and DASM-alpha\_l showed a decreasing trend.

### Relaxation Regression Results

**Table 5** shows the ACC results of the different training models in each group. From the average accuracy results of the different training models in each group, it could be observed that the accuracy of XGB and LGBM reached above 80%, and LGBM got the best result of 80.42% on average. While using LGBM to train the model, Group 1 performed the best, with the accuracy of 80.69%.

Mean absolute error result is shown in **Table 6**. It could be found that in using LGBM to train Group 1, the lowest value of 1.00494 was obtained. While comparing the results of each group, it could be found that in general, using 30 s of EEG data (Groups 1, 2, and 3) to train the model would get better results than 60 s of EEG data (Groups 4, 5, 6, and 7). Furthermore, the model stacking method was applied to train Group 1. The result is shown in **Table 7**. It could be found that using stacking increased the accuracy of the predictions by approximately 1%



**TABLE 5 |** Relaxation model training ACC results.

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	AVE
LR	0.79934	0.80281	0.79370	0.70145	0.63516	0.41623	0.34172	0.64149
SVM	0.78001	0.78299	0.78501	0.77675	0.78107	0.78072	0.781061	0.78102
RF	0.80029	0.79645	0.80021	0.80219	0.80079	0.79871	0.79849	0.79959
AdaBoost	0.78393	0.78701	0.78949	0.78143	0.78367	0.78859	0.79119	0.78647
Bagging	0.80111	0.80163	0.79669	0.80190	0.79709	0.80195	0.79637	0.79953
GB	0.79695	0.79989	0.80222	0.79602	0.79792	0.79721	0.79539	0.79794
XGB	0.80199	0.80354	0.80199	0.80317	0.80138	0.79588	0.79755	0.80079
LGBM	<b>0.80692</b>	0.80431	0.80253	0.80519	0.80538	0.80277	0.80237	<b>0.80421</b>

ACC, mean relative accuracy; LR, linear regression; SVM, support vector machine; RF, random forest; GB, gradient boosting; XGB, eXtreme gradient boosting; LGBM, light gradient boosting. The bold values indicate the best experimental results.

**TABLE 6 |** Relaxation model training MAE results.

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6	Group 7	AVE
LR	1.02831	1.02005	1.09353	1.65507	2.10446	3.56199	4.04282	2.07232
SVM	1.07132	1.06300	1.05260	1.08742	1.06601	1.06833	1.06866	1.06819
RF	1.03068	1.05419	1.04110	1.02747	1.03011	1.03923	1.05331	1.03944
AdaBoost	1.18253	1.12698	1.09482	1.20856	1.18663	1.12918	1.10883	1.14822
Bagging	1.02489	1.03342	1.06349	1.03192	1.04542	1.04085	1.07164	1.04452
GB	1.04648	1.03242	1.03109	1.05666	1.05017	1.05481	1.06062	1.04746
XGB	1.05373	1.06005	1.07795	1.04851	1.06283	1.09146	1.09887	1.07049
LGBM	<b>1.00494</b>	1.01620	1.03201	1.01564	1.01933	1.02982	1.03269	<b>1.02152</b>

MAE, mean absolute error; LR, linear regression; SVM, support vector machine; RF, random forest; GB, gradient boosting; XGB, eXtreme gradient boosting; LGBM, light gradient boosting. The bold values indicate the best experimental results.

and decreased the MAE values by 1, which indicated that the model was optimized. Moreover, using SVM at the second level got better results than LR.

## APPLICATION

The VR relaxation scenes were applied to assist in the treatment of patients with depression in the Guangzhou First People's Hospital. Twenty-two patients with first-episode depression including six men and 16 women (age ranging from 19 to 50 years) volunteered for the VR treatment. Each patient was asked to watch only one VR scene, the Great Wall, to avoid discomfort caused by watching VR for a long time. EEG was also acquired during the procedure, and patients were asked to verbally answer how they felt after watching. **Figure 9** shows the data collection settings for the patients, who were asked to sit and watch the VR scene wearing the VR glasses in front of a table on which there was a computer monitor and the EEG acquisition device. Written consent was obtained from each participant before the experiment.

Patients' subjective answer results are shown in **Table 8**. It could be seen that most patients with depression felt relaxed after watching the VR scene except for two patients.

The EEG datasets of the depression patients were preconditioned in the same manner as those of normal college students. After preprocessing, the last 30-s EEG data while watching the relaxation scene were put into the stacking

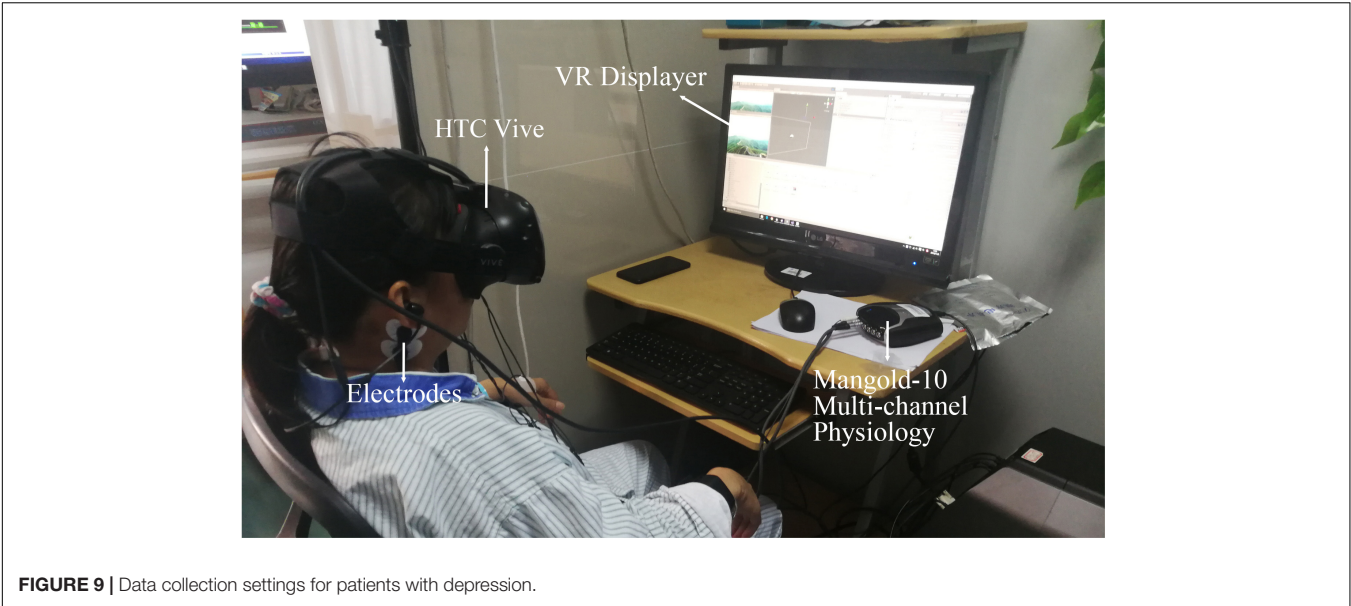
**TABLE 7 |** Stacking model results.

Second level model	MAE	ACC
LR	0.98942	0.81216
SVM	<b>0.46846</b>	<b>0.81462</b>

MAE, mean absolute error; ACC, mean relative accuracy; LR, linear regression; SVM, support vector machine. The bold values indicate the best experimental results.

regression model whose second layer was SVM to predict the R-state to demonstrate the effectiveness of the VR relaxation scene to depression patients. The method of EEG data progress was the same as that of Group 1, of which the R-state prediction result was the best. The predicted R-state results of disorder patients are shown in **Table 9**, demonstrating that all the prediction values were over 5, and the average of the predicted R-state was 6.54, which was close to the subjective rating value (6.71). These results confirmed that the VR relaxation scene has a positive effect on the relaxation therapy of patients with depression.

The EEG features of patients with depression were analyzed in the same manner as those of normal participants. Five features had significant consistency differences after watching the VR relaxation scenes, including E-delta/beta<sub>l</sub>, E-delta/beta<sub>h</sub>, E-alpha<sub>l</sub>/gamma, E-alpha<sub>h</sub>/gamma, and E-beta<sub>l</sub>/gamma, as shown in **Figure 10**. All of these changing features were on a downward trend. Moreover, E-alpha<sub>h</sub>/gamma had the same trend as that of normal participants.



**TABLE 8 |** Relaxation model training MAE results.

Relaxation state	Subject serial number	Total number	Male number	Female number
Very relaxed	Subject 1, 3, 4, 5–10, 15, 20, 22	12	1	11
A little relaxed	Subject 2, 11, 13, 14, 17, 18, 19, 21	8	4	4
No relaxed	Subject 12, 16	2	1	1

MAE, mean absolute error.

**TABLE 9 |** Predicted R-state of each patient.

Subject number	Predicted R-state	Subject number	Predicted R-State
Subject 1	7.15	Subject 12	5.76
Subject 2	6.32	Subject 13	7.13
Subject 3	6.55	Subject 14	6.64
Subject 4	7.36	Subject 15	6.63
Subject 5	7.11	Subject 16	5.66
Subject 6	5.92	Subject 17	5.68
Subject 7	7.09	Subject 18	6.18
Subject 8	6.64	Subject 19	6.53
Subject 9	6.82	Subject 20	5.51
Subject 10	7.01	Subject 21	6.56
Subject 11	6.68	Subject 22	6.85
Average result		6.54	

## DISCUSSION

The subjective evaluations in the normal participants and patients with depression demonstrated that the sightseeing-relaxation VR scenes with new age music had relaxing effects. However, it was hard to say whether the visual scene or the auditory music had a greater effect on relaxation. Some studies have found that auditory stimulations aroused emotions much better than visual materials. Therefore, the impact of relaxation VR scenes and relaxing background music must be explored.

Different data processing methods were used to analyze the EEG datasets for R-state study in this study. The length of chosen EEG was 30 s including sections during watching VR scenes and during the 2-min relaxing period. In the previous study, 60 s of EEG data of one section before and after watching the VR scenes were used (Zhu et al., 2019). However, since the participant’s emotion was easily influenced by other psychological activities during 2 min of relaxation before and after watching the VR scene, the 60 s of EEG data was relatively long. While training the R-state model, using 30 s of EEG data also showed

Feature	Pre → End	Pre → Post	Begin → End
E-delta/alpha_l	↓		↓
E-alpha_h/gamma	↓		↓
E-alpha_l/beta_h_ASM	↓	↓	
EE-beta_l	↓	↓	
EE-beta_h	↓	↓	

**FIGURE 10 |** Summary of changes in electroencephalogram (EEG) characteristics of patients with depression after the relaxation scenes.

higher accuracy and lower MAE than that at 60 s, consistent with Kumar's research (Nitin et al., 2016).

From the EEG feature analysis results, it could be found that most of the distinctive features were theta, alpha, beta, and gamma waves, which were consistent with Cahn's research (Baruch et al., 2016). The appearance of the beta wave was associated with mental tension and emotional excitement. When people felt relaxed, the energy and entropy of the beta wave should go down. As a result, the feature values of EE-beta\_h and EE-beta\_l went down. Moreover, since many studies have found that meditation and relaxation could increase gamma wave, the value of E-alpha\_l/gamma and E-alpha\_h/gamma increased after watching relaxation scenes. The significant variance in gamma-related features also indicated that the relaxation effects of the sightseeing scenes used in the experiment might be similar to those of meditation.

Moreover, it could be easily detected from **Figure 10** that all the features with significant differences among the group of patients with depression were energy ratios, most of which were beta- and gamma-related features. It was probably because depression varied widely among individuals, and the ratio-related features could neutralize some of the individual differences (Kan and Lee, 2015). Moreover, since patients with depression felt stressed more easily, beta wave, which was associated more with anxiety, would more likely to be affected by relaxation scenes. As per Smith's theory, reducing stress enhanced relaxation (Smith, 1988). Many studies have proved that patients with depression had increased alpha (Hosseini et al., 2013) and beta power (Clark et al., 2016). The decreased features of E-alpha\_l/gamma, E-alpha\_h/gamma, and E-beta\_l/gamma might also have confirmed the effectiveness of VR-relaxation therapy in treating depression. When comparing the changes of EEG features between normal people and patients with depression, it could be found that there existed one feature, E-alpha\_h/gamma, having the same trend as that of normal participants. Related studies have found that the gamma wave and alpha wave of

normal people and patients with depression are relatively sensitive (Grey et al., 2010), which might cause the same significant changes. Although the findings have not been medically proven yet, these results have provided a reference for future studies.

It could be found from **Table 9** that all the EEG data of depression patients predicted a level of R-state greater than 5, which meant that the emotions of all the patients were predicted to be relaxed. However, two patients were not relaxed in their subjective assessment. This difference might be due to the fact that the relaxation model was based on EEG data from normal people. Since the R-state model was built by datasets of normal people, and there existed some differences between the EEG of normal individuals and patients with depression (Davidson et al., 2002; Acharya et al., 2015), using the R-state model to predict people with depression might not be particularly accurate. However, one feature with the same trend when watching the relaxation scenes in these two groups was found, and the predicted results of the R-state still have reference value. Therefore, it is necessary to train the relaxation model using the EEG data and relaxation label of patients with depression. In further research, more experiments would be conducted covering college students, patients with depression, and some other groups of people, to verify the application scopes of the relaxation model.

## CONCLUSION

In this study, VR relaxation scenes were used to promote the R-state for college students. Some EEG features were found to have a consistent significant trend of variance among different participants while watching the relaxation scenes, including EE-gamma, E-alpha\_h/gamma, and DASM. These significantly changed features provide a reference for optimizing the relaxation prediction model and relaxation interaction system research based on EEG in the future. Eight machine learning models including LR, SVM, and LGBM were conducted to train

the R-state regression model, and the LOSO method for cross-validation was used to evaluate the results. The mean accuracy reached approximately 80.42% using the LGBM model. Model stacking methods were then applied to optimize the model. The mean accuracy of the framework achieved approximately 81.46%, which increased by approximately 1%. The VR relaxation scenes were then used to help with the treatment of patients with depression, which have received good results. This work provides an objective index reference for the evaluation and treatment of depression using VR relaxation scenes and also explores the feasibility of VR relaxation scenes in the adjuvant treatment of depression.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Guangzhou First People's Hospital (202002030262, on April 1 2020). The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

YZ and LLZ were responsible for the entire study, including the study concept, the study design, and the application. YZ, HH, JJ, and LQZ contributed to the VR scene design. YZ, JJ, LS, and XX

were responsible for the EEG collection design and data analysis. FK and YL helped with the EEG collection. All authors listed have made a direct and intellectual contribution to the work.

## FUNDING

This work was supported in part by the Major Science and Technology Projects in Guangdong Province under Grant 2016B010108008, in part by the Technology Program of Guangzhou under Grants 202002030354 and 202002030262, in part by the Science and Technology Project of Zhongshan under Grants 2019AG024 and 2020B2053, in part by the Natural Science Foundation of Guangdong Province under Grant 2018A030310407, in part by the Guangzhou Key Laboratory of Body Data Science under Grant 201605030011, and in part by the National Natural Science Foundation of China under Grants 61972163 and 61806210.

## ACKNOWLEDGMENTS

We would also like to thank all participants and thank the Guangzhou First People's Hospital for medical support. We would also like to thank Editage ([www.editage.com](http://www.editage.com)) for English-language editing.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2021.719869/full#supplementary-material>

## REFERENCES

- Acharya, U. R., Sudarshan, V. K., Adeli, H., Santhosh, J., Koh, J. E., and Adeli, A. (2015). Computer-aided diagnosis of depression using EEG signals. *Eur. Neurol.* 73, 329–336.
- Alexandros, L., Christos, K., Dimitris, S., Michalis, X., and Nikos, K. (2015). "Recognizing emotions in human computer interaction: studying stress using skin conductance," in *Proceedings of the 15th Human-Computer Interaction (INTERACT)*, Bamberg, 255–262.
- Allison, P. A., Michael, D. M., Abigail, M. F., Devin, R. C., Mark, T. H., and Jay, C. B. (2017). Relaxation with immersive natural scenes presented using virtual reality. *Aerosp. Med. Hum. Perform.* 88, 520–526.
- Andersen, T., Anisimovaite, G., Christiansen, A., Hussein, M., Lund, C., Nielsen, T. L., et al. (2017). A preliminary study of users' experiences of meditation in virtual reality *Paper Presented at the 2017 IEEE Virtual Reality (VR)* (Los Angeles, CA: IEEE), 343–344.
- Baruch, R. C., Arnaud, D., and John, P. (2016). Occipital gamma activation during vipassana meditation. *Cogn. Process.* 11, 39–56.
- Bohdan, M. P. (2020). "Using bayesian regression for stacking time series predictive models," in *Proceedings of the 2020 IEEE 3rd International Conference on Data Stream Mining & Processing (DSMP)* (Lviv: IEEE), 305–309.
- Bradley, M. M., and Lang, P. J. (1994). Measuring emotion: the self-assessment manikin and the semantic differential. *J. Behav. Ther. Exp. Psychiatry* 25, 49–59.
- Claas, L., Rainer, S., Peter, H., Joram, R., Moritz, M., Thomas, L., et al. (2008). Brief relaxation versus music distraction in the treatment of dental anxiety: a randomized controlled clinical trial. *J. Am. Dent. Assoc.* 139, 317–324.
- Clark, D. L., Brown, E. C., Ramasubbu, R., and Kiss, Z. H. T. (2016). Intrinsic local beta oscillations in the subgenual cingulate relate to depressive symptoms in treatment-resistant depression. *Biol. Psychiatry* 80, 93–94.
- Cohen, S., Kamarck, T., and Mermelstein, R. (1983). A global measure of perceived stress. *J. Health Soc. Behav.* 24, 385–396.
- David, H. W. (1992). Stacked generalization. *Neur. Net.* 5, 241–259.
- Davidson, R. J., Pizzagalli, D., Nitschke, J. B., and Putnam, K. (2002). Depression: perspectives from affective neuroscience. *Annu. Rev. Psychol.* 53, 545–574.
- DeBerry, S. (1982). The effects of meditation-relaxation on anxiety and depression in a geriatric population. *Psychother. Theory Res. Prac.* 19, 512–521.
- Deng, X., Yang, M., and An, S. (2021). Differences in frontal EEG asymmetry during emotion regulation between high and low mindfulness adolescents. *Biol. Psychol.* 158:107990.
- Dombrowe, I., and Hilgetag, C. C. (2014). Occipitoparietal alpha-band responses to the graded allocation of top-down spatial attention. *J. Neurophysiol.* 112, 1307–1316.
- Du, R., and Lee, H. (2015). "Frontal alpha asymmetry during the audio emotional experiment revealed by event-related spectral perturbation," in *Proceedings of the 2015 8th International Conference on Biomedical Engineering and Informatics (BMEI)* (Shenyang: IEEE), 531–536.
- Ernst, E., and Kanji, N. (2000). Autogenic training for stress and anxiety: a systematic review. *Complement Ther. Med.* 8, 106–110.
- Freeman, D., Reeve, S., Robinson, A., Ehlers, A., Clark, D., Spanlang, B., et al. (2017). Virtual reality in the assessment, understanding, and treatment of mental health disorders. *Psychol. Med.* 47, 2393–2400.



- Grey, J. S., Ruth, C., Michael, E. T., Matcheri, K., and Stuart, R. S. (2010). Sustained gamma-band EEG following negative words in depression and schizophrenia. *Int. J. Psychophysiol.* 75, 107–118.
- Hosseinfard, B., Moradi, M. H., and Rostami, R. (2013). Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. *Comput. Methods Programs Biomed.* 109, 339–345.
- Hou, H., Zhang, X., and Hao, M. (2020). Odor-induced emotion recognition based on average frequency band division of EEG signals. *J. Neurosci. Methods* 334:108599.
- Ilse, V. D., Karen, V., Andre, E. A., Devy, W., Debora, V., and Elke, V. (2014). Inhalation/exhalation ratio modulates the effect of slow breathing on heart rate variability and relaxation. *Appl. Psychophysiol. Biofeedback* 39, 171–180.
- Imran, M., Mohd, E. R., Azmi, M. Y., Mohd, Z. M. Y., and Ahmad, R. R. Z. (2014). “Study of immersion effectiveness in VR-based stress therapy,” in *Proceedings of the 6th International Conference on Information Technology and Multimedia* (Putrajaya: IEEE), 380–384.
- Jenke, R., Peer, A., and Buss, M. (2014). Feature extraction and selection for emotion recognition from EEG. *IEEE Trans. Aff. Comp.* 5, 327–339.
- Joseph, W., Brianna, S. S., Antoine, L., and Richard, J. D. (2016). Long-term mindfulness training is associated with reliable differences in resting respiration rate. *Sci. Rep.* 6:27533.
- Kan, D. P. X., and Lee, P. (2015). “Decrease alpha waves in depression: an electroencephalogram (EEG) study,” in *Proceedings of the 2015 International Conference on BioSignal Analysis, Processing and Systems (ICBAPS)* (Kuala Lumpur: IEEE), 156–161.
- Kiehl, N., Figas, P., and Bichlmeier, C. (2018). “Effects of graphical styles on emotional states for VR-supported psychotherapy,” in *Proceedings of the 2018 10th International Conference on Virtual Worlds and Games for Serious Applications (VS-Games)*, Würzburg, 1–4.
- Kim, C., You, S. C., Reps, J. M., Cheong, J. Y., and Park, R. W. (2020). Machine-learning model to predict the cause of death using a stacking ensemble method for observational data. *J. Am. Med. Inform. Assoc.* 28, 1098–1107.
- Knott, V., Bakish, D., Lusk, S., and Barkely, J. (1997). Relaxation-induced EEG alterations in panic disorder patients. *J. Anxiety Disord.* 11, 365–376.
- Kwon, O. Y., Kam, S. C., Choi, J. H., Do, J. M., and Hyun, J. S. (2011). Effects of sertraline on brain current source of the high beta frequency band: analysis of electroencephalography during audiovisual erotic stimulation in males with premature ejaculation. *Int. J. Impot. Res.* 23, 213–219.
- Lazar, S. W., Bush, G., Gollub, R. L., Fricchione, G. L., Khalsa, G., and Benson, H. (2000). Functional brain mapping of the relaxation response and meditation. *Neuroreport* 11, 1581–1585.
- Leo, B. (1996). Stacked regressor. *Mach. Learn.* 24, 49–64.
- Li, Y., Zheng, W., Cui, Z., and Zong, Y. (2019). EEG emotion recognition based on graph regularized sparse linear regression. *Neural Process. Lett.* 49, 555–571.
- Linda, G., Stefan, L., and Maic, M. (2020). “Playing in virtual nature: improving mood of elderly people using VR technology,” in *Proceedings of the MuC '20 Conference on Mensch und Computer*, Munich, 155–164.
- Lolak, S., Connors, G. L., Sheridan, M. J., and Wise, T. N. (2008). Effects of progressive muscle relaxation training on anxiety and depression in patients enrolled in an outpatient pulmonary rehabilitation program. *Psychother. Psychosom.* 77, 119–125.
- Narendra, J., Ramchandra, M., and Yashwant, J. (2017). Effect of meditation on emotional response: an EEG-based study. *Biomed. Signal Process. Control* 34, 101–113.
- Nitin, K., Kaushikee, K., and Shyamanta, M. H. (2016). Bispectral analysis of EEG for emotion recognition. *Procedia Comp. Sci.* 84, 31–35.
- Olga, S., Yisi, L., and Minh, K. N. (2011). Real-time EEG-based emotion recognition for music therapy. *J. Multimodal User Interfaces* 5, 27–35.
- O’Sullivan, R. J. (1991). A musical road to recovery: music in intensive care. *Intensive Care Nurs.* 7, 160–163.
- Patil, S., and Shirley, T. (2006). Effects of two yoga based relaxation techniques on Heart Rate Variability (HRV). *Int. J. Stress. Manag.* 13, 460–475.
- Randy, P. A., Corina, B., Ronny, B., David, E., Eirini, K., and Ronald, C. K. (2019). The World Health Organization world mental health international college student initiative: an overview. *Int. J. Methods Psychiatr. Res.* 28:e1761.
- Shu, L., Xie, J., Yang, M., Li, Z., Li, Z., Liao, D., et al. (2018). A review of emotion recognition using physiological signals. *Sensors* 18:2074.
- Shu, L., Yu, Y., Chen, W., Hua, H., Li, Q., Jin, J., et al. (2020). Wearable emotion recognition using heart rate data from a smart bracelet. *Sensors (Basel)* 20:718.
- Sidra, J. G., Avshalom, C., Honalee, H., Sean, H., Shyamala, N., Richie, P., et al. (2013). Suicide attempt in young people a signal for long-term health care and social needs. *JAMA Psychiatry* 71, 119–127.
- Smith, J. C. (1988). Steps toward a cognitive-behavioral model of relaxation. *Biofeedback Self Regul.* 13, 307–329.
- Smith, J. C. (2005). *Relaxation, Meditation, and Mindfulness: A Mental Health Practitioner's Guide to New and Traditional Approaches*. New York, NY: Springer Publishing Co.
- Soraia, M. A., and Manuel, J. F. (2019). Emotions recognition using EEG signals: a survey. *IEEE Trans. Aff. Comp.* 10, 374–393.
- Suranjita, G., and Rajesh, S. (2019). Electrode channel selection for emotion recognition based on EEG signal Paper Presented at the 2019 IEEE 5th International Conference for Convergence in Technology (I2CT) (Bombay: IEEE), 1–4.
- Theresa, M. M., and Hilary, B. (1992). The development of a six-item short-form of the state scale of the Spielberger State-Trait Anxiety Inventory (STAI). *Br. J. Clin. Psychol.* 31, 301–306.
- Tommaso, M., Marinazzo, D., and Stramaglia, S. (2006). The measure of randomness by leave-one-out prediction error in the analysis of EEG after laser painful stimulation in healthy subjects and migraine patients. *Clin. Neurophysiol.* 116, 2775–2782.
- Tortella-Feliu, M., Morillas-Romero, A., Balle, M., Llabrés, J., Bornas, X., and Putman, P. (2014). Spontaneous EEG activity and spontaneous emotion regulation. *Int. J. Psychophysiol.* 94, 365–372.
- Wesley, E. S., and Douglas, E. D. (1977). Effect of EMG biofeedback and progressive muscle relaxation training on awareness of frontalis muscle tension. *Psychophysiology* 14, 522–530.
- WHO (2020). *E. coli. Mental Health: Global Strategic Direction*. Available online at: <https://www.who.int/observatories/global-observatory-on-health-research-and-development/analyses-and-syntheses/mental-health/global-strategic-direction> (accessed December 28, 2020).
- Xing, X., Li, Z., Xu, T., Shu, L., Hu, B., and Xu, X. (2019). SAE+LSTM: a new framework for emotion recognition from multi-channel EEG. *Front. Neurobot.* 13:37. doi: 10.3389/fnbot.2019.00037
- Zheng, W., and Lu, B. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Trans. Auton. Ment. Dev.* 7, 162–175.
- Zhu, L., Tian, X., Xu, X., and Shu, L. (2019). “Design and evaluation of the mental relaxation VR scenes using forehead EEG features,” in *Proceedings of the 2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)* (Piscataway, NJ: IEEE), 1–4.

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships 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.

Copyright © 2021 Zhang, Zhang, Hua, Jin, Zhu, Shu, Xu, Kuang and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



# Monitoring and Evaluation of Emotion Regulation by Aerobic Exercise and Motor Imagery Based on Functional Near-Infrared Spectroscopy

Peng Ding<sup>1,2,3</sup>, Fawang Wang<sup>1,2,3</sup>, Siyu Li<sup>1,2,3</sup>, Wei Zhang<sup>4</sup>, Hongquan Li<sup>1,2,3</sup>, Zhuangfei Chen<sup>2,3</sup>, Lei Zhao<sup>2,5</sup>, Anmin Gong<sup>6\*</sup> and Yunfa Fu<sup>1,2\*</sup>

<sup>1</sup> Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming, China, <sup>2</sup> Brain Cognition and Brain-Computer Intelligence Integration Innovation Group, Kunming University of Science and Technology, Kunming, China, <sup>3</sup> Brain Science and Visual Cognition Research Center, Medical School of Kunming University of Science and Technology, Kunming, China, <sup>4</sup> School of Rehabilitation, Kunming Medical University, Kunming, China, <sup>5</sup> Faculty of Science, Kunming University of Science and Technology, Kunming, China, <sup>6</sup> Information Engineering College, Engineering University of People's Armed Police, Xi'an, China

## OPEN ACCESS

### Edited by:

Jane Zhen Liang,  
Shenzhen University, China

### Reviewed by:

Ping Xie,  
Yanshan University, China  
Xiaotian Wang,  
Xidian University, China

### \*Correspondence:

Anmin Gong  
gonganmincapf@163.com  
Yunfa Fu  
fyf@ynu.edu.cn

**Received:** 16 August 2021

**Accepted:** 23 September 2021

**Published:** 27 October 2021

### Citation:

Ding P, Wang F, Li S, Zhang W, Li H, Chen Z, Zhao L, Gong A and Fu Y (2021) Monitoring and Evaluation of Emotion Regulation by Aerobic Exercise and Motor Imagery Based on Functional Near-Infrared Spectroscopy. *Front. Comput. Neurosci.* 15:759360. doi: 10.3389/fncom.2021.759360

**Objective:** We sought to effectively alleviate the emotion of individuals with anxiety and depression, and explore the effects of aerobic exercise on their emotion regulation. Functional near-infrared spectroscopy (fNIRS) brain imaging technology is used to monitor and evaluate the process of aerobic exercise and imagination that regulates emotion.

**Approach:** Thirty participants were scored by the state-trait anxiety inventory (STAI) and profile of mood states (POMS), and fNIRS images were collected before, after, and during aerobic exercise and motor imagery. Then, the oxygenated hemoglobin (HbO), deoxygenated hemoglobin (HbR), and total hemoglobin (HbT) concentrations and their average value were calculated, and the ratio of HbO concentration in the left and right frontal lobes was determined. Spearman's correlation coefficient was used to calculate the correlation between variations in the average scores of the two scales and in blood oxygen concentrations.

**Results:** In comparison with motor imagery, STAI, and POMS scores decreased after 20 min of aerobic exercise. The prefrontal cortex had asymmetry and laterality (with the left side being dominant in emotion regulation). The increase in hemoglobin concentration recorded by fNIRS was negatively correlated with STAI and POMS scores. Aerobic exercise has a good effect on emotion regulation.

**Significance:** The study showed that portable fNIRS could be effectively used for monitoring and evaluating emotion regulation by aerobic exercise. This study is expected to provide ideas for constructing fNIRS-based online real-time monitoring and evaluation of emotion regulation by aerobic exercise.

**Keywords:** functional near-infrared spectroscopy, aerobic exercise, emotion regulation, motor imagery, state-trait anxiety inventory, profile of mood states

## INTRODUCTION

The rapid development of society has placed a certain amount of pressure on individuals or groups, which can trigger different degrees of anxiety or depression. If individuals do not pay attention to emotion regulation, physical, and mental illnesses may occur, with severe ones capable of leading to mental dysfunction or adverse social events (Veerapa et al., 2020). Aerobic exercise is one option to improve the mood and promote the generation of positive emotions (Brush et al., 2020) that can be used by people with severe anxiety for emotion regulation (Tempest and Parfitt, 2013). However, there is currently a lack of monitoring and evaluation of aerobic exercise in mood regulation. Meanwhile, the pleasant, comfortable, or energetic experience of motor imagery may also further promote the regulation of emotion (Tempest and Parfitt, 2013), but its regulatory effect on motion is still unclear. Therefore, in the study, functional near-infrared spectroscopy (fNIRS) brain imaging was used to monitor and evaluate the effect of aerobic exercise and motor imagery on emotion regulation (Jiang et al., 2017; Veerapa et al., 2020).

Improving the emotional health of people with anxiety is very important. We hypothesize that corresponding aerobic exercise imagination may also help to improve the mood of individuals with anxiety, just as the memory of a good experience can enhance their mood. Motor imagery based on aerobic exercise was designed to verify our hypothesis. In general, aerobic exercise of an individual is familiar, easy, and can be carried out naturally, habitually, and automatically. In this context, the brain does not need to recruit too many nervous system resources. Therefore, individuals can allocate certain psychological resources to carry out aerobic exercise imagination when they perform aerobic exercise automatically (Tempest and Parfitt, 2013). The present study used fNIRS to monitor and evaluate emotion regulation by aerobic exercise and motor imagery.

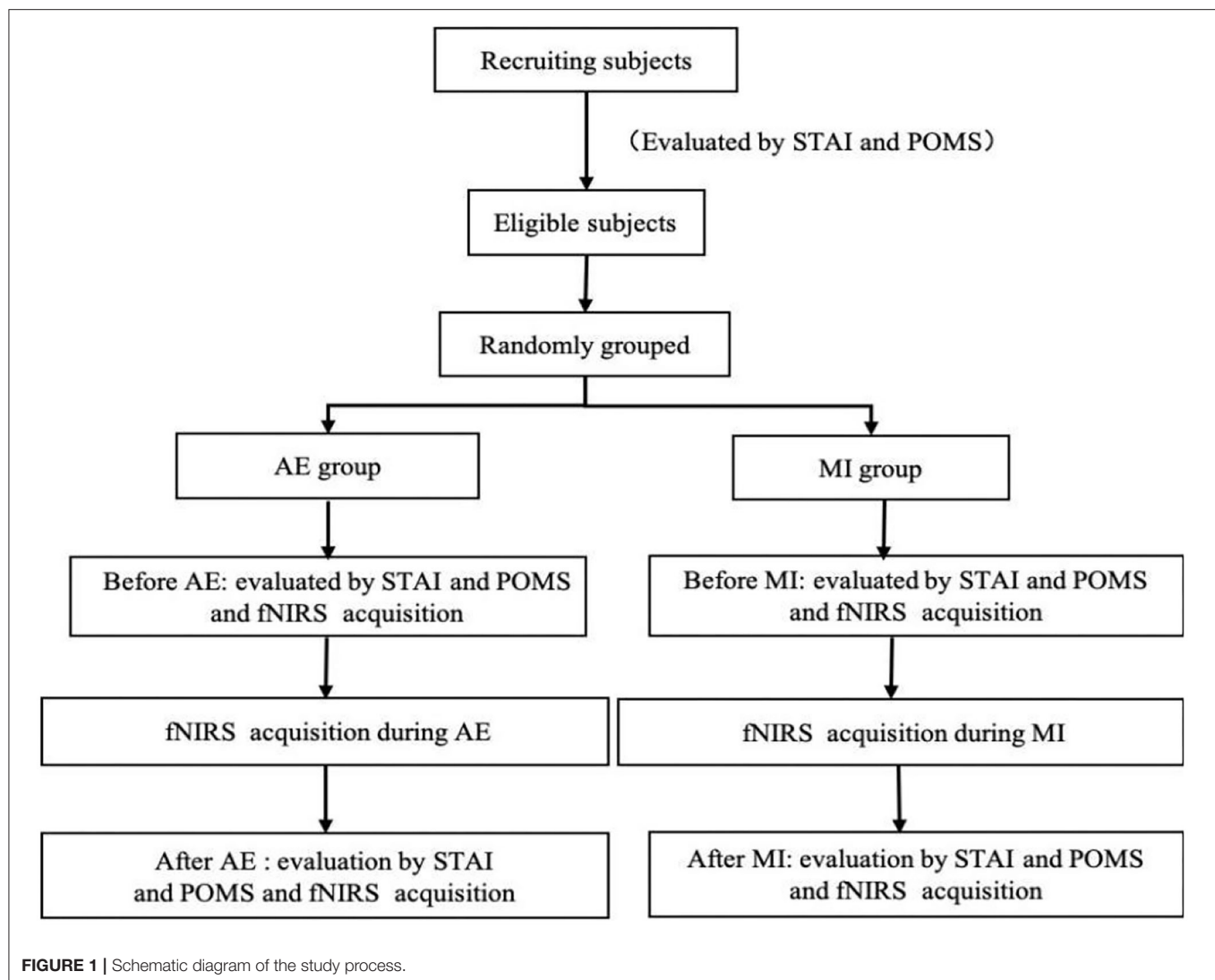
Sports medicine shows that aerobic exercise is a type of physical exercise that can involve the full exchange of oxygen to achieve physiological balance. At the same time, the heart rate needs to reach 150 bpm before aerobic exercise; therefore, aerobic exercise generally includes moderate- or high-intensity activities (60–80% of the maximum heart rate), such as jogging, walking, fast running, cycling, swimming, or rope skipping. These exercises are considered aerobic exercise, which is conducive to the health of the body (Cheng, 2007). Practice and sports medicine studies also have suggested that aerobic exercise can regulate emotion, alter mood, improve negative emotions, and promote the production of positive emotions (Ekkekakis et al., 2013; Tempest et al., 2014; Bernstein and McNally, 2015; Bernstein and McNally, 2017; Edwards et al., 2017, 2018; Brush et al., 2020). However, the changes of blood oxygen metabolism in brain tissue under aerobic exercise and the relationship between these changes and emotional state need to be further discussed.

In comparison with electroencephalography (EEG), fNIRS is less sensitive to motion artifacts, has a good ecological effect, and can tolerate a certain degree of exercise interference (Sitaram et al., 2007; Cui et al., 2011; Naseer and Hong, 2013). Moreover, fNIRS can measure the blood oxygen metabolism

(HbO and HbR) of brain tissue during aerobic exercise, while EEG measures the discharge activities of central neurons. The spatial resolution and spatial positioning accuracy of fNIRS are better than those of EEG. fNIRS is also non-invasive and portable. Thus, relative to EEG, fNIRS may be more suitable for monitoring and evaluating the emotion-regulation effects of aerobic exercise. In addition, as compared with fNIRS, functional magnetic resonance imaging and magnetoencephalography are bulky, not portable, and expensive, and are not suitable for monitoring and evaluating brain activity during aerobic exercise emotion regulation (Weiskopf et al., 2004; Goldin et al., 2013).

Some research has examined the influence of aerobic exercise on changes in the HbO concentration as measured by fNIRS (Chen et al., 2017), but few investigators have evaluated its impact on changes in HbR and HbT, as measured by fNIRS. In the present study, three characteristics of fNIRS (HbO, HbR, and HbT) were extracted to evaluate the concentration changes that occurred between before and after exercise and the ratio of HbO concentration changes between the left and right prefrontal lobes before and after aerobic exercise, and motor imagery was calculated to investigate the lateralization of the activation of brain regions. Blackhart et al. used a questionnaire for pre- and post-test assessments to verify that the degree of left frontal EEG can predict the symptoms of anxiety and depression (Blackhart et al., 2006), while Smith et al. used a model to calculate the risk of anxiety and depression and revealed the correlation between the degree of left frontal EEG and said risk (Smit et al., 2007), which also supported the conclusions of Blackhart et al. At the same time, patients with anxiety and depression also exhibit decreased left frontal lobe activity. Therefore, the degree of left lateralization of frontal EEG can be used as an indicator of anxiety and depression to a certain extent and to predict the development of symptoms. However, fNIRS has not been used to observe frontal lobe asymmetry nor has it been applied to detect the effects of aerobic exercise on emotion regulation. Therefore, we chose to use fNIRS to observe these two brain regions and analyze the relationship between the changes in neural mechanisms and emotion regulation in these two brain regions to confirm the hypothesis of this experiment, which is as follows: fNIRS can be used as a means of monitoring and evaluating emotion regulation. We present a three-dimensional topographic map of the dynamic changes in HbO concentration that occur during aerobic exercise. In addition, previous studies have used either the profile of mood states (POMS) (Chen et al., 2017) or the state-trait anxiety inventory (STAI) (Chen et al., 2019; Clemente-Suárez, 2020) for assessing the emotion regulation of individuals with state-trait anxiety. To more comprehensively evaluate the effects of emotion regulation by aerobic exercise, POMS and STAI were used in this study, and an ANOVA was used to analyze aerobic exercise and motor imagery, focusing on the significance of changes in the two scores between before and after aerobic exercise.

In addition to the above-mentioned assessment of emotion regulation by aerobic exercise, few people have studied whether motor imagery can regulate emotion, and especially few people have monitored and evaluated it using fNIRS. Francesca et al. showed that motor imagery can promote or inhibit related neural



activities and then regulate individual anxiety (Fardo et al., 2015). Shafir et al. also reported that individuals can regulate their own emotions by imagining aerobic exercise or by means conducive to regulating their anxiety (Tal, 2016). In addition, until now, few people have explored the correlation between STAI and POMS scores and the changes in HbO, HbR, and HbT concentrations based on fNIRS. In this study, Spearman's correlation coefficient was used to analyze the correlation between STAI and POMS scores and changes in the HbO, HbR, and HbT concentrations. The emotion subscale has a certain participativity in the evaluation of emotional changes (Knapen et al., 2009; Szabo et al., 2015; Subramaniapillai et al., 2016; Bernstein and McNally, 2018). In addition to using STAI and the POMS emotion subscale to monitor and evaluate emotional changes, fNIRS technology was introduced to monitor and evaluate changes in cerebral blood oxygen metabolism before, during, and after aerobic exercise, which is expected to improve the objectivity of emotion monitoring and evaluation. The present study is expected to provide ideas for developing fNIRS-based online real-time

monitoring and evaluation of emotion regulation by aerobic exercise and motor imagery, which can be used to monitor and evaluate individual state-trait anxiety and mood states.

## MATERIALS AND METHODS

### Research Scheme

**Figure 1** shows a schematic diagram of the scheme of this study, which will be described in detail henceforth. In this study, eligible subjects were randomly divided into an aerobic exercise group and exercise imagination group, and then STAI and POMS were evaluated, and fNIRS was collected before and after aerobic exercise and exercise imagination tasks. Perform fNIRS acquisition during the mission.

### Study Participants

A total of 30 participants (21 men, average age:  $23 \pm 2.0$  years, age range: 21–26 years; nine women, average age:  $23 \pm 2.0$  years) were asked to complete STAI and POMS to evaluate whether



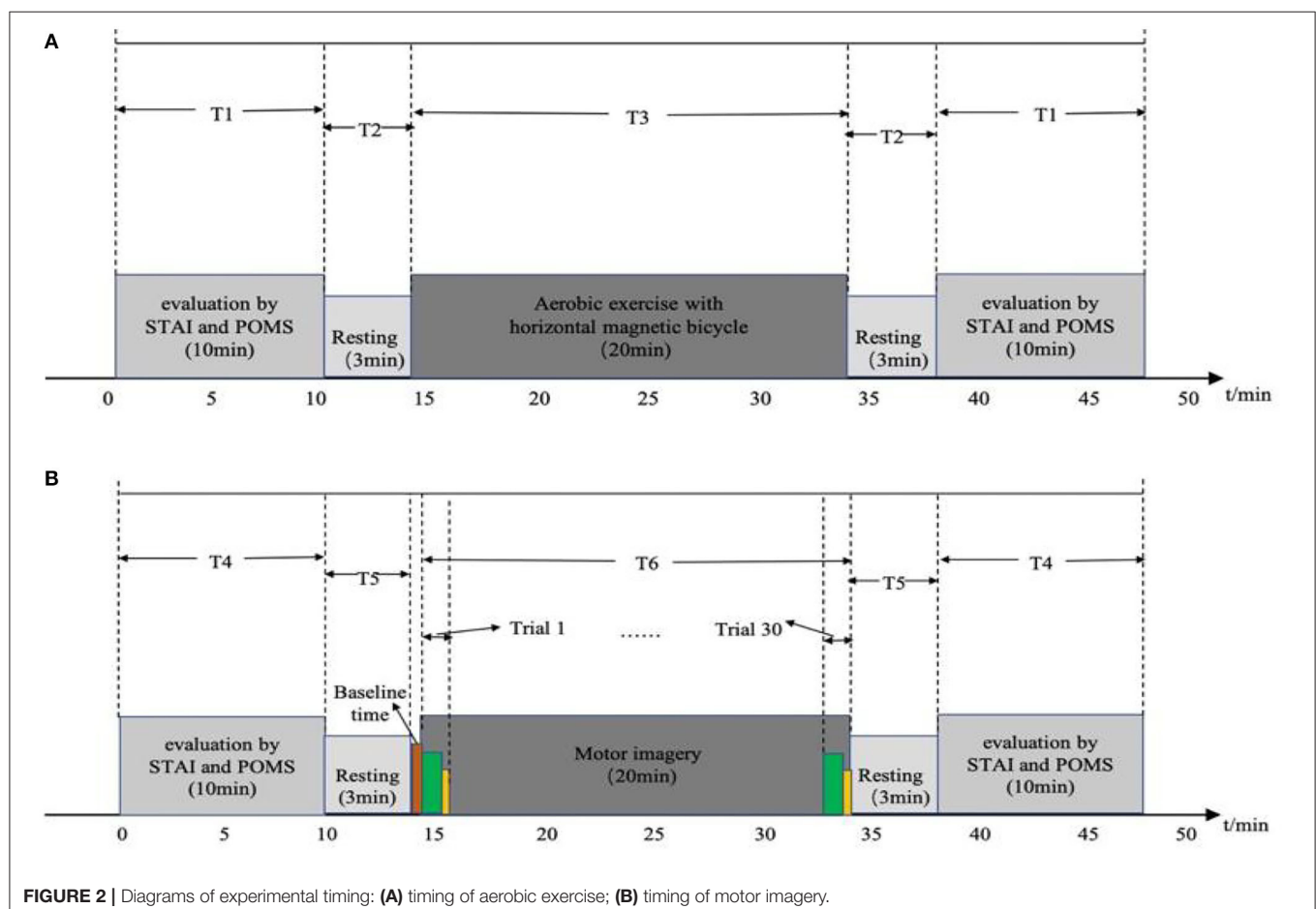
they met the inclusion criteria (STAI score of 40–59 points and POMS score of 110–140 points). The selected participants were divided into two groups ( $n = 15$  each), an aerobic exercise group (S1–S15) and a motor imagery group (S16–S30), according to their height, body shape, sex, age, and other factors; there was no significant difference in these factors between the two groups. All of the participants were undergraduates or graduate students, were right-handed, had no history of mental, neurological, or musculoskeletal disease or drug abuse, had normal or corrected vision, and had no color blindness. All of the participants signed the experimental informed consent form, and this study was approved by the medical ethics committee of the Kunming University of Science and Technology School of Medicine.

## Participant Training

The aerobic exercise group performed warm-up exercises for 5 min, and then aerobic exercise (Perini et al., 2016) (refer to section Experimental Equipment and Data Collection for exercise requirements) using a horizontal magnetic bicycle (the resistance was adjusted to four levels of medium resistance) for 1 min to adapt to the machine, then rested for 3 min (walking and relaxing), after which point a computer voice prompted emotion regulation by aerobic exercise. The experiment began and the participants performed aerobic exercise for 20 min until the end of the voice prompt experiment. Participants in the

motor imagery group first performed the aerobic exercise with the horizontal magnetic bicycle (resistance adjusted to four levels of medium resistance) for 1 min to experience the actual aerobic exercise process and then were asked to rehearse or feel the aerobic exercise process from the first-person perspective—but no actual movement occurred (i.e., kinesthetic imagery) (Proske and Gandevia, 2018). Specifically, they repeated the actual moving process for 1 min and, at the same time, used motor imagery to evoke the pleasure or comfort brought by the movement, then rested for 3 min before the computer voice prompted the motor imagery experiment to begin and the participants performed the motor imagery experiment for a continued 20 min until the end of the voice prompt experiment. Before and after the experiment, the participants were required to fill in the STAI and POMS questionnaires.

Aerobic exercise involves moderate- or high-intensity exercise, so recumbent cycling can be divided into eight levels of intensity, with level 4 representing moderate-intensity exercise; thus, it is necessary for this study to ensure that exercise was performed above this level. During the experiment, the heart rate and exercise time of the participants were recorded by the recumbent cycle. To accurately grasp the  $HR_{max}$  of the study participants, before the beginning of the experiment, they were asked to exercise continuously for 5 min, performing recumbent cycling at level 8 to measure the  $HR_{max}$  (Wallert and Madison,



2014). The participants understood the whole process of the experiment and performed preexperiment training.

## Experimental Design and Process

The diagram for experiment timing is shown in **Figure 2**. **Figure 2A** is a time diagram of the experiment of aerobic exercise regulating emotion. During the T1 period, participants filled out the STAI and POMS questionnaires for 10 min; then, a computer voice prompted them to stay awake and relaxed for 3 min, which is the T2 period. At the end of this rest, the voice prompted the emotion regulation by aerobic exercise test to begin and the participant performed aerobic exercise on the horizontal magnetic bicycle for a continued 20 min, which is the T3 time period. Finally, after the end of the aerobic exercise period, the voice prompts the participant to rest for 3 min, then complete the STAI and POMS questionnaires again.

**Figure 2B** is the timing diagram of the motor imagery experiment. The participants filled in the STAI and POMS questionnaires during the T4 time period for 10 min; then, a computer voice prompted them to stay at rest for 3 min, which is the T5 time period, before beginning the baseline period, in which they were asked to stay awake and relax for 1 min, without performing any mental tasks. At the end of the baseline state, a voice and picture prompted the start of the motor imagery experiment, which lasts for 2 s, before the participant imagines doing aerobic exercise with a horizontal magnetic bicycle for 30 s. During this period, the computer screen was blank. After the imagery task is over, the participant was asked to rest for 10 s; this constitutes the end of a trial. A total of 30 trials, 20 min in length, composed the T6 time period. Then a voice and picture prompted the participant to rest, asking them to stay awake and relaxed for 3 min, and then fill out the STAI and POMS questionnaires again. The timing of the experiment was implemented by MATLAB Psychtoolbox-3 (R2018a; MathWorks, Natick, MA, USA).

## Experimental Equipment and Data Collection

The fNIRS device used in this experiment was a portable Nir Smart [two wavelengths: 760 and 850 nm, 16 channels (eight light sources and eight detectors); Danyang Huichuang Medical Equipment Co., Ltd., Danyang, China]. According to the 10–20 international standard lead system, the fNIRS helmet was placed on the head of the participant such that the light poles covered the left and right prefrontal areas of the brain, including eight channels of each of the left and right prefrontal lobes ( $3 \times 4$  array of emitter and detection light poles). The left and right areas were symmetrical and the left and right prefrontal medial channels were located at Fp1 and Fp2, respectively. The emission and detection light poles were arranged as shown in **Figure 3A**.

The fNIRS sampling rate was 20 Hz, the single wavelength power of the light source was  $>20$  mW, the time resolution was 100 Hz, the dynamic range was  $>110$  dB, and the digital-to-analog conversion accuracy was 24-bit. Data collection was completed according to the experimental timing and requirements in **Figure 2A**.

The aerobic exercise equipment used in the experiment was a horizontal magnetic bicycle, model JTH-735RS-1

(size:  $120 \times 50 \times 122$  CM; two 6-KG two-way rotating flywheels, resistance: magnetic control eight-speed resistance adjustment, foot distance: 56–72 CM, load-bearing: 120 kg; Guangzhou Jintong Fitness Equipment Co., Ltd., Guangzhou, China). The equipment had a digital dashboard, which can display parameters such as exercise time, speed, mileage, heart rate, and calories burned. The real experimental scene is shown in **Figure 3B**. There were 16 channels in total, of which channels one to eight were located in the right prefrontal cortex (PFC) and channels nine to 16 were located in the left PFC.

Before aerobic exercise, the average heart rate of 15 participants in the aerobic exercise group was 75 bpm. Heart rate is the most direct indicator of the effect and intensity of aerobic exercise (the appropriate heart rate for aerobic exercise is 120–135 bpm, as determined by sports medicine) (Ekkekakis et al., 2013). The aerobic exercise in the study was set at a moderate exercise intensity (60–80% of the maximum heart rate). During aerobic exercise, the average heart rate was 128 bpm, the average number of calories burned was 101 kcal, and the average exercise mileage was 3.39 kilometers.

## Data Processing

### NIRS Data Preprocessing and Feature Extraction

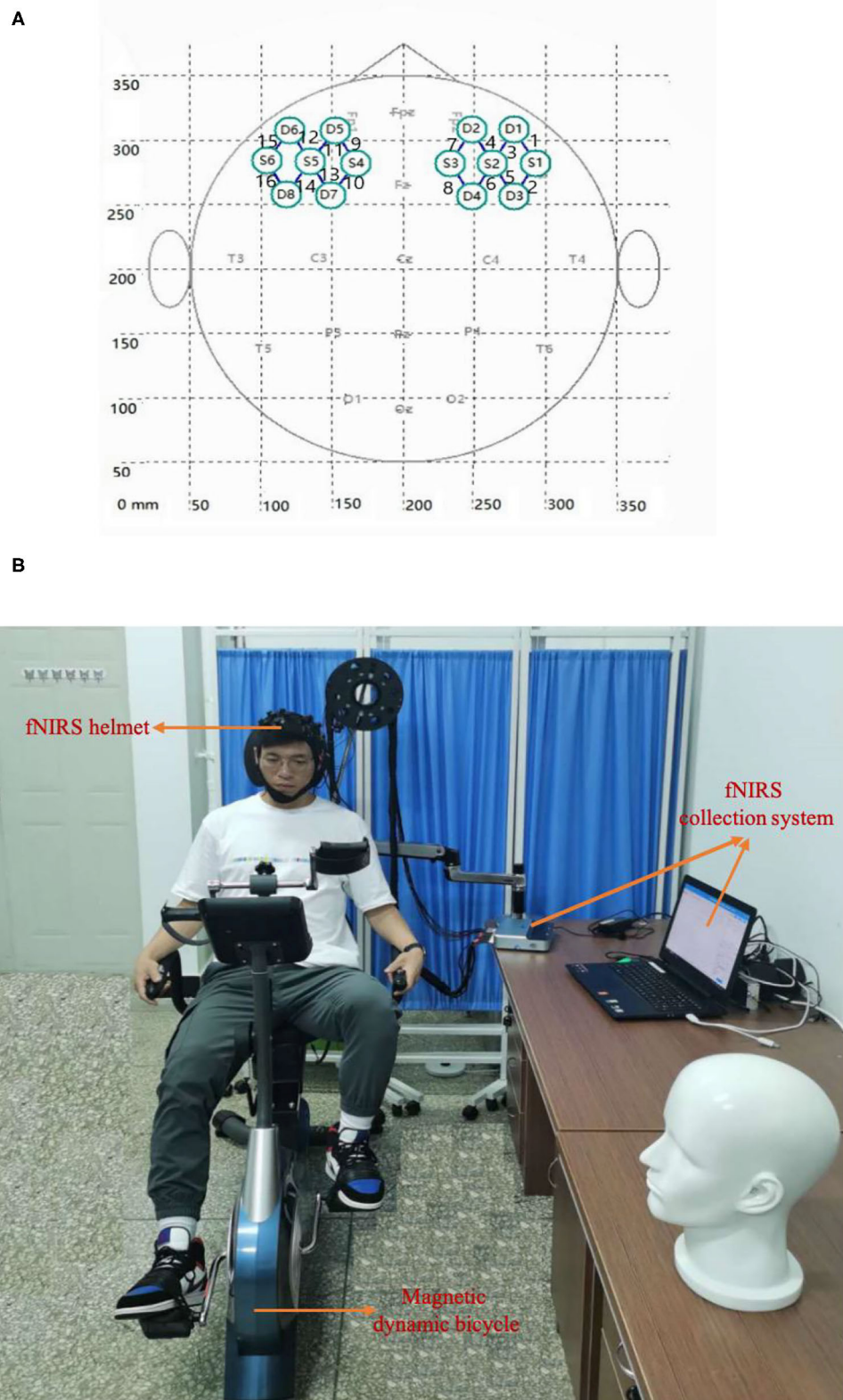
The fNIRS signal collected in the experiment was the original light-intensity signal, which needed to be converted using the improved Lambert-Beer law to discern HbO and HbR concentrations, which are denoted by the relative change values of  $\Delta\text{Oxy-Hb}$  and  $\Delta\text{Deoxy-Hb}$  (Cui et al., 2010).

After the fNIRS data were Butterworth band-pass filtered and corrected for baseline drift, the HbO, HbR, and HbT signals were extracted, respectively, and the means of the three signals across all participants and the average concentration changes before and after aerobic exercise, and motor imagery were calculated.

The study also calculated the ratio of HbO concentration in the left and right prefrontal lobes before and after aerobic exercise and motor imagery.

### Evaluation of STAI and POMS Scores

State-trait anxiety inventory was compiled by Charles Spielberger in 1977 (X version) and revised in 1983 (Y version) (Spielberger et al., 1970; Spielberger, 1983). This scale is characterized by simplicity, high validity, and easy analysis. It can intuitively reflect the participative feelings of anxious individuals, especially the current S-AI with T-AI differentiation. Each item of STAI has four (1–4) grades. The grading standards of S-AI are as follows: 1 = not at all, 2 = some, 3 = moderate, 4 = very obvious, while the grading standard of T-AI are: 1 = almost none, 2 = some, 3 = often, 4 = almost always. Positive emotion items (1, 2, 5, 8, 10, 11, 15, 16, 19, 20, 21, 23, 24, 26, 27, 30, 33, 34, 36, and 39; items are scored with a single superscript \*) are reverse scored—, that is, they are rated as four, three, two, and one point(s) in the above order and negative emotions are scored positively. The minimum score of the two scales is 20 points and the maximum is 80 points; the higher the score, the higher the degree of anxiety. The degree of anxiety is divided into four levels: no anxiety ( $\leq 20$  points), mild anxiety (21–39), moderate anxiety (40–59), and severe anxiety (60–80).



**FIGURE 3 |** Experimental setup. **(A)** The arrangement of the light source and detector probe. S and D denote the light source probe and the detector probe, respectively. The connecting line between the light source probe and the detector probe denotes the channel, and the number (1–16) denotes the channel identifier. **(B)** Real experiment scene with horizontal magnetic bicycle and fNIRS collection system.

**TABLE 1** | Changes in STAI and POMS scores before and after aerobic exercise and motor imagery\*.

	Pre-AE	Post-AE	F	p	Pre-MI	Post-MI	F	p
STAI (M ± SD)	43.27 ± 2.49	36 ± 1.90	75.47	$p < 0.01$	43.87 ± 5.67	42.87 ± 5.80	0.80	0.38
POMS (M ± SD)	123.67 ± 4.54	96.67 ± 3.52	309.48	$p < 0.01$	123.40 ± 4.63	121.73 ± 3.70	1.11	0.30

\*Two-way ANOVA test showed significant results.

AE, aerobic exercise; M ± SD, mean ± standard deviation; MI, motor imagery; POMS, Profile of Mood States; STAI, State-Trait Anxiety Inventory.

Profile of mood states is a scale for the evaluation of positive and negative emotion (Curran et al., 1995), which consists of 40 adjectives (corresponding to 40 items), and is rated from zero (not at all) to four (very) points according to the feelings of the participant (Grove and Prapavessis, 1992; Zhu, 1995). The 40 items of the scale correspond to the scores of seven subscales: tension ( $n = 6$  items), anger ( $n = 7$  items), fatigue ( $n = 5$  items), depression ( $n = 6$  items), energy ( $n = 6$  items), panic ( $n = 5$  items), and self-esteem ( $n = 5$  items). The total mood disturbance (TMD) score = (tension score + anger score + fatigue score + depression score + confusion score) – (energy score + emotional score related to self-esteem) + 100 (Andrykowski et al., 1990, 1993). Higher TMD scores indicate that the emotional state of the participants is negative.

An STAI score of 40–59 points indicates that an individual has moderate anxiety, while a POMS score of 110–140 points indicates that an individual is in a negative mood. These two scales were limited to these score intervals to screen the participants who are in line with the experiment.

### Calculation of Spearman's Correlation Coefficient Between Changes in the Average STAI and POMS Scores Before and After Emotion Regulation and Changes in Blood Oxygen Concentration

To gain a more accurate grasp of the emotion regulation of the study participants before and after aerobic exercise and motor imagination, the change of HbO signal was selected as the judgment standard, and the three stages before, during, and after aerobic exercise and motor imagination were selected to draw a real-time dynamic diagram of HbO.

In this study, Spearman's correlation coefficient (Fieller and Pearson, 1961) was used to measure the dependence of the two variables. The correlation coefficient was defined as Pearson's correlation coefficient. The correlation coefficient is used to calculate the correlation between changes in the average STAI and POMS scores and changes in the HbO, HbR, and HbT concentrations before and after aerobic exercise and motor imagery emotion regulation. In this study, MATLAB\_2018a (MathWorks) was used to calculate Spearman's correlation coefficient.

## RESULTS

### STAI and POMS Evaluation and fNIRS Data Used Two-Way ANOVA Results

Statistics in Tables 1, 2 revealed that two-way ANOVA was used to test the group factors (aerobic exercise group and motor imagery group) and intervention factors (pretest and posttest) of the participants, respectively. In the pretest and posttest factor

**TABLE 2** | Mean and standard deviation values of HbO concentration changes in the left and right prefrontal areas before and after aerobic exercise and motor imagery\*.

	Pre (M ± SD)	Post (M ± SD)	Degrees of freedom	F	P
<b>Aerobic exercise group</b>					
Left prefrontal cortex	−0.032 ± 0.308	0.100 ± 1.025	1.8	25.67	$p < 0.01$
Right prefrontal cortex	−0.048 ± 0.373	−0.142 ± 1.089		5.16	0.04
HbO activation concentration ratio	1.404				
<b>Motor imagery group</b>					
Left prefrontal cortex	−0.039 ± 0.365	0.016 ± 0.391	1.8	4.5	0.06
Right prefrontal cortex	−0.034 ± 0.379	−0.011 ± 0.405		3.7	0.07
HbO activation concentration ratio	2.391				

\*Two-way ANOVA test showed significant results.

M ± SD, mean ± standard deviation.

analysis ( $p < 0.01$ ), posttest anxiety was significantly lower than pretest anxiety. Among the groups ( $p < 0.01$ ), the anxiety degree of the aerobic group was significantly lower than that of the motor imagery group. The laterality ratio was the left and right PFC fNIRS concentration ratio, similar to the laterality score gleaned when using EEG to assess PFC asymmetry (Palmiero and Piccardi, 2017).

Table 3 presents the mean and standard deviation values of the HbO, HbR, and HbT concentration changes in the prefrontal area between before and after aerobic exercise and motor imagery. The average changes in the concentrations of HbO, HbR, and HbT in the prefrontal lobe area increased after aerobic exercise; meanwhile, considering the prefrontal lobe area after motor imagery, the average changes in HbO, HbR, and HbT concentrations also increased to a certain extent, but the range was small.

### Results of Spearman's Correlation Coefficient Analysis

Table 4 presents the Spearman correlation coefficients calculated using the average changes of the STAI and POMS scores between before and after aerobic exercise and the changes in HbO, HbR,



and HbT concentrations. The results showed that the average STAI and POMS scores before and after aerobic exercise were negatively correlated with the concentrations of HbO, HbR, and HbT.

## Change in HbO Concentration

For the convenience of discussion, the start time of aerobic exercise was specified as  $t = 0$  min ( $t = -1$  means 1 min before aerobic exercise,  $t = 10$  means 10 min after the start of aerobic exercise,  $t = 20$  means the end time of aerobic exercise, and  $t = 21$  means 1 min after the end of aerobic exercise). We selected four time periods ( $-1$  to  $0$  min,  $0$ – $10$  min,  $10$ – $20$  min, and  $20$ – $21$  min) and calculated the HbO concentration change in the corresponding period, as shown in **Figures 4A–D**.

A minute before the start of aerobic exercise, the change in HbO concentration tends to be flat, without a prominent peak signal, and the participant may still be in a state of depression and anxiety. Ten minutes after the start of aerobic exercise, the HbO concentration has a peak signal, with a largely positive change, and the depression and anxiety mood of the participant may transform into a positive mood. Twenty minutes after the start of aerobic exercise, the positive change in HbO concentration has increased compared relative to during the previous 10 min, and the depression and anxiety mood of the participant may continue to transform into a positive mood. Finally, 1 min after the end of aerobic exercise, the HbO concentration evolves to its peak, and the effect of aerobic exercise on improving negative emotion continues.

**TABLE 3 |** The mean and standard deviation of HbO, HbR, and HbT concentration changes in the prefrontal area before and after aerobic exercise and motor imagery.

	Aerobic exercise group			Motor imagery group		
	HbO	HbR	HbT	HbO	HbR	HbT
Before aerobic exercise or motor imagery, $M \pm SD$	$-0.039 \pm 1.141$	$-0.03 \pm 0.562$	$-0.01 \pm 1.643$	$-0.036 \pm 0.414$	$-0.002 \pm 0.134$	$-0.013 \pm 0.26$
After aerobic exercise or motor imagery, $M \pm SD$	$0.121 \pm 1.058$	$0.047 \pm 0.514$	$0.168 \pm 1.502$	$0.013 \pm 0.398$	$0.002 \pm 0.136$	$-0.012 \pm 0.266$

$M \pm SD$ , mean  $\pm$  standard deviation.

For the convenience of discussion, the start time of motion imagery was specified as  $t = 0$  min ( $t = -1$  means 1 min before motion imagery,  $t = 10$  means 10 min after the start of aerobic exercise,  $t = 20$  means the end time of aerobic exercise, and  $t = 21$  means 1 min after the end of aerobic exercise). We selected four time periods ( $-1$  to  $0$  min,  $0$ – $10$  min,  $10$ – $20$  min, and  $20$ – $21$  min) and calculated the HbO concentration change in the corresponding period, as shown in **Figure 5**.

A minute before the start of the motor imagery, the HbO concentration change tends to be flat, without a prominent peak signal, and the mood of the participant may still be in a state of depression and anxiety. Ten minutes after the start of the motor imagery, the HbO concentration changes little, yet the overall trend is still flat, although depression and anxiety may show a small conversion to positive emotion. Twenty minutes after the start of motor imagery, the HbO concentration changes little relative to in the first 10 min and still tends to be flat. The depression and anxiety of the participants may still experience a small conversion to positive emotion. Finally, 1 min after the end of motor imagery, the HbO concentration changes with a small peak, and the effect of motor imagery on improving negative emotion continues, but the improvement may be small.

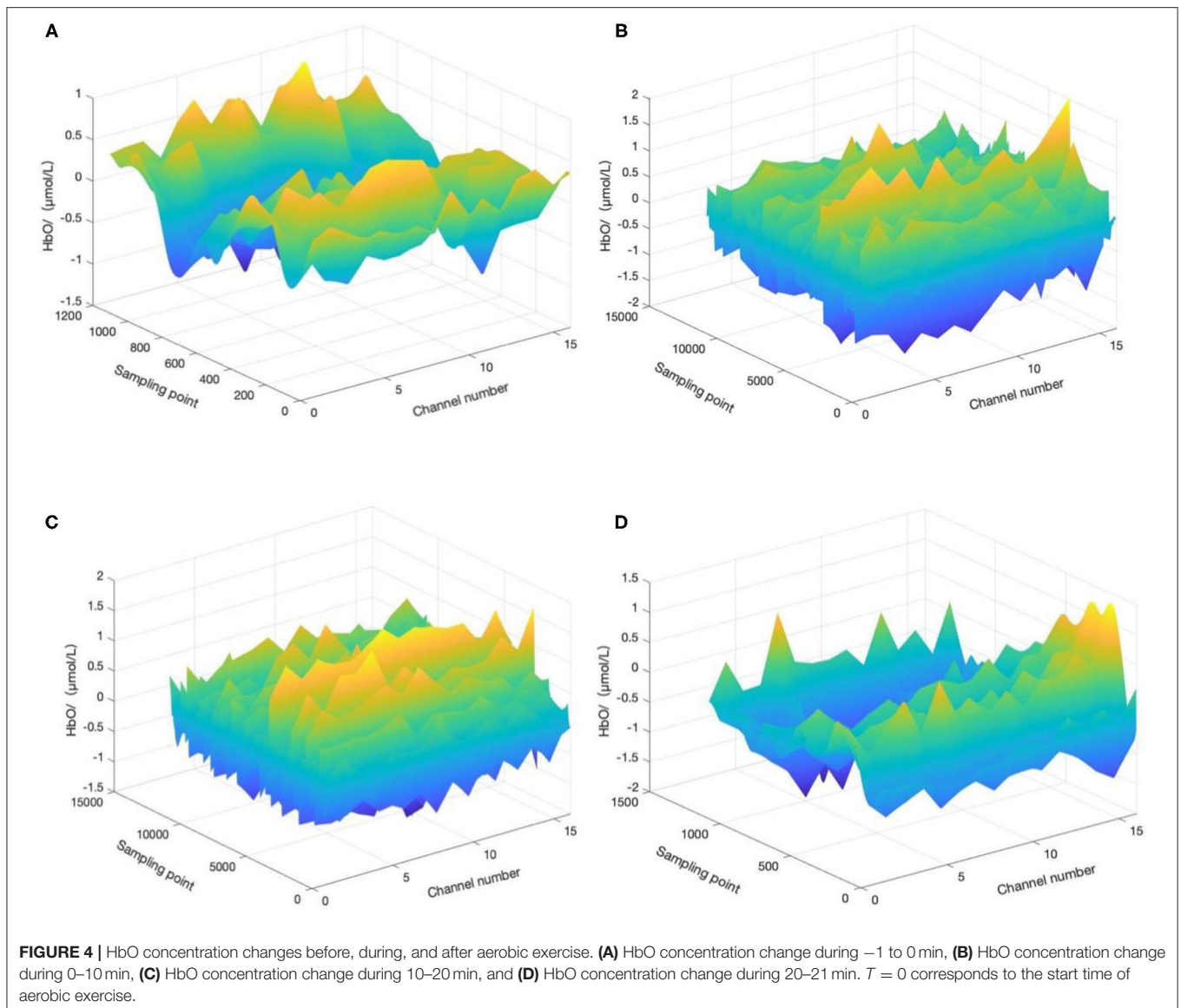
## DISCUSSION

To further verify the above results, ANOVA was used to analyze the changes in STAI and POMS scores before and after aerobic exercise and motor imagery. The results showed that the scores of the two scales changed significantly between before and after emotion regulation by aerobic exercise, and depression and anxiety emotions transformed into positive emotions. Before and after motor imagery emotion regulation, the scores of the two scales changed, but not significantly, and the transformation of depression and anxiety into positive emotion was not significant.

To reveal the relationship between changes in the blood oxygen concentration (Hb) and changes in the STAI and POMS scores, Spearman's correlation coefficient analysis revealed that the changes in the average STAI and POMS scores before and after aerobic exercise and motor imagery emotion regulated the changes in HbO, HbR, and HbT concentrations in a negatively correlated manner, which indicates that the decrease in average STAI and POMS scores before and after aerobic exercise and motor imagery emotion regulation decrease (due to the conversion of negative emotion to positive emotion) corresponds to an increase in HbO, HbR, and HbT concentrations.

**TABLE 4 |** Spearman correlation coefficients of changes in the average scores of STAI and POMS scales before and after aerobic exercise and changes in HbO, HbR, and HbT concentrations.

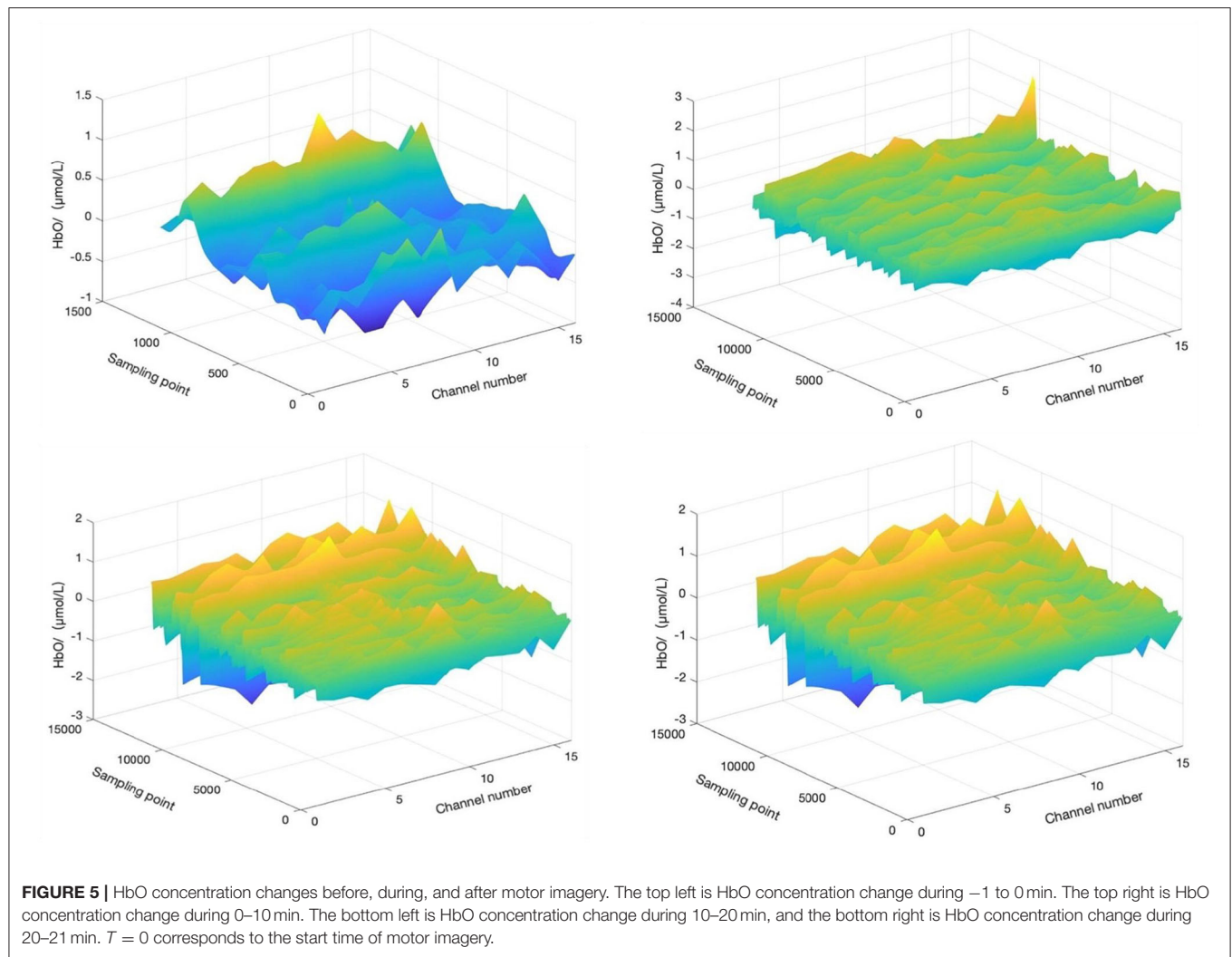
	STAI and POMS scores		Aerobic exercise group		
	STAI	POMS	HbO	HbR	HbT
Before aerobic exercise	$43.27 \pm 2.49$	$123.67 \pm 4.54$	$-0.039 \pm 1.141$	$-0.03 \pm 0.562$	$-0.01 \pm 1.643$
After aerobic exercise	$36 \pm 1.90$	$96.67 \pm 3.52$	$0.121 \pm 1.058$	$0.047 \pm 0.514$	$0.168 \pm 1.502$
Spearman's correlation coefficient calculation			$-1$	$-1$	$-1$



**Table 5** shows the comparison between this study and other related research. Bernstein et al. (Bernstein and McNally, 2018) concluded that, as compared with a stretching exercise group, a bicycle exercise group attained a more significant effect on emotion regulation. Meanwhile, Subramaniapillai et al. (Subramaniapillai et al., 2016) reported that adolescents with bipolar disorder would also feel the positive emotional benefits brought about by exercise. Szabo et al. (2015) concluded that, after aerobic exercise, positive emotions increased, and negative emotions decreased. Knapen et al. (2009) showed that the state of anxiety and negative emotions are not affected by aerobic exercise type. When study participants chose their exercise intensity, they influenced the happiness and fatigue changes brought about by exercise in positive and negative ways. The results of Zhang et al. (2018) showed that short-term aerobic exercise significantly improved the executive function and emotional regulation ability of female college students with anxiety, where the lower the

level of aerobic adaptability, the better the improvement effect of short-term aerobic exercise on unconscious cognition.

Compared with the above-mentioned studies, the present study uses relatively new fNIRS brain function imaging technology to quantitatively monitor and evaluate changes in brain tissue blood oxygen concentrations (HbO, HbR, and HbT) during emotion regulation by aerobic exercise and motor imagery. Here, participants' emotions were regulated by horizontal magnetic bicycle and motor imagination, and STAI and POMS scores were used to evaluate the emotional changes in participants before and after aerobic exercise and motor imagery. The results showed that, during the period of aerobic exercise emotion regulation, HbO can represent the metabolic activity of oxyhemoglobin in brain tissue and indirectly reflect the activity of the neuron group according to the neurovascular coupling relationship. Previous studies have found that when individuals evolve from negative emotion to



positive emotion, the concentration of HbO in the left PFC increases, while that in the right PFC decreases (Ochsner et al., 2004; Kim and Hamann, 2007). Two-way ANOVA was used to test the group factors (aerobic exercise group and motor imagery group) and intervention factors (pretest and posttest) of the participants, respectively. Considering the pretest and posttest factors ( $p < 0.01$ ), posttest anxiety was significantly lower than pretest anxiety. Considering group factors ( $p < 0.01$ ), the anxiety degree of the aerobic group was significantly lower than that of the motor imagery group. The increase in the left PFC HbO concentration and functional activation and the decrease in the right PFC HbO concentration and functional deactivation may be related to the reduction in STAI and POMS scores. Furthermore, Spearman's correlation coefficient of the STAI and POMS scores and HbO concentration were calculated, and the result was  $-1$ , indicating that they were completely negatively correlated.

The above results show that fNIRS can be effectively applied to the monitoring and evaluation of emotion regulation by aerobic exercise. In comparison with aerobic exercise, motor imagery has

no significant effect on state-trait anxiety and mood state, but it may play an auxiliary role in regulating emotion.

The disadvantage of this study is that single-mode fNIRS brain imaging technology records single brain activity, and its spatial and temporal resolutions are limited. Emotion regulation is a medium- and long-term change process, but this study involved just 20 min of aerobic exercise. Future research should incorporate long-term tracking of aerobic exercise to regulate emotion.

## CONCLUSION

This study showed that the portable fNIRS could be effectively used for monitoring and evaluating emotion regulation by aerobic exercise on a horizontal magnetic bicycle. The effects of aerobic exercise on emotion regulation were more significant than those of motor imagery, and the effect of motor imagery on emotion regulation was limited, although carrying out motor imagery on the basis of aerobic exercise may be beneficial to enhance the effect of emotion regulation. This study was expected

**TABLE 5 |** Comparison of this and other related studies.

Study	Exercise manner	Scale	Features extracted	Analytical method(s)
Bernstein and McNally (2018)	Bicycle exercise Stretching exercises	SES DAS	Heart rate SES score DAS score	Independent-samples <i>t</i> -test, linear regression, maximum likelihood estimation
Subramaniapillai et al. (2016)	Bicycle Dynamometer	EFI RPE	EFI score RPE score	Independent-samples <i>t</i> -test
Szabo et al. (2015)	Exercise bicycle	PANAS	Heart rate PANAS mean and standard deviation values	Wilcoxon signed-rank test
Knapen et al. (2009)	Electronic brake bicycle force gauge	STAI SEES	Heart rate STAI score SEES score	SAS program, mixed repeated measurement variance analysis
Zhang et al. (2018)	International affective picture system, power bicycle	PANAS SAS	PANAS score SAS score	Independent-samples <i>t</i> -test, repeated measures ANOVA
This study	Horizontal magnetic bicycle, motor imagery	STAI POMS	Heart rate STAI score POMS score Before and after aerobic exercise and motor imagery; HbO, HbR, and HbT mean concentrations	ANOVA, Spearman's correlation coefficient, fNIRS analysis

to provide ideas for constructing fNIRS-based online real-time monitoring and the evaluation of emotion regulation by aerobic exercise and motor imagery, which could be used to monitor and evaluate individual state–trait anxiety and mood states.

## 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

The studies involving human participants were reviewed and approved by Medical Ethics Committee of Kunming

University of Science and Technology School of Medicine. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

## FUNDING

This work was supported by Fund projects: NSFC (81771926, 61763022, 81470084, 61463024, 62006246, and 32060196).

## REFERENCES

- Andrykowski, M. A., Altmaier, E. M., Barnett, R. L., Otis, M. L., Gingrich, R., and Henslee-Downey, P. J. (1990). Quality of life in adult survivors of allogeneic bone marrow transplantation: correlates and comparison with matched renal transplant recipients. *Transplantation* 50, 399–406. doi: 10.1097/00007890-199009000-00009
- Andrykowski, M. A., Brady, M. J., and Hunt, J. W. (1993). Positive psychosocial adjustment in potential bone marrow transplant recipients: cancer as a psychosocial transition. *Psycho-Oncology* 2, 261–276. doi: 10.1002/pon.2960020406
- Bernstein, E. E., and McNally, R. J. (2015). Acute aerobic exercise helps overcome emotion regulation deficits. *Cogn. Emot.* 31, 834–843. doi: 10.1080/02699931.2016.1168284
- Bernstein, E. E., and McNally, R. J. (2017). Acute aerobic exercise hastens emotional recovery from a subsequent stressor. *Health Psychol.* 36, 560–567. doi: 10.1037/hea0000482
- Bernstein, E. E., and McNally, R. J. (2018). Exercise as a buffer against difficulties with emotion regulation: a pathway to emotional wellbeing. *Behav. Res. Ther.* 109, 29–36. doi: 10.1016/j.brat.2018.07.010
- Blackhart, G. C., Minnix, J. A., and Kline, J. P. (2006). Can EEG asymmetry patterns predict future development of anxiety and depression? A preliminary study. *Biol. Psychol.* 72, 46–50. doi: 10.1016/j.biopsycho.2005.06.010
- Brush, C. J., Foti, D., Bocchine, A. J., Muniz, K. M., and Alderman, B. L. (2020). Aerobic exercise enhances positive emotional reactivity in individuals with depressive symptoms: evidence from neural responses to reward and emotional content. *Ment. Health Phys. Act.* 19:100339. doi: 10.1016/j.mhpa.2020.100339
- Chen, T., Yue, G. H., Tian, Y. X., and Jiang, C. H. (2017). Baduanjin mind-body intervention improves the executive control function. *Front. Psychol.* 7:2015. doi: 10.3389/fpsyg.2016.02015
- Chen, Y. C., Chen, C., Martínez, R. M., Etnier, J. L., and Cheng, Y. (2019). Habitual physical activity mediates the acute exercise-induced modulation of anxiety-related amygdala functional connectivity. *Sci. Rep.* 9:19787. doi: 10.1038/s41598-019-56226-z



- Cheng, L. (2007). Retrospect and prospect on some theoretic issues in aerobic exercise. *Bull. Sport Sci. Technol.* 3, 35–38. doi: 10.3969/j.issn.1005-0256.2007.03.017
- Clemente-Suárez, V. J. (2020). Multidisciplinary intervention in the treatment of mixed anxiety and depression disorder. *Physiol. Behav.* 219:112858. doi: 10.1016/j.physbeh.2020.112858
- Cui, X., Bray, S., Bryant, D. M., Glover, G. H., and Reiss, A. L. (2011). A quantitative comparison of nirs and fmri across multiple cognitive tasks. *Neuroimage* 54, 2808–2821. doi: 10.1016/j.neuroimage.2010.10.069
- Cui, X., Bray, S., and Reiss, A. L. (2010). Functional near infrared spectroscopy (nirs) signal improvement based on negative correlation between oxygenated and deoxygenated hemoglobin dynamics. *Neuroimage* 49, 3039–3046. doi: 10.1016/j.neuroimage.2009.11.050
- Curran, S. L., Andrykowski, M. A., and Studts, J. L. (1995). Short form of the profile of mood states (poms-sf): psychometric information. *Psychol. Assess.* 7, 80–83. doi: 10.1037/1040-3590.7.1.80
- Edwards, M. K., Rhodes, R. E., and Loprinzi, P. D. (2017). A randomized control intervention investigating the effects of acute exercise on emotional regulation. *Am. J. Health Behav.* 41, 534–543. doi: 10.5993/AJHB.41.5.2
- Edwards, M. K., Rhodes, R. E., Mann, J. R., and Loprinzi, P. D. (2018). Effects of acute aerobic exercise or meditation on emotional regulation. *Physiol. Behav.* 186, 16–24. doi: 10.1016/j.physbeh.2017.12.037
- Ekkekakis, P., Hargreaves, E. A., and Parfitt, G. (2013). Invited guest editorial: envisioning the next fifty years of research on the exercise–effect relationship. *Psychol. Sport Exerc.* 14, 751–758. doi: 10.1016/j.psychsport.2013.04.007
- Fardo, F., Allen, M., Jegindø, E. M., Angrilli, A., and Roepstorff, A. (2015). Neurocognitive evidence for mental imagery-driven hypoalgesic and hyperalgesic pain regulation - ScienceDirect. *Neuroimage* 120, 350–361. doi: 10.1016/j.neuroimage.2015.07.008
- Fieller, E. C., and Pearson, E. S. (1961). Tests for rank correlation coefficients ii. *Biometrika* 48, 29–40. doi: 10.1093/biomet/48.1-2.29
- Goldin, P., Ziv, M., Jazaieri, H., Hahn, K., and Gross, J. J. (2013). Mbsr vs aerobic exercise in social anxiety: fmri of emotion regulation of negative self-beliefs. *Soc. Cogn. Affect. Neurosci.* 8, 65–72. doi: 10.1093/scan/nss054
- Grove, J. R., and Prapavessis, H. (1992). Preliminary evidence for the reliability and validity of an abbreviated profile of mood states. *Int. J. Sport Psychol.* 23, 93–109. doi: 10.1007/BF00636229
- Jiang, D., Zhang, D., Chen, Y., He, Z., Gao, Q., Gu, R., et al. (2017). Trait anxiety and probabilistic learning: behavioral and electrophysiological findings. *Biol. Psychol.* 132, 17–26. doi: 10.1016/j.biopsycho.2017.10.010
- Kim, S. H., and Hamann, S. (2007). Neural correlates of positive and negative emotion regulation. *J. Cogn. Neurosci.* 19, 776–798. doi: 10.1162/jocn.2007.19.5.776
- Knapen, J., Sommerijns, E., Vancampfort, D., Sienaert, P., Pieters, G., and Haake, P., et al. (2009). State anxiety and subjective well-being responses to acute bouts of aerobic exercise in patients with depressive and anxiety disorders. *Br. J. Sports Med.* 43, 756–759. doi: 10.1136/bjsm.2008.052654
- Naseer, N., and Hong, K. S. (2013). Classification of functional near-infrared spectroscopy signals corresponding to the right- and left-wrist motor imagery for development of a brain-computer interface. *Neuroence Lett.* 553, 84–89. doi: 10.1016/j.neulet.2013.08.021
- Ochsner, K. N., Ray, R. D., Robertson, E. R., Cooper, J. C., Chopra, S., Gabrieli, J. D., et al. (2004). For better or for worse: neural systems supporting the cognitive down- and up regulation of negative emotion. *Neuroimage* 23, 483–499. doi: 10.1016/j.neuroimage.2004.06.030
- Palmiero, M., and Piccardi, L. (2017). Frontal EEG asymmetry of mood: a mini-review. *Front. Behav. Neurosci.* 11:224. doi: 10.3389/fnbeh.2017.00224
- Perini, R., Bortoletto, M., Capogrosso, M., Fertonani, A., and Miniussi, C. (2016). Acute effects of aerobic exercise promote learning. *Sci. Rep.* 6:25440. doi: 10.1038/srep25440
- Proske, U., and Gandevia, S. C. (2018). Kinesthetic senses. *Comprehen. Physiol.* 3, 1157–1183. doi: 10.1002/cphy.c170036
- Sitaram, R., Zhang, H., Guan, C., Thulasidas, M., Hoshi, Y., Ishikawa, A., et al. (2007). Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface. *Neuroimage* 34, 1416–1427. doi: 10.1016/j.neuroimage.2006.11.005
- Smit, D. J. A., Posthuma, D., Boomsma, D. I., and De Geus, E. J. C. (2007). The relation between frontal EEG asymmetry and the risk for anxiety and depression. *Biol. Psychol.* 74, 26–33. doi: 10.1016/j.biopsycho.2006.06.002
- Spielberger, C. D. (1983). *Manual for the State-Trait Anxiety Inventory (Form)*. Palo Alto, CA: Consulting Psychologists Press. doi: 10.1037/t06496-000
- Spielberger, C. D., Gorsuch, R., and Lushene, R. E. (1970). *Manual for the State-Trait Anxiety Inventory (Self-Evaluation Questionnaire)*. Palo Alto, CA: Consulting Psychologists Press.
- Subramaniapillai, M., Goldstein, B. I., MacIntosh, B. J., Korczak, D. J., Ou, X., Scavone, A., et al. (2016). Characterizing exercise-induced feelings after one bout of exercise among adolescents with and without bipolar disorder. *J. Affect. Disord.* 190, 467–473. doi: 10.1016/j.jad.2015.10.018
- Szabo, A., Gáspár, Z., Kiss, N., and Radványi, A. (2015). Effect of spinning workouts on affect. *J. Mental Health* 24, 145–149. doi: 10.3109/09638237.2015.1019053
- Tal, S. (2016). Using movement to regulate emotion: neurophysiological findings and their application in psychotherapy. *Front. Psychol.* 7:1451. doi: 10.3389/fpsyg.2016.01451
- Tempest, G., and Parfitt, G. (2013). Imagery use and affective responses during exercise: an examination of cerebral hemodynamics using near-infrared spectroscopy. *J. Sport Exerc. Psychol.* 35, 503–513. doi: 10.1123/jsep.35.5.503
- Tempest, G. D., Eston, R. G., Gaynor, P., and Amanda, B. (2014). Prefrontal cortex hemodynamics and affective responses during exercise: a multi-channel near infrared spectroscopy study. *PLoS ONE* 9:e95924. doi: 10.1371/journal.pone.0095924
- Veerapa, E., Grandgenevre, P., El Fayoumi, M., Vinnac, B., Haelewyn, O., Szafrarczyk, S., et al. (2020). Attentional bias towards negative stimuli in healthy individuals and the effects of trait anxiety. *Sci. Rep.* 10, 2045–2322. doi: 10.1038/s41598-020-68490-5
- Wallert, J., and Madison, G. (2014). Recovery after aerobic exercise is manipulated by tempo change in a rhythmic sound pattern, as indicated by autonomic reaction on heart functioning. *Front. Human Neurosci.* 8:738. doi: 10.3389/fnhum.2014.00738
- Weiskopf, N., Mathiak, K., Bock, S. W., Scharnowski, F., Veit, R., Grodd, W., et al. (2004). Principles of a brain-computer interface (bci) based on real-time functional magnetic resonance imaging (fmri). *Biomed. Eng. IEEE Transact.* 51, 966–970. doi: 10.1109/TBME.2004.827063
- Zhang, Y., Gong, Y., Tang, D., and Guan, Y. (2018). The effect of short-term aerobic exercise on the emotional regulation ability of anxious female college students: mechanism and influencing factors. *J. Tianjin Instit. Phys. Educ.* 33, 31–37.
- Zhu, B. L. (1995). Brief introduction of POMS scale and its model for China. *J. Tianjin Instit. Phys. Educ.* 1, 35–37. doi: 10.13297/j.cnki.issn1005-0000.1995.01.007

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships 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.

Copyright © 2021 Ding, Wang, Li, Zhang, Li, Chen, Zhao, Gong and Fu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.

## APPENDIX

Changes of HBO concentration in each channel before and after aerobic exercise and motor imagery.

Channel number	1	2	3	4	5	6	7	8
Pre-AE	0.003 ± 0.39	-0.016 ± 0.25	-0.014 ± 0.34	-0.043 ± 0.32	-0.052 ± 0.32	-0.096 ± 0.45	-0.042 ± 0.34	-0.119 ± 0.50
Post-AE	0.067 ± 0.37	0.171 ± 0.33	0.129 ± 0.43	0.210 ± 0.41	0.337 ± 0.39	0.055 ± 0.32	0.115 ± 0.36	0.050 ± 0.42
Pre-MI	-0.036 ± 0.32	-0.034 ± 0.36	-0.044 ± 0.33	-0.039 ± 0.35	-0.033 ± 0.39	-0.024 ± 0.35	-0.022 ± 0.38	-0.040 ± 0.41
Post-MI	0.018 ± 0.29	0.005 ± 0.19	-0.007 ± 0.28	0.017 ± 0.31	0.013 ± 0.35	0.046 ± 0.27	-0.004 ± 0.24	-0.003 ± 0.21
Channel number	9	10	11	12	13	14	15	16
Pre-AE	-0.031 ± 0.42	-0.087 ± 0.28	-0.062 ± 0.32	-0.023 ± 0.37	-0.058 ± 0.28	-0.023 ± 0.23	0.007 ± 0.31	0.025 ± 0.17
Post-AE	0.156 ± 0.45	0.127 ± 0.38	0.324 ± 0.33	0.071 ± 0.51	0.171 ± 0.43	0.021 ± 0.33	0.182 ± 0.35	0.210 ± 0.31
Pre-MI	-0.035 ± 0.42	-0.008 ± 0.33	-0.039 ± 0.36	-0.063 ± 0.32	-0.027 ± 0.36	-0.036 ± 0.32	-0.066 ± 0.35	-0.036 ± 0.43
Post-MI	0.010 ± 0.36	-0.012 ± 0.34	0.020 ± 0.28	0.040 ± 0.23	-0.005 ± 0.37	0.011 ± 0.32	0.051 ± 0.29	0.013 ± 0.22



# Efficacy Evaluation of Neurofeedback-Based Anxiety Relief

Chao Chen<sup>1,2†</sup>, Xiaolin Xiao<sup>1†</sup>, Abdelkader Nasreddine Belkacem<sup>3</sup>, Lin Lu<sup>4</sup>, Xin Wang<sup>2</sup>, Weibo Yi<sup>5</sup>, Penghai Li<sup>2</sup>, Changming Wang<sup>6,7,8\*</sup>, Sha Sha<sup>8</sup>, Xixi Zhao<sup>8</sup> and Dong Ming<sup>1\*</sup>

<sup>1</sup> Academy of Medical Engineering and Translational Medicine, Tianjin University, Tianjin, China, <sup>2</sup> Key Laboratory of Complex System Control Theory and Application, Tianjin University of Technology, Tianjin, China, <sup>3</sup> Department of Computer and Network Engineering, College of Information Technology, United Arab Emirates University, Al Ain, United Arab Emirates, <sup>4</sup> Zhonghuan Information College, Tianjin University of Technology, Tianjin, China, <sup>5</sup> Beijing Machine and Equipment Institute, Beijing, China, <sup>6</sup> Department of Neurosurgery, Xuanwu Hospital, Capital Medical University, Beijing, China, <sup>7</sup> Brain-Inspired Intelligence and Clinical Translational Research Center, Xuanwu Hospital, Capital Medical University, Beijing, China, <sup>8</sup> Beijing Key Laboratory of Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China

## OPEN ACCESS

### Edited by:

Jane Zhen Liang,  
Shenzhen University, China

### Reviewed by:

Banghua Yang,  
Shanghai University, China  
Yan Wang,  
Institute of Psychology, Chinese  
Academy of Sciences (CAS), China

### \*Correspondence:

Changming Wang  
superwcm@163.com  
Dong Ming  
richardming@tju.edu.cn

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

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

**Received:** 13 August 2021

**Accepted:** 22 September 2021

**Published:** 28 October 2021

### Citation:

Chen C, Xiao X, Belkacem AN,  
Lu L, Wang X, Yi W, Li P, Wang C,  
Sha S, Zhao X and Ming D (2021)  
Efficacy Evaluation  
of Neurofeedback-Based Anxiety  
Relief. *Front. Neurosci.* 15:758068.  
doi: 10.3389/fnins.2021.758068

Anxiety disorder is a mental illness that involves extreme fear or worry, which can alter the balance of chemicals in the brain. This change and evaluation of anxiety state are accompanied by a comprehensive treatment procedure. It is well-known that the treatment of anxiety is chiefly based on psychotherapy and drug therapy, and there is no objective standard evaluation. In this paper, the proposed method focuses on examining neural changes to explore the effect of mindfulness regulation in accordance with neurofeedback in patients with anxiety. We designed a closed neurofeedback experiment that includes three stages to adjust the psychological state of the subjects. A total of 34 subjects, 17 with anxiety disorder and 17 healthy, participated in this experiment. Through the three stages of the experiment, electroencephalography (EEG) resting state signal and mindfulness-based EEG signal were recorded. Power spectral density was selected as the evaluation index through the regulation of neurofeedback mindfulness, and repeated analysis of variance (ANOVA) method was used for statistical analysis. The findings of this study reveal that the proposed method has a positive effect on both types of subjects. After mindfulness adjustment, the power map exhibited an upward trend. The increase in the average power of gamma wave indicates the relief of anxiety. The enhancement of the wave power represents an improvement in the subjects' mindfulness ability. At the same time, the results of ANOVA showed that  $P < 0.05$ , i.e., the difference was significant. From the aspect of neurophysiological signals, we objectively evaluated the ability of our experiment to relieve anxiety. The neurofeedback mindfulness regulation can effect on the brain activity pattern of anxiety disorder patients.

**Keywords:** neurofeedback, anxiety disorder, EEG signal, anxiety assessment, efficacy evaluation

## INTRODUCTION

Anxiety is an emotional response to a potential future threat or danger that, depending on intensity and duration, can cause symptoms of negative emotional, physical, behavioral, and cognitive components. While "normal" anxiety is adaptive to make the body alert and prepare it for potential threats, it is considered pathological when it becomes maladaptive, permanent, and out of control.

Furthermore, it is associated with serious social and occupational harm, other comorbidities, and an increased risk of suicide (Nepon et al., 2010). The classification of anxiety disorders has a long history (Crocq, 2015). According to the International Statistical Classification of Diseases and Related Health Problems (ICD-10), anxiety disorders are classified into generalized anxiety disorder, phobias, social anxiety disorder, post-traumatic stress disorder (PTSD), panic disorder with/without agoraphobia, and obsessive-compulsive disorder (OCD) (Kogan et al., 2016; Reed et al., 2019). Vulnerability to the development of anxiety disorders (Otowa et al., 2016; Gottschalk and Domschke, 2018) usually begins in childhood or adolescence (Kalin, 2017) and becomes a chronic condition that persists into adulthood (Bandelow and Michaelis, 2015; Craske et al., 2017). In the western world, the lifetime prevalence of these diseases in the general population is about 20–30%, making it the most common neuropsychiatric disorder, with women more susceptible than men (Revicki et al., 2012; Remes et al., 2016; GBD 2017 DALYs and Hale Collaborators, 2018). In summary, anxiety disorders impose a staggering burden on public health and global economy, highlighting the dire need to develop a more comprehensive understanding of the underlying mechanisms (Mokdad et al., 2018).

Current treatment options are mainly on psychotherapy and medication, which has proven effective in anxiety disorders (Carpenter et al., 2018). Psychological therapy is time-consuming and requires extensive training of therapists. Non-compliance, non-response, or incomplete response, and relapse are still major issues in patients receiving treatment (Taylor et al., 2012; Roy-Byrne, 2015). Currently available drug treatments for anxiety include selective serotonin reuptake inhibitors and serotonin and norepinephrine reuptake inhibitors, and benzodiazepines are most suitable for short-term and adjuvant antianxiety therapy. Traditional Chinese medicine injections and oral contraceptives are effective, but tolerance-related problems restrict their usage. It is encouraging that new mechanical compounds targeting glutamate, neuropeptides, and the endocannabinoid system are also being developed; however, there is insufficient information regarding the role of the glutamate system in the pathogenesis and persistence of anxiety disorders (Bandelow et al., 2016), and cannabis itself increases the risk of anxiety attacks (Grunberg et al., 2015; Mammen et al., 2018). In addition to the compounds covered in the current review, other potentially promising areas for future research include components of the neurotrophic signaling, renin-angiotensin, acetylcholine, and even the opioid system (Morrison and Ressler, 2014). In conclusion, there is still an urgent need to develop novel methods to treat anxiety disorders and related diseases (Griebel and Holmes, 2013). In a recent review, Markiewicz (2017) showed that neurofeedback is effective in many psychiatric disorders that affect psychological variables such as stress and anxiety. To avoid the side effects of drugs, from the perspective of anxiety-reducing technology (Pintado and Llamazares, 2014), neurofeedback therapy is a promising new method with stable and lasting therapeutic effects no side effects.

A number of studies in the extant literature have affirmed that in the treatment of anxiety disorders, neurofeedback focuses on the central nervous system and the brain (Fovet et al., 2015) to improve neuroregulation and stability. Among them, the regulation of brain activity can affect behavioral changes (Marzbani et al., 2016; Van der Kolk et al., 2016). Neurofeedback uses computer technology to train patients to improve poorly regulated brain wave patterns (Micoulaud-Franchi et al., 2015). Current imaging modes of neurofeedback include real-time magnetic resonance imaging (RT-MRI), functional near-infrared spectroscopy (fNIRS), and electroencephalography (EEG). For example, Lori-Ann et al. (2013) explored frontal lobe asymmetry using fNIRS. To assess the prefrontal asymmetry of female college students with the highest and lowest percentile scores in the high and low anxiety groups on social challenge tasks *in vivo*, the results showed that the high anxiety group exhibited a non-significant trend toward greater right frontal activity than the low anxiety group but only to assess the prefrontal cortex. For example, Morgenroth et al. (2020) assigned 32 participants with high trait anxiety to either an experimental group to undergo RT-MRI or a control group to receive a false feedback. The results showed that RT-fMRI neurofeedback training led to a reduction in anxiety levels and the feasibility of altering activation in the wider network. However, there was no group difference in Stroop's task performance. In studies such as Sachs et al. (2004) using quantitative EEG to compare participants with a healthy control group in a state of rest and alertness (the participants would sound an alarm if drowsiness occurred), Sachs and his colleagues observed population differences in beta frequencies in the frontal lobe and right central region. Although no statistical analysis of hemispheric data was performed, the beta acceleration appeared to be predominantly in the right hemisphere. Subjects with high or low trait anxiety used alpha feedback to increase and decrease their EEG alpha activity. Changes in alpha were strongly associated with changes in anxiety but only in subjects with a high level of anxiety (for whom anxiety decreased linearly with an increase in alpha and increased linearly with an increase in alpha inhibition). These results suggest that long-term alpha feedback training (at least 5 h) may be helpful in anxiety management.

This study is based on an evaluation of the efficacy of an anxious state classification described in Chao et al. (2021), where in EEG signals were used to study neural changes, and the results showed that the support vector machine classifier was able to classify and recognize two psychological states (anxiety and no anxiety) using power spectral density as a model. In this paper, we design a neurofeedback system based on the alpha band oscillation (frequency power) of EEG signals. Subjects with anxiety disorders and healthy subjects were recruited to participate in the experiment where EEG signals were recorded and analyzed. The findings revealed that the activity of alpha, theta, and gamma waves of anxious subjects increased significantly. After the adjustment of mindfulness, the observation power graph showed an increasing trend. At the same time, the analysis of variance (ANOVA) showed that  $P < 0.05$ , i.e., the difference was significant, and the anxiety symptoms of the subjects could be relieved from the perspective of neurophysiology.



## MATERIALS AND METHODS

### Subjects

In this study, a total of 34 subjects, 17 with anxiety disorder ( $37 \pm 7.61$  years old) and 17 healthy ( $24.41 \pm 1.49$  years old) participated in this experiment. All the subjects had normal hearing and never received mindfulness recording therapy or training. These anxiety disorders were judged by professional psychiatrists from Beijing Anding Hospital affiliated with Capital Medical University. All procedures performed in studies involving human participants were in accordance with the ethical standards of ethics committee of Beijing Anding Hospital, Capital Medical University (ZYLX201607). Healthy subjects are graduate students. Prior to the experiment, the subjects were instructed to read and sign the informed consent form and detailed personal information. The subjects were classified as healthy or anxious. The anxiety targets were pure anxiety patients. All the subjects participated in this experiment in a psychiatric hospital with consent. The selection criteria were decided by professional psychiatrists to appraise the subjects' eligibility to participate in this experiment.

### Experiment Paradigm and Data Recording

The subjects were asked to sit in a chair facing the desktop computer. As shown in **Figure 1**, the example includes three stages. In the first stage, the subjects were asked to remain emotionally stable for 5 min. The mindfulness recording was played in the second stage; the subjects followed the mindfulness (Schwartz et al., 2003; Sever et al., 2003) recording and gave an 8 min voice prompt to adjust their mental state. Finally, the subjects continued to return to a static state for 5 min. In this experiment, all subjects were required to keep their eyes open. In three small experimental phases, they needed to complete a self-assessment of anxiety. Each subject was required to fill in a visual analog scale comprising a scale axes marked with numbers 0–10. The number from 0 to 4 are defined as non-anxious, from 5 to 7 as moderate anxiety, and 8 to 10 as severe anxiety.

During the experiment, a 32-channel EEG signal was recorded from the subjects' scalp (Brain Products, Germany). According to the International 10–20 system, the EEG signals were recorded through 19 electrodes, namely Fp1, Fp2, F7, F3, F4, F8, T7, C3, C4, T8, P7, P3, Pz, P4, P8, O1, O2, Fz, and Cz as shown in **Figure 2**. The electrode Cz were chosen as reference electrodes. During data recording, the impedance of each electrode was kept below 5 K $\Omega$ . The EEG data were collected at a sampling rate of 500 Hz.

Throughout this experiment, the alpha band power of the electrodes in the left and right frontal lobes (Harmon-Jones and Allen, 1997) was calculated in real time and displayed as feedback to establish a neurofeedback system. The energy of the signal is shown as red and green bars to depict the energy asymmetry in the frontal lobe. The red and green bar graphs represent the energy values of the alpha wave on the left and right sides of the frontal lobe, respectively. Subjects can see the changes in the visual bar and try to adjust their mental state during mindfulness

training. As shown in **Figure 3**, prior to the mindfulness training, the left and right strengths are different, while subsequent to the training, the strength difference of some subjects decreased.

### Electroencephalography Data Processing

Because there are many noise artifacts in EEG signals, such as electrocardiogram (ECG), electromyography (EMG), and power frequency interference, it is necessary to preprocess the original EEG signal to obtain a relatively pure brain signal. In this study, independent component analysis (Arnaud, 2004) was used to eliminate eye movement artifacts. Our preprocessed EEG data took 4 s as the step size, calculated the power spectral density (Ahani et al., 2014; Cai et al., 2018), and obtained the alpha, theta, and gamma wave power values. A total of 4 min of data were calculated. Periodogram method is a method to estimate the power spectral density directly by Fourier transform of the sampled data  $X(n)$  of the signal. It is assumed that the finite length random signal sequence is  $x(n)$ . Its Fourier transform and power spectral density have the following relationship:

$$\tilde{x}_x(f) = \frac{1}{N} |x(f)|^2 \quad (1)$$

where  $N$  is the length of the random signal sequence  $x(n)$ . At discrete frequency points  $f = k\Delta f$ . There are:

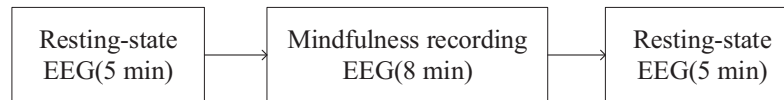
$$\tilde{x}_x(k) = \frac{1}{N} |X(k)|^2 = \frac{1}{N} |FFT[x(n)]|^2 \quad k = 0, 1, \dots, N-1 \quad (2)$$

where,  $FFT[x(n)]$  is Fast Fourier Transform of the sequence  $x(n)$ . Because the period of  $FFT[x(n)]$  is  $N$ , the power spectrum estimation obtained took  $N$  as the period.

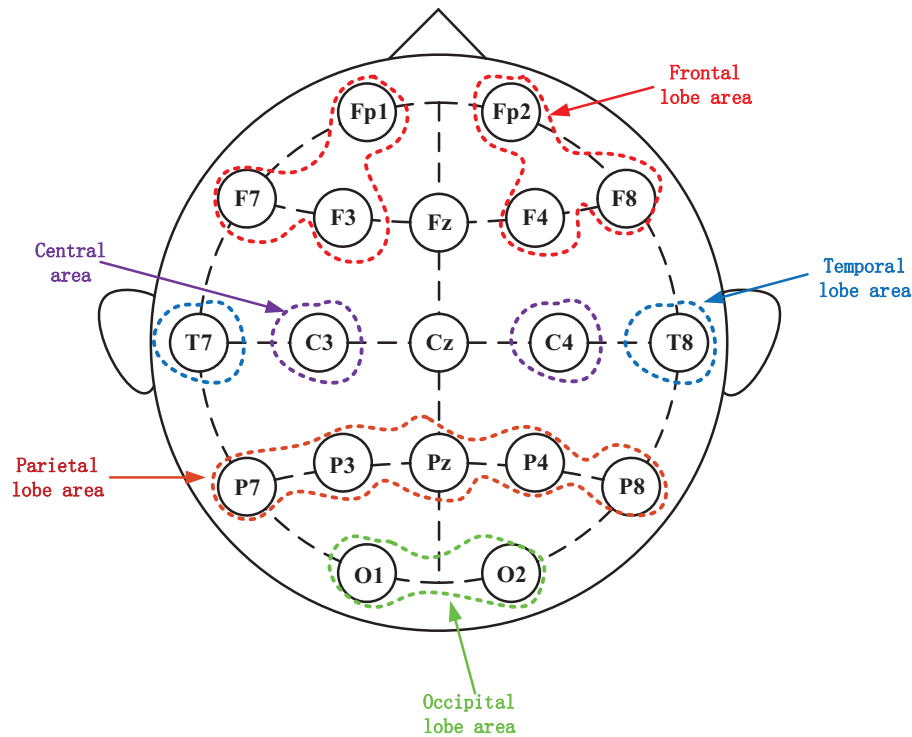
Finally, the average power of each electrode was calculated. For statistical analysis, 16 electrodes were selected to be divided into the following 10 brain regions: right occiput (O2), left occiput (O1), right parietal (P8, P4), left parietal (P7, P3), right central (C4), left center (C3), right frontal lobe (F8, F4, and Fp2), left frontal lobe (F7, F3, and Fp1), right temporal lobe (T8), and left temporal lobe (T7). Under each zone, the power values of the constituent electrodes are averaged, and the process is repeated for the alpha, theta, and gamma frequency bands.

## EXPERIMENTAL RESULTS

To verify the impact of mindfulness adjustment on the EEG signals of the subjects, we conducted ANOVA using Statistical Package for the Social Sciences (SPSS) software (Wang et al., 2015). ANOVA (Harne and Hiwale, 2018) included the influence of mindfulness adjustment on frequency bands and the influence of different frequency bands in different brain regions. In neurofeedback, the changes in alpha, theta, and gamma waves are usually used as evaluation indicators, and corresponding improvement and treatment are executed by strengthening these waves. These are rhythmic waveforms produced by the brain during some activities. The characteristics of alpha wave are: it is easy to observe when



**FIGURE 1** | Experimental paradigm of the proposed affective brain-computer interface.



**FIGURE 2** | Electroencephalography electrode placement based on the international 10–20 system. Scalp potential and reference point distribution of left and right frontal lobes.

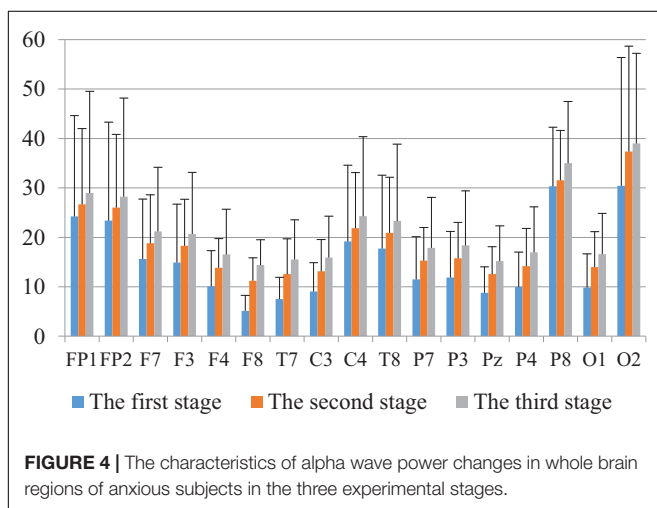
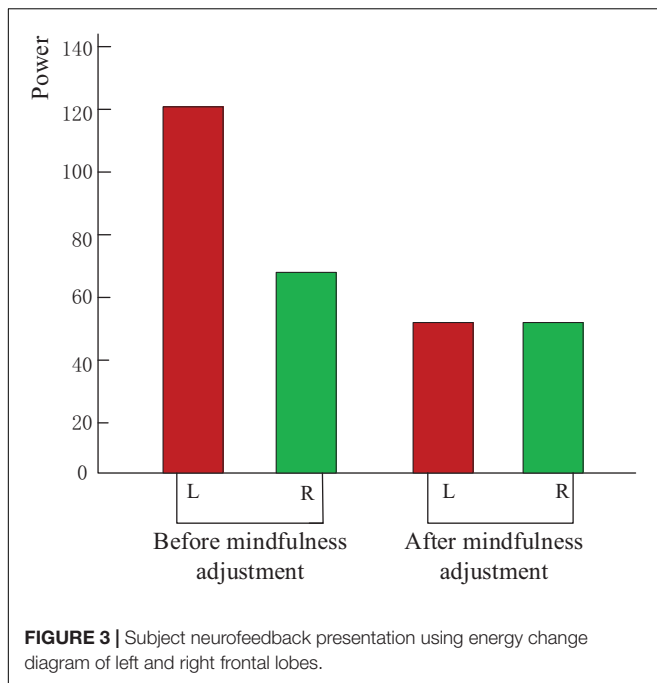
people are in a relaxed, calm but awake state; Theta wave is characterized by its low frequency when people are sleepy; Gamma wave is characterized by: when the brain is engaged in a cognitive task, it connects neurons that have not been connected before to create a new working loop – popular understanding is that when creative thinking, creativity, and ideas suddenly appear at this time, gamma waves can be observed. Furthermore, related research shows that the main EEG indicators that are sensitive to mindfulness are alpha, theta, and gamma waves. Combining with previous research, the EEG indicators we chose while performing ANOVA were alpha, theta, and gamma waves.

First, we analyzed the change characteristics of the average power values of the alpha, theta, and gamma waves in the three stages of the experiment for anxious subjects and healthy subjects. It can be seen from **Figures 4–6** that anxious subjects exhibited a characteristic that the power of alpha, theta, and gamma waves was generally very low prior to mindfulness adjustment, and the power gradually increased with this adjustment. This trend of

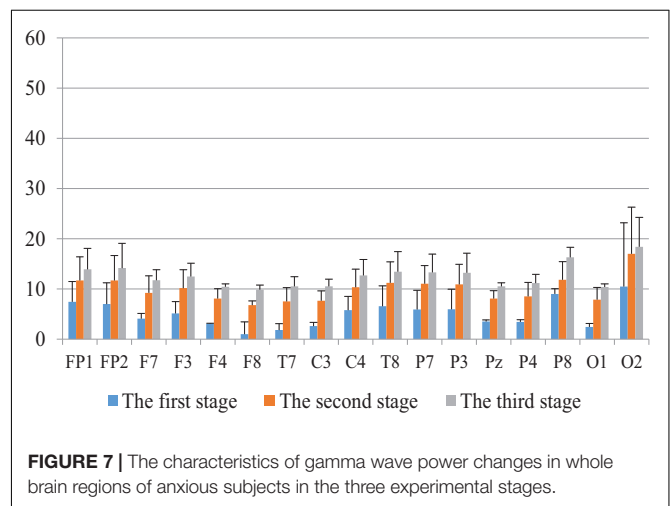
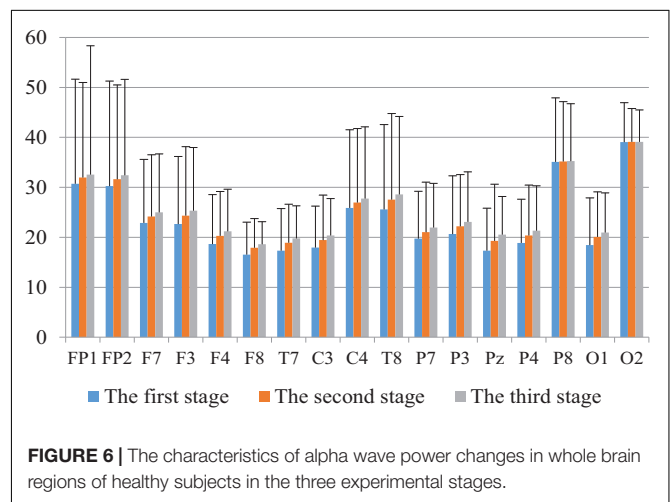
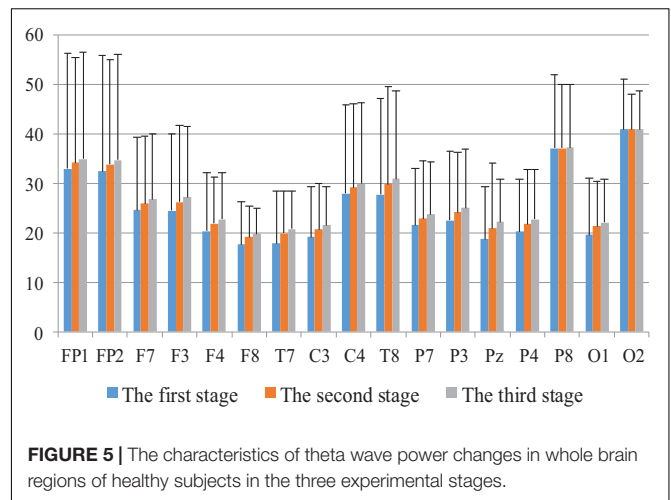
change is consistent on each electrode, and the magnitude of the change is obvious. Mindfulness adjustment activates higher alpha, theta, and gamma waves. For healthy subjects, although the alpha wave also changed during the three stages of the experiment, this change is not evident in **Figure 7**, and not all electrodes show the same changes as the anxious subjects. For P8 and O2, the electrode changes can be seen as flat. **Figures 8, 9** also show similar features to **Figure 7**.

Through the comparison between **Figures 4–9**, we can also see that the power value of the alpha wave of each lead of anxious subjects is lower than the power value of the alpha wave of healthy subjects. For theta and gamma waves, an analogous pattern was observed.

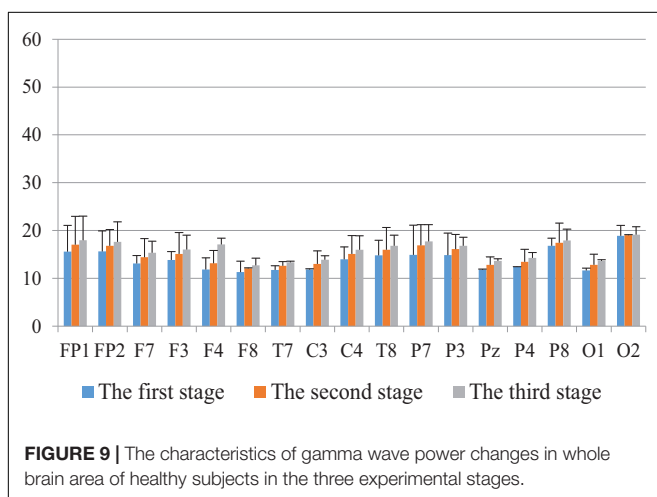
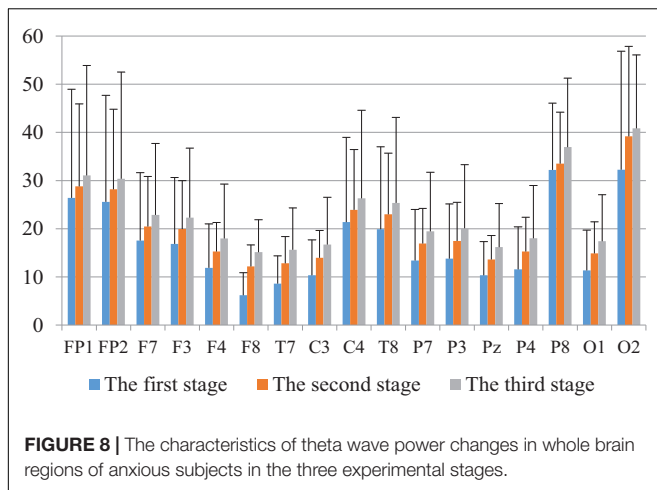
To assess the connection between the mindfulness regulation hemisphere, region, and condition, the wave bands were divided into hemispheres (left and right), regions (frontal, central, parietal, occipital, and temporal lobe), and condition (before and after mindfulness adjustment) as factors for three-way repeated ANOVA.



In **Tables 1, 2**, alpha, theta, and gamma waves were analyzed under anxiety and health conditions experimental results. Under the main effect, the number of treatment groups ( $k = 2$ ), the number of samples ( $n = 390$ ), the degree of freedom between groups was  $k - 1$ , and the degree of freedom within each group was  $n - k$ . In **Table 1**,  $F(1,338) = 127.65$ , which means that the degree of freedom of components is 1. The degree of freedom within the group is 338, and  $P = 2.50 \times 10^{-25} < 0.05$ , which means that the difference between groups is significant. In condition  $\times$  region interaction effect, it is divided into regions (frontal lobe, central lobe, parietal lobe, occipital lobe, and temporal lobe), number of treatment groups ( $k = 5$ ), number of samples ( $n = 390$ ), degree of freedom between groups is  $k - 1$ , degree of freedom within group is  $n(k - 1)$ ,  $F(4,339) = 20.52$  in **Table 1**, indicating that the degree of freedom of component is 1, and degree of



freedom within group is 339.  $P = 4.55 \times 10^{-15} < 0.05$ , indicating a significant difference between groups. In condition  $\times$  hemispheric interaction effect, it is divided into hemispheres (left and right),



number of processing groups ( $k = 2$ ), number of samples ( $n = 390$ ), degree of freedom between groups is  $k - 1$ , and degree of freedom within group is  $n(k - 1)$ . In **Table 1**,  $F(1,339) = 14.62$ , indicating that the degree of freedom between groups is 1, and the degree of freedom within group is 339. Similarly,  $P < 0.05$ , i.e., the difference between groups was significant, which was statistically significant.

The repeated ANOVA results in **Tables 1, 2** indicated that both healthy and anxious subjects have significant main effects under the conditions of alpha, theta, and gamma zones. However, this main effect is more prominent in the latter. We can thus say that the analysis of our experimental results shows that neurofeedback can alleviate the anxiety of the subjects, and the alleviating effect on anxious subjects is evidently stronger than that on healthy subjects.

The interaction effects (hemisphere  $\times$  region  $\times$  condition, region  $\times$  condition, and hemisphere  $\times$  condition) are also obvious for alpha, theta, and gamma bands, and this significant effect is more intense in healthy subjects. To analyze the interaction between hemispheres, regions, and conditions in more detail, we performed a paired  $t$ -test on the alpha, theta, and gamma power before and after each region condition.

We made a significant analysis of the changes in the average power of alpha, theta, and gamma waves of EEG before and after the mindfulness adjustment of anxious subjects and healthy subjects in five brain regions. It can be seen from **Tables 3–8** that regardless of subject, the average power of alpha, theta, and gamma waves in the five brain regions show very significant changes before and after mindfulness. In addition, in healthy subjects, there were observed higher significant changes than anxious subjects. It can be concluded that mindfulness adjustment made the brain's electrical alpha, theta, and gamma waves of different subjects become more active, and this change was more significant among healthy subjects.

As shown in **Figures 4–6** for anxious subjects, the average power of the EEG biomarkers alpha, theta, and gamma waves of the subjects before and after the neurofeedback mindfulness adjustment increased significantly, indicating that our experiment activated higher alpha, theta, and gamma activities of the subjects, and this change is significant in each brain area. **Tables 1, 3–5** can illustrate this significance. Healthy subjects did not manifest any tendency toward anxiety. They appeared to be in control throughout the experiment. **Figures 7–9** depict that although healthy subjects were tested before and after neurofeedback mindfulness adjustment, the changes in average power of the EEG biomarkers alpha, theta, and gamma waves are not as large as that of anxious subjects. However, they also show a trend of power increase, indicating that our experiment also provided a certain relief to healthy subjects' mood, the role of auxiliary regulation. **Tables 2, 6–8** show that this subtle change is also significant.

## DISCUSSION

In previous studies, the assessment of anxiety relief chiefly relied on some anxiety scales (Zafeiri et al., 2019). Whether it is before or after adjustment, anxiety is relieved and is too one-sided based on the scores of a single scale. Subjects are often not aware of their own situation. For some subjects, the anxiety scale is obscure and difficult to understand, and they appear anxious or even fidgeting during evaluation. It is unclear what such evaluation results can indicate. Although the participant's scale score can explain the relief of anxiety symptoms, the subjectivity of the scale evaluation is still too strong, and thus it needs to be evaluated from an objective perspective to reflect the true state of the participant, and the scale evaluation can be used as an auxiliary evaluation means.

This study was aimed to assess the relieving effect of neurofeedback mindfulness regulation on subjects' anxiety, and to objectively evaluate this relieving effect through neurophysiological signals. The effect on anxiety relief was judged by analyzing the change characteristics of the average power of alpha, theta, and gamma waves of the brain's electrical signal 4 min before and after the subject's neurofeedback mindfulness adjustment.

The alpha, theta, and gamma band power were evaluated before and after neurofeedback mindfulness regulation. For anxiety disorder patients, the average alpha, theta, and gamma



**TABLE 1 |** The average power variance analysis results of the two hemispheres, five regions, and two conditions in the alpha, theta, and gamma bands for anxious subjects before and after the experiment.

Frequency band	Condition main effect	Condition × region interaction effect	Condition × hemisphere interaction effect	Condition × region × hemisphere interaction effect
Alpha	$F(1,338) = 127.65$ $P = 2.50 \times 10^{-25}$	$F(4,339) = 20.52$ $P = 4.55 \times 10^{-15}$	$F(1,339) = 14.62$ $P = 1.50 \times 10^{-4}$	$F(4,339) = 18.42$ $P = 1.21 \times 10^{-13}$
Theta	$F(1,338) = 110.84$ $P = 1.31 \times 10^{-22}$	$F(4,339) = 19.87$ $P = 1.26 \times 10^{-14}$	$F(1,339) = 16.86$ $P = 5.10 \times 10^{-5}$	$F(4,339) = 18.75$ $P = 7.20 \times 10^{-14}$
Gamma	$F(1,338) = 633.73$ $P = 1.66 \times 10^{-79}$	$F(4,339) = 6.51$ $P = 4.70 \times 10^{-5}$	$F(1,339) = 3.74$ $P = 0.0439$	$F(4,339) = 5.65$ $P = 2.08 \times 10^{-4}$

Significant ( $P < 0.0439$ ).**TABLE 2 |** The average power variance analysis results of the two hemispheres, five regions, and two conditions in the alpha, theta, and gamma bands for healthy subjects before and after the experiment.

Frequency band	Condition main effect	Condition × region interaction effect	Condition × hemisphere interaction effect	Condition × region × hemisphere interaction effect
Alpha	$F(1,338) = 9.93$ $P = 0.0017$	$F(4,339) = 37.97$ $P = 5.33 \times 10^{-26}$	$F(1,339) = 50.00$ $P = 9.69 \times 10^{-12}$	$F(4,339) = 48.24$ $P = 9.89 \times 10^{-32}$
Theta	$F(1,338) = 9.78$ $P = 0.0019$	$F(4,339) = 43.24$ $P = 4.25 \times 10^{-29}$	$F(1,339) = 56.34$ $P = 6.08 \times 10^{-13}$	$F(4,339) = 18.42$ $P = 6.86 \times 10^{-25}$
Gamma	$F(1,338) = 77.13$ $P = 8.13 \times 10^{-17}$	$F(4,339) = 41.26$ $P = 6.95 \times 10^{-28}$	$F(1,339) = 113.49$ $P = 6.95 \times 10^{-23}$	$F(4,339) = 36.62$ $P = 3.31 \times 10^{-25}$

Significant ( $P < 0.0017$ ).**TABLE 3 |** The results of paired *t*-test for average alpha power before and after each regional condition for anxious subjects.

Hemisphere	Region	Conditions	Paired differences				
			Mean ± SD	SE mean	95% confidence interval of the difference		
					Lower	Upper	P
Left	Occipital	Before–after	$-7.69 \pm 3.09$	0.75	−9.28	−6.10	$1.94 \times 10^{-8}$
	Parietal	Before–after	$-7.57 \pm 3.22$	0.78	−9.22	−5.91	$4.28 \times 10^{-8}$
	Central	Before–after	$-5.77 \pm 4.05$	0.98	−7.85	−3.69	$2.30 \times 10^{-5}$
	Frontal	Before–after	$-6.40 \pm 3.65$	0.88	−8.28	−4.52	$2.00 \times 10^{-6}$
	Temporal	Before–after	$-8.89 \pm 2.90$	0.70	−10.38	−7.39	$9.83 \times 10^{-10}$
Right	Occipital	Before–after	$-26.51 \pm 14.09$	3.41	−33.75	−19.26	$8.23 \times 10^{-7}$
	Parietal	Before–after	$-6.41 \pm 1.65$	0.40	−7.26	−5.56	$2.91 \times 10^{-11}$
	Central	Before–after	$-5.93 \pm 4.13$	1.00	−8.10	−3.85	$2.00 \times 10^{-5}$
	Frontal	Before–after	$-7.90 \pm 3.25$	0.78	−9.58	−6.23	$2.64 \times 10^{-8}$
	Temporal	Before–after	$-6.83 \pm 3.62$	0.87	−8.69	−4.96	$2.12 \times 10^{-7}$

power is generally very low before the regulation of mindfulness, and gradually increases with the regulation of mindfulness. For healthy subjects, it can be observed that the power of alpha, theta, and gamma bands increased not obviously, compared with anxiety disorder patients. Additionally, the power of each band of patients was lower than healthy subjects. The statics analysis showed the significant effect of brain activities after neurofeedback.

In the past, neurofeedback has been used to regulate brain activity and reduce alpha asymmetry to improve anxiety in patients with depression and anxiety. However, different types of intervention are required for different patients with different

duration of training, and the sample size is too small. The data used in the evaluation process is relatively simple, and there are incorrect experimental data, which will affect the results of the experiment. For example, Dias divided 87 patients with major depressive disorder and anxiety into alpha-asymmetry neurofeedback (ALAY), high-beta down-training, or control groups. Both neurofeedback groups received 10 sessions of neurofeedback (Dias and van Deusen, 2011) and had reduced symptoms of depression and anxiety. Compared with the other groups, the BETA group was more effective at reducing the high beta power in the parietal cortex, but it may take more than 10 repetitions of training to reach the neurofeedback

**TABLE 4 |** The results of paired *t*-test for average theta power before and after each regional condition for anxious subjects.

Hemisphere	Region	Conditions	Paired differences				
			Mean $\pm$ SD	SE mean	95% confidence interval of the difference		
					Lower	Upper	<i>P</i>
Left	Occipital	Before–after	$-7.15 \pm 3.32$	0.80	-8.86	-5.45	$1.37 \times 10^{-7}$
	Parietal	Before–after	$-7.38 \pm 3.37$	0.81	-9.12	-5.65	$1.11 \times 10^{-7}$
	Central	Before–after	$-5.75 \pm 4.08$	0.99	-7.85	-3.65	$2.70 \times 10^{-5}$
	Frontal	Before–after	$-6.24 \pm 3.79$	0.92	-8.19	-4.29	$4.00 \times 10^{-6}$
	Temporal	Before–after	$-8.03 \pm 2.94$	0.71	-9.54	-6.52	$5.11 \times 10^{-9}$
Right	Occipital	Before–after	$-27.09 \pm 14.45$	3.50	-34.53	-19.66	$8.64 \times 10^{-7}$
	Parietal	Before–after	$-6.19 \pm 1.75$	0.42	-7.09	-5.28	$1.20 \times 10^{-10}$
	Central	Before–after	$-5.97 \pm 4.21$	1.02	-8.14	-3.81	$2.40 \times 10^{-5}$
	Frontal	Before–after	$-7.78 \pm 3.47$	0.84	-9.57	-6.00	$8.15 \times 10^{-8}$
	Temporal	Before–after	$-6.75 \pm 3.74$	0.90	-8.67	-4.82	$1.00 \times 10^{-6}$

**TABLE 5 |** The results of paired *t*-test for average gamma power before and after each regional condition for anxious subjects.

Hemisphere	Region	Condition	Paired differences				
			Mean $\pm$ SD	SE mean	95% confidence interval of the difference		
					Lower	Upper	<i>P</i>
Left	Occipital	Before–after	$-8.56 \pm 2.16$	0.52	-9.68	-7.45	$2.11 \times 10^{-11}$
	Parietal	Before–after	$-8.46 \pm 3.13$	0.76	-10.07	-6.85	$6.05 \times 10^{-9}$
	Central	Before–after	$-7.19 \pm 2.55$	0.62	-8.50	-5.87	$3.39 \times 10^{-9}$
	Frontal	Before–after	$-7.86 \pm 2.42$	0.58	-9.10	-6.61	$4.21 \times 10^{-10}$
	Temporal	Before–after	$-9.41 \pm 2.34$	0.56	-10.62	-8.20	$1.77 \times 10^{-11}$
Right	Occipital	Before–after	$-14.80 \pm 7.35$	1.78	-18.58	-11.02	$3.41 \times 10^{-7}$
	Parietal	Before–after	$-7.89 \pm 1.48$	0.36	-8.66	-7.13	$2.39 \times 10^{-13}$
	Central	Before–after	$-8.01 \pm 2.68$	0.65	-9.39	-6.63	$1.40 \times 10^{-9}$
	Frontal	Before–after	$-8.46 \pm 2.25$	0.54	-9.62	-7.30	$4.69 \times 10^{-11}$
	Temporal	Before–after	$-7.82 \pm 2.84$	0.68	-9.28	-6.36	$4.52 \times 10^{-9}$

**TABLE 6 |** The results of paired *t*-test for average alpha power before and after each regional condition for healthy subjects.

Hemisphere	Region	Conditions	Paired differences				
			Mean $\pm$ SD	SE mean	95% confidence interval of the difference		
					Lower	Upper	<i>P</i>
Left	Occipital	Before–after	$-2.54 \pm 0.33$	0.08	-2.71	-2.36	$1.08 \times 10^{-15}$
	Parietal	Before–after	$-2.35 \pm 0.27$	0.06	-2.49	-2.21	$1.25 \times 10^{-16}$
	Central	Before–after	$-2.06 \pm 0.28$	0.06	-2.21	-1.92	$1.51 \times 10^{-15}$
	Frontal	Before–after	$-2.29 \pm 0.26$	0.06	-2.43	-2.15	$1.36 \times 10^{-16}$
	Temporal	Before–after	$-2.48 \pm 0.35$	0.08	-2.66	-2.29	$3.53 \times 10^{-15}$
Right	Occipital	Before–after	$-0.01 \pm 0.001$	0.0003	-0.02	-0.01	$4.07 \times 10^{-20}$
	Parietal	Before–after	$-1.32 \pm 0.09$	0.02	-1.37	-1.27	$7.60 \times 10^{-20}$
	Central	Before–after	$-2.22 \pm 0.34$	0.08	-2.39	-2.04	$9.19 \times 10^{-15}$
	Frontal	Before–after	$-2.30 \pm 0.31$	0.07	-2.47	-2.14	$1.86 \times 10^{-15}$
	Temporal	Before–after	$-3.04 \pm 0.39$	0.09	-3.24	-2.83	$5.81 \times 10^{-16}$

goal. In addition, Cheon et al. (2015) modified the 8-week ALAY neurofeedback regimen to increase the beta power of the left frontal cortex (F3) and decrease the alpha power, while

increasing the theta (alpha/theta ratio) of the parietal cortex in depressed patients. The results demonstrated that within 8 weeks, depression and anxiety symptoms were significantly reduced, as

**TABLE 7 |** The results of paired *t*-test for average theta power before and after each regional condition for healthy subjects.

Hemisphere	Region	Conditions	Paired differences				
			Mean $\pm$ SD	SE mean	95% confidence interval of the difference		
					Lower	Upper	P
Left	Occipital	Before–after	$-2.58 \pm 0.47$	0.11	-2.82	-2.33	$1.84 \times 10^{-13}$
	Parietal	Before–after	$-2.39 \pm 0.32$	0.07	-2.55	-2.22	$1.13 \times 10^{-16}$
	Central	Before–after	$-2.11 \pm 0.29$	0.07	-2.26	-1.96	$1.85 \times 10^{-15}$
	Frontal	Before–after	$-2.32 \pm 0.29$	0.07	-2.47	-2.17	$5.24 \times 10^{-16}$
	Temporal	Before–after	$-2.98 \pm 0.58$	0.14	-3.28	-2.68	$4.03 \times 10^{-13}$
Right	Occipital	Before–after	$-0.01 \pm 0.001$	0.0004	-0.01	-0.01	$3.18 \times 10^{-17}$
	Parietal	Before–after	$-1.31 \pm 0.12$	0.03	-1.38	-1.25	$6.14 \times 10^{-18}$
	Central	Before–after	$-2.26 \pm 0.35$	0.08	-2.44	-2.08	$1.13 \times 10^{-14}$
	Frontal	Before–after	$-2.34 \pm 0.38$	0.09	-2.54	-2.14	$3.38 \times 10^{-14}$
	Temporal	Before–after	$-3.15 \pm 0.41$	0.10	-3.36	-2.29	$8.00 \times 10^{-16}$

**TABLE 8 |** The results of paired *t*-test for average gamma power before and after each regional condition for healthy subjects.

Hemisphere	Region	Conditions	Paired differences				
			Mean $\pm$ SD	SE mean	95% confidence interval of the difference		
					Lower	Upper	P
Left	Occipital	Before–after	$-1.99 \pm 0.07$	0.02	-2.03	-1.96	$8.98 \times 10^{-25}$
	Parietal	Before–after	$-2.39 \pm 0.33$	0.08	-2.57	-2.22	$1.94 \times 10^{-15}$
	Central	Before–after	$-2.40 \pm 0.22$	0.05	-2.51	-2.29	$3.40 \times 10^{-18}$
	Frontal	Before–after	$-2.26 \pm 0.13$	0.03	-2.33	-2.19	$4.61 \times 10^{-21}$
	Temporal	Before–after	$-1.61 \pm 0.05$	0.01	-1.63	-1.58	$1.39 \times 10^{-25}$
Right	Occipital	Before–after	$-0.28 \pm 0.003$	0.0008	-0.28	-0.27	$6.31 \times 10^{-32}$
	Parietal	Before–after	$-1.48 \pm 0.02$	0.007	-1.50	-1.47	$5.71 \times 10^{-29}$
	Central	Before–after	$-1.97 \pm 0.12$	0.03	-2.04	-1.91	$1.26 \times 10^{-20}$
	Frontal	Before–after	$-1.86 \pm 0.09$	0.02	-1.91	-1.82	$2.35 \times 10^{-22}$
	Temporal	Before–after	$-2.00 \pm 0.09$	0.02	-2.05	-1.95	$1.31 \times 10^{-22}$

was the clinical severity of psychiatric symptoms. The 15 patients in the feedback group were given the neurofeedback training of alpha enhancement, theta enhancement, and beta3 reduction, three times a week for 4 weeks. The fake feedback group did not give real feedback, but simply played back previous training data from other people for the same amount of time as the neurofeedback group. Patients in both groups were treated with the same drug (duloxetine hydrochloride 60 mg once daily). Results After training, the alpha and theta amplitude of the feedback group were significantly higher than that of the false feedback group, and the beta3 amplitude had a downward trend; however, there was no statistical difference (*P*-value was 0.004, 0.038, and 0.818, alpha, theta, and beta3, respectively). However, the feedback group had the function of helping to improve the anxiety of patients with generalized anxiety disorder.

This experiment has made some progress in the evaluation of anxiety state, which is only a small step forward, and there is still a lot of room for improvement in the accuracy of the evaluation. Due to the limited research time and small sample size, there may be many methods that can be applied to the assessment of anxiety state, and the future research prospects are broad. In the near

future, we will need to optimize our experiments to improve the relief level of anxiety symptoms of anxiety subjects to the level of healthy subjects.

To sum up, the nervous feedback can effectively control the brain wave patterns and achieve cure, and possesses the advantages of non-invasive, less adverse reaction, the characteristics of being simple, safe, and convenient (Vernon et al., 2003). Through the analysis of neurophysiological signals, it can be concluded that our experiment can alleviate the anxiety symptoms of the subjects. In the current period of new crown epidemic most people suffer from anxiety, whether it is healthy people or patients with anxiety disorders. We hope that our experiment can provide people with relief from their anxiety.

## DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the experiment data is not available online for further research, but available on reasonable request according to the policy of Tianjin University, Capital Medical University, and

Tianjin University of Technology. Requests to access the datasets should be directed to CC, cccovb@hotmail.com.

## ETHICS STATEMENT

All procedures performed in studies involving human participants were in accordance with the ethical standards of Ethics Committee of Beijing Anding Hospital, Capital Medical University (ZYLX201607). This ethics committee's responsibilities, composition, function, operations, and records are fully compliant with ICH-GCP, and related regulation and law of China. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

CC, XX, AB, and CW recorded the original experiment data, analyzed the experiment data, and wrote parts of the manuscript.

SS and XZ completed the ethic files of this experiment. LL, XW, WY, and PL wrote parts of the manuscript. CW, CC, and DM designed the experiment and revised the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was partly financially supported by the National Natural Science Foundation of China (61806146, 61971118, and 81901860), Research Plan for Innovation in Clinical Technology by Beijing Hospitals Authority (XMLX201805), Beijing Municipal Administration of Hospitals Incubating Program (PX2018063), Scientific Special Commissioner Foundation of Tianjin City (19JCTPJC56000), Graduate Research and Innovation Project of Tianjin City (2019YJSS052), Beijing Key Laboratory of Mental Disorders (Code: 2020JSJB04), and Natural Science Foundation of Tianjin City (18JCYBJC95400).

## REFERENCES

- Ahani, A., Wahbeh, H., and Hooman, N. (2014). Quantitative change of EEG and respiration signals during mindfulness meditation. *J. Neuroeng. Rehabil.* 11:87. doi: 10.1186/1743-0003-11-87
- Arnaud, D. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009
- Bandelow, B., Baldwin, D., Abelli, M., Bolea-Alamanac, B., Bourin, M., Chamberlain, S. R., et al. (2016). Biological markers for anxiety disorders, OCD and PTSD: a consensus statement. Part II: neurochemistry, neurophysiology and neurocognition. *World J. Biol. Psychiatry* 18, 162–214. doi: 10.1080/15622975.2016.1190867
- Bandelow, B., and Michaelis, S. (2015). Epidemiology of anxiety disorders in the 21st century. *Dialogues Clin. Neurosci.* 17, 327–335.
- Cai, H., Han, J., Chen, Y., Sha, X., Wang, Z., Hu, B., et al. (2018). A pervasive approach to EEG-based depression detection. *Complexity* 3:5238028.
- Carpenter, J. K., Andrews, L. A., Witcraft, S. M., Powers, M. B., Smits, J. A. J., and Hofmann, S. G. (2018). Cognitive behavioral therapy for anxiety and related disorders: a meta-analysis of randomized placebo-controlled trials. *Depress. Anxiety* 35, 502–514. doi: 10.1002/da.22728
- Chao, C., Xuecong, Y., Belkacem, A. N., Lu, L., Li, P., Zhang, Z., et al. (2021). EEG-based anxious states classification using affective BCI-based closed neurofeedback system. *J. Med. Biol. Eng.* 2, 1–10. doi: 10.1007/s40846-020-00596-7
- Cheon, E. J., Koo, B. H., and Choi, J. H. (2015). The efficacy of neurofeedback in patients with major depressive disorder: an open labeled prospective study. *Appl. Psychophys. Biofeedback* 41, 103–110. doi: 10.1007/s10484-015-9315-8
- Craske, M. G., Stein, M. B., Eley, T. C., Milad, M. R., Holmes, A., Rapee, R. M., et al. (2017). Anxiety disorders. *Nat. Rev. Dis. Primers* 3:17100.
- Crocq, M. A. (2015). A history of anxiety: from Hippocrates to DSM. *Dialogues Clin. Neurosci.* 17, 319–325.
- Dias, A. M., and van Deussen, A. (2011). A new neurofeedback protocol for depression. *Span. J. Psychol.* 14, 374–384. doi: 10.5209/rev.SJOP.2011.v14.n1.34
- Fovet, T., Jardri, R., and Linden, D. (2015). Current issues in the use of fMRI-based neurofeedback to relieve psychiatric symptoms. *Curr. Pharm. Des.* 21, 3384–3394.
- GBD 2017 DALYs and Hale Collaborators (2018). Global, regional, and national disability-adjusted life years (DALYs) for 359 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *Lancet* 392, 1859–1922. doi: 10.1016/S0140-6736(18)32335-3
- Gottschalk, M. G., and Domschke, K. (2018). Oxytocin and anxiety disorders. *Curr. Top. Behav. Neurosci.* 35, 467–498. doi: 10.3190/01674829809044221
- Griebel, G., and Holmes, A. (2013). 50 years of hurdles and hope in anxiolytic drug discovery. *Nat. Rev. Drug Discov.* 12, 667–687. doi: 10.1038/nrd4075
- Grunberg, V. A., Cordova, K. A., Bidwell, L. C., and Ito, T. A. (2015). Can marijuana make it better? Prospective effects of marijuana and temperament on risk for anxiety and depression. *Psychol. Addict. Behav.* 29, 590–602. doi: 10.1037/adb0000109
- Harmon-Jones, E., and Allen, J. J. B. (1997). Behavioral activation sensitivity and resting frontal EEG asymmetry: COVRiation of putative indicators related to risk for mood disorders. *J. Abnorm. Psychol.* 106, 159–163. doi: 10.1037/0021-843X.106.1.159
- Harne, B. P., and Hiwale, A. S. (2018). EEG spectral analysis on OM mantra meditation: a pilot study. *Appl. Psychophysiol. Biofeedback* 43, 123–129. doi: 10.1007/s10484-018-9391-7
- Kalin, N. H. (2017). Mechanisms underlying the early risk to develop anxiety and depression: a translational approach. *Eur. Neuropsychopharmacol.* 27, 543–553. doi: 10.1016/j.euroneuro.2017.03.004
- Kogan, C. S., Stein, D. J., Maj, M., First, M. B., Emmelkamp, P. M., and Reed, G. M. (2016). The classification of anxiety and fear-related disorders in the ICD-11. *Depress. Anxiety* 33, 1141–1154. doi: 10.1002/da.22530
- Lori-Ann, T., James, D. H., Forman, E. M., Juarascio, A. S., Izzetoglu, M., and Schultheis, M. (2013). Exploring frontal asymmetry using functional near-infrared spectroscopy: a preliminary study of the effects of social anxiety during interaction and performance tasks. *Brain Imaging Behav.* 7, 140–153. doi: 10.1007/s11682-012-9206-z
- Mammen, G., Rueda, S., Roercke, M., Bonato, S., Lev-Ran, S., and Rehm, J. (2018). Association of cannabis with long-term clinical symptoms in anxiety and mood disorders: a systematic review of prospective studies. *J. Clin. Psychiatry* 79:17r11839. doi: 10.4088/JCP.17r11839
- Markiewicz, R. (2017). The use of EEG biofeedback/neurofeedback in psychiatric rehabilitation. *Psychiatr. Polska* 51, 1095–1106. doi: 10.12740/PP/68919
- Marzbani, H., Marateb, H. R., and Mansourian, M. (2016). Neurofeedback: a comprehensive review on system design, methodology and clinical applications. *Basic Clin. Neurosci.* 7, 143–158. doi: 10.15412/J.BCN.03070208
- Micoulaud-Franchi, J. A., McGonigal, A., Lopez, R., Daudet, C., Kotwas, I., and Bartolomei, F. (2015). Electroencephalographic neurofeedback: level of evidence in mental and brain disorders and suggestions for good clinical practice. *Neurophysiol. Clin.* 45, 423–433. doi: 10.1016/j.neucli.2015.10.077



- Mokdad, A. H., Ballestros, K., Echko, M., Glenn, S., Olsen, H. E., Mullany, E., et al. (2018). The state of us health, 1990–2016: burden of diseases, injuries, and risk factors among us states. *JAMA* 319, 1444–1472. doi: 10.1001/jama.2018.0158
- Morgenroth, E., Saviola, F., Gilleen, J., Allen, B., Lührs, M., Eysenck, M. W., et al. (2020). Using connectivity-based real-time FRMI neurofeedback to modulate attentional and resting state networks in people with high trait anxiety. *Neuroimage Clin.* 25:102191. doi: 10.1016/j.nicl.2020.102191
- Morrison, F. G., and Ressler, K. J. (2014). From the neurobiology of extinction to improved clinical treatments. *Depress. Anxiety* 31, 279–290. doi: 10.1002/da.22214
- Nepon, J., Belik, S. L., Bolton, J., and Sareen, J. (2010). The relationship between anxiety disorders and suicide attempts: findings from the national epidemiologic survey on alcohol and related conditions. *Depress. Anxiety* 27, 791–798. doi: 10.1002/da.20674
- Otowa, T., Hek, K., Lee, M., Byrne, E. M., Mirza, S. S., Nivard, M. G., et al. (2016). Meta-analysis of genome-wide association studies of anxiety disorders. *Mol. Psychiatry* 21, 1391–1399. doi: 10.1016/j.euroneuro.2016.09.604
- Pintado, I. S., and Llamazares, M. (2014). Description of the general procedure of a stress inoculation program to cope with the test anxiety. *Psychology* 5, 956–965. doi: 10.4236/psych.2014.58106
- Reed, G. M., First, M. B., Kogan, C. S., Hyman, S. E., Gureje, O., Gaebel, W., et al. (2019). Innovations and changes in the ICD-11 classification of mental, behavioural and neurodevelopmental disorders. *World Psychiatry* 18, 3–19. doi: 10.1002/wps.20611
- Remes, O., Brayne, C., van der Linde, R., and Lafortune, L. (2016). A systematic review of reviews on the prevalence of anxiety disorders in adult populations. *Brain Behav.* 6:e00497. doi: 10.1002/brb3.497
- Revicki, D. A., Travers, K., Wyrwich, K. W., Svedsater, H., Locklear, J., Mattera, M. S., et al. (2012). Humanistic and economic burden of generalized anxiety disorder in North America and Europe. *J. Affect. Disord.* 140, 103–112. doi: 10.1016/j.jad.2011.11.014
- Roy-Byrne, P. (2015). Treatment-refractory anxiety; definition, risk factors, and treatment challenges. *Dialogues Clin. Neurosci.* 17, 191–206. doi: 10.1002/jclp.1035
- Sachs, G., Anderer, P., Dantendorfer, K., and Saletu, B. (2004). EEG mapping in patients with social phobia. *Psychiatry Res. Neuroimaging* 131, 237–247. doi: 10.1016/j.pscychresns.2003.08.007
- Schwartz, G. G., Olsson, A. G., Ezekowitz, M. D., Ganz, P., Oliver, M. F., Waters, D., et al. (2003). Effects of atorvastatin on early recurrent ischemic events in acute coronary syndromes: the MIRACL study: a randomized controlled trial. *JAMA* 285, 1711–1718. doi: 10.1016/j.amjcard.2003.06.009
- Sever, S. P., Dahlöf, B., Poulter, N. R., Wedel, H., Beevers, G., Caulfield, M., et al. (2003). Prevention of coronary and stroke events with atorvastatin in hypertensive patients who have average or lower-than-average cholesterol concentrations, in the Anglo-Scandinavian Cardiac Outcomes Trial-Lipid Lowering Arm (ASCOT-LLA): a multicentre randomised controlled trial. *Lancet* 361, 1149–1158. doi: 10.1016/S0140-6736(03)12948-0
- Taylor, S., Abramowitz, J. S., and McKay, D. (2012). Non-adherence and non-response in the treatment of anxiety disorders. *J. Anxiety Disord.* 26, 583–589. doi: 10.1016/j.janxdis.2012.02.010
- Van der Kolk, B. A., Hodgdon, H., Gapen, M., Musicaro, R., Suvak, M. K., Hamlin, E., et al. (2016). A randomized controlled study of neurofeedback for chronic PTSD. *PLoS One* 11:e0166752. doi: 10.1371/journal.pone.0166752
- Vernon, D., Egner, T., Cooper, N., Compton, T., Neilands, C., Sheri, A., et al. (2003). The effect of training distinct neurofeedback protocols on aspects of cognitive performance. *Int. J. Psychophysiol.* 47, 75–85. doi: 10.1016/S0167-8760(02)00091-0
- Wang, F., Wang, C., Yin, Q., Wang, K., Li, D., Mao, M., et al. (2015). Reappraisal writing relieves social anxiety and may be accompanied by changes in frontal alpha asymmetry. *Front. Psychol.* 6:1604. doi: 10.3389/fpsyg.2015.01604
- Zafeiri, E., Kandylaki, A., Zyga, S., Zargiannis, I., and Panoutsopoulos, G. I. (2019). The contribution of biofeedback brain boy method to the treatment of anxiety disorders. *Materia. Socio-Medica* 31, 105–109. doi: 10.5455/msm.2019.31.105-109

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships 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.

Copyright © 2021 Chen, Xiao, Belkacem, Lu, Wang, Yi, Li, Wang, Sha, Zhao and Ming. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



# Hierarchical Spatiotemporal Electroencephalogram Feature Learning and Emotion Recognition With Attention-Based Antagonism Neural Network

Pengwei Zhang, Chongdan Min, Kangjia Zhang, Wen Xue and Jingxia Chen\*

School of Electronic Information and Artificial Intelligence, Shaanxi University of Science and Technology, Xi'an, China

## OPEN ACCESS

### Edited by:

Jane Zhen Liang,  
Shenzhen University, China

### Reviewed by:

Ke Liu,  
Chongqing University of Posts  
and Telecommunications, China  
Ying Yang,  
Facebook, United States

### \*Correspondence:

Jingxia Chen  
chenjx\_sust@foxmail.com

### Specialty section:

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

**Received:** 08 July 2021

**Accepted:** 29 October 2021

**Published:** 02 December 2021

### Citation:

Zhang P, Min C, Zhang K, Xue W  
and Chen J (2021) Hierarchical  
Spatiotemporal  
Electroencephalogram Feature  
Learning and Emotion Recognition  
With Attention-Based Antagonism  
Neural Network.  
Front. Neurosci. 15:738167.  
doi: 10.3389/fnins.2021.738167

Inspired by the neuroscience research results that the human brain can produce dynamic responses to different emotions, a new electroencephalogram (EEG)-based human emotion classification model was proposed, named R2G-ST-BiLSTM, which uses a hierarchical neural network model to learn more discriminative spatiotemporal EEG features from local to global brain regions. First, the bidirectional long- and short-term memory (BiLSTM) network is used to obtain the internal spatial relationship of EEG signals on different channels within and between regions of the brain. Considering the different effects of various cerebral regions on emotions, the regional attention mechanism is introduced in the R2G-ST-BiLSTM model to determine the weight of different brain regions, which could enhance or weaken the contribution of each brain area to emotion recognition. Then a hierarchical BiLSTM network is again used to learn the spatiotemporal EEG features from regional to global brain areas, which are then input into an emotion classifier. Especially, we introduce a domain discriminator to work together with the classifier to reduce the domain offset between the training and testing data. Finally, we make experiments on the EEG data of the DEAP and SEED datasets to test and compare the performance of the models. It is proven that our method achieves higher accuracy than those of the state-of-the-art methods. Our method provides a good way to develop affective brain-computer interface applications.

**Keywords:** EEG, emotion recognition, spatiotemporal features, attention, antagonism neural network, BiLSTM

## INTRODUCTION

Emotion plays an important role in human life (Picard and Picard, 1997). Positive emotions may help improve the efficiency of our daily work, while negative emotions may affect our decision making, attention, and even health (Picard and Picard, 1997). Although it is easier for us to recognize emotions of other people from their facial expression or voice, it is still difficult for machines to do that (Li et al., 2019). In the past few years, emotion recognition by computer has

attracted more and more researchers, and it has become a hot research topic in the field of affective computing and pattern recognition (Purnamasari et al., 2017). The emotion recognition methods can be based on speech signals, facial expression images, and physiological signals (Chen et al., 2019a). In recent years, EEG-based emotion recognition algorithms have been increasingly focused on by researchers.

While researching emotion recognition with EEG, we usually face two difficulties. One is how to obtain a discriminative feature representation method from original EEG signals, and the other is how to build an effective model to better improve the performance of emotion classification. Technically, EEG features can be extracted from the time domain, frequency domain, and time–frequency domain (Jenke et al., 2014). For example, Zhang and Lee (2010) regarded the amplitude difference of symmetric electrodes in the time domain as the EEG feature of emotion recognition. Lin et al. (2010) studied the relationship between emotional state and brain activity, and extracted power spectral density, differential asymmetric power, and reasonable asymmetric power separately as features of EEG signals. Duan et al. (2013) extracted features by calculating the correlation coefficient between the features of each frequency band and their emotional labels. In the aspect of models, Garcia-Martinez et al. (2019) summarized the research results of applying nonlinear methods to EEG signal analysis in recent years. Li et al. (2018b) proposed a graph-regularized sparse linear regression model to make emotion classification and achieved better recognition performance. Zheng and Lu (2015) studied the key frequency bands and key brain regions of EEG signals, and proposed to use group sparse canonical correlation analysis algorithm (Zheng, 2017) for multichannel EEG-based emotion recognition.

With the development of artificial intelligence, deep learning has become very popular, and emotion classification based on deep learning has also continuously improved the performance of emotion recognition and, thus, has gradually become the dominant method. Alhagry et al. (2017) proposed an end-to-end LSTM-RNN network to learn the time dependence of EEG signals. Li et al. (2018a) considered the area shift of EEG data and used deep neural network to learn the difference between left and right hemispheres to narrow the distribution shift. Song et al. (2018) established a graph relationship based on multichannel EEG data, adjacency matrix to build the internal relationship between different EEG channels, and then used dynamical graph convolution network to extract features for emotion classification. Salama et al. (2018) used a three-dimensional convolutional neural network (3D-CNN) to recognize emotions from multichannel EEG data. The author of this paper has also proposed a deep CNN model (Chen et al., 2019c) to learn high-level discriminative feature representations from the combined features of the EEG signal in the time-frequency domain. In Chen et al. (2019b), a hierarchical bidirectional LSTM model based on attention mechanism was proposed to reduce the influence of long-term instability of EEG sequences on emotion recognition.

Although many EEG emotion recognition methods have emerged recently, there are still some problems that needs to be further studied. One of the problems is how to obtain effective high-level features from the original EEG signals automatically.

Most researchers often extract some time or frequency statistical EEG features manually combined with classic machine learning algorithms to make emotion classification. However, feature engineering needs to consume a lot of computation resources and time. It is expected to automatically learn more prominent spatiotemporal features with less feature engineering. The second question is which brain area contributes more to human emotion recognition, and how to use the distribution information of different brain areas to improve recognition performance. The latest researches (Etkin et al., 2011; Lindquist and Barrett, 2012) have shown that human emotions are closely related to multiple areas of the cerebral cortex, such as the orbitofrontal cortex, ventromedial prefrontal cortex, amygdala, and so on. Therefore, the contribution of EEG signals associated with each brain area is different. If the spatial information of different brain regions can be used, it is expected to provide help in understanding human emotions (Heller and Nitscke, 1997; Davidson, 2000; Lindquist et al., 2012). The third question is how to enhance the emotion recognition performance by using time series information in each brain area, as EEG signals are dynamic time series carrying important emotion dynamics, which is effective to identify human emotions.

Literature (Lin et al., 2010; Zhang and Lee, 2010; Duan et al., 2013; Zheng and Lu, 2015; Zheng, 2017; Li et al., 2018b; Garcia-Martinez et al., 2019) has proven that EEG signals in different brain regions have different contributions to emotion recognition. Literature (Alhagry et al., 2017; Li et al., 2018a; Salama et al., 2018; Song et al., 2018; Chen et al., 2019b) found that either deep CNN model or the bidirectional long- and short-term memory (BiLSTM) model combined with attention mechanism could hierarchically extract deep temporal and spatial context of EEG signals. Inspired by these two aspects and neuroscience research basis (Heller and Nitscke, 1997; Davidson, 2000; Etkin et al., 2011; Lindquist and Barrett, 2012; Lindquist et al., 2012), this paper proposes a new emotion computing model called R2G-ST-BiLSTM, which is used to solve the above three main problems. Its core idea is to extract the EEG spatial temporal dynamics associated with human emotions from local and global brain areas. Specifically, the R2G-ST-BiLSTM model contains two two-layer neural networks, in space and time domain, respectively, and features are learned hierarchically from region to global (R2G) to grasp more discriminative spatiotemporal EEG features related to human emotions. The proposed R2G-ST-BiLSTM model consists of three parts:

- (1) Feature learning module. It uses the bidirectional long- and short-term memory (BiLSTM) network to learn the hierarchical spatiotemporal EEG characteristics within and between each brain region. In order to better judge the effect of different brain regions on emotion recognition, this paper introduces the regional attention mechanism to learn a set of weights, which represent the contributions of different brain regions.
- (2) Emotion classifier. The purpose of this module is to predict emotion category based on EEG spatiotemporal features obtained by feature learning module. At the same time,

it also guides the whole neural network to learn more discriminative EEG features for emotion classification.

- (3) Domain discriminator. This module aims to decrease the domain offset between the training EEG data and the testing EEG data through introducing a discriminator, so that the hierarchical feature learning module can produce EEG features with more emotional discrimination and stronger domain adaptability.

Through collaborative work of the above three modules, the R2G-ST-BiLSTM model can learn EEG features with better discrimination ability and domain robustness simultaneously, thus, further improving human emotion recognition performance. Overall, there are three main contributions in our work:

- Inspired by neuroscience, we propose a new hierarchical spatiotemporal EEG feature learning model, which obtains spatiotemporal emotional information from EEG data within and between each cerebral region.
- Proposes an attention weighted model to estimate the contribution of each cerebral region to the different affections of humans. The influence of the most dependent cerebral region is enhanced by the learned weight, and the impact of the less dependent region was reduced as well.
- Proposes a domain discriminator to work on antagonism with the classifier to improve the adaptability of the R2G-ST-BiLSTM model.

## METHOD BASED ON THE R2G-ST-BiLSTM MODEL

Traditional one-way LSTM network (Hochreiter and Schmidhuber, 1997) has a special structure that is different from the simple recurrent neural network (RNN) (Graves et al., 2013) and is more capable of dealing with the frequent dependence of the sample sequence. Its special “gate” structures enable LSTM to retain significant data information and forget unnecessary redundant information (Yan et al., 2017). However, one disadvantage of this network is that it only uses the context-related information that happened before. The BiLSTM network can process data by using separate hidden layers in two directions (Bottou, 2010). Because the BiLSTM network can obtain long-term contextual information in both forward and backward orientation, it is better than the traditional one-way LSTM network for modeling time series. Because EEG data related to each channel in each brain region are in time series with the same dimension, therefore, BiLSTM can be used to extract the deep spatiotemporal context features of EEG data from the local brain regions to the global brain.

In this section, we will introduce the framework of the R2G-ST-BiLSTM model in detail and explain the specific application of EEG signals for emotion recognition methods and procedures. **Figure 1** shows the framework of the R2G-ST-BiLSTM model. It consists of three main modules, which are feature extractors, classifiers, and discriminators.

## Spatial Feature Extraction

First, we divide the EEG sequence into several equal-length segments. Then a set of manual features is extracted from the EEG segments corresponding to each electrode. For example, the differential entropy feature (DE feature) is extracted from  $\delta(1\sim4\text{ Hz})$ ,  $\theta(5\sim8\text{ Hz})$ ,  $\alpha(9\sim14\text{ Hz})$ ,  $\beta(15\sim30\text{ Hz})$ , and  $\gamma(31\sim50\text{ Hz})$  (Zheng and Lu, 2015). In addition, to capture dynamic time information from input EEG sequence, every five adjacent EEG segments make up one EEG sample, and each EEG sample is represented by a tensor of its manual feature.

Let  $S = [s_1, s_2, \dots, s_{T-1}, s_T] \in \mathbb{R}^{d \times n \times T}$  represent an EEG sample, where  $s_i$  represents the feature data extracted from the divided  $i$ -th segment of EEG, shown in the bottom blue rectangle of **Figure 1**,  $d$  is the number of EEG features per channel,  $n$  is the number of channels, and  $T$  is the number of segments per EEG sample. **Figure 1** shows that when extracting spatial features, each sample includes a regional feature extraction layer and a global feature extraction layer to gradually learn high-level semantic features from local to global.

**Figure 2** shows the specific feature learning process of the EEG data  $s_i$ . At first, the channels of  $s_i$  are grouped into different areas according to the spatial position of the brain electrodes. The number of electrodes in each brain area varies due to the different functions of each brain area, thereby generating a set of regional manual feature vectors in each brain region. Then these manual feature vectors are input into the equal number of BiLSTM networks to learn the local abstract features of each region. After learning the regional deep features, the region attention mechanism is introduced to learn a set of weights that represent the significance of each region. Finally, at the top of **Figure 1**, the extracted weighted feature of each region is input into another set of BiLSTM networks to further extract the global emotional semantic features.

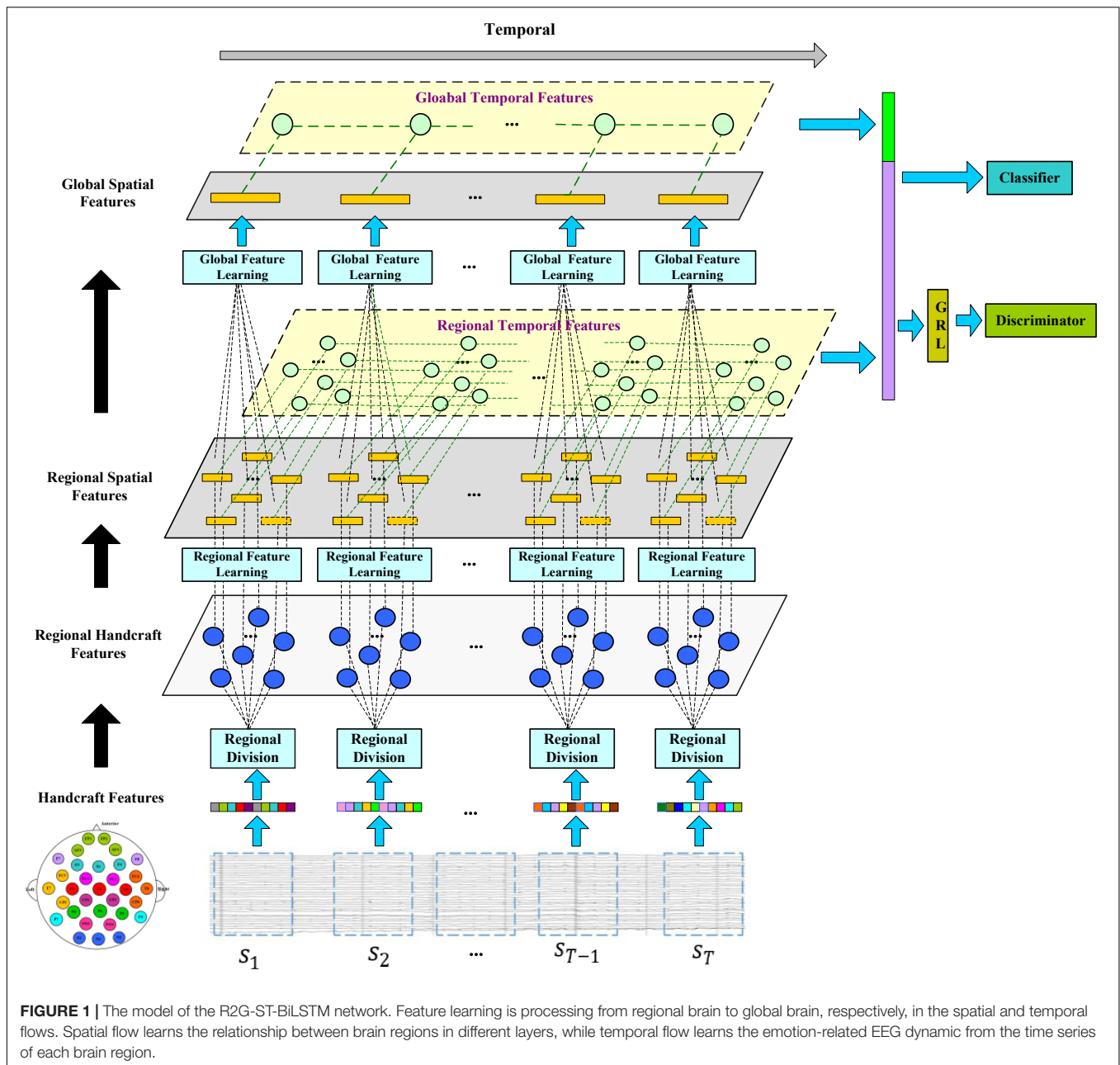
(1) *Regional feature extraction layer.* Let  $x_{ij}$  represent the manual feature vector of the  $j$ -th EEG channel, so  $s_i = [x_{i1}, x_{i2}, \dots, x_{in}] \in \mathbb{R}^{d \times n}$ . Then according to the related electrodes,  $n$  channels of  $s_i$  are divided into different groups: each group of channels belongs to a cerebral area, and each area is expressed as: brain area 1:  $R_i^1 = [x_{i1}^1, x_{i2}^1, \dots, x_{in_1}^1]$ , brain area 2:  $R_i^2 = [x_{i1}^2, x_{i2}^2, \dots, x_{in_2}^2]$ , brain area  $n$ :  $R_i^N = [x_{i1}^N, x_{i2}^N, \dots, x_{in_N}^N]$ , where  $N$  is the quantity of cerebral areas,  $n_j$  is the quantity of the  $j$ -th cerebral area of the channels, and  $n_1 + \dots + n_N = n$ . Furthermore, we adjust the column order of  $s_i$ , which is represented as a new matrix  $\hat{s}_i = [R_i^1, \dots, R_i^N]$ . The submatrix  $R_i^j (j = 1, \dots, N)$  represents cerebral area  $j$ , and per column of  $R_i^j$  corresponds to an EEG channel of this area. The spatial relationship of the brain area can be modeled by a BiLSTM working on the  $R_i^j$  matrix to extract the advanced features of each region, which process is expressed as:

$$\mathcal{F}(R_i^1) = [\tilde{h}_{i1}^1, \tilde{h}_{i2}^1, \dots, \tilde{h}_{in_1}^1] \in \mathbb{R}^{2d_r \times n_1}, \quad (1)$$

$$\mathcal{F}(R_i^N) = [\tilde{h}_{i1}^N, \tilde{h}_{i2}^N, \dots, \tilde{h}_{in_N}^N] \in \mathbb{R}^{2d_r \times n_N}, \quad (2)$$

where  $\mathcal{F}(\cdot)$  represents the BiLSTM operation,  $\tilde{h}_{ik}^j \in \mathbb{R}^{2d_r}$  represents hidden vectors output by the  $k$ th forward and backward hidden





units of the BiLSTM, and  $d_r$  represents the dimensions of the output state vectors of each hidden unit in the BiLSTM. At last, the state vector outputs by the last hidden unit of each BiLSTM are connected as the local deep features of all regions, which are expressed as follows:

$$\tilde{H}_i^r = [\tilde{h}_{in_1}^1, \tilde{h}_{in_2}^1, \dots, \tilde{h}_{in_N}^N] \in \mathbb{R}^{2d_r \times N}. \quad (3)$$

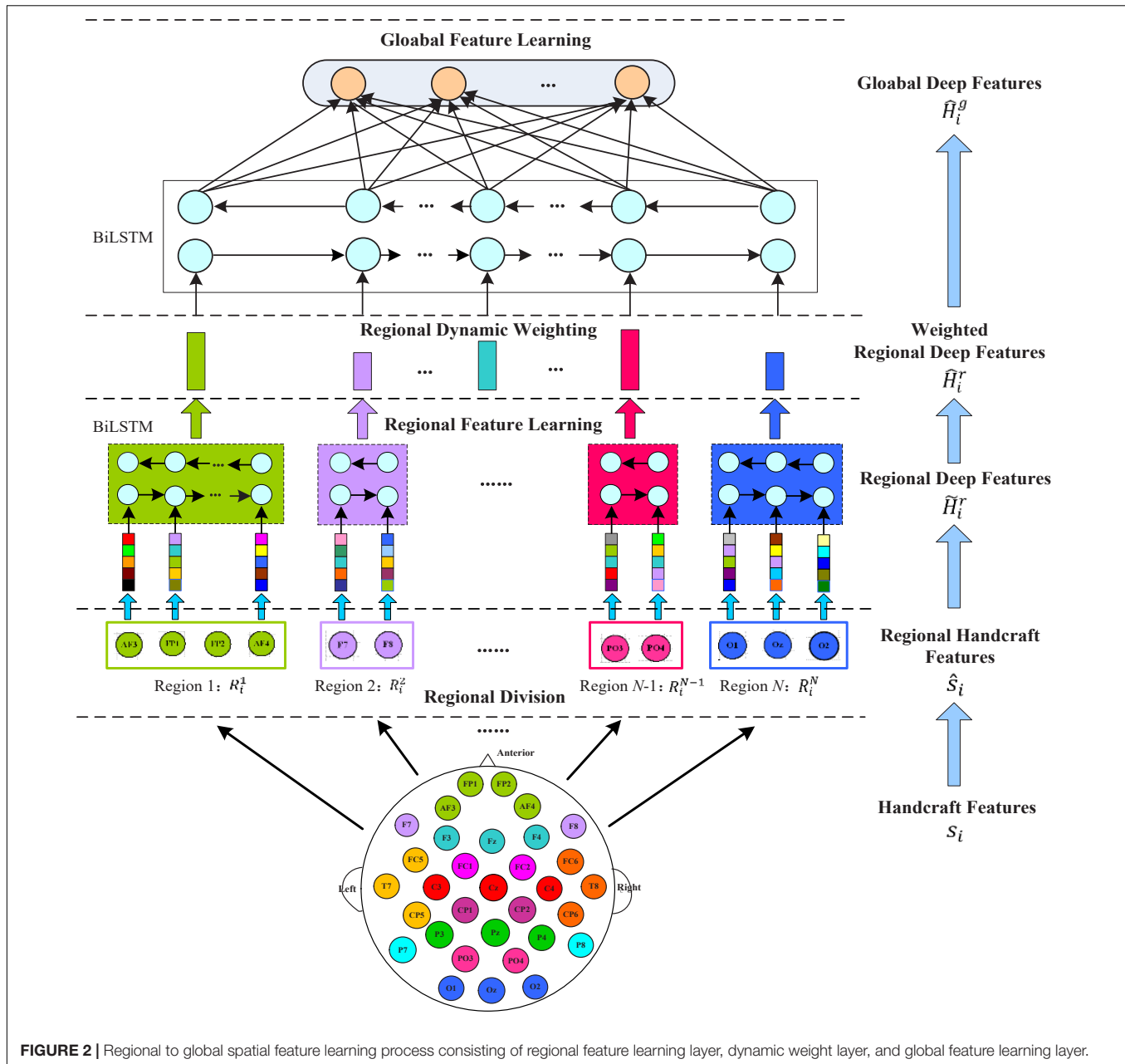
For simplicity, each BiLSTM model in this part is initialized and fit jointly, and the hyperparameters are shared with each other.

(2) *Attention-based brain region weighting layer.* Neuroscience-related research shows that different brain

areas respond to different types of emotions. Therefore, EEG signals from diverse brain areas have different contributions to emotion classification. To emphasize the role of the different brain area electrodes in EEG emotion recognition, we introduce a weighting layer based on attention mechanism. Expressed by  $W = \{w_{ij}\}$ , it can characterize the significance of channels in different areas. After that, the local deep features of all areas are expressed by  $\hat{H}_i^r$  as follows:

$$\hat{H}_i^r = \tilde{H}_i^r W, \quad (4)$$

$$W = (U \tanh(V\tilde{H}_i^r + b^r e^T))^T, \quad (5)$$



$$w_{ij} = \frac{\exp(w_{ij})}{\sum_{k=1}^N \exp(w_{ik})}, \quad (6)$$

where  $U$  and  $V$  are learnable transpositional matrices,  $b^r$  represents the deviation, and  $e$  represents an  $N$ -dimensional vector whose elements are 1, that is,  $e = [1, 1, \dots, 1]^T$ . The matrix  $W$  is normalized across the columns, so that its values are limited to non-negative by formula (6). The larger the  $w_{ij}$  value obtained, the more important the  $j$ -th brain area is for emotion recognition.

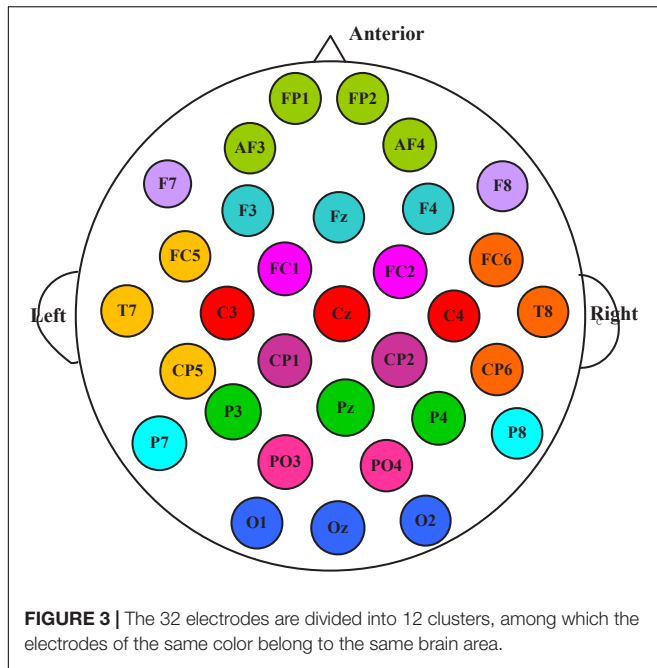
(3) *Global feature extraction layer.* To further capture the potential global structural information on the basis of the learned local deep feature  $\hat{H}_i^r$ , we use another BiLSTM network with  $N$

hidden units to extract global spatial features.

$$\mathcal{F}(\hat{H}_i^r) = [\tilde{h}_{i1}^g, \tilde{h}_{i2}^g, \dots, \tilde{h}_{iN}^g] \in \mathbb{R}^{2d_g \times N}, \quad (7)$$

where,  $\tilde{h}_{ik}^g$  represents the hidden vector output by the  $k$ -th forward and backward hidden unit of the BiLSTM network, and  $d_g$  is the dimension of the output state vector of each hidden unit. Next, input the vector sequence  $\tilde{h}_{i1}^g, \dots, \tilde{h}_{iN}^g$  into a fully connected layer to learn a new compressed feature vector with the following formulas:

$$\hat{h}_{jK}^g = \sigma \left( \sum_{j=1}^N P_{jk}^g \tilde{h}_{ij}^g + b^g \right), \quad k = 1, 2, \dots, K, \quad (8)$$



$$\hat{H}_i^g = [\hat{h}_{i1}^g, \hat{h}_{i2}^g, \dots, \hat{h}_{iK}^g], \quad (9)$$

where  $P^g = [P_{jk}^g]_{N \times K}$  denotes a projection matrix,  $b^g$  denotes deviation,  $K$  denotes the length of the compressed sequence, and  $\sigma(\cdot)$  is a nonlinear function. Thus, the global deep feature  $\hat{H}_i^g$  related to the manual feature matrix  $s_i$  of the  $i$ -th EEG segment is finally obtained.

## Temporal Feature Extraction

Let  $\tilde{h}_i^j(j = 1, \dots, N)$  represent the state vector output by the  $j$ -th brain region of the  $i$ -th EEG manual feature matrix  $s_i$  through the last hidden unit of the BiLSTM network, then the time series of each brain region feature can be expressed as:

$$\tilde{H}^{r1} \triangleq [\tilde{h}_1^1, \tilde{h}_2^1, \dots, \tilde{h}_T^1], \quad (10)$$

$$\tilde{H}^{rN} \triangleq [\tilde{h}_1^N, \tilde{h}_2^N, \dots, \tilde{h}_T^N] \quad (11)$$

In this way, the columns of the feature matrix  $\tilde{H}^{rj}(j = 1, \dots, N)$  constitute the time series of the feature vectors related to the  $j$ -th brain region. Therefore, a BiLSTM network can be applied again to learn the temporal context between these eigenvector sequence:

$$\begin{aligned} Z^{rt} &= [\mathcal{F}(\tilde{H}^{r1}), \dots, \mathcal{F}(\tilde{H}^{rN})] = [(z_{11}^{rt}, \dots, z_{1T}^{rt}), \dots, (z_{N1}^{rt}, \dots, z_{NT}^{rt})] \\ &= [(z_1^{rt}, \dots, z_N^{rt})], \end{aligned} \quad (12)$$

where  $Z_j^{rt} = [z_{j1}^{rt}, \dots, z_{jT}^{rt}] \in \mathbb{R}^{2d_{rt} \times T}$  represents the regional temporal feature matrix related to the  $j$ -th brain region, and  $d_{rt}$  is the dimension of each hidden unit state vector in the regional

temporal BiLSTM network. Take the output  $z_{jT}^{rt}$  of the last hidden unit of the BiLSTM network in each brain area as the learned temporal feature of this brain area, and then get the final temporal feature  $z^{rt}$  of all brain areas, which is expressed as:

$$z^{rt} = [(z_{1T}^{rt})^T, (z_{2T}^{rt})^T, \dots, (z_{NT}^{rt})^T]^T. \quad (13)$$

In addition, to explore the time context on the basis of matrix  $\hat{H}_i^g$ , we convert the columns of  $\hat{H}_i^g$  to a new sequence, which is represented by  $\hat{h}_i^g$ :

$$\hat{h}_i^g = [(\hat{h}_{i1}^g)^T, (\hat{h}_{i2}^g)^T, \dots, (\hat{h}_{iK}^g)^T]^T. \quad (14)$$

Set up  $Z^g = [\hat{h}_1^g, \dots, \hat{h}_T^g]$ . Then a BiLSTM network with  $T$  hidden units is used to learn the global temporal feature  $Z^{gt}$ :

$$Z^{gt} = \mathcal{F}(Z^g) = [z_1^{gt}, \dots, z_T^{gt}] \in \mathbb{R}^{2d_{gt} \times T}, \quad (15)$$

where  $d_{gt}$  denotes the size of the hidden state vector of the global temporal BiLSTM network, and the output  $z_T^{gt}$  of the last hidden unit is taken as the learned global temporal feature. Finally, by concatenating  $z^{rt}$  with  $z_T^{gt}$ , the optimal feature vector  $z^{rg}$  of the EEG sample  $S$  (composed of  $T$  EEG fragments) is obtained, which contains complex temporal context information, and its expression is:

$$z^{rg} = [(z_{1T}^{rt})^T, (z_{2T}^{rt})^T, \dots, (z_{NT}^{rt})^T, (z_T^{gt})^T]^T. \quad (16)$$

## Classifier and Discriminator

For the final eigenvector  $z^{rg}$  input to this layer, a simple linear transformation method can be used to recognize the human emotional type of the input EEG data  $S$  as the following formula:

$$Y = Qz^{rg} + b_c = [y_1, y_2, \dots, y_c], \quad (17)$$

where  $Q$  and  $b_c$ , respectively, denotes the projection matrix and deviation.  $C$  is number of emotional categories. The element of the transformation result  $Y$  is input into a softmax function to predict the emotion category:

$$P(c | S) = \max \left\{ \frac{\exp(y_k)}{\sum_{i=1}^C \exp(y_i)} | k = 1, \dots, C \right\}, \quad (18)$$

where  $P(c | X)$  represents the probability the input EEG data  $S$  is predicted to be the emotion of type  $c$ .

Supposing the training set of the model is composed of  $M$  EEG data, which is expressed by matrix  $S_i^S(i = 1, \dots, M)$ . The loss function of the emotion classifier can be expressed as:

$$\mathcal{L}_c(S_1^S, \dots, S_M^S; \theta_f, \theta_c) = \sum_{i=1}^M \sum_{c=1}^C -\varphi(l_i, c) \times \log P(c | S_i^S) \quad (19)$$

where  $l_i$  represents the real label of the  $S_i^S$  sample, and  $\theta_f$  and  $\theta_c$  represent the learning parameters.  $\varphi(l_i, c)$  is expressed as:

$$\varphi(l_i, c) = \begin{cases} 1, & \text{if } l_i = c, \\ 0, & \text{otherwise.} \end{cases} \quad (20)$$

**TABLE 1** | Electroencephalogram electrodes and data size associated with each brain area.

Brain region	DEAP dataset		SEED dataset	
	Electrode name	Data size (d × n <sub>f</sub> )	Electrode name	Data size (d × n <sub>f</sub> )
Pre-frontal	FP1, FP2, AF3, AF4	4 × 4	AF3, FP1, FPZ, FP2, AF4	4 × 5
Frontal	F3, FZ, F4	4 × 3	F3, F1, FZ, F2, F4	4 × 5
Bilateral frontal	F7, F8	4 × 2	F7, F5, F6, F8	4 × 4
Left temporal	FC5, T7, CP5	4 × 3	FT7, FC5, T7, C5, TP7, CP5	4 × 6
Right temporal	FC6, T8, CP6	4 × 3	FT8, FC6, T8, C6, TP8, CP6	4 × 6
Frontal central	FC1, FC2	4 × 2	FC3, FC1, FCZ, FC2, FC4	4 × 5
Central	C3, CZ, C4	4 × 3	C3, C1, CZ, C2, C4	4 × 5
Central parietal	CP1, CP2	4 × 2	CP3, CP1, CPZ, CP2, CP4	4 × 5
Bilateral parietal	P7, P8	4 × 2	P7, P5, P6, P8	4 × 4
Parietal	P3, PZ, P4	4 × 3	P3, P1, PZ, P2, P4	4 × 5
Parietal occipital	PO3, PO4	4 × 2	PO5, PO3, POZ, PO4, PO6	4 × 5
Occipital	O1, OZ, O2	4 × 3	CB1, O1, OZ, O2, CB2	4 × 5

From formulas (19) and (20), it can be concluded that by minimizing the loss function  $\mathcal{L}_c(S_1^S, S_2^S, \dots, S_M^S; \theta_f, \theta_c)$ , the emotion category of each training sample can be correctly predicted to the maximum extent.

Let  $S_{test}$  represent a test sample, and the emotion label of  $S_{test}$  is determined by the formula:

$$l_{test} = \underset{c}{\operatorname{argmax}} \{P(c|S_{test}) | c = 1, \dots, C\}, \quad (21)$$

where  $l_{test}$  represents the predicted label of the test sample  $S_{test}$ .

When performing prediction, the EEG samples for the training and testing data may be from various subjects and even different experiments. Based on this, the recognition model learned by using the training data may not have a high recognition accuracy for the test data. To optimize the generalization ability of the model, a discriminator is introduced to work collaboratively with the classifier to learn features with strong emotion discrimination and domain invariance.

Specifically, suppose that  $D^S = \{S_1^S, \dots, S_{M_1}^S\}$  denotes the dataset of the source domain, and  $D^T = \{S_1^T, \dots, S_{M_2}^T\}$  denotes the dataset of the target domain, where  $M_1$  and  $M_2$  are their sample number. To alleviate the domain difference, the loss function of the discriminator is defined as:

$$\mathcal{L}_d(S_i^S, S_j^T; \theta_f, \theta_d) = - \sum_{i=1}^{M_1} \log P(0 | S_i^S) - \sum_{j=1}^{M_2} \log P(1 | S_j^T). \quad (22)$$

Here,  $P(0 | S_i^S)$  is the probability that EEG sample  $S_i^S$  is classified into the source domain,  $P(1 | S_j^T)$  is the probability that EEG sample  $S_j^T$  is classified into the target domain, and  $\theta_d$  is the parameter. The discriminator enables this model to learn the domain-invariant features gradually.

## Optimization of the Bidirectional Long- and Short-Term Memory Neural Network From Region to Global Brain Model

The previous description indicates that through minimizing formula (19) and maximizing formula (22), domain difference

can be reduced and better domain invariant characteristics can be learned. Therefore, we redefine the total loss function of R2G-ST-BiLSTM model as:

$$\mathcal{L}(S^S, S^T | \theta_f, \theta_c, \theta_d) = \mathcal{L}_c(S^S; \theta_f, \theta_c) - \mathcal{L}_d(S^S, S^T; \theta_f, \theta_d). \quad (23)$$

To optimize our model, we need to find the best parameters that minimize the new loss function  $\mathcal{L}(S^S, S^T | \theta_f, \theta_c, \theta_d)$ . By minimizing  $\mathcal{L}_c(S^S; \theta_f, \theta_c)$  and maximizing  $\mathcal{L}_d(S^S, S^T; \theta_f, \theta_d)$  synchronously and iteratively, the optimal parameters of  $\mathcal{L}(S^S, S^T | \theta_f, \theta_c, \theta_d)$  can be obtained. Specifically, the stochastic gradient descent (SGD) algorithm (Yu et al., 2015) is used to find the optimal model parameters:

$$(\hat{\theta}_f, \hat{\theta}_c) = \underset{\theta_f, \theta_c}{\operatorname{argmin}} \mathcal{L}_c(S^S, (\theta_f, \theta_c), \hat{\theta}_d), \quad (24)$$

$$\hat{\theta}_d = \underset{\theta_d}{\operatorname{argmax}} \mathcal{L}_d(S^S, S^T(\hat{\theta}_f, \hat{\theta}_c), \theta_d), \quad (25)$$

The feature extractor can learn to obtain emotional discriminative features by minimizing the loss function  $\mathcal{L}_c$ . Meanwhile, it extracts domain invariant features by maximizing the loss function  $\mathcal{L}_d$ . When obtaining the optimal parameters of the R2G-ST-BiLSTM model, we also introduced a gradient reverse layer (GRL) (Ganin et al., 2016), which performs gradient sign reversal when performing backward propagation operation and enables the discriminator to transform the maximization problem into a minimization problem, so that SGD can be used for parameter optimization. The parameter updating can be expressed as:

$$\theta_d \leftarrow \theta_d - \alpha \frac{\partial \mathcal{L}_d}{\partial \theta_d}, \theta_f \leftarrow \theta_f + \alpha \frac{\partial \mathcal{L}_d}{\partial \theta_f}, \quad (26)$$

where  $\alpha$  is the learning rate.



**TABLE 2 |** Implementation details of 10 benchmark models.

Benchmark models	Input data size		Implementation details
	DEAP dataset	SEED dataset	
Support vector machine (SVM) (Suykens and Vandewalle, 1999)	[32 × 5, sample_size]	[Features, sample_size]	Kernel = 'rbf', gamma = 8, c = 0.05
Bagging tree (BT) (Chuang et al., 2012)	[32 × 5, sample_size]	[64 × 5, sample_size]	Method: bag, nLearn:100, weak learner: tree, type: classification
Random forest (RF) (Breiman, 2001)	[32 × 5, sample_size]	[64 × 5, sample_size]	n_estimators = 50, max_depth = 10, max_features = 8, min_samples_split = 20, min_samples_leaf = 10, oob_score = true, random_sate = 10
Deep confidence network (DBN) (Zheng and Lu, 2015)	[Batch_size, feature_size]: [60, 32 × 5]	[Batch_size, feature_size]: [60, 64 × 5]	hidden_layers = 3, hidden_size = 64, batch_size = 60, learning_rate = 0.04, dropout = 0.5, epochs = 200
Long- and short-term memory (LSTM) (Alhagry et al., 2017)	[Batch_size, seq_len, channels]: [120, 5, 32]	[Batch_size, seq_len, channels]: [120, 5, 64]	hidden_layers = 2, hidden_size = 64, seq_len = 5, batch_size = 120, learning_rate = 0.03, dropout = 0.5, num_directions = 2, epochs = 100 ×
Two-dimensional convolutional neural network (2D-CNN) (Chen et al., 2019c)	[Batch_size, seq_len × band_size, channels]: [60, 5 × 4, 32]	[Batch_size, seq_len × band_size, channels]: [60, 5 × 4, 64]	conv_layers = 2, max_pool_layers = 2, full_conn_layers = 2 (hidden_size = 128), conv_kernels = [32, 64], kernel_size = [5 × 5, 3 × 3], pool_size = (2,2), batch_size = 60, learning_rate = 0.05, dropout = 0.7, epochs = 300, padding = 0, stride = 1
Three-dimensional convolutional neural network (3D-CNN) (Salama et al., 2018)	[Batch_size, band_size, seq_len, channels]: [80, 4, 5, 32]	[Batch_size, band_size, seq_len, channels]: [80, 4, 5, 64]	conv_layers = 2, max_pool_layers = 1, full_conn_layers = 1 (hidden_size = 128), conv_kernels = [8, 16], kernel_size = [3 × 3 × 7, 2 × 2 × 5], pool_size = (2,2), batch_size = 80, learning_rate = 0.01, dropout = 0.6, epochs = 200, padding = 0, stride = 1
Hierarchical bidirectional GRU network based on attention mechanism (H-ATT-BGRU) (Chen et al., 2019b)	[Batch_size, band_size, seq_len, channels]: [60, 4, 5, 32]	[Batch_size, band_size, seq_len, channels]: [60, 4, 5, 64]	hidden_layers = 2, hidden_size = 64, seq_len = 5, batch_size = 60, learning_rate = 0.06, dropout = 0.4, num_directions = 2, epochs = 400
Domain adaptive neural network (DANN) (Ganin et al., 2016)	[Batch_size, feature_size]: [30, 32 × 5]	[Batch_size, feature_size]: [30, 64 × 5]	hidden_layers = 2, hidden_size = 128, batch_size = 30, L2-weight-regularization = 0.003, learning_rate = 0.02, dropout = 0.5, epochs = 500, momentum = 0.05, MMD regularization constant $\gamma = 10e3$
convolutional recurrent neural network (Casc-CNN-LSTM) (Chen et al., 2020)	[Batch_size, seq_len × band_size, channels]: [80, 5 × 4, 32]	[Batch_size, seq_len × band_size, channels]: [80, 5 × 4, 64]	CNN: conv_layers = 3, max_pool_layers = 3, full_conn_layers = 2 (hidden_size = 256), conv_kernels = [32, 64, 128], kernel_size = [3 × 3, 3 × 3, 3 × 3], pool_size = (2,2), batch_size = 80, learning_rate = 0.05, dropout = 0.5, epochs = 500, padding = 0, stride = 1 LSTM: hidden_layers = 2, hidden_size = 128, seq_len = 256, batch_size = 80, learning_rate = 0.05, dropout = 0.5, num_directions = 2, epochs = 500

## Configuration and Training of Bidirectional Long- and Short-Term Memory Neural Network From Region to Global Brain Model

The proposed model is implemented in TensorFlow framework on a NVIDIA Titan × Pascal GPU-equipped work station and trained from scratch in a fully supervised manner. When training the whole model, we define a search space to find the optimal model parameters. The search space includes the hidden\_layers (one to three layers), hidden\_size (32, 64, 128, and 256), batch\_size (30, 60, 80, and 120), learning\_rate (0.1, 0.01, 0.001, and 0.0001), dropout (0.5, 0.6, and 0.7), and epochs (100, 200, 300, and 500). The search space was defined to balance the trade-off between a deeper architecture and limited training samples. For simplicity, each BiLSTM

model is initialized and fit jointly, their hyperparameters are shared with each other, the hidden\_size of the single-layer perception network used to learn the attention weight of each brain region is 128, the hidden\_size of the full connection layer for learning the compressed global brain feature is 64, all hidden layers use ReLU activation function for faster approximation, all BiLSTM models are trained using SGD with AdaGrad optimizer, and the maximum training iteration was set to be 10,000. For searching each hyperparameter, we only adjust one hyperparameter in a defined search space and fix others each time. When observing that there is no growth trend of the accuracy on training and validation sets, we can judge to stop the training process in advance, as shown in **Figures 4, 6**. Finally, we select the best model that produces the highest accuracy on the validation dataset.

Through this fine-tuning process, the selected best hidden\_layers is 2, the hidden\_size of  $d_r$ ,  $d_g$ ,  $d_{rt}$ , and  $d_{gt}$  is consistently 64, learning rate is 0.001, batch-size is 120, and epochs is 200. All parameters and offsets are initialized with randomly assigned nonzero regularization float. For cross-subject experiment on DEAP dataset, the total number of parameters in the whole model is about 50,156, which is larger than the total number of training samples. To prevent the overfitting of the model, a dropout layer is added after the first full connection layer of each BiLSTM, and the selected optimal dropout is 0.7.

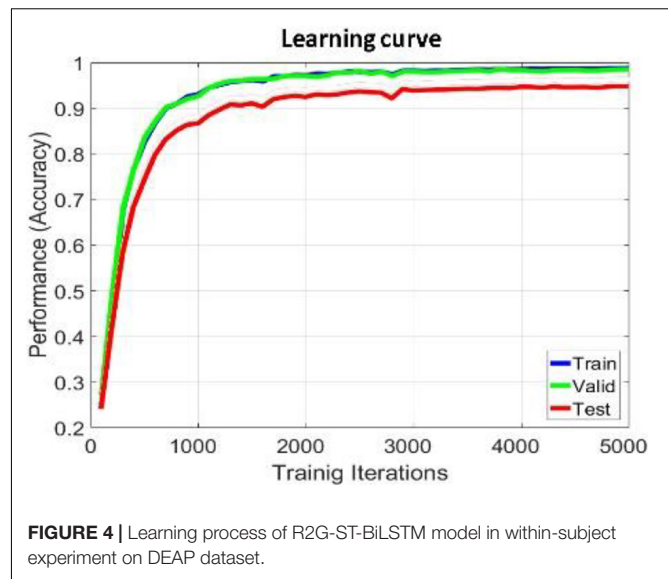
## EXPERIMENTS AND RESULTS

### Dataset and Preprocessing

To evaluate the proposed method, we make extensive experiments on the DEAP (Koelstra et al., 2012) dataset, which come from the Queen Mary University of London and is publicly and freely available for research on emotion recognition. This dataset records EEG, EMG, ECG, and other types of physiological signals induced by 32 subjects watching 40 music videos with different emotional tendencies. The emotion labels are evaluated with 1–9 consecutive values in four emotional dimensions of arousal, valence, preference, and dominance. In our research, we just take the EEG signals of each subject in 32 channels and 60 s from the DEAP dataset for study. The electrodes are positioned according to the 10–20 system. The sampling frequency is reduced to 128 Hz. For other artifacts, a 4- to 45-Hz bandpass filter is used for data filtering, and then blind source separation is used to remove the electro-oculogram (EOG) interference.

According to the spatial distribution of EEG electrodes, 32 electrodes are divided into 12 regions, that is, the number of brain regions  $N$  is 12, and each region contains at least two electrodes. We divided the 32 electrodes into 12 clusters or brain regions, where the electrodes of the same color belong to the same region, as shown in **Figure 3**. The electrodes contained in each brain region and the size of the corresponding manual feature set are listed in **Table 1**. In the DEAP database, there are 32 subjects, and each subject takes a 40-trial EEG data acquisition experiment. Each experiment collects  $60 \times 128 = 7,680$  EEG records and emotional labels induced by watching videos for 60 s.

To balance the samples of three kinds of emotion labels in DEAP, the values 4 and 7 are used as the threshold to distinguish the positive, neutral, and negative emotion labels. As a result, for the total 32 subjects of the DEAP dataset, the number of positive, neutral, and negative trials are 373, 540, and 367, respectively. The proportion of samples in positive, neutral, and negative class is about 29%, 42%, and 29%, respectively. In this way, 40 trials were collected for each subject including 2,400-s EEG records, which is segmented according to 1 s, including 2,400 EEG segments. Each segment corresponds to three types of emotional labels: positive, neutral, and negative, of which there are about 800 segments of each



type of emotional label. In this way, each subject has a total of  $40 \text{ trials} \times 60 \text{ s} = 2,400 \text{ s}$  of EEG records, which were segmented by 1 s and contained a total of 2,400 EEG segments. Each segment corresponds to three types of emotions: positive, neutral, or negative tags. Then all segments are divided into 480 EEG samples according to  $T = 5$ . That means each sample contains five EEG segments, and DE features of four bands are extracted from 32 electrodes of each segment, so that each EEG sample is expressed by a manual feature tensor of  $4 \times 32 \times 5$ , and the size of each EEG dataset is  $480 \times 4 \times 32 \times 5$ . The size of the 32-subject EEG dataset is  $15,360 \times 4 \times 32 \times 5$ .

To further prove the performance of our proposed model and make the conclusion more convincing, we also conducted serial comparison experiments on the SEED dataset (Zheng and Lu, 2017). The dataset collected EEG records related to emotional stimulation from 64 channels of 15 subjects (7 men and 8 women). The emotional labels fed back by the subjects were divided into positive, neutral, and negative. The dataset has been preprocessed, and DE features for each subject were extracted. On the SEED dataset, we also used trial-wise randomization method to construct cross validation sets for within-subject experiments and used the same leave-one-subject-out (LOSO) method as that used on the DEAP dataset to construct cross validation sets for cross-subject experiments. As for brain area division, to facilitate comparison, we removed the PO7 and PO8 electrodes and divided the remaining 62 electrodes into 12 brain areas. **Table 1** shows the detailed brain area division method on the DEAP and SEED datasets.

### Benchmark Methods

For comparison, we use the following benchmark methods to perform within-subject and cross-subject emotion classification experiments on the same dataset.

**TABLE 3 |** The results of within-subject experiment on the DEAP dataset.

Method	SVM (Suykens and Vandewalle, 1999)	BT (Chuang et al., 2012)	RF (Breiman, 2001)	DBN (Zheng and Lu, 2015)	LSTM (Alhagry et al., 2017)	2D-CNN (Chen et al., 2019c)
Average classification accuracy (ACC)(%)/standard deviation (STD)	80.72/7.67	84.65/8.93	78.87/11.32	82.83/9.54	84.51/10.06	85.63/8.72
p-Value	0.0005	0.0004	0.0006	0.0008	0.0023	0.0019
Method	3D-CNN (Salama et al., 2018)	H-ATT-BGRU (Chen et al., 2019b)	DANN (Ganin et al., 2016)	Casc-CNN-LSTM (Chen et al., 2020)	R2G-ST-BiLSTM	
ACC (%) /STD	87.21/10.57	87.89/8.94	88.54/9.26	93.95/7.88	94.69/9.81	
p-Value	0.0052	0.0066	0.0074	0.0089		

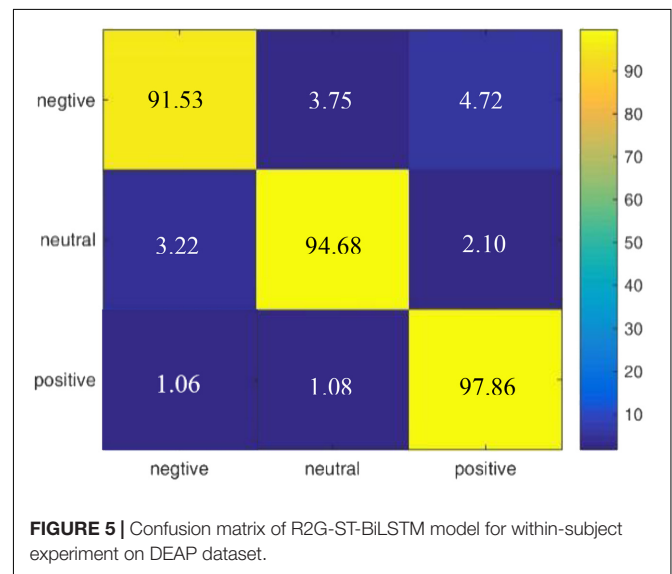
**TABLE 4 |** The results of within-subject experiment on the SEED dataset.

Method	SVM (Suykens and Vandewalle, 1999)	BT (Chuang et al., 2012)	RF (Breiman, 2001)	DBN (Zheng and Lu, 2015)	LSTM (Alhagry et al., 2017)	2D-CNN (Chen et al., 2019c)
ACC (%) /STD	80.14/9.27	83.72/8.68	77.95/9.32	82.58/11.26	83.92/9.44	84.64/7.98
p-Value	0.0006	0.0002	0.0004	0.0007	0.0009	0.0008
Method	3D-CNN (Salama et al., 2018)	H-ATT-GRU (Chen et al., 2019b)	DANN (Ganin et al., 2016)	Casc-CNN-LSTM (Chen et al., 2020)	R2G-ST-BiLSTM	
ACC /STD	87.31/11.14	86.38/9.56	88.96/10.45	92.72/9.33	93.57/8.52	
p-Value	0.0021	0.0035	0.0052	0.0098		

The three traditional learning methods are the following: support vector machine (SVM) (Suykens and Vandewalle, 1999), bagging tree (BT) (Chuang et al., 2012), and random forest (RF) (Breiman, 2001).

The Seven deep learning methods are the following: deep confidence network (DBN) (Zheng and Lu, 2015), deep LSTM recurrent neural network (Alhagry et al., 2017), 2D-CNN (Chen et al., 2019c), 3D-CNN (Salama et al., 2018), hierarchical bidirectional GRU network based on attention mechanism (H-ATT-BGRU) (Chen et al., 2019b), domain adaptive neural network (DANN) (Ganin et al., 2016), and cascaded convolutional recurrent neural network (Casc-CNN-LSTM) (Chen et al.).

In order to horizontally compare the advantages of the proposed model, the input features of the benchmark models are also DE features extracted from four bands of EEG data in the DEAP and SEED datasets, which are consistent with those of our proposed model. The feature extraction method is the same as that stated in the experiment part of section “Within-Subject Experiment of Electroencephalogram Emotion Recognition.” *Classifier and discriminator*. However, the specific format of the input EEG features needs to be reshaped according to the interface of each model. Some key implementation details of these 10 benchmark models are listed in Table 2. The selection of model hyperparameters is also the result of fine-tuning experiments in the same search space mentioned in section “Discussion About Several Variants of the Proposed Model.” *Configuration and training of the bidirectional long- and short-term memory neural network from region to global brain model of part II*.

**FIGURE 5 |** Confusion matrix of R2G-ST-BiLSTM model for within-subject experiment on DEAP dataset.

## Within-Subject Experiment of Electroencephalogram Emotion Recognition

We apply within-subject EEG emotion recognition method like that in literature (Li et al., 2018a) to evaluate our proposed model. To make the experiment result convincing, we use trial-wise randomization to construct the validation dataset. Specifically, we first picked out the subjects with a relatively balanced number

**TABLE 5 |** The results of four frequency bands in within-subject experiment.

Methods	The results of ACC (%) (STD)							
	DEAP				SEED			
	$\theta$	$\alpha$	$\beta$	$\gamma$	$\theta$	$\alpha$	$\beta$	$\gamma$
SVM (Suykens and Vandewalle, 1999)	60.90 (8.76)	62.16 (10.49)	72.75 (7.87)	74.28 (11.13)	57.64 (9.93)	63.19 (7.55)	76.85 (10.19)	72.26 (12.31)
BT (Chuang et al., 2012)	65.44 (7.82)	67.31 (9.65)	77.52 (10.04)	76.06 (8.28)	61.65 (11.42)	62.74 (8.16)	75.58 (9.37)	78.83 (6.96)
RF (Breiman, 2001)	62.38 (12.52)	64.54 (7.39)	72.06 (9.77)	71.87 (10.45)	62.79 (7.58)	62.85 (12.32)	69.10 (11.14)	71.72 (10.56)
DBN (Zheng and Lu, 2015)	61.19 (8.97)	62.74 (11.25)	73.73 (6.80)	75.05 (7.74)	58.32 (9.44)	62.56 (10.28)	70.47 (13.51)	74.29 (8.84)
LSTM (Alhagry et al., 2017)	64.98 (9.15)	77.66 (10.57)	79.13 (7.22)	80.29 (8.83)	60.57 (11.89)	70.14 (12.67)	76.35 (10.23)	78.81 (9.50)
2D-CNN (Chen et al., 2019c)	65.73 (8.89)	68.45 (6.35)	79.96 (9.84)	81.42 (10.77)	67.22 (6.73)	69.36 (8.65)	77.24 (11.38)	80.58 (12.72)
3D-CNN (Salama et al., 2018)	65.26 (7.34)	70.17 (9.89)	82.51 (11.43)	83.68 (12.05)	64.54 (7.42)	71.09 (12.16)	78.67 (9.93)	82.11 (10.64)
H-ATT-BiGRU (Chen et al., 2019b)	66.27 (8.11)	68.58 (6.92)	81.96 (11.15)	84.25 (9.32)	65.03 (9.19)	67.15 (11.54)	81.58 (10.26)	85.34 (8.81)
DANN (Garin et al., 2016)	68.39 (12.56)	70.87 (10.75)	85.73 (11.62)	86.92 (9.19)	67.56 (11.04)	72.42 (7.75)	79.96 (8.42)	85.47 (9.73)
Casc-CNN-LSTM (Chen et al., 2020)	70.07 (7.44)	73.25 (8.81)	88.54 (9.69)	89.18 (11.23)	69.21 (8.12)	75.88 (9.93)	85.25 (10.36)	89.53 (7.39)
R2G-ST-BiLSTM	<b>71.46</b> <b>(10.73)</b>	<b>75.82</b> <b>(9.55)</b>	<b>90.57</b> <b>(7.36)</b>	<b>91.38</b> <b>(8.92)</b>	<b>71.35</b> <b>(8.28)</b>	<b>87.14</b> <b>(6.67)</b>	<b>86.72</b> <b>(9.81)</b>	<b>90.86</b> <b>(11.92)</b>

*Bold values represent the better results obtained by the proposed method, highlighting the comparison.*

**TABLE 6 |** The results of cross-subject experiment on the DEAP dataset.

Method	SVM (Suykens and Vandewalle, 1999)	BT (Chuang et al., 2012)	RF (Breiman, 2001)	DBN (Zheng and Lu, 2015)	LSTM (Alhagry et al., 2017)	2D-CNN (Chen et al., 2019c)
ACC/STD	56.32/10.25	58.49/8.76	51.74/11.13	59.01/7.88	64.66/11.40	65.25/9.37
Method	3D-CNN (Salama et al., 2018)	H-ATT-BGRU (Chen et al., 2019b)	DANN (Garin et al., 2016)	Casc-CNN-LSTM (Chen et al., 2020)	R2G-ST-BiLSTM	
ACC/STD	68.13/14.07	77.82/10.12	75.24/8.59	82.36/7.15	84.51/9.26	

**TABLE 7 |** The results of cross-subject experiment on the SEED dataset.

Method	CM (Suykens and Vandewalle, 1999)	BT (Chuang et al., 2012)	RF (Breiman, 2001)	DBN (Zheng and Lu, 2015)	LSTM (Alhagry et al., 2017)	2D-CNN (Chen et al., 2019c)
ACC/STD	56.73/16.29	51.23/14.82	69.00/10.89	61.28/14.62	63.54/15.47	71.31/14.09
Method	3D-CNN (Salama et al., 2018)	H-ATT-BGRU (Chen et al., 2019b)	DANN (Garin et al., 2016)	Casc-CNN-LSTM (Chen et al., 2020)	R2G-ST-BiLSTM	
ACC/STD	69.13/13.07	76.31/15.89	79.95/9.02	83.28/9.60	85.49/7.96	

of three types of trials. These 13 selected subjects include sub05, sub10, sub12, sub13, sub15, sub21, sub22, sub24, sub25, sub26, sub28, sub29, sub32. For each of these selected subjects, we randomly selected all segments of about 10% of the trials from each type as the test set, then randomly selected all segments of about another 10% of the trials from the rest of each type as the validation set, and at last take all segments of the remaining 80% of the trials as the training set. In this division process, we will make sure all segments belonging to one trial is allocated either as the training set, test set, or validation set to avoid “data leakage.”

Then the proposed R2G-ST-BiLSTM model is used for feature learning and emotion classification. The learning process on the DEAP dataset is shown in **Figure 4**.

We use the average classification accuracy (ACC) and standard deviation (STD) of all subjects to evaluate the model performance. For comparison, we also use the abovementioned benchmark methods to make experiments on the equal dataset. We use paired t-test against the benchmark methods to show the difference between them. *T*-test is a test method for the difference between two mean values of small samples



(sample size less than 30). It uses *t*-distribution theory to infer the probability of difference, to judge whether the difference is significant. The significance of the classification performance of the proposed method against each benchmark method is calculated with paired *t*-test. For all the paired *t*-tests, we used Bonferroni criteria (Genovese et al., 2002) and the implementation method (Weisstein, 2004) to make *p*-value correction for multiple hypothesis testing to limit false discovery rate (FDR). The results of within-subject experiment on the DEAP and SEED datasets are shown, respectively, in **Tables 3, 4**. The *p*-Value indicates the corrected results of paired *t*-test. A value of  $p < 0.05$  means the difference is significant.

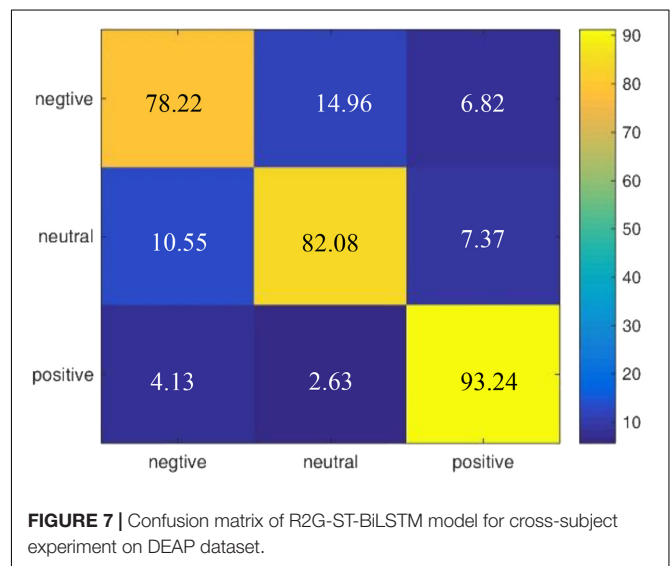
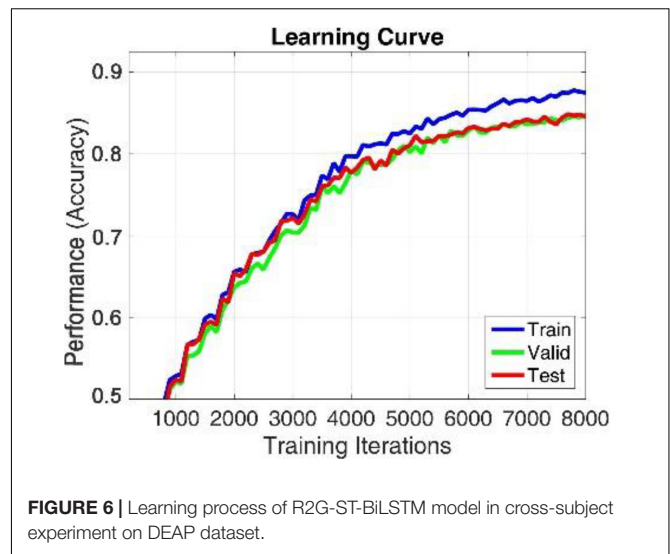
It can be seen from **Tables 3, 4** that the average accuracy of the R2G-ST-BiLSTM method achieves 94.69% on DEAP and 93.57% on SEED, which is best among the above methods. From a statistical point of view, the performance of the proposed model is significantly better than the benchmark models. This result is largely because our R2G-ST-BiLSTM model explores both temporal and spatial context information of the different brain regions of EEG signals.

According to the experimental result of our proposed model on DEAP, we draw a confusion matrix for the three categories of emotions in **Figure 5**. It is found that compared with neutral emotions, positive and negative affections are less likely to be confused.

We also used a method-like reference (Zheng, 2017) to conduct some additional experiments to test the classification performance of different frequency bands of EEG data. Specifically, the DE features are extracted from four frequency bands  $\theta$ ,  $\alpha$ ,  $\beta$ , and  $\gamma$  related to the original signal, and then the EEG emotion recognition experiment is performed based on these DE features of the four bands. We can see the experimental results on DEAP and SEED datasets in **Table 5**, which indicate that on both datasets, the recognition performance in the higher frequency bands of  $\beta$  and  $\gamma$  is better than those in the lower frequency bands of  $\theta$  and  $\alpha$ . This result is consistent with the neurophysiology research in literature (Mauss and Robinson, 2009).

## Cross-Subject Experiment of Electroencephalogram Emotion Recognition

In this section, we use the cross-subject and the leave-one-subject-out (LOSO) cross-validation strategy similar to that in Zheng and Lu (2016); Li et al. (2018a) to evaluate the proposed method, in which the training and testing data are selected from different subjects. The EEG data of one subject is selected as test data, and the EEG data of all the rest of the subjects are used as training data. After each subject is rounded, the average prediction accuracy and standard deviation are calculated as the results. To better compare the performance of the proposed method, we use the abovementioned methods as benchmark. The comparison results of various methods on DEAP and SEED are illustrated in **Tables 6, 7**, respectively. On both datasets, our R2G-ST-BiLSTM model also performs



better. The learning process on the DEAP dataset is shown in **Figure 6**.

We also draw a confusion matrix in **Figure 7** according to the results of our model on the DEAP dataset, which shows that positive emotion is easier to be recognized than the negative and neutral emotions.

We also compared the influence of the different frequency bands on cross-subject emotion recognition. The experimental results on DEAP and SEED datasets are shown in **Table 8**, from which it can be seen that, on both datasets, the classification performance in the higher frequency bands of  $\beta$  and  $\gamma$  are better than those in the lower frequency bands of  $\theta$  and  $\alpha$ , and the R2G-ST-BiLSTM method achieves the best performance on the four frequency bands.

To prove the influence of the different brain regions on emotion recognition, we visualize the weight distribution of brain regions based on the weighting matrix  $W$  defined in

**TABLE 8 |** The results of four frequency bands in cross-subject experiment.

Methods	The results (%) of ACC (STD)							
	DEAP				SEED			
	$\theta$	$\alpha$	$\beta$	$\gamma$	$\theta$	$\alpha$	$\beta$	$\gamma$
SVM (Suykens and Vandewalle, 1999)	41.91 (8.39)	44.73 (7.56)	48.66 (10.21)	51.32 (9.08)	40.62 (9.83)	42.05 (12.65)	47.97 (12.47)	50.06 (10.48)
BT (Chuang et al., 2012)	45.17 (9.38)	48.29 (14.77)	53.95 (8.54)	54.48 (7.43)	45.98 (9.70)	48.63 (10.28)	49.79 (12.41)	54.07 (6.87)
RF (Breiman, 2001)	41.50 (8.57)	41.86 (4.52)	47.31 (12.02)	47.72 (10.05)	40.07 (6.50)	42.09 (13.34)	48.29 (12.77)	48.98 (12.82)
DBN (Zheng and Lu, 2015)	44.36 (11.82)	46.15 (8.98)	55.94 (6.01)	56.81 (9.27)	45.76 (10.98)	48.43 (9.75)	56.66 (6.58)	56.62 (6.84)
LSTM (Alhagry et al., 2017)	47.92 (6.45)	48.69 (10.40)	59.02 (7.83)	59.16 (11.62)	48.63 (10.29)	51.59 (11.83)	62.13 (7.73)	59.37 (10.75)
2D-CNN (Chen et al., 2019c)	48.33 (7.61)	49.74 (13.26)	62.18 (9.90)	62.09 (11.31)	48.36 (10.31)	50.60 (8.30)	62.04 (6.74)	62.19 (7.62)
3D-CNN (Salama et al., 2018)	51.81 (9.79)	53.46 (9.84)	65.15 (11.32)	64.97 (8.46)	52.60 (11.84)	54.95 (10.45)	64.47 (13.69)	64.47 (14.69)
H-ATT-BiGRU (Chen et al., 2019b)	63.44 (12.50)	61.52 (7.07)	70.39 (12.14)	72.63 (5.28)	64.47 (14.96)	59.81 (12.43)	71.03 (10.48)	73.55 (8.80)
DANN (Ganin et al., 2016)	56.98 (5.33)	58.06 (11.80)	67.70 (8.65)	70.46 (12.17)	55.47 (9.80)	56.72 (10.79)	67.14 (7.17)	71.03 (10.14)
Casc-CNN-LSTM (Chen et al., 2020)	61.27 (8.02)	62.83 (6.56)	73.59 (10.54)	73.55 (8.69)	62.04 (6.64)	63.31 (11.96)	73.25 (9.12)	74.29 (7.98)
R2G-ST-BiLSTM	<b>64.03</b> <b>(14.41)</b>	<b>66.26</b> <b>(5.99)</b>	<b>74.64</b> <b>(9.38)</b>	<b>75.02</b> <b>(10.10)</b>	<b>66.14</b> <b>(8.10)</b>	<b>67.14</b> <b>(7.05)</b>	<b>74.85</b> <b>(8.02)</b>	<b>75.89</b> <b>(8.15)</b>

*Bold values represent the better results obtained by the proposed method, highlighting the comparison.*

formula (5) and learned in our cross-subject experiment on DEAP, where the sum of each row of  $W$  matrix represents the contribution of corresponding brain region. **Figure 8** shows the weighted map of the brain areas, where the darker the color of the region, the more significant contribution of the corresponding brain region. It can be seen from **Figure 6** that EEG signals in the frontal lobe are very important for human emotion recognition, which is consistent with the results of the cognitive observations of biological psychology in the literature (Coan and Allen, 2004).

## Discussion About Several Variants of the Proposed Model

Various experiments on the DEAP dataset demonstrates that the proposed R2G-ST-BiLSTM model is more effective than the other methods, which is largely due to our R2G-ST-BiLSTM model utilizing both regional weighting layer and regional to global time layer. In order to confirm that, we obtained the following three simplified models by removing some layers from the R2G-ST-BiLSTM network, and use them to make within-subject and cross-subject experiments on the DEAP dataset. These three simplified models are described as follows:

- (1) R2G-ST-BiLSTM-V1—removes the dynamic regional weighting layer and regional temporal feature learning layers;
- (2) R2G-ST-BiLSTM-V2—only uses global temporal feature as the final input feature to classify;

- (3) R2G-ST-BiLSTM-V3—does not change the original structure of the R2G-ST-BiLSTM, except that the weight of each brain region is set to 1, which means all brain regions are of the same importance to emotion classification.

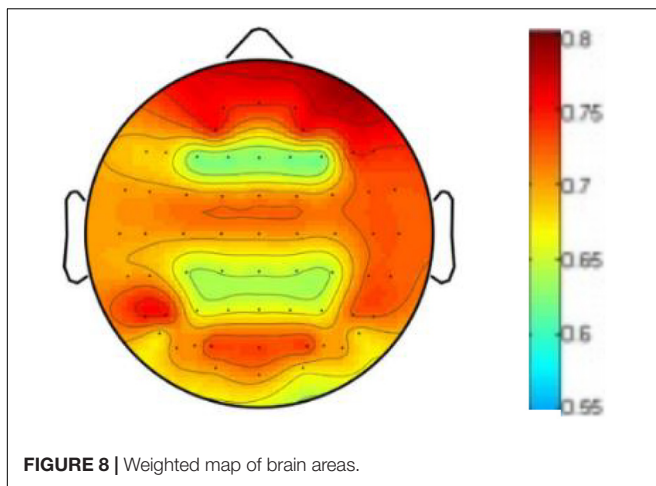
**Table 9** demonstrates the comparison outcome of the above four variant models. The comparison relationship is as follows:

R2G-ST-BiLSTM-V1 < R2G-ST-BiLSTM-V2 < R2G-ST-BiLSTM-V3 < R2G-ST-BiLSTM, (27)

The significance of the regional weighting layer and the regional temporal feature learning layer has been proven by the above comparisons, which shows that these two parts play important roles in enhancing the capability of our R2G-ST-BiLSTM model.

To further discuss whether the different components of R2G-ST-BiLSTM are necessary to outperform other models, we modified it according to the following methods to obtain its several variants:

- (1) R2G-ST-CNN-V1: replaces all BiLSTM modules used for learning spatial and temporal features of local and global brain regions with two-layer 2D-CNN modules.
- (2) R2G-ST-CNN-V2: only the BiLSTM modules used for learning temporal features of local and global brain regions are replaced with two-layer 2D-CNN modules.
- (3) R2G-ST-CNN-V3: only the BiLSTM modules used for learning spatial features of local and global brain regions are replaced with two-layer 2D-CNN modules.



- (4) R2G-ST-BiLSTM-V4: only remove the domain discriminator from the proposed model.

The structure and parameter configuration of the 2D-CNN here are consistent with those in literature (Chen et al., 2019c). These four variant models are used to make within-subject and cross-subject emotion classification experiments on DEAP. The comparison results are shown in **Table 10**.

It can be seen from **Table 10** that the classification performance of the proposed model is significantly better than that of the four variant models. Specifically, the proposed model outperforms the R2G-ST-CNN-V2, which indicates that the BiLSTM components can extract more discriminative time-context features from EEG sequences than 2D-CNN. The performance of our proposed model is better than that of R2G-ST-CNN-V3, which shows that BiLSTM components can better cooperate with the attention mechanism of brain regions and extract more spatial context-dependent features than 2D-CNN. The proposed model significantly outperforms R2G-ST-CNN-V1, which further proves that BiLSTM has obvious advantages over 2D-CNN in learning deep temporal and spatial features in our proposed hierarchical framework. The proposed model significantly outperforms R2G-ST-CNN-V4, especially in cross-subject experiment, which illustrates that the domain discriminator is indeed helpful to extract more discriminative EEG features with small differences between subjects and, therefore, improve the adaptability of the model. In general, the components of BiLSTM and the domain discriminator play very important roles on the whole performance of the proposed model and are necessary to outperform other models.

## DISCUSSION

Although the proposed model has achieved high classification accuracy, there are still some limitations to study and overcome in the future.

At first, the model is complex and lacks interpretability. The model proposed in this paper is a combined hierarchical deep neural network composed of multiple bidirectional LSTM

**TABLE 9** | Comparison results of four models on the DEAP dataset.

Methods	Within-subject experiment	Cross-subject experiment
	ACC (%)	ACC (%)
R2G-ST-BiLSTM-V1	90.43	80.32
R2G-ST-BiLSTM-V2	91.58	81.15
R2G-ST-BiLSTM-V3	93.72	83.96
R2G-ST-BiLSTM	<b>94.69</b>	<b>84.51</b>

*Bold values represent the better results obtained by the proposed method, highlighting the comparison.*

**TABLE 10** | Comparison results of five models on the DEAP dataset.

Methods	Within-subject experiment	Cross-subject experiment
	ACC (%)	ACC (%)
R2G-ST-CNN-V1	85.26	75.64
R2G-ST-CNN-V2	88.75	80.72
R2G-ST-CNN-V3	90.83	81.47
R2G-ST-BiLSTM	<b>94.69</b>	<b>84.51</b>
R2G-ST-BiLSTM-V4	92.14	78.39

*Bold values represent the better results obtained by the proposed method, highlighting the comparison.*

models with attention mechanism. Although the principle and learning process of the model is clear, the decision making and intermediate process made by the model are difficult to understand and interpreted. It is hard to explain the correlation and the interaction among input data, learned features, and output class. At present, researchers have put forward some specific deep model interpretation methods including activation maximization, gradient-based interpretation, class activation mapping (CAM), and so on. The interpretation result of the activation maximization is more accurate and can help people understand the internal working logic of DNN, but the data containing some noise generated in the optimization process makes it difficult to interpret the input (Dong et al., 2017). The gradient-based interpretation methods include deconvolution, guided backpropagation, integrated gradients, and smooth gradients, which aim to use backpropagation to calculate the gradient of specific output relative to input to derive feature importance. This gradient information can only be used to locate important features, but not to quantify the contribution of each feature to the classification results. The CAM method (Jorg, 2019) can locate the objects from the learned features by the excellent ability of the last convolution layer in CNN, which could only provide coarse-grained interpretation results for various CNN models. Additionally, there are some model-agnostic (MA) explanations, such as LIME and knowledge distillation, and causal interpretable method. Although many methods have been proposed in the interpretability research for deep models, there are still many problems to be solved, such as the lack of unified indicators for evaluating interpretation methods, the balance

between model accuracy and interpretability, and the balance between data privacy protection and model interpretability, which will be one of our future research directions to improve the performance of the model.

Second, the complex model and limited amount of data make the model prone to overfitting. We use the EEG data of the DEAP and SEED datasets, which include 32 and 15 subjects, respectively, to make our experiments. In cross-subject experiments on the DEAP dataset, the number of the model parameters reaches about 50,156, which exceeds the number of training samples at 15,360. Compared with the complexity of the model, the training dataset is small, which makes the model prone to overfitting. At present, researchers usually use methods such as expanding dataset, removing features, regularization, and terminating training in advance to prevent model overfitting (Sanjar et al., 2020). Data enhancement is a way to increase training data, which can be realized by flipping, translation, rotation, scaling, and generation methods. Removing features is to reduce the complexity of the model by removing some layers or some neurons from it. Through monitoring the performance of each training iteration and when the loss on the verification set tends to increase, we could stop the training process to prevent the model from overfitting. The regularization method reduces the complexity of the model by punishing the loss function with L1 or L2 paradigm. In our work, we use L2 regularization and dropout method to suppress the overfitting problem, but we still face the challenge of insufficient data. In the future, we will design experiments or ask some medical institutions or hospitals to collect more EEG data for the study. We will also explore to use the generated antagonism network (GAN) to generate a large number of artificial EEG data to make up for this deficiency.

Third, the proposed model is so complex that it needs to consume a lot of computation resources and time to train the model, and it is hard to quickly verify and improve the model, as well as make real-time prediction. In the future, we will try to further simplify the structure of the model without changing its performance, and make deep research on accelerating the speed of model training and real-time application.

## CONCLUSION

Based on the discovery of neuroscience that each region of the human brain can produce different dynamic responses to emotions, we suggest a new hierarchical EEG feature learning method by using attention mechanism and bidirectional LSTM neural network from region to global brain. A large number of experiments and verification are carried out on the DEAP and SEED datasets. The results show that the proposed R2G-ST-BiLSTM model achieves the best performance in subject-dependent and subject-independent EEG-based emotion recognition. Through experiments on several variants of the model, we compare and analyze the impact of different components of the model on its overall performance, and summarize the following advantages of the proposed model:

- (1) The BiLSTM networks are used to hierarchically learn the deep spatial correlation features within and cross each brain region. The attention mechanism is combined to weigh the contribution of each brain region to the emotion classification, which could enhance the influence of the brain region with more contribution and reduce the influence of the brain region with less contribution.
- (2) The BiLSTM networks are used to hierarchically learn the deep temporal correlation features from the EEG time sequence of each local brain region and global brain. The learned deep temporal and spatial features are connected to make the features more discriminative.
- (3) By introducing the domain discriminator, the feature difference between different subjects is reduced, and the robustness and adaptability of the model are improved.

Although the proposed model shows some advantages, there are still some problems to be solved. For example, the model is more complicated, which costs much time and computing resource for training. The whole proposed model still works as a black box, and it is difficult to explain the physical meaning represented by the learned abstract features. The complex model and limited amount of data make the model prone to overfitting. Therefore, in the future, we will further study how to improve the interpretability of the proposed model, simplify the structure of the model, and further improve the robustness and domain adaptability of the model.

## 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.

## AUTHOR CONTRIBUTIONS

PZ designed and implemented the R2G-ST-BiLSTM model and participated in drafted the manuscript. CM carried out the cross-subject EEG-based emotion classification experiments with R2G-ST-BiLSTM model, analyzed the experimental results, and drafted the manuscript. KZ carried out the within-subject experiments and analyzed the experimental results. WX was responsible for literature review, EEG data preprocessing and manual feature extraction. JC conceived of the study, participated in its design, and revised and proofread the manuscript. All authors read and approved the final manuscript.

## FUNDING

This work was supported by the National Natural Science Foundation of China under the Project Agreement No. 61806118 and the Research Startup Fund Project of Shaanxi University of Science and Technology under Project No. 2020bj-30.



## REFERENCES

- Alhagry, S., Aly, A., and Reda, A. (2017). Emotion recognition based on EEG using LSTM recurrent neural network. *Int. J. Adv. Comput. Sci. Appl.* 8, 134–143. doi: 10.14569/IJACSA.2017.081046
- Bottou, L. (2010). “Large-scale machine learning with stochastic gradient descent,” in *Proceedings of COMPSTAT'2010* (Berlin: Springer), 177–186. doi: 10.1007/978-3-7908-2604-3\_16
- Breiman, L. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324
- Chen, J. X., Jiang, D. M., and Zhang, Y. N. (2019b). A hierarchical bidirectional GRU Model with attention for EEG-based emotion classification. *IEEE Access* 7, 118530–118540. doi: 10.1109/ACCESS.2019.2936817
- Chen, J. X., Jiang, D. M., and Zhang, Y. N. (2019a). A common spatial pattern and wavelet packet decomposition combined method for EEG-based emotion recognition. *J. Adv. Computat. Intell. Intell. Informat.* 23, 274–281. doi: 10.20965/jaciii.2019.p0274
- Chen, J. X., Zhang, P. W., Mao, Z. J., Huang, Y. F., Jiang, D. M., and Zhang, Y. N. (2019c). Accurate EEG-based emotion recognition on combined features using deep convolutional neural networks. *IEEE Access* 7, 4107–4115. doi: 10.1109/ACCESS.2019.2908285
- Chen, J. X., Jiang, D. M., Zhang, Y. N., and Zhang, P. W. (2020). Emotion recognition from spatiotemporal EEG representations with hybrid convolutional recurrent neural networks via wearable multi-channel headset [J]. *Comput. Commun.* 154, 58–65.
- Chuang, S. W., Ko, L. W., Lin, Y. P., Huang, R. S., Jung, T. P., Lin, C. T., et al. (2012). Co-modulatory spectral changes in independent brain processes are correlated with task performance. *Neuroimage* 62, 1469–1477. doi: 10.1016/j.neuroimage.2012.05.035
- Coan, J. A., and Allen, J. J. (2004). Frontal EEG asymmetry as a moderator and mediator of emotion. *Biol. Psychol.* 67, 7–50. doi: 10.1016/j.biopsycho.2004.03.002
- Davidson, R. J. (2000). Affective style, psychopathology, and resilience: brain mechanisms and plasticity. *Am. Psychol.* 55:1196. doi: 10.1037/0003-066X.55.11.1196
- Dong, Y., Su, H., Zhu, J., and Bao, F. (2017). Towards interpretable deep neural networks by leveraging adversarial examples[EB/OL]. *ArXiv* [Preprint]. ArXiv:1901.09035,
- Duan, R. N., Zhu, J. Y., and Lu, B. L. (2013). “Differential entropy feature for EEG-based emotion classification,” in *Proceedings of the International IEEE EMBS Conference on Neural Engineering* (San Diego, CA: IEEE), 81–84. doi: 10.1109/NER.2013.6695876
- Etkin, A., Egner, T., and Kalisch, R. (2011). Emotional processing in anterior cingulate and medial prefrontal cortex. *Trends Cogn. Sci.* 15, 85–93. doi: 10.1016/j.tics.2010.11.004
- Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., et al. (2016). Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* 17, 1–35.
- Garcia-Martinez, B., Martinez-Rodrigo, A., Alcaraz, R., and Fernandez-Caballero, A. (2019). A review on nonlinear methods using electroencephalographic recordings for emotion recognition [J]. *IEEE Trans. Affect. Comput.* 12, 801–820. doi: 10.1109/TAFFC.2018.2890636
- Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 15, 870–878. doi: 10.1006/nimg.2001.1037
- Graves, A., Mohamed, A. R., and Hinton, G. (2013). “Speech recognition with deep recurrent neural networks,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, Vol. 38, Vancouver, BC, 6645–6649. doi: 10.1109/ICASSP.2013.6638947
- Heller, W., and Nitsche, J. B. (1997). Regional brain activity in emotion: a framework for understanding cognition in depression. *Cogn. Emot.* 11, 637–661. doi: 10.1080/026999397379845a
- Hochreiter, S., and Schmidhuber, J. (1997). Long short-term memory. *Neural Computat.* 9, 1735–1780. doi: 10.1162/neco.1997.9.8.1735
- Jenke, R., Peer, A., and Buss, M. (2014). Feature extraction and selection for emotion recognition from EEG [J]. *IEEE Trans. Affect. Comput.* 5, 327–339. doi: 10.1109/TAFFC.2014.2339834
- Jorg, W. (2019). “Interpretable and fine-grained visual explanations for convolutional neural network,” in *Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR)*, 9097–9107.
- Koelstra, S., Muhl, C., Soleymani, M., Lee, J. S., Yazdani, A., Ebrahimi, T., et al. (2012). DEAP: a database for emotion analysis using physiological signals. *IEEE Trans. Affect. Comput.* 3, 18–31. doi: 10.1109/T-AFFC.2011.15
- Li, Y., Zheng, W., Cui, Z., Zhang, T., and Zong, Y. (2018a). “A novel neural network model based on cerebral hemispheric asymmetry for EEG emotion recognition,” in *Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI)*, Sweden, (Palo Alto, CA: AAAI Press), 1561–1567. doi: 10.24963/ijcai.2018/216
- Li, Y., Zheng, W., Cui, Z., Zong, Y., and Ge, S. (2018b). EEG emotion recognition based on graph regularized sparse linear regression. *Neural Process. Lett.* 49, 555–571. doi: 10.1007/s11063-018-9829-1
- Li, Y., Zheng, W., Wang, L., Zong, Y., and Cui, Z. (2019). “From regional to global brain: a novel hierarchical spatial-temporal neural network model for EEG emotion recognition,” in *Proceedings of the IEEE Transactions on Affective Computing*. doi: 10.1109/TAFFC.2019.2922912
- Lin, Y. P., Wang, C. H., Jung, T. P., Wu, T. L., Jeng, S. K., Duann, J. R., et al. (2010). EEG-based emotion recognition in music listening. *IEEE Trans. Biomed. Eng.* 57, 1798–1806. doi: 10.1109/TBME.2010.2048568
- Lindquist, K. A., and Barrett, L. F. (2012). A functional architecture of the human brain: emerging insights from the science of emotion. *Trends Cogn. Sci.* 16, 533–540. doi: 10.1016/j.tics.2012.09.005
- Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E., and Barrett, L. F. (2012). The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* 35, 121–143. doi: 10.1017/S0140525X11000446
- Maus, I. B., and Robinson, M. D. (2009). Measures of emotion: a review. *Cogn. Emot.* 23, 209–237. doi: 10.1080/02699930802204677
- Picard, R. W., and Picard, R. (1997). *Affective Computing*, Vol. 252. Cambridge: MIT press. doi: 10.1037/e526112012-054
- Purnamasari, P. D., Ratna, A. A. P., and Kusumoputro, B. (2017). Development of filtered bispectrum for EEG signal feature extraction in automatic emotion recognition using artificial neural networks. *Algorithms* 10:63. doi: 10.3390/a10020063
- Salama, E. S., El-Khoribi, R. A., Shoman, M. E., and Shalaby, M. A. (2018). EEG based emotion recognition using 3D convolutional neural networks. *Int. J. Adv. Comput. Sci. Appl.* 9, 329–337. doi: 10.14569/IJACSA.2018.090843
- Sanjar, K., Rehman, A., Paul, A., and JeongHong, K. (2020). “Weight dropout for preventing neural networks from overfitting,” in *Proceedings of the 2020 8th International Conference on Orange Technology (ICOT) Daegu, South Korea*. IEEE, 2020, 1–4. doi: 10.1109/ICOT51877.2020.9468799
- Song, T., Zheng, W., Song, P., and Cui, Z. (2018). EEG emotion recognition using dynamical graph convolutional neural networks [J]. *IEEE Trans. Affect. Comput.* 11, 532–541. doi: 10.1109/TAFFC.2018.2817622
- Suykens, J. A., and Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Process. Lett.* 9, 293–300. doi: 10.1023/A:1018628609742
- Weisstein, E. W. (2004). *Bonferroni Correction*[J]. Available online at: [https://mne.tools/dev/generated/mne.stats.bonferroni\\_correction.html](https://mne.tools/dev/generated/mne.stats.bonferroni_correction.html).
- Yan, X., Zheng, W. L., Liu, W., and Lu, B. L. (2017). “Investigating gender differences of brain areas in emotion recognition using LSTM neural network,” in *Proceedings of the Neural Information Processing* (Cham: Springer), 820–829. doi: 10.1007/978-3-319-70093-9\_87
- Yu, Z., Ramanarayanan, V., Suendermann-Oeft, D., Wang, X., Zechner, K., Chen, L., et al. (2015). “Using bidirectional LSTM recurrent neural networks to learn high-level abstractions of sequential features for automated scoring of non-native spontaneous speech,” in

- Proceedings of the Automatic Speech Recognition and Understanding (ASRU)* (Scottsdale, AZ), 338–345. doi: 10.1109/ASRU.2015.7404814
- Zhang, Q., and Lee, M. (2010). A hierarchical positive and negative emotion understanding system based on integrated analysis of visual and brain signals. *Neurocomputing* 73, 3264–3272. doi: 10.1016/j.neucom.2010.04.001
- Zheng, W. (2017). Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis. *IEEE Trans. Cogn. Dev. Syst.* 9, 281–290. doi: 10.1109/TCDS.2016.2587290
- 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
- Zheng, W.-L., and Lu, B.-L. (2016). “Personalizing EEG-based affective models with transfer learning,” in *Proceedings of the 25th International Joint Conference on Artificial Intelligence (IJCAI)* (Palo Alto, CA: AAAI Press), 2732–2738.
- Zheng, W. L., and Lu, B. L. (2017). A multimodal approach to estimating vigilance using EEG and forehead EOG. *J. Neural Eng.* 14:026017. doi: 10.1088/1741-2552/aa5a98
- Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships 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.
- Copyright © 2021 Zhang, Min, Zhang, Xue and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



# Measuring and Modeling the Effect of Audio on Human Focus in Everyday Environments Using Brain-Computer Interface Technology

Aia Haruvi, Ronen Kopito, Noa Brande-Eilat, Shai Kalev, Eitan Kay and Daniel Furman\*

Arctop Inc., San Francisco, CA, United States

## OPEN ACCESS

### Edited by:

Jane Zhen Liang,  
Shenzhen University, China

### Reviewed by:

Yuxiao Yang,  
University of Central Florida,  
United States  
Hiroshi Higashi,  
Kyoto University, Japan

### \*Correspondence:

Daniel Furman  
df@arctop.com

**Received:** 18 August 2021

**Accepted:** 17 December 2021

**Published:** 27 January 2022

### Citation:

Haruvi A, Kopito R, Brande-Eilat N, Kalev S, Kay E and Furman D (2022) Measuring and Modeling the Effect of Audio on Human Focus in Everyday Environments Using Brain-Computer Interface Technology. *Front. Comput. Neurosci.* 15:760561. doi: 10.3389/fncom.2021.760561

The goal of this study was to investigate the effect of audio listened to through headphones on subjectively reported human focus levels, and to identify through objective measures the properties that contribute most to increasing and decreasing focus in people within their regular, everyday environment. Participants ( $N = 62$ , 18–65 years) performed various tasks on a tablet computer while listening to either no audio (silence), popular audio playlists designed to increase focus (pre-recorded music arranged in a particular sequence of songs), or engineered soundscapes that were personalized to individual listeners (digital audio composed in real-time based on input parameters such as heart rate, time of day, location, etc.). Audio stimuli were delivered to participants through headphones while their brain signals were simultaneously recorded by a portable electroencephalography headband. Participants completed four 1-h long sessions at home during which different audio played continuously in the background. Using brain-computer interface technology for brain decoding and based on an individual's self-report of their focus, we obtained individual focus levels over time and used this data to analyze the effects of various properties of the sounds contained in the audio content. We found that while participants were working, personalized soundscapes increased their focus significantly above silence ( $p = 0.008$ ), while music playlists did not have a significant effect. For the young adult demographic (18–36 years), all audio tested was significantly better than silence at producing focus ( $p = 0.001–0.009$ ). Personalized soundscapes increased focus the most relative to silence, but playlists of pre-recorded songs also increased focus significantly during specific time intervals. Ultimately we found it is possible to accurately predict human focus levels *a priori* based on physical properties of audio content. We then applied this finding to compare between music genres and revealed that classical music, engineered soundscapes, and natural sounds were the best genres for increasing focus, while pop and hip-hop were the worst. These insights can enable human and artificial intelligence

composers to produce increases or decreases in listener focus with high temporal (millisecond) precision. Future research will include real-time adaptation of audio for other functional objectives beyond affecting focus, such as affecting listener enjoyment, drowsiness, stress and memory.

**Keywords:** artificial intelligence, audio, brain-computer interface, focus, human

## INTRODUCTION

### The Effect of Sound on Human Experience

Sounds are all around us, from natural sounds like the wind, to engineered sounds like music. It is well-established that sounds have a major influence on the human brain and consequently, human experience (Levitin, 2006; Sacks, 2010). Sounds can reduce stress (Davis and Thaut, 1989), support learning and memory formation (Hallam et al., 2002), improve mood (Chanda and Levitin, 2013), and increase motivation (Salimpoor et al., 2015). Sounds can also do the opposite and create aversive experiences (Schreiber and Kahneman, 2000; Zald and Pardo, 2002; Kumar et al., 2012). One of the most significant effects of sounds is to impact focus. Focus is commonly demanded by tasks of daily living and work, and in these areas sounds experienced as audio through headphones, earbuds or speakers offer a safe way to increase focus levels and productivity. However, sounds can be both beneficial or distracting and previous results have been inconclusive in determining the reasons why (de la Mora Velasco and Hirumi, 2020).

For example, it has been found that listening to music with lyrics while reading or working can decrease concentration or cognitive performance (Shih et al., 2012; Liu et al., 2021), while several studies have shown oppositely that natural-occurring sounds such as white noise, or highly composed sound such as classical music, can be beneficial for increasing focus and can even improve learning outcomes (Davies, 2000; Chou, 2010; Angwin et al., 2017; Gao et al., 2020). Therefore, one interesting question that emerges is: what are the specific properties of an audio experience that affect human focus levels the most? Additionally, studies have shown that the effect of audio is often subjective, where whether one likes a given sound or not is a key factor in its effect on their experience (Cassidy and Macdonald, 2009; Huang and Shih, 2011; Mori et al., 2014). Although this finding about the subjectivity of experience of audio reappears across many studies, psychophysical thresholds are known to exist and there are clearly natural laws governing much of the way humans hear and experience sound (Levitin et al., 2012; Nia et al., 2015; Washburne, 2020).

The potential of audio alone to increase focus, and the consumer demand for non-pharmaceutical tools that enable individuals to enhance their own ability to focus has recently led several companies (including Focus@Will, Endel, Brain.fm, Mubert, Enophone, Melodia, AIVA, and others) to develop audio content that is dedicated to increasing focus “on-demand.” These new audio forms include elements of white noise, music, and other sonic properties that are functionally

combined to increase a listener’s focus and maintain high levels of focus over a long duration of time. One of the challenges in this field is to figure out the physical properties of sound that contribute to human experience the most so that design principles can be defined correctly to create audio that reliably achieves the goal of increasing focus, opposed to the inverse of causing distractions and impairing an individual’s ability to focus. Insights about audio properties therefore have been sought by commercial groups alongside academic groups in order to learn how to optimally enhance human focus.

Many scientific studies have explored this question and looked for the relationship between sound, music, and human experience using objective measures that empirically assess properties of audio and their emotional correlates. For example, Cheung et al. (2019) found that pleasure from music depends on states of expectation, such as a skipped rhythmic beat, which can either be pleasurable or discomfoting depending on the listener’s specific circumstance. *Sweet Anticipation* (Huron, 2006) similarly maps how music evokes emotions within a theory of expectation and describes psychological mechanisms that are responsible for many people’s mixed responses to audio of various types. Other studies used machine learning methods to map from features of audio signals to emotions (Yang et al., 2008; Vempala and Russo, 2012; Brotzer et al., 2019; Cunningham et al., 2020; Hizlisoy et al., 2021). These machine learning studies to date have, however, only aimed to predict emotions based on the limited valence-arousal circumplex model, and as far as we know, no attempts have been made to predict human focus levels exclusively based on audio signal analysis.

One persistent obstacle to the field’s understanding has been studies that rely on data with a low temporal resolution. Since audio content and emotions can change fast, on the order of tens of milliseconds, the current lack of modeling tools capable of capturing quick, transient changes in human experience that accompany changes in sound is a major hindrance to progress (Larsen and Diener, 1992; Cowen and Keltner, 2017). Commonly, for example, reports are based on data where there is a single emotional label per song, while the song lasts 2–3 min and throughout it there are emotional dynamics that change dramatically. This mismatch of data can lead to conclusions being drawn from inadequately small amounts of samples, and worse than that, inaccurate emotional labels.

### Attention and Emotion Decoding From Brain Signal

Brain decoding technology offers an exceptional opportunity to solve this issue, since it enables an estimation for the experience



dynamics at the same time resolution as focus phenomena occur. Using electroencephalogram (EEG) sensor data, which contains electrical brain activity measured from the scalp (non-invasive) on the order of hundreds of measurements per second, many studies have established that it is possible to capture fast changes in human emotions and experience, such as stress (Perez-Valero et al., 2021), arousal (Faller et al., 2019), fatigue (Hu and Yang, 2017), and happiness (Lin et al., 2017). Several studies have similarly shown the ability to capture focus and attentional state changes, affirming that this information is present in EEG sensor data (Jung et al., 1997; Hamadicharef et al., 2009; Micoulaud-Franchi et al., 2014; Tuckute et al., 2021). While brain decoding technology has been applied widely to study the effects of different types of stimuli (e.g., visual, tactile, and auditory) on human experience within a laboratory environment (Bhatti et al., 2016; Shahabi and Moghimi, 2016; Asif et al., 2019), as far as we know, it has not been applied to study the joint effects of audio and focus at the high temporal resolution needed to explain both phenomena as they occur in people's natural, everyday environments.

In recent years, progress in the development of consumer brain-computer interface wearable technology such as non-clinical, non-invasive EEG sensors (such as Muse, NeuroSky, Emotiv, Bitbrain, etc.), which are intended for personal use, has led to new research paradigms. Now it is possible to use comfortable, affordable, wireless, and easy-to-use at-home measurement devices to collect neuroscientific data “in-the-wild” at a large scale, which opens up for the first time the opportunity to measure brain responses from diverse audiences within their natural habitats. Many of the wearable brain-computer interface devices offer real-time decoding outputs that are derived from the raw electrophysiological sensor data. These “off-the-shelf” decoding outputs include attention, relaxation, and other states (Rebolledo-Mendez et al., 2009; Liu et al., 2013; González et al., 2015; Abiri et al., 2019; Bird et al., 2019). It is important to note, however, that although decoder algorithms have existed in the market for consumer uses for several years, verifying their reliability to accurately capture attention, valence, arousal, stress, and other attributes of human experience at a high temporal resolution, advanced research quality has remained a challenge.

## Combining Brain-Computer Interface Technology With Audio Tests to Decode Focus

In the current study, we used a brain-computer interface algorithm package, Neuos<sup>TM</sup> Software Development Kit (Neuos SDK from Arctop Inc.), for processing data from portable fabric EEG headbands (Muse-S from Interaxon Inc.) in order to measure human focus levels in individuals performing tasks at home while they listened to different audio content through headphones. The ground truth focus levels we used were based on each individual's subjective, self-report. Since the Neuos SDK product is a relatively new technology, we first evaluated the validity of the focus outputs within the experimental conditions. Then, once the algorithm outputs were found to be reliable and accurate in this context, we use the focus data to compare effects

of different sound stimuli on individuals as they carried out different tasks.

Next, we exploited the high temporal resolution of the decoded data to map between raw audio signals and the focus dynamics. Based on this mapping, we built a model that takes as an input an audio file and predicts from the properties of sound in the audio the corresponding focus levels that human listeners will experience. This high resolution modeling enables us to compare between new songs, various sounds, and between genres to gain additional insights about the nature of audio stimuli that drive human focus the most. These insights can help produce optimal playlists to increase focus for general audiences, improve design of custom soundscapes for work and learning environments, and even adapt audio in real-time based on an individual's focus levels to allow them precise influence over their own mental state.

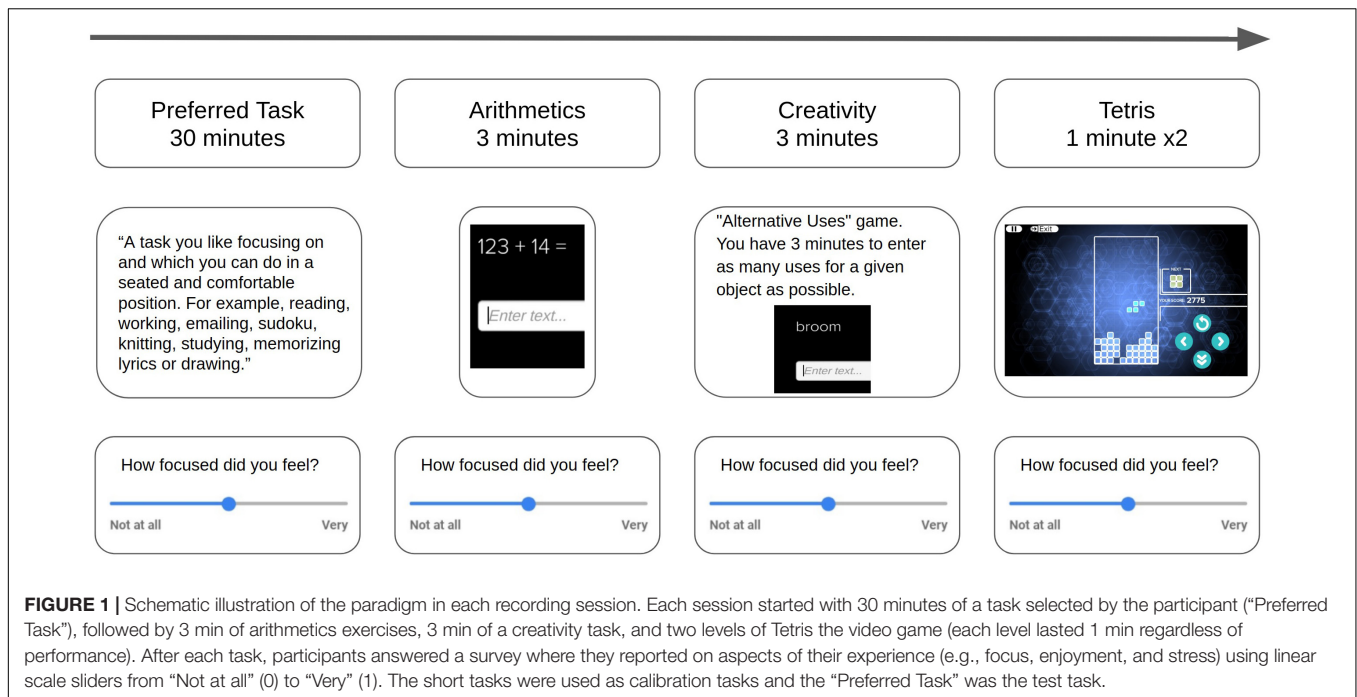
## MATERIALS AND METHODS

### Participants

Sixty-two participants (40 males, 22 females, 18–65 years), completed four sessions over a single week at their own home. All participants were recruited from an opt-in screening panel and were distributed approximately evenly across the five major regions of the continental United States (Northeast, Southwest, West, Southeast, and Midwest). Only participants who reported normal hearing, normal vision, or vision that was corrected to normal with contact lenses, were included. We excluded volunteers who reported using medication that might influence the experiment and who reported neurological or psychiatric conditions that could influence results. Participants were native English speakers and a written informed consent was obtained from each participant prior to their participation. Participants received compensation for their time.

### Paradigm Tasks

Participants performed various tasks within a mobile Android app (“Neuos Central” by Arctop Inc.) while listening to one of three types of audio and wearing a brain signal measuring headband (four-channel EEG Muse-S device by Interaxon Inc.). Each participant received a kit at their home by mail that included all the equipment needed to participate, including over-ear (Sony Group Corporation) headphones, headband (Interaxon Inc.) and tablet computer (Samsung Electronics Co., Ltd.) with the mobile app installed. Participants recorded four 1 h long sessions, while listening to different audio types. Sessions included 30 min of a “Preferred Task” – a task chosen by the participant – followed by short tasks (“calibration tasks”). These short tasks included video games (Tetris), math problems (Arithmetics), and word problems (Creativity) that were used to calibrate the sensors to the individual. Participants were assigned to groups according to a pseudorandom schedule that controlled for potential sequence effects of the tasks and different audio types (**Figure 1**). The short tasks calibrated the Neuos SDK decoding algorithms to a validated performance level for each participant and afterwards



the individually validated model was used to measure each participant's focus level across the Preferred Task.

Participants were instructed to choose a Preferred Task that they could perform in a seated position while listening to audio through the headphones, and which they would be happy to repeat in all four sessions. For example, Preferred Tasks that were chosen included working, reading, knitting and solving Sudoku puzzles. At the end of each task the participants self-reported their experience through a survey in the app which used linearly scaled slider buttons to quantify experience along several dimensions (e.g., focus level, enjoyment, stress, motivation, etc.). For the Preferred Task, the survey included reporting on their focus level during the first and second half of the task separately, resulting in six self-reported quantitative focus labels per session (Preferred Task: two labels, arithmetics: one label, creativity: one label, tetris: two labels).

### Audio Stimuli

Each participant experienced four audio conditions over the 4 days of the study: two music playlists by leading digital service providers Spotify and Apple (downloaded September 2020), one personalized soundscape engineered by Endel, and silence (no audible sounds). We selected Spotify's "Focus Flow" playlist and Apple Music's "Pure Focus" playlist to represent the category of pre-recorded audio designed to increase listener focus. For soundscapes we selected the mobile app Endel to represent the category of real-time, engineered audio that contains a mixture of noise and musical properties. The Endel app "Focus" soundscape was used by each participant on their own mobile device. All audio conditions were instrumental and did not include singing or any audible lyrics. For the condition of silence, participants wore headphones exactly as they did in the audio conditions but

no music or audible sounds of any kind were played and no soundscape was generated – participants simply completed the session in a quiet environment.

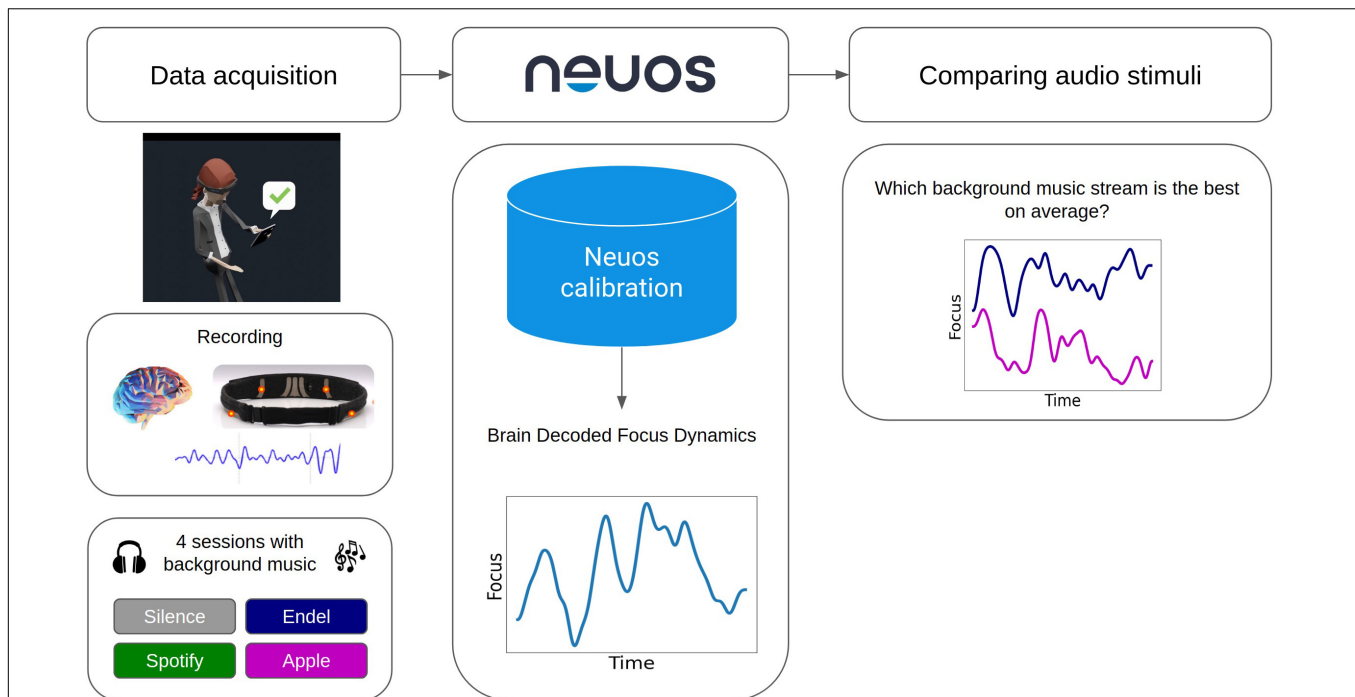
## Data Processing

### Data Acquisition

While participants were listening to audio stimuli and engaging in the experimental tasks, their electrical brain activity was recorded using a fabric electroencephalograph (EEG) headband that was wireless and weighed 41 g (Muse-S device by Interaxon Inc.). The headband included four dry EEG sensors (sampling rate: 256 Hz), photoplethysmography (PPG) sensors (for heart rate) and motion sensors (gyroscope, accelerometer). The brain-measuring EEG sensors were located on the scalp at two frontal channels (AF7 and AF8) and two temporal channels (TP9 and TP10), with the reference channel at Fpz. The headbands were put on by participants themselves with the assistance of a quality control screen in the app that started each session by giving participants real-time feedback on the signal quality of their headband and made it easy for them to adjust the headband appropriately to acquire the optimal signal (**Figure 2**). No technicians or other support staff assisted in the placement of the headbands – the process was completely automated by the in-app prompts within the "Neuos Central" app, freeing the participants to complete sessions at any time or place of their choosing.

### Brain Based Models of Focus

Brain decoding algorithms designed for real-time brain-computer interfacing (Neuos SDK) was used to transform raw sensor data into predicted focus dynamics with a time resolution of 5 Hz (**Figure 2**). The short tasks (games, word and math



**FIGURE 2 |** Schematic illustration of the data processing pipeline. Data acquisition included at-home recordings of four sessions, each with a different background sound type. Arctop's Neuos SDK brain decoding technology package was used to predict the focus dynamics at a rate of 5 Hz. Obtaining the brain decoded focus dynamics synchronously with the sound content enables comparison of focus levels correlated with different physical properties of sound.

problems) were used for calibration of the sensors to the focus dynamics of each individual, and then a calibrated model per participant was applied on the Preferred Task data across all days. The full analysis procedure included data exclusion, preprocessing, feature extraction, and applying machine learning models to transform raw data features to decoded focus dynamics is explained below.

### Data Exclusion

Since brain activity and survey data was collected in participants' own homes, as a quality control step before preprocessing we first validated the data with respect to headband positioning (to confirm it was correctly placed), survey responses (to make sure participants followed the instructions properly) and internet issues (which occasionally resulted in missing measurements). This procedure led to 11 participants data being excluded from further analysis due to the following reasons:

1. Three participants were excluded due to misplacement of the headband which caused excessive noise in their recorded data. To identify the misplacement we simply extracted the standard deviation of the raw signal ( $SD > 500$  reflects a misplaced channel). **Supplementary Figure 1** shows three examples of the raw data of problematic participants vs. three examples of valid participants.
2. Two participants were excluded for not following the instructions correctly during the Preferred Task.

3. Six participants were excluded due to persistent internet issues which caused missing or disrupted data.

After data exclusion, a total of 51 participants (mean age = 36,  $SD = 8$ , 17 females and 34 males) were included in the experimental analysis.

### Preprocessing

A band-pass filter (0.5–70 Hz) was applied to each channel together with a notch filter (60 Hz) to remove line noise. During the performed tasks, 5 s of EEG data segments were extracted from the filtered signal using a sliding window with a stride of 200 ms (96% overlap) to obtain a time resolution of 5 Hz of accurate focus measurements reliably across each task. Headband motion sensor (gyroscope) data were used to detect the motion state of each segment (static, medium or high movement) and segments with substantial movements (medium or high) during the short calibration tasks were automatically excluded. During the Preferred Tasks, all segments were included regardless of movement state in order to obtain continuous dynamics for the full 30 min.

To validate that the differences in responses to each audio stream were not due to differences in movement patterns evoked by the audio, we compared the motion statistics between audio types. **Supplementary Figure 2C** shows that during the Preferred Task, participants were stationary 91% of the time and similarly for all audio types (**Supplementary Figure 2D**). To address eye blinks, which are a normal human function that can corrupt EEG data, we calculated the number of blinks in each EEG segment.

**Supplementary Figure 2A** shows the histogram of the blink rate across participants (average blinks per minute =  $16 \pm 6$ ). **Supplementary Figure 2B** shows similar rates of eye blinks for all audio streams, ruling out the possibility that the differences in the effects of audio were due to differences in eye blink patterns which may have introduced artifacts to the decoded data.

### Feature Extraction

From each EEG segment (epoch) a total of 124 features were extracted, then to handle outliers and avoid extreme values each feature underwent a programmatic trimming procedure that denoised high and low values (extreme values were defined as above or below 2 SD from the mean). The following features were used:

- Average power spectrum features – each segment was transformed to the frequency domain using Welch method, and for each channel, the average power in different frequency bands was calculated (0.5–4, 4–8, 8–12.5, 12.5–30, 30–47, 52–70, and 30–70 Hz) – a total of 4 channels  $\times$  7 bands = 28 features.
- Ratios between average power for spatial symmetric channels (frontal:  $\frac{AF7}{AF8}$  and temporal:  $\frac{TP9}{TP10}$ ) – a total of 2 pairs  $\times$  7 bands = 14 features.
- Power spectrum interactions – the power spectrum ratio between bands ( $\frac{\alpha}{\delta}, \frac{\beta}{\theta}, \frac{\theta}{\alpha}$ ) (Barachant, 2017) and engagement index ( $\frac{\beta}{\alpha+\theta}$ ) (Pope et al., 1995) – a total of 4 channels  $\times$  4 interactions = 16 features.
- Pairwise Pearson correlations between channels in the above frequency bands – 6 pairs  $\times$  7 bands = 42 features.
- Time domain features – for each channel, the first four moments (average, standard deviation, skewness, and kurtosis), entropy and number of zero-crossing points – total of 4 channels  $\times$  6 types = 24 features.

### Machine Learning Models

Average features were calculated across each short task (games, word and math problems) for all valid participants and from all days, resulting in 816 focus-ranked tasks (51 participants  $\times$  4 sessions  $\times$  4 ranked subtasks per session). Then, in a cross validation procedure, multiple random forest regression models provided by the Neuos SDK software package were trained on random subsets of participants (80%) to predict the self-reported focus based on the computed features. For each participant, from the subset of models for which their data were not part of training, the single best model was selected based on the Pearson correlation between the model prediction and the self-reported focus by that participant during the short calibration tasks. The selected regression model was then applied to EEG segment data during their Preferred Task 30 min recordings to get a continuous brain-decoded gradient of focus dynamics that was accurate.

A Gaussian filter was used to smooth the dynamics of the brain-decoded focus gradient and all of the presented results and statistical analysis in this paper are projections of the Gaussian filtered model outputs on the Preferred Task which was not part of the training and selection process for each participant. **Figure 3** shows the resulting brain decoded focus levels of two

representative participants across all four sessions during the Preferred Task period. Model performance was evaluated using Pearson correlation coefficient between the self-reported focus and the brain decoded focus values after thresholding the values, with the area under the ROC curve for binary classification of low/high focus (**Figure 5**).

Electroencephalogram signals are non-stationary and can change dramatically over time (Haartsen et al., 2020; Padilla-Buritica et al., 2020; Yang et al., 2021). To validate that the obtained focus dynamics were not influenced by the non-stationarity of the EEG signal or other forms of signal drift that can occur with electrophysiological measures, we compared the averaged focus levels across all sessions during the first 15 min of the Preferred Task to the last 15 min (**Supplementary Figure 3**). We found there was no significant difference between the segments and concluded that the signal processing methods were robust to this form of signal artifact.

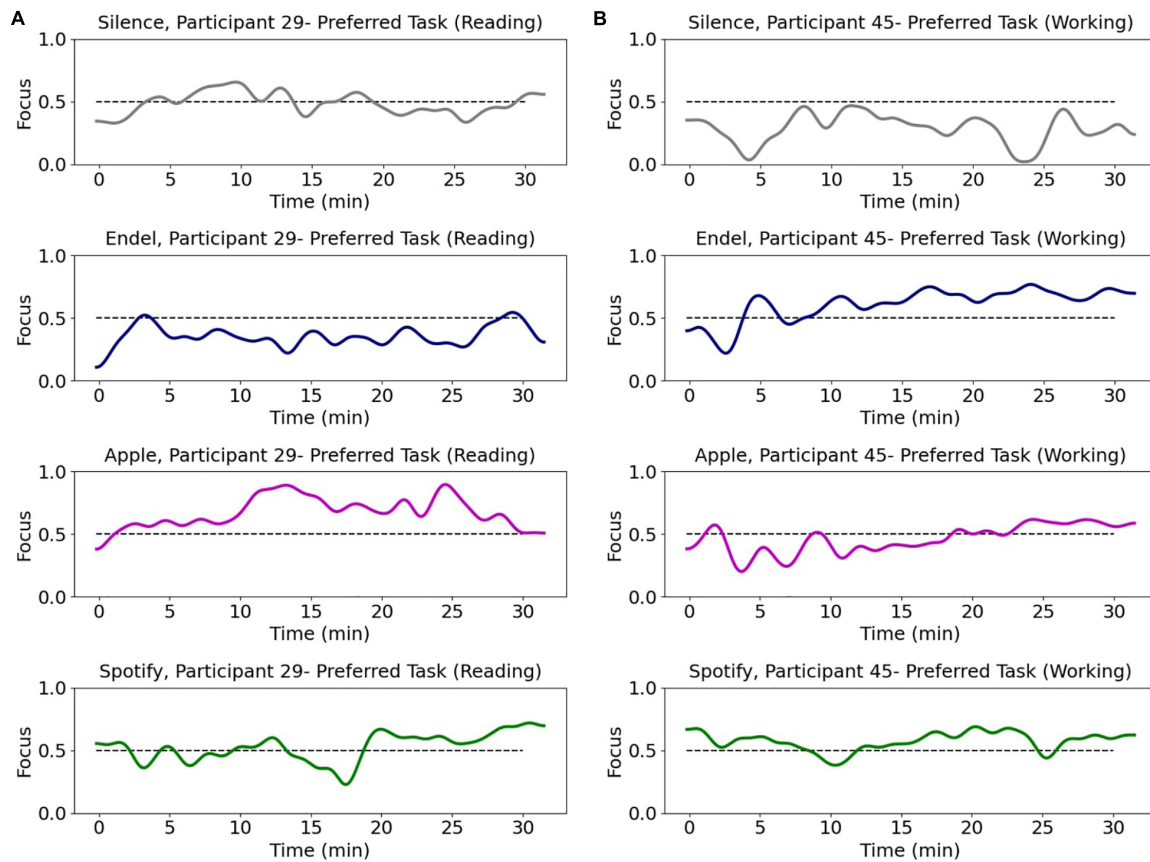
### Statistical Methods

For comparisons between average focus levels in response to the different audio streams, we calculated for each participant ( $N = 51$ ) the median focus level while performing the Preferred Task and conducted a one-way repeated measures ANOVA (analysis of variance) test. Then, if  $p < 0.05$ , paired  $t$ -tests were applied *post hoc* to compare between pairs of audio streams using the Holm–Bonferroni correction. Time series statistical tests were applied to compare focus level dynamics and discover specific time periods where there was significant difference. A paired  $t$ -test was applied to each second between focus levels of two audio streams and the  $p$ -values were then corrected for multiple comparisons by setting a threshold for a minimum significant sequential time-samples. The threshold was determined by random permutations (1000 iterations) of participants' conditions and repeating the statistical test, resulting in a distribution of significant sequential time samples. The threshold was set as the 95% percentile of the resultant distribution (Broday-Dvir et al., 2018).

### Audio Signal Decomposition and Feature Extraction

The pre-recorded music playlist conditions (Apple and Spotify) provided raw audio data that we used to obtain sound property dynamics in the time and frequency domain. These dynamics could then be correlated with the obtained focus dynamics as averaged across participants. Soundscape audio content and silence conditions were not used in this analysis because the soundscapes were produced in real-time personally for each participant, which limited the ability to apply sound property analysis appropriately across the data set, and the silent condition yielded no sound features (no microphones were used during the session from which miscellaneous sounds might have otherwise been extracted). The audio features were calculated for each playlist using Python's library pyAudioAnalysis (Giannakopoulos, 2015), for example, the sound signal energy, spectral entropy, and chroma coefficients were extracted. The features were calculated in short-time windows of 50 ms with a sliding window of 25 ms. Basic statistics were then calculated over the sound features in windows of 30 s





**FIGURE 3 |** Brain data based focus model dynamics of two representative participants during the Preferred Task performed at each of the four sessions. Each row represents a session with a different sound stream playing in the background as participants perform their chosen task. Each session included 30 min (x-axis = time in minutes) of a “Preferred Task” over which their focus level (y-axis = decoded focus) was measured. Participant 29 (**A**) was reading while Participant 45 (**B**) was working.

(e.g., mean and SD), resulting in 136 sound properties (link to full list). To enable mapping of audio features to the brain model, the brain decoded focus levels were averaged across participants and averaged in corresponding 30 s windows (**Figure 4**) to obtain a singular collective dynamic that could be used to predict focus from audio content with the same number of samples as the audio feature properties.

### Obtaining the Sound Decoded Focus Model

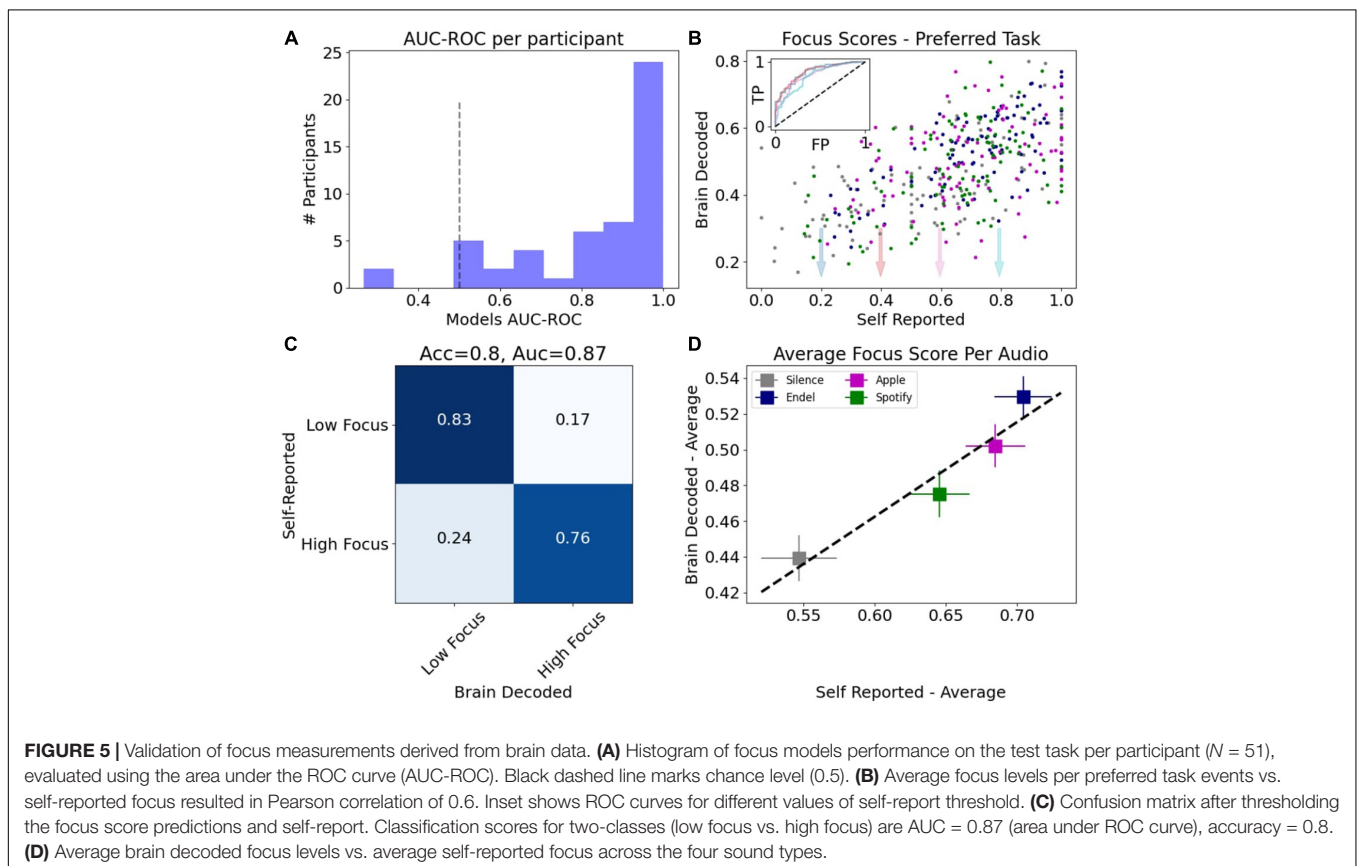
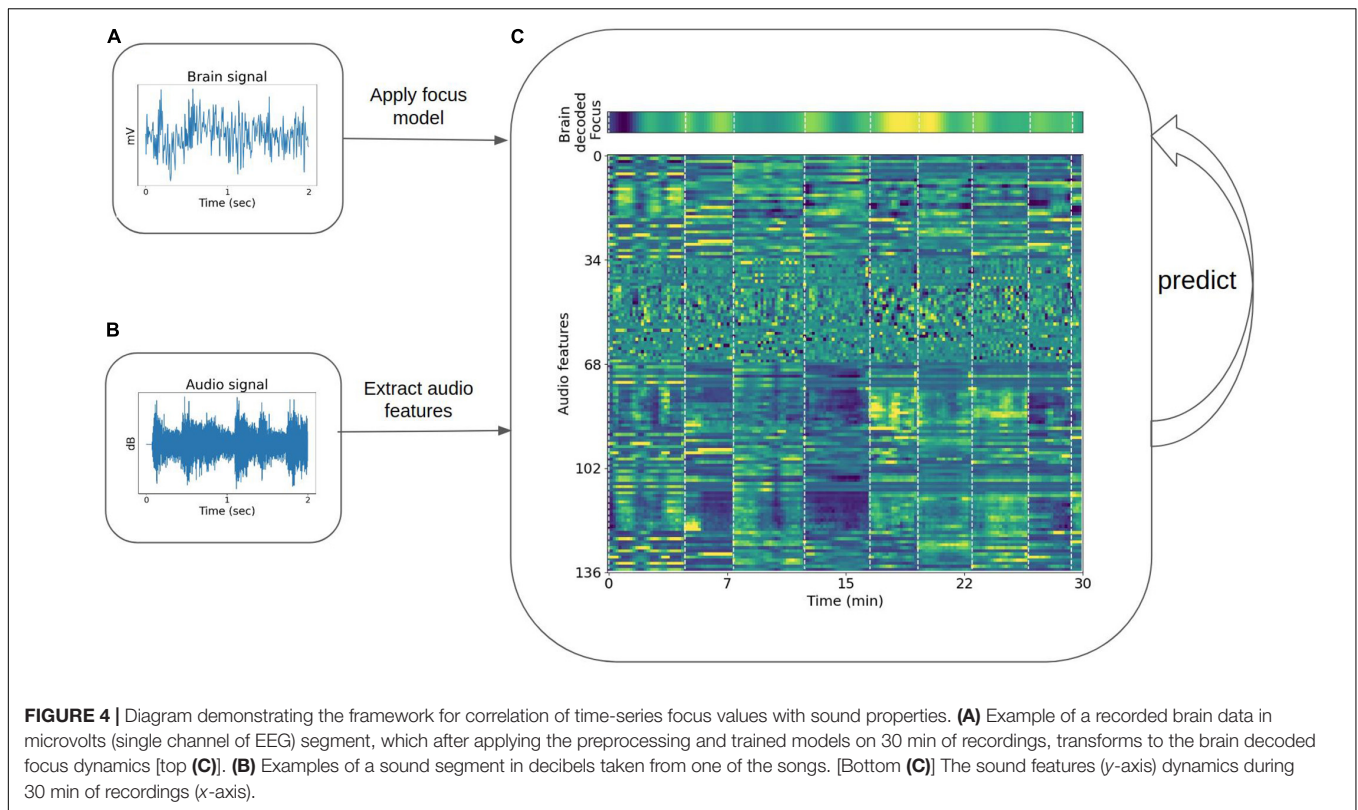
To map the relationship between properties of the audio heard and focus levels measured directly from the brain, we applied principal component analysis (PCA) to reduce the dimensionality of the audio features (using 33 dimensions ultimately, which explained 95% of the data variance). We then trained a linear regression model to map between the transformed audio features and the averaged brain decoded focus levels. The training was done using backward features elimination where in each iteration the component with the smallest weight on average was eliminated. To evaluate model performance, training was done with a stratified cross validation procedure in which we divided the data set to training and validation according to the songs played (to avoid time

dependency issues between the sound features). A total of 18 different songs were played during the pre-recorded playlists (8 songs for Apple, 10 songs for Spotify). In each iteration, 14 songs were used in training and 4 as validation (77/23%). For each sample and song, the audio decoded score is the average model predictions calculated across the models it was part of in the validation set.

## RESULTS

### Brain-Measured Focus Levels Accurately Reflect Self-Reported Focus Levels

After calibration tasks established an initial model for each participant, and before comparing focus levels elicited by the different audio types, we validated that the underlying brain decoding technology was accurate and correctly calibrated by comparing between the brain-based focus predictions and the self-reported focus values during the test task (the “Preferred Task”). **Figure 5A** shows a histogram of the model performance per participant. The model is evaluated based on the AUC score (of the ROC curve) for prediction of self-reported focus



**TABLE 1** | Results of a one-way repeated measures ANOVA performed on each subgroup, comparing the average brain decoded focus levels of each sound stream during the Preferred Task.

Group	N	F	p
All	51	4.28 (3,150)	0.006
Working	26	3.74 (3,75)	0.014
Not working	25	1.91 (3,72)	0.14
Age >36	26	1.81 (3,75)	0.15
Age <36	25	6.97 (3,72)	<0.001

Sound most significantly affected those below 36 years old.

during the Preferred Task (low-high focus) where the chance guessing level is 0.5 (black dashed line). The average result across participants obtained was  $\langle \text{auc} \rangle = 0.83$  ( $N = 51$ ,  $SD = 0.19$ ), a strong validation of the brain-measured focus accuracy.

When aggregating tasks from all participants, the Pearson correlation between the brain decoded focus model and the self-reported focus was  $\text{Corr}(414) = 0.6$ ,  $p < 10^{-4}$  (Figure 5B). The inset in Figure 5B shows the ROC curves for different values of self-reported threshold and the confusion matrix for one of these thresholds (0.4) resulted in an accuracy score of 0.8 (Figure 5C). Figure 5D shows the average brain decoded focus level for each audio type vs. the average self-reported score.

## Soundscapes Induce Higher Focus Levels Compared to Silence

Using the validated focus models which output five measurements per second (5 Hz), we then compared between the average focus levels elicited by the audio listed to during the Preferred Task. The background audio condition was found to have a significant effect (top row in Table 1,  $F(3,150) = 4.28$ ,  $p = 0.006$ , statistical methods for details) on the elicited focus level and the *post hoc* tests [Holm–Bonferroni correction] revealed that streaming soundscapes (Endel app) were significantly higher compared to silence [Figure 6A1 and Supplementary Table 1;  $M = 0.090$ ,  $SE = 0.027$ ,  $t(50) = -3.38$ ,  $p = 0.008$ ], while streaming music using Apple or Spotify did not have an effect [Apple:  $t(50) = -2.37$ ,  $p = 0.11$ , Spotify:  $t(50) = -1.24$ ,  $p = 0.65$ ].

For 35.3% of participants the soundscape session produced their highest focus level, while for 27.5% of participants the Apple playlist produced their personal highest focus level. For 19.6% of participants Spotify was best for producing focus and for 17.6% silence was (Figure 6A2, the details sorted focus levels per participant are shown in Supplementary Figure 4). To gain a better understanding of the conditions where audio affected focus, we next split the participants into subgroups of interest and repeated the statistical analysis. We first asked whether the focus level difference is task dependent. During the Preferred Task, 51% of the participants (26) chose to work, while the remainder (49%) read a book (29.4%), played games (9.8%), or performed other various tasks (e.g., knitting, 9.8%). To assess the effect of audio on focus levels during these different tasks, we split the participants to the ones who worked and those that did other tasks. We found that for the “working” group, the focus level elicited by Endel soundscapes was higher compared to silence

[Figure 6B;  $M = 0.12$ ,  $SE = 0.04$ ,  $t(25) = 3.26$ ,  $p = 0.017$ ], while for the “not-working” group there was no difference [Figure 6C and Supplementary Table 1;  $M = 0.06$ ,  $SE = 0.04$ ,  $t(24) = 1.552$ ,  $p = 0.447$ ]. These results suggest that the focus level differences between Endel and Silence are task-dependent, where audio was particularly beneficial for specific types of tasks, namely, “working.”

We next split the participants into two age groups according to the median age (36 years). We found that for the younger participants (age < 36,  $N = 25$ ), all audio types were superior to silence for producing elevated focus levels [Figure 6E and Supplementary Table 1;  $M = 0.14$ , 0.13, 0.12,  $SE = 0.04$ , 0.03, 0.03,  $t(24) = 3.79$ , 4.49, 3.67,  $p = 0.004$ , 0.001, 0.005 for Endel, Apple, and Spotify, respectively] while for the older participants (Figure 6D; age > 36,  $N = 26$ ), there was no difference between audio and silence. The focus level differences were therefore found to also be age-dependent.

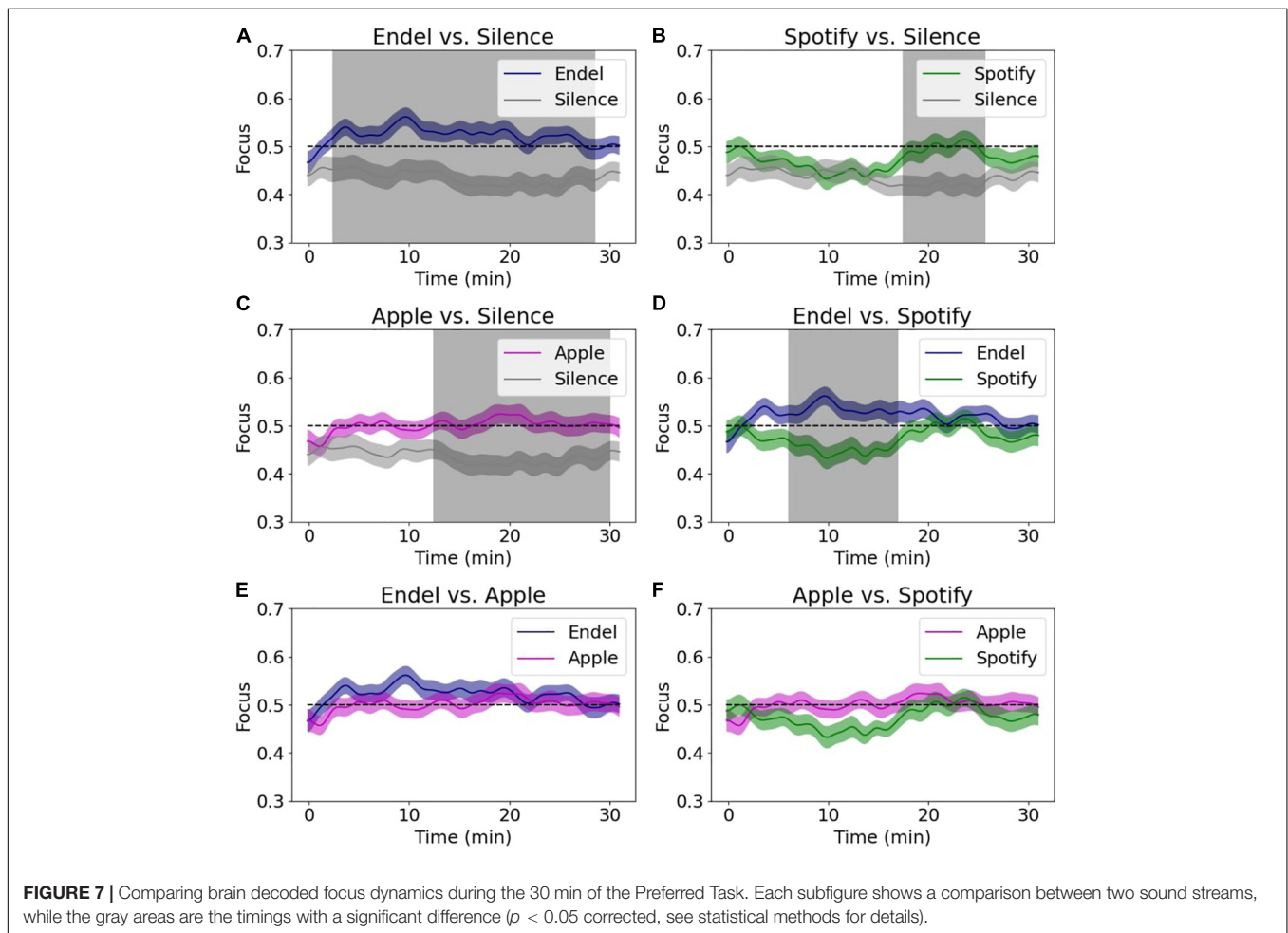
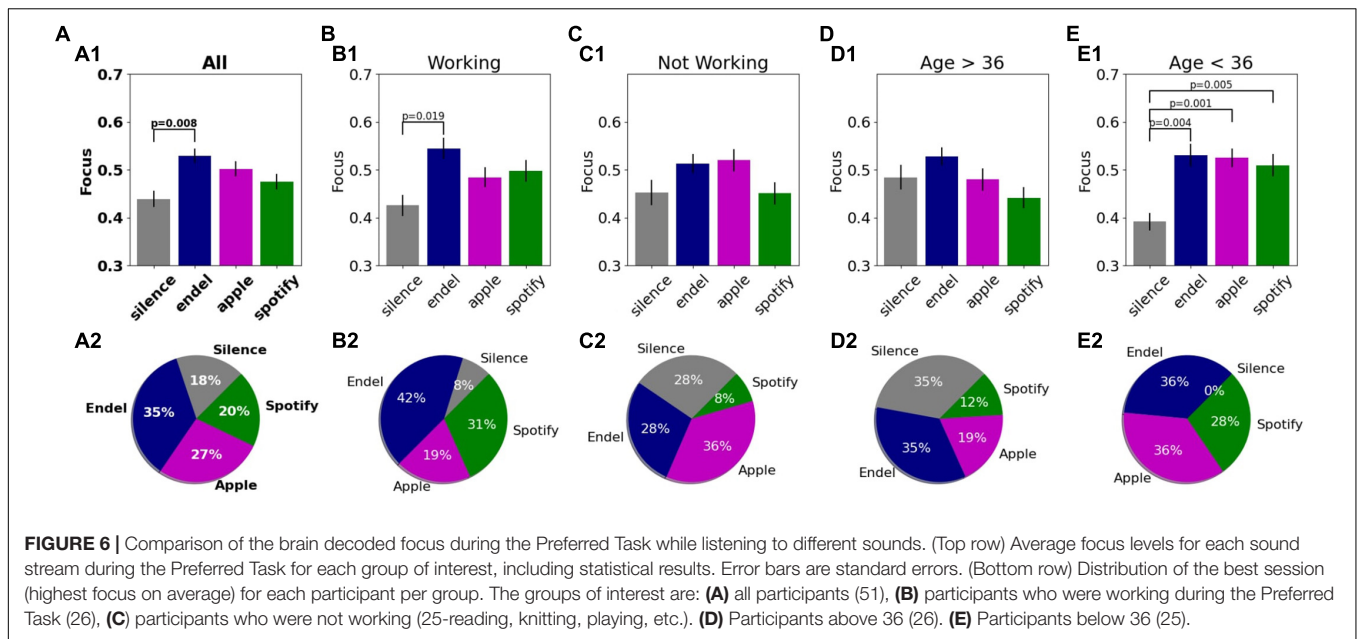
## Time Series Analysis of Focus Dynamics Reveal Differences Between Audio and Silence

Exploiting the high temporal resolution of the focus measurements, we compared the focus dynamics to each audio stream that played during the 30 min of the Preferred Task (Figure 7 and Table 2). When comparing Endel’s soundscapes vs. Silence (Figure 7A), we found that the focus level elicited by Endel’s soundscape was higher 87% of the time, a separation whose significance started after 2.5 min of listening. In addition, although on average there was not a significant difference, the focus level elicited by Apple’s playlist was higher than Silence 60% of the time, starting at 12.5 min (Figure 7C), and the focus level elicited by Spotify’s playlist was higher than Silence 27% of the time, starting at 17 min (Figure 7B). Focus elicited by Endel’s soundscape was higher than Spotify’s playlist in 37% of the time, starting at 6 min (Figure 7D).

## Focus Levels Can Be Predicted by Properties of Audio

Seeing that background sound had an effect on focus levels, we go further and ask whether music and soundscapes can be composed according to a formula to increase focus levels. Meaning, can we understand which audio properties drive focus well enough to predict focus levels from exclusively an analysis of the properties of the sounds within the audio content?

Leveraging the high temporal resolution of the brain measurements, we generated a prediction model which predicts the brain-based focus level from features extracted from the audio signal alone. Raw audio files containing the Apple and Spotify sessions were used to extract different sound properties with a running sliding window of 30 s. The personalized soundscape session (Endel) was not used in this analysis because the real-time streaming did not allow saving the raw audio files that were consistent across participants. We combined multiple audio features to generate an audio data based model that predicts focus levels (see section “Materials and Methods”).





**TABLE 2 |** Summary of focus time dynamics comparison, showing for each pair the percentage of time and time segments with significant difference (where 100% = 30 min).

Pair	Significant difference (% session)	Significant segments (minutes)
Endel-Silence	87	2.5–28
Apple-Silence	60	12.5–30
Spotify-Silence	27	17.5–25.5
Endel-Apple	0	
Endel-Spotify	37	6–17
Spotify-Apple	0	

**Figure 8** shows the audio model performance in predicting the brain decoded focus levels. As explained in section “Materials and Methods,” **Figure 8A** shows the validation correlation and the training correlation for the backward elimination procedure, showing the best correlation on the validation set ( $\langle \text{corr} \rangle = 0.72$ ) is with four PCA components (PC1, PC2, PC9, and PC16). In addition, using only a single component (PC1) yielded a very close result ( $\langle \text{Corr} \rangle = 0.71$ ). The distribution of these correlations can be seen in **Figure 8B**. Comparing the brain decoded focus scores to the average predicted scores of each sample (across models the sample was part of the validation set), yielded a correlation coefficient of  $\text{Corr}(274) = 0.68$ ,  $p < 1e-5$ , and  $\text{Corr}(16) = 0.79$ ,  $p < 1e-4$  when averaging the samples within each song (black points, **Figure 8C**). Using the songs scores, one can apply these prediction models to assemble more successful playlists for enhancing focus based on existing songs. **Figure 8D** shows that if we threshold the sample scores to output a binary prediction (low/high focus), the audio model reaches 87% accuracy in predicting the brain based focus (area under ROC curve = 0.91).

For visualization of the decoded dynamics during the 30 min of the Preferred Task, we next trained the audio model using all 18 songs (without cross validation) and the four PCA components as features and projected it on the Apple (**Figure 9A**) and Spotify (**Figure 9B**) sessions.

## Analysis of Audio Properties Can Be Used to Understand Song Performance

To gain additional insights about the effects that different audio types have on human focus, we used the trained audio model to infer focus values for songs and sounds which were not played during the brain recording experiment. Meaning, we obtained a focus score and dynamics for chosen songs based solely on the properties of the sounds they contained. Here we selected audio examples that challenged the validity of the audio model based on their categorical exclusion from the brain recording experiment. A future approach can include these different genres as controls for further brain measurement validation studies. For example, soundscapes which are not personalized (taken from the playlist: “Focus: Calm Clear Morning”), natural sounds which are commonly used for increasing focus (such as white noise, waves, rain, taken from: <https://mc2method.org/white-noise/>), and popular songs from other music genres (classical music, electronic, pop, rock, jazz, and hip-hop) were used.

org/white-noise/), and popular songs from other music genres (classical music, electronic, pop, rock, jazz, and hip-hop) were used.

**Figure 10A** shows the predicted focus score based on the audio model which took into account only the properties of the audio itself. Songs are sorted from the highest focus evoking song (Endel – Three No Paradoxes) to the lowest (Dr. Dre – What’s The Difference). The top two songs are Endel soundscapes which are not personalized, a finding which strengthens our main result since it implies that the high focus scores elicited by Endel’s soundscape were not solely a byproduct of personalization but also related to the core audio content the personalized compositions were created from. **Figure 10B** shows the sorted focus scores averaged across genres, where notably sounds from classical music and natural sounds contained properties that predicted the highest focus levels. In contrast, pop and hip-hop songs predicted relatively low focus scores. Although we do not have ground truth focus labels for these songs based on real human brain data, given the relatively high scores of the audio which were known to have generated increased focus objectively in the experimental data, we can conclude that there is a consistent validity to the model. Future research can gather ground truth labels for these songs and evaluate the model mathematically in this context.

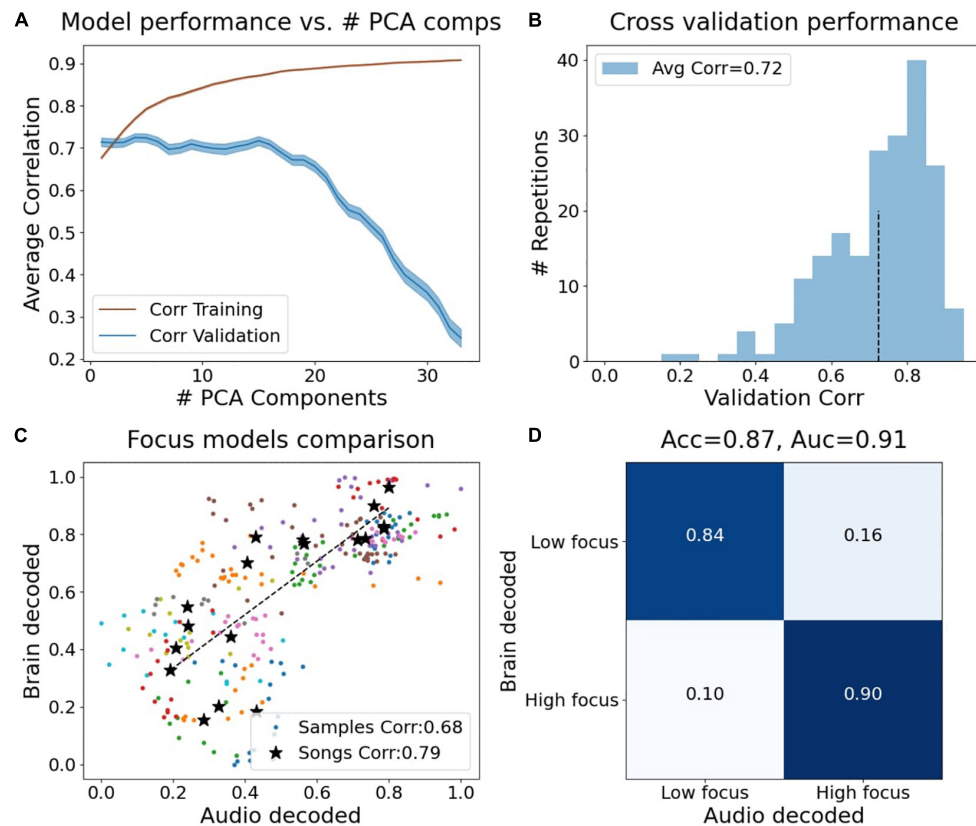
Analyzing the average within-song variance across different genres revealed that the model predicts the largest variance on average for electronic sounds (**Figure 10C**), while the lowest variance was found for natural sounds. The variance can be interpreted as a range of focus dynamics, where the focus dynamics of the electronic sounds are observed to change dramatically during a given song (**Supplementary Figure 5**), confirming the preference for a tool which outputs dynamics with a high temporal resolution when studying such audio content and not the oversimplification of post-song surveys or other low resolution methods. **Figures 10D,E** show the focus dynamics for the song with the lowest focus evoking score and the highest. The dynamics for all songs can be seen in **Supplementary Figure 5**.

## DISCUSSION

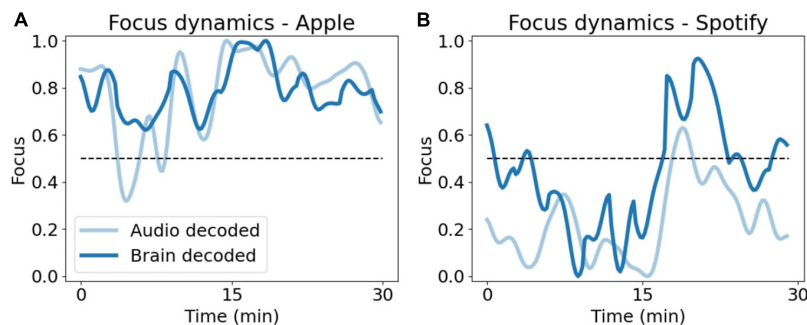
“The soundscape of the world is changing. Modern man is beginning to inhabit a world with an acoustical environment radically different from any he has hitherto known” said the composer R. Murray Schafer, presaging the time we live in now when the sounds available to us continue to multiply by the day. As we have an increasing number of options to modulate our auditory lives by, a handful of take-aways from this study stand out:

### Objective, Brain-Based Measurement of Focus Is Possible in Everyday Environments

Although the effects of audio on the human brain can be subtle in measured brain signals when judging by the changes produced in raw electromagnetic currents, they are robust and highly quantifiable with effectively trained algorithms, as shown here.



**FIGURE 8 |** Results of predicting brain decoded focus from audio features. **(A)** Training and validation correlations vs. number of PCA components used as audio features, using backward elimination in each iteration the component with the smallest weight was eliminated. **(B)** Histogram of the validation correlations using four PCA components (PC1, PC2, PC9, and PC16). The average validation score is 0.72. **(C)** Brain decoded focus (y-axis) vs. audio decoded focus (x-axis) for all samples and average per song. The audio decoded scores were calculated across iterations they were part of the validation songs. **(D)** Confusion matrix after thresholding the focus predictions to classify between low and high focus. Classification accuracy obtained: 88% (area under ROC curve: 0.91).

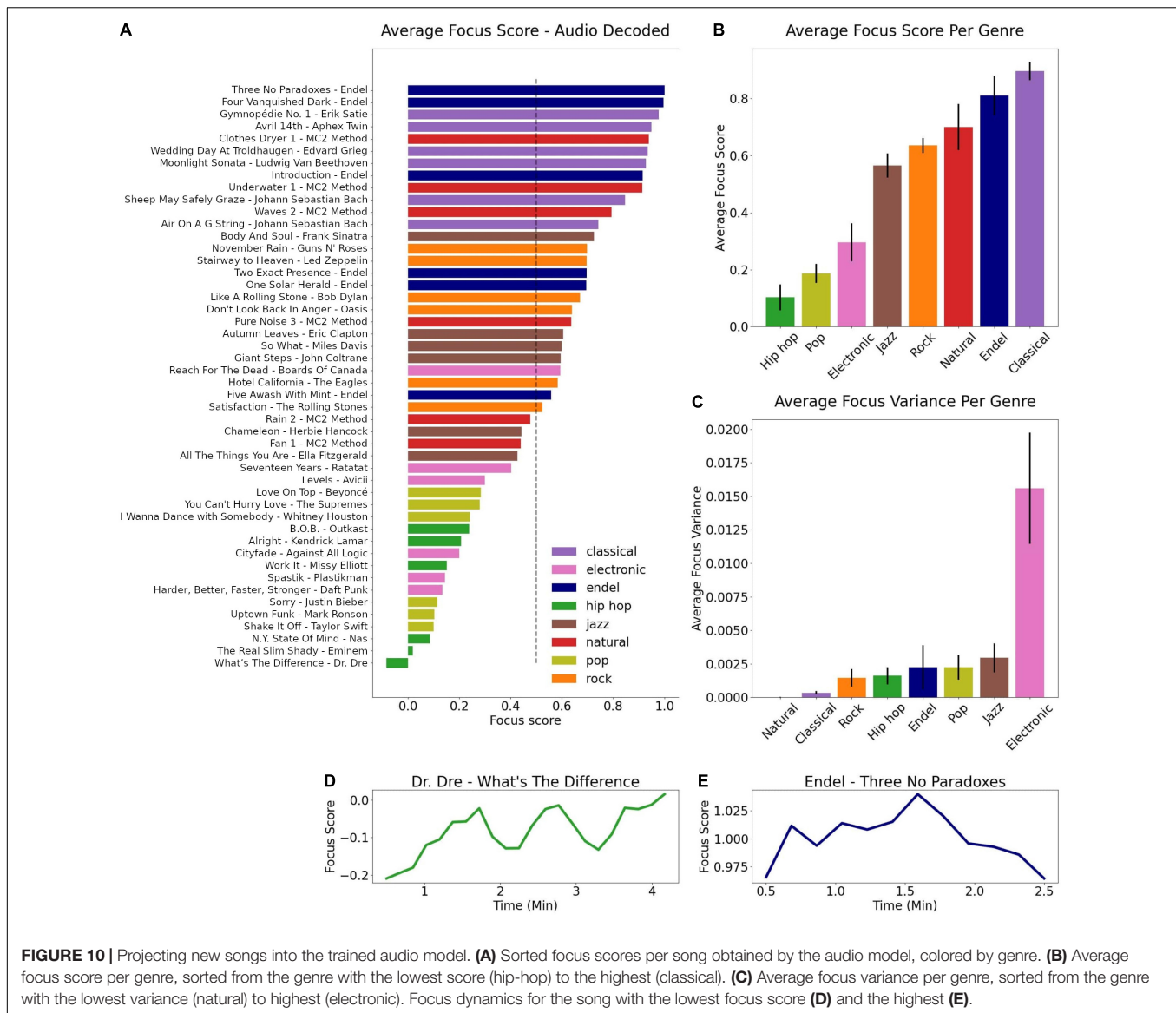


**FIGURE 9 |** Comparison of focus model dynamics. Smoothed dynamics of brain decoded focus (dark blue) and audio decoded focus (light blue), during 30 min of the Preferred Task for Apple **(A)** and Spotify **(B)**. The audio decoded dynamics here was obtained using a model trained on all data (all songs).

Classifying emotional and attentional responses is particularly useful when done at the sub-second temporal resolution since it allows one to track dynamics continuously over time at the same timescale as the brain functions that impact perception and behavior. Furthermore, sub-second resolution into reactions that occur within ecologically valid conditions, like a participant's home, model the real world in an everyday manner that is missing

from experiments that take place within laboratories or other controlled test locations.

In this study we demonstrated that brain decoding algorithms processing data from a non-invasive, consumer brain-computer interface device, are able to deliver sub-second temporal resolution with a high degree of accuracy (approximately 80% match to self-report, **Figure 5C**) at people's own homes.



Since there are inherent biases in subjective self-reporting for experience (Kahneman et al., 1999; Mauss and Robinson, 2009), when mapping physiological signals to self-reported experiences, as done here, there is an upper boundary for accuracy beyond which any model must be deemed to over fit self-reported values and incorrectly represent the information observed in physiological signals. According to a recent review (Larradet et al., 2020) which summarizes multiple peer-reviewed studies that predict self-reported emotions from physiological signals, the average accuracy reported was  $\sim 82\%$ . Given this average and the experimental conditions here – a small number of sensors, at home recordings, simple self-report scales – the achieved accuracy was satisfactory for drawing deeper conclusions on properties of audio since it aligns with state-of-the-art emotion recognition accuracies in the context of audio as a stimulus used elsewhere in controlled laboratory environments (Tripathi

et al., 2017, 81.41 and 73.35% for two classes of Valence and Arousal, respectively).

A key benefit of the current approach is that this method of high temporal resolution brain measurement can be performed reliably outside of traditional laboratories. In this current study not a single laboratory or facility was used for data acquisition. Instead, 18–65 years olds across the United States received a kit in the mail that included a head-wearable device, and they experienced music playlists and personalized soundscapes while they recorded their own brain signals in the comfort of their own home at the times of their choosing. In other words, all participants were in their natural habitat, wearing headphones and a headband that did not interfere with their experience, and they went through the study at their own pace, factors which altogether lend the research a rare degree of ecological validity.

## Focus Is Increased Most by Personalized Soundscapes

Within the at-home environment of this study, personalized, engineered soundscapes were found to be the best at increasing participant focus levels (**Figure 6A**). After 2.5 min, on average, listeners of the personalized soundscapes experienced a meaningful increase in their focus level, while for music playlists it took approximately 15 min to gain a similarly appreciable increase (**Figure 7**). The audio effect on focus levels was found to be task dependent, where soundscapes increased focus levels most in participants who were working (**Figure 6B**). For participants who were not working, no significant difference was found. This result suggests that willful orientation of attention toward work tasks may have created a brain context especially suited to modification by audio. While engaged in work, participants may also have been more prone to distraction and thus more impacted by the positive uplift of audio compared to when engrossed in reading or playing a game which may have contained more intrinsic motivation to stay focused on.

One limitation of this current study is that it did not allow us to disentangle the effects of personalization of sounds on the listener, since pre-recorded soundscapes were not tested. Equivalently, a comparison of personalized soundscapes to personalized music playlists, where audiences either made their own playlist for focus or were allowed to skip songs whenever they wanted, will likely contribute to a more complete understanding of how audio properties correlate with emotion and attention changes. Follow-up research will incorporate these variables. An additional limitation was the inability to reach conclusions regarding gender-dependent effects which was at least partially due to this study's slightly imbalanced data set. Despite efforts to recruit a balanced group of participants, which included even outreach to all genders, enrollment was done on a rolling basis as necessitated by the data collection timeline for the research and ultimately the female subgroup was statistically underpowered in the analysis.

In future research, especially for closed loop, real-time testing, balanced participant sets will be important for reaching more detailed conclusions. Future research should also address whether any effects were introduced by the current study design's sequence of tasks, since here we did not randomize the order of Preferred Task and validation tasks. The Preferred Task was always first and validation tasks after it intentionally, in order to allow for randomization of the background sound stimuli during the Preferred Task session which was done across four groups in this study. In future research it will be helpful to randomize the task order also to compare how different audio affects focus levels on different tasks according to a given task's place within a sequence of tasks.

## Audio Preferences and Focus Effects Vary Between People

It is important to emphasize that the results reported here are audio effects on the average focus levels across a United States based population, and that there was a large variance in this effect between participants. Evidence for this large variance can

be seen in **Supplementary Figure 4** and in the age dependency effect (**Figures 6D,E**), where for the younger audience, all sounds increased focus while for the older audience, the sounds did not have any effect. These results are consistent with other studies showing personal preferences are critical for the improvements possible by audio (Cassidy and Macdonald, 2009; Huang and Shih, 2011; Mori et al., 2014). Due to this variety observed together with the highest focus being elicited by the personalized soundscapes, a next step will include closed-loop selections of sounds, where iterative sound testing is used per person to identify the significant parameters for maximizing focus for that person.

Personalized soundscapes specifically, and personalized audio in general, should be investigated further for their capacity to increase productivity, creativity and well-being as these attributes of human experience are associated with one's ability to focus. For clinical populations as well, for example children with ADHD, the tailoring of sounds for this purpose of increased focus can be particularly impactful. It is possible that the seamlessness of the personalized soundscapes tested here, which played continuously without gaps in the sound like the music playlists had between songs, was also critical part of the observed effect on focus. At every juncture of the experience there is more to be learned, but at a high level, a main finding of this study is that there is a strong need for personalization of audio in order to most effectively achieve functional goals like increasing focus.

## Brain Decoded Focus Data Enabled a New Predictive Model Based on Audio Data Alone

Leveraging the high temporal resolution of the brain decoded dynamics, a focus prediction model based on the physical properties of audio was successfully trained, resulting in an accuracy score of 88% in predicting the brain decoded focus score from an audio decomposition that assessed 136 different properties of sounds as unique features (**Figure 8**). This model enabled a further examination of how sounds and different genres effects focus and allowed testing additional conditions, such as pre-recorded soundscapes and commonly used background sounds (e.g., white noise), as well as other genres (pop, rock, jazz, etc.). We found that the model predicted the highest focus scores for classical music, followed by engineered soundscapes and natural sounds. These results complement previous studies which showed natural sounds and classical music to be beneficial for learning and concentration (Davies, 2000; DeLoach et al., 2015; Angwin et al., 2017; Liu et al., 2021).

In contrast, the models predicted that genres such as pop and hip-hop produce lower focus levels (**Figures 10A,B**). It is possible that these sounds contain more distractors that attract attention away from other objects of attention, or that they contain types of sounds that the brain requires more resources to process (depending on familiar patterns, surprises, and more), leading to less resources available to perform other tasks. Sounds in these genres may also activate the reward system differently (Salimpoor et al., 2015; Gold et al., 2019), which can increase motivation to listen intently to the songs themselves rather than orient toward



other tasks. Understanding the brain mechanisms that underlay the modified focus from these genres is beyond the scope of this current research, but the mapping found here can provide fruitful avenues for future brain imaging experiments that may be equipped to answer these questions.

The analysis here demonstrates a process in which we utilize the temporal resolution of brain-computer interface technology to generate a product where the neurotechnology is eventually out of the loop, resulting in a stand alone audio model which takes as an input a raw audio file and outputs a predicted focus score. This model can be used independently to generate focus playlists or to compose optimal soundscapes, and can further be improved by expanding to populations outside the United States and different age groups. In this way, the current research hearkens back to Pythagoras, who first identified the mathematical connection between a string's length and its pitch and believed the whole cosmos was a form of musical composition (James, 1995). We too see the rich mathematical models obtained in this study, by mapping audio properties to human experience, as a glimpse into the natural laws governing how we feel and think. The better these laws can be understood, the more empowered individuals will be to modulate their environments to suit their goals and states of mind. There remains much to figure out: while we as a species continue to cause a "shift in the sensorium," we simultaneously experience that shift all over daily life and it is not clear where we as a species are headed. This study showed that audio has a distinct effect on our focus levels, and paves the way for designing sounds to help us focus better in the future.

## CONCLUSION

We studied the effects of audio on human focus levels using noninvasive brain decoding technology to gain a better understanding of the optimal audio properties for increasing focus levels in listeners. We combined a custom app ("Neuos Central"), portable fabric EEG-measuring headbands, and brain decoding technology (Neuos SDK) to enable us to obtain high temporal resolution focus dynamics from participants at home. Using the brain decoded focus dynamics, we then analyzed how various properties of audio impacted focus levels in different tasks.

We found that while performing a self-paced task for a long period of time (such as working), personalized soundscapes increased focus the most relative to silence. Curated playlists of pre-recorded songs by Apple and Spotify also increased focus during specific time intervals, especially for the youngest audience demographic. Large variance in response profiles across participants, together with task and age dependent effects, suggest that personalizing audio content in real-time may be the best strategy for producing focus in a given listener.

Finally, we generated an audio property based focus model which successfully predicts brain decoded focus scores from audio file alone as an input. Using this model, we extracted predicted focus scores from new songs based on audio decomposition and performed a genre analysis to develop new

intuitions about the experimental findings and the sources of focus-producing audio content. We found that based on our model, engineered soundscapes and classical music are the best for increasing focus, while pop and hip-hop music are the worst.

The approach taken here can be adapted to include other emotions (e.g., enjoyment, anxiety, happiness, etc.), attentional parameters ("Flow state," memory formation, etc.) and can be used to assess additional content as well (e.g., visual, ambient, olfactory, etc.), including interactive gaming and e-learning experiences where personalization and high temporal resolution measures of brain responses may be especially beneficial.

## DATA AVAILABILITY STATEMENT

The dataset for this study is available through an open Git repository (<https://bit.ly/3HbZd8n>). Data includes the brain decoded focus dynamics for each participant together with scripts that run the statistical tests.

## ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

AH, RK, NB-E, EK, and DF designed the experiment. AH analyzed the data. RK, NB-E, and DF advised on data analysis and statistics. SK, EK, and DF developed the app and software platform for data collection. AH and DF wrote the manuscript. RK and EK revised the manuscript. All authors approved the work for publication.

## FUNDING

This study received funding from Arctop Inc., and Endel Sound GmbH. Endel Sound GmbH was not involved in data collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

## ACKNOWLEDGMENTS

We would like to thank Warner Music, Sony Music, Endel Sound, and Universal Music for providing audio content, data, and support in conducting this study and advancing theoretical aspects of the research.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2021.760561/full#supplementary-material>

## REFERENCES

- Abiri, R., Borhani, S., Jiang, Y., and Zhao, X. (2019). Decoding attentional state to faces and scenes using EEG brainwaves. *Complexity* 2019:e6862031. doi: 10.1155/2019/6862031
- Angwin, A. J., Wilson, W. J., Arnott, W. L., Signorini, A., Barry, R. J., and Copland, D. A. (2017). White noise enhances new-word learning in healthy adults. *Sci. Rep.* 7:13045. doi: 10.1038/s41598-017-13383-3
- Asif, A., Majid, M., and Anwar, S. M. (2019). Human stress classification using EEG signals in response to music tracks. *Comput. Biol. Med.* 107, 182–196.
- Barachant, A. (2017). *Muse LSL [Python]*. Available online at: <https://github.com/alexandrebarachant/muse-lsl> (accessed May 25, 2019).
- Bhatti, A. M., Majid, M., Anwar, S. M., and Khan, B. (2016). Human emotion recognition and analysis in response to audio music using brain signals. *Comput. Hum. Behav.* 65, 267–275. doi: 10.1016/j.chb.2016.08.029
- Bird, J. J., Ekart, A., Buckingham, C. D., and Faria, D. R. (2019). “Mental emotional sentiment classification with an eeg-based brain-machine interface,” in *Proceedings of The International Conference on Digital Image and Signal Processing (DISP'19)*, Oxford University, UK.
- Broday-Dvir, R., Grossman, S., Furman-Haran, E., and Malach, R. (2018). Quenching of spontaneous fluctuations by attention in human visual cortex. *Neuroimage* 171, 84–98. doi: 10.1016/j.neuroimage.2017.12.089
- Brotzer, J. M., Mosqueda, E. R., and Gorro, K. (2019). Predicting emotion in music through audio pattern analysis. *IOP Conf. Ser. Mater. Sci. Eng.* 482:012021. doi: 10.1088/1757-899X/482/1/012021
- Cassidy, G., and Macdonald, R. (2009). The effects of music choice on task performance: a study of the impact of self-selected and experimenter-selected music on driving game performance and experience. *Music. Sci.* 13, 357–386. doi: 10.1177/102986490901300207
- Chanda, M. L., and Levitin, D. J. (2013). The neurochemistry of music. *Trends Cogn. Sci.* 17, 179–193. doi: 10.1016/j.tics.2013.02.007
- Cheung, V. K. M., Harrison, P. M. C., Meyer, L., Pearce, M. T., Haynes, J.-D., and Koelsch, S. (2019). Uncertainty and surprise jointly predict musical pleasure and amygdala, hippocampus, and auditory cortex activity. *Curr. Biol.* 29, 4084.e–4092.e. doi: 10.1016/j.cub.2019.09.067
- Chou, P. T.-M. (2010). Attention drainage effect: how background music effects concentration in Taiwanese college students. *J. Scholarsh. Teach. Learn.* 10, 36–46.
- Cowen, A. S., and Keltner, D. (2017). Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proc. Natl. Acad. Sci. U.S.A.* 114, E7900–E7909.
- Cunningham, S., Ridley, H., Weinel, J., and Picking, R. (2020). Supervised machine learning for audio emotion recognition. *Pers. Ubiquit. Comput.* 25, 637–650.
- Davies, M. A. (2000). Learning ... the beat goes on. *Childh. Educ.* 76, 148–153. doi: 10.1080/00094056.2000.10522096
- Davis, W. B., and Thaut, M. H. (1989). The influence of preferred relaxing music on measures of state anxiety, relaxation, and physiological responses I. *J. Music Ther.* 26, 168–187. doi: 10.1093/jmt/26.4.168
- de la Mora Velasco, E., and Hirumi, A. (2020). The effects of background music on learning: a systematic review of literature to guide future research and practice. *Educ. Technol. Res. Dev.* 68, 2817–2837. doi: 10.1007/s11423-020-09783-4
- DeLoach, A. G., Carter, J. P., and Braasch, J. (2015). Tuning the cognitive environment: sound masking with “natural” sounds in open-plan offices. *J. Acoust. Soc. Am.* 137, 2291–2291. doi: 10.1121/1.4920363
- Faller, J., Cummings, J., Saproo, S., and Sajda, P. (2019). Regulation of arousal via online neurofeedback improves human performance in a demanding sensory-motor task. *Proc. Natl. Acad. Sci. U.S.A.* 116, 6482–6490.
- Gao, C., Fillmore, P., and Scullin, M. K. (2020). Classical music, educational learning, and slow wave sleep: a targeted memory reactivation experiment. *Neurobiol. Learn. Mem.* 171:107206. doi: 10.1016/j.nlm.2020.107206
- Giannakopoulos, T. (2015). pyAudioAnalysis: an open-source python library for audio signal analysis. *PLoS One* 10:e0144610. doi: 10.1371/journal.pone.0144610
- Gold, B. P., Mas-Herrero, E., Zeighami, Y., Benovoy, M., Dagher, A., and Zatorre, R. J. (2019). Musical reward prediction errors engage the nucleus accumbens and motivate learning. *Proc. Natl. Acad. Sci. U.S.A.* 116, 3310–3315. doi: 10.1073/pnas.1809855116
- González, V. M., Robbes, R., Góngora, G., and Medina, S. (2015). “Measuring concentration while programming with low-cost BCI devices: differences between debugging and creativity tasks,” in *Foundations of Augmented Cognition*, eds D. D. Schmorrow and C. M. Fidopiastis (New York, NY: Springer International Publishing), 605–615. doi: 10.1007/978-3-319-20816-9\_58
- Haartsen, R., van der Velde, B., Jones, E. J. H., Johnson, M. H., and Kemner, C. (2020). Using multiple short epochs optimises the stability of infant EEG connectivity parameters. *Sci. Rep.* 10:12703. doi: 10.1038/s41598-020-68981-5
- Hallam, S., Price, J., and Katsarou, G. (2002). The effects of background music on primary school pupils' Task Performance. *Educ. Stud.* 28, 111–122. doi: 10.1080/03055690220124551
- Hamadicharef, B., Zhang, H., Guan, C., Wang, C., Phua, K. S., Tee, K. P., et al. (2009). “Learning EEG-based spectral-spatial patterns for attention level measurement,” in *Proceedings of the 2009 IEEE International Symposium on Circuits and Systems*, Taipei, Taiwan. 1465–1468.
- Hizlisoy, S., Yildirim, S., and Tufekci, Z. (2021). Music emotion recognition using convolutional long short term memory deep neural networks. *Eng. Sci. Technol. Int. J.* 24, 760–767. doi: 10.1016/j.jestech.2020.10.009
- Hu, X., and Yang, Y.-H. (2017). Cross-dataset and cross-cultural music mood prediction: a case on western and Chinese pop songs. *IEEE Trans. Affect. Comput.* 8, 228–240. doi: 10.1109/TAFFC.2016.2523503
- Huang, R.-H., and Shih, Y.-N. (2011). Effects of background music on concentration of workers. *Work* 38, 383–387. doi: 10.3233/WOR-2011-1141
- Huron, D. B. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press.
- James, J. (1995). *The Music of the Spheres: Music, Science, and the Natural Order of the Universe*. Copernicus. Available online at: <https://www.springer.com/gp/book/9780387944746> (accessed July 5, 2021).
- Jung, T.-P., Makeig, S., Stensmo, M., and Sejnowski, T. J. (1997). Estimating alertness from the EEG power spectrum. *IEEE Trans. Biomed. Eng.* 44, 60–69.
- Kahneman, D., Diener, E., and Schwarz, N. (1999). *Well-being: Foundations Of Hedonic Psychology*. New York, NY: Russell Sage Foundation.
- Kumar, S., von Kriegstein, K., Friston, K., and Griffiths, T. D. (2012). Features versus feelings: dissociable representations of the acoustic features and valence of aversive sounds. *J. Neurosci.* 32, 14184–14192.
- Larradet, F., Niewiadomski, R., Barresi, G., Caldwell, D. G., and Mattos, L. S. (2020). Toward emotion recognition from physiological signals in the wild: approaching the methodological issues in real-life data collection. *Front. Psychol.* 11:1111. doi: 10.3389/fpsyg.2020.01111
- Larsen, R. J., and Diener, E. (1992). *Promises And Problems With The Circumplex Model Of Emotion*. Thousand Oaks, CA: Sage Publications.
- Levitin, D. J. (2006). *This Is Your Brain On Music: The Science Of A Human Obsession*. London: Penguin.
- Levitin, D. J., Chordia, P., and Menon, V. (2012). Musical rhythm spectra from Bach to Joplin obey a 1/f power law. *Proc. Natl. Acad. Sci. U.S.A.* 109, 3716–3720. doi: 10.1073/pnas.1113828109
- Lin, Y.-P., Jao, P.-K., and Yang, Y.-H. (2017). Improving cross-day EEG-based emotion classification using robust principal component analysis. *Front. Comput. Neurosci.* 11:64. doi: 10.3389/fncom.2017.00064
- Liu, H., He, H., and Qin, J. (2021). Does background sounds distort concentration and verbal reasoning performance in open-plan office? *Appl. Acoust.* 172:107577. doi: 10.1016/j.apacoust.2020.107577
- Liu, N.-H., Chiang, C.-Y., and Chu, H.-C. (2013). Recognizing the degree of human attention using EEG signals from mobile sensors. *Sensors* 13, 10273–10286. doi: 10.3390/s130810273
- Mauss, I. B., and Robinson, M. D. (2009). Measures of emotion: a review. *Cogn. Emot.* 23, 209–237. doi: 10.1080/02699930802204677
- Micoulaud-Franchi, J.-A., Geoffroy, P. A., Fond, G., Lopez, R., Bioulac, S., and Philip, P. (2014). EEG neurofeedback treatments in children with ADHD: an updated meta-analysis of randomized controlled trials. *Front. Hum. Neurosci.* 8:906. doi: 10.3389/fnhum.2014.00906
- Mori, F., Naghsh, F. A., and Tezuka, T. (2014). “The effect of music on the level of mental concentration and its temporal change,” in *Proceedings of the 6th International Conference on Computer Supported Education*, Setubal, Portugal, 34–42. doi: 10.5220/0004791100340042
- Nia, H. T., Jain, A. D., Liu, Y., Alam, M.-R., Barnas, R., and Makris, N. C. (2015). The evolution of air resonance power efficiency in the violin and its ancestors. *Proc. R. Soc. A Math., Phys. Eng. Sci.* 471:20140905. doi: 10.1098/rspa.2014.0905

- Padilla-Buritica, J. I., Ferrandez-Vicente, J. M., Castaño, G. A., and Acosta-Medina, C. D. (2020). Non-Stationary group-level connectivity analysis for enhanced interpretability of oddball tasks. *Front. Neurosci.* 14:446. doi: 10.3389/fnins.2020.00446
- Perez-Valero, E., Vaquero-Blasco, M. A., Lopez-Gordo, M. A., and Morillas, C. (2021). Quantitative assessment of stress through EEG during a virtual reality stress-relax session. *Front. Comput. Neurosci.* 15:684423.
- Pope, A. T., Bogart, E. H., and Bartolome, D. S. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biol. Psychol.* 40, 187–195. doi: 10.1016/0301-0511(95)05116-3
- Rebolledo-Mendez, G., Dunwell, I., Martínez-Mirón, E. A., Vargas-Cerdán, M. D., de Freitas, S., Liarokapis, F., et al. (2009). "Assessing neurosky's usability to detect attention levels in an assessment exercise," in *Human-Computer Interaction. New Trends*, ed. J. A. Jacko (New York, NY: Springer), 149–158. doi: 10.1007/978-3-642-02574-7\_17
- Sacks, O. (2010). *Musicophilia: Tales Of Music And The Brain*. New York, NY: Vintage.
- Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., and McIntosh, A. R. (2015). Predictions and the brain: how musical sounds become rewarding. *Trends Cogn. Sci.* 19, 86–91. doi: 10.1016/j.tics.2014.12.001
- Schreiber, C. A., and Kahneman, D. (2000). Determinants of the remembered utility of aversive sounds. *J. Exp. Psychol. Gen.* 129:27.
- Shahabi, H., and Moghimi, S. (2016). Toward automatic detection of brain responses to emotional music through analysis of EEG effective connectivity. *Comput. Hum. Behav.* 58, 231–239. doi: 10.1016/j.chb.2016.01.005
- Shih, Y.-N., Huang, R.-H., and Chiang, H.-Y. (2012). Background music: effects on attention performance. *Work* 42, 573–578. doi: 10.3233/WOR-2012-1410
- Tripathi, S., Acharya, S., Sharma, R. D., Mittal, S., and Bhattacharya, S. (2017). "Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset," in *Proceedings of the Twenty-Ninth IAAI Conference*, San Francisco, California, USA.
- Tuckute, G., Hansen, S. T., Kjaer, T. W., and Hansen, L. K. (2021). Real-Time decoding of attentional states using closed-loop EEG neurofeedback. *Neural Comput.* 33, 967–1004. doi: 10.1162/neco\_a\_01363
- Vempala, N. N., and Russo, F. A. (2012). "Predicting emotion from music audio features using neural networks," in *Proceedings of the 9th International Symposium on Computer Music Modeling and Retrieval (CMMR)*, Queen Mary University of London, 336–343.
- Washburne, C. (2020). "More Cowbell": *Latin Jazz In The Twenty-First Century In Latin Jazz*. Oxford: Oxford University Press. doi: 10.1093/oso/9780195371628.003.0007
- Yang, Y., Ahmadipour, P., and Shanechi, M. M. (2021). Adaptive latent state modeling of brain network dynamics with real-time learning rate optimization. *J. Neural Eng.* 18:036013. doi: 10.1088/1741-2552/abcefd
- Yang, Y.-H., Lin, Y.-C., Su, Y.-F., and Chen, H. H. (2008). A regression approach to music emotion recognition. *IEEE Trans. Audio Speech Lang. Process.* 16, 448–457. doi: 10.1109/TASL.2007.911513
- Zald, D. H., and Pardo, J. V. (2002). The neural correlates of aversive auditory stimulation. *Neuroimage* 16, 746–753.

**Conflict of Interest:** The authors were employed by the company Arctop Inc. This study received funding from Arctop Inc., and Endel Sound GmbH. The funders had the following involvement with the study: Arctop Inc., was involved in the study design, collection, analysis, interpretation of data, the writing of this article, and the decision to submit it for publication. Endel Sound GmbH was involved in the study design and provided audio stimuli used in the experiment.

**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.

Copyright © 2022 Haruvi, Kopito, Brande-Eilat, Kalev, Kay and Furman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



# Is Mate Preference Recognizable Based on Electroencephalogram Signals? Machine Learning Applied to Initial Romantic Attraction

Guangjie Yuan<sup>1,2,3</sup>, Wenguang He<sup>4</sup> and Guangyuan Liu<sup>1,2,3,5\*</sup>

<sup>1</sup> College of Electronic and Information Engineering, Southwest University, Chongqing, China, <sup>2</sup> Institute of Affective Computing and Information Processing, Southwest University, Chongqing, China, <sup>3</sup> Chongqing Key Laboratory of Nonlinear Circuits and Intelligent Information Processing, Southwest University, Chongqing, China, <sup>4</sup> College of Psychology, Qufu Normal University, Qufu, China, <sup>5</sup> Key Laboratory of Cognition and Personality, Ministry of Education, Faculty of Psychology, Southwest University, Chongqing, China

## OPEN ACCESS

### Edited by:

Jane Zhen Liang,  
Shenzhen University, China

### Reviewed by:

Hugo F. Posada-Quintero,  
University of Connecticut,  
United States  
Onder Aydemir,  
Karadeniz Technical University, Turkey  
Noor Kamal Al-Qazzaz,  
University of Baghdad, Iraq

### \*Correspondence:

Guangyuan Liu  
liugy@swu.edu.cn

### Specialty section:

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

**Received:** 07 December 2021

**Accepted:** 24 January 2022

**Published:** 11 February 2022

### Citation:

Yuan G, He W and Liu G (2022) Is  
Mate Preference Recognizable Based  
on Electroencephalogram Signals?  
Machine Learning Applied to Initial  
Romantic Attraction.  
*Front. Neurosci.* 16:830820.  
doi: 10.3389/fnins.2022.830820

Initial romantic attraction (IRA) refers to a series of positive reactions toward potential ideal partners based on individual preferences; its evolutionary value lies in facilitating mate selection. Although the EEG activities associated with IRA have been preliminarily understood; however, it remains unclear whether IRA can be recognized based on EEG activity. To clarify this, we simulated a dating platform similar to Tinder. Participants were asked to imagine that they were using the simulated dating platform to choose the ideal potential partner. Their brain electrical signals were recorded as they viewed photos of each potential partner and simultaneously assessed their initial romantic attraction in that potential partner through self-reported scale responses. Thereafter, the preprocessed EEG signals were decomposed into power-related features of different frequency bands using a wavelet transform approach. In addition to the power spectral features, feature extraction also accounted for the physiological parameters related to hemispheric asymmetries. Classification was performed by employing a random forest classifier, and the signals were divided into two categories: IRA engendered and IRA unengendered. Based on the results of the 10-fold cross-validation, the best classification accuracy 85.2% (SD = 0.02) was achieved using feature vectors, mainly including the asymmetry features in alpha (8–13 Hz), beta (13–30 Hz), and theta (4–8 Hz) rhythms. The results of this study provide early evidence for EEG-based mate preference recognition and pave the way for the development of EEG-based romantic-matching systems.

**Keywords:** aesthetic preference, mate choice, physiological signals, frequency band, hemispheric asymmetries

## INTRODUCTION

Finding an ideal partner is a prerequisite for achieving high-quality romantic relationships. However, finding an ideal partner in real life can be extremely challenging (Spielmann et al., 2013; Joel et al., 2017). Because mate selection is not only a multivariate process involving the integration and trade-offs of multiple preferences but is also influenced by many factors, such as gender, culture, and personal experience (Buston and Emlen, 2003; Thomas et al., 2020). However, opportunities



always coexist with these challenges. This is precisely because of the challenge of this task, which has created a huge economic market for matchmaking services (Joel et al., 2017). In this market, matchmaking agencies strive to provide customers with “tailored” romantic matching services and earn huge returns on this. The success of such a business model hinges on finding key features from appropriate signals that can effectively identify a user’s initial romantic interest toward a potential partner, as this largely determines the effectiveness of a matching service and consequently whether a user is willing to pay for it (Joel et al., 2017).

The current mainstream approach taken by matchmaking companies is that when users register for romantic matching services, they are required to fill in a series of questionnaires about their own characteristics and preferences based on their subjective feelings; these answers will then be fed into the matching algorithm as features to match suitable potential partners for users. Many matchmaking companies claim that effective romantic pairing can be achieved in this manner (Joel et al., 2017). However, Joel et al. (2017) demonstrated that it was impossible to predict initial romantic desire using any combination of traits and preferences reported prior to dating. In other words, effective romantic pairing cannot be achieved using this method. For matchmaking companies that take this as the core selling point, this conclusion is undoubtedly very destructive. However, from the perspective of psychology, this conclusion is undoubtedly reasonable, because the self-reported data are easily affected by subjective consciousness and the surrounding environment, which makes many characteristics of the input matching algorithm invalid, thereby invalidating the matching algorithm (Lin et al., 2010; Alarcao and Fonseca, 2019).

The essence of initial romantic attraction (IRA) is a series of positive responses to potential ideal partners based on individual preferences, including positive emotional responses (such as feelings of exhilaration and craving for emotional union) (Fisher, 1998; Fisher et al., 2002, 2005; Gerlach and Reinhard, 2018; Yuan and Liu, 2021; Yuan et al., 2021). An individual’s internal emotional reaction can be revealed not only through subjective self-reports but also through internal expression (i.e., physiological signals) (Gunes et al., 2011; Alarcao and Fonseca, 2019). Moreover, physiological signals have many advantages over self-reported data, one of which is that they are less susceptible to subjective consciousness and environmental factors (Lin et al., 2010; Alarcao and Fonseca, 2019). Thus, these signals open up new possibilities for identifying users’ emotional responses and preferences for potential partners. For instance, Zhang et al. (2021) successfully identified participants’ initial romantic interest to potential partners based on the features extracted from electrocardiogram signals, while Lu et al. (2020) successfully detected participants’ initial romantic desire to potential romantic partners based on the information extracted from photoplethysmogram signals. These results demonstrate that IRA, as an important part of human emotion, can be recognized on the basis of periphery physiological signals (Lu et al., 2020; Zhang et al., 2021).

In addition to periphery physiological signals, signals captured from the central nervous system, such as EEG, functional magnetic resonance imaging, or positron emission tomography, have also been proved to provide informative information for emotion recognition (Lin et al., 2010; Alarcao and Fonseca, 2019). Furthermore, among the many biosignals recorded over the brain, EEG is considered to a preferred method in studying the brain’s response to emotional stimuli due to its characters of high temporal resolution, non-invasive, inexpensive and convenient (Niemic, 2004; Alarcao and Fonseca, 2019). Therefore, in the field of neurophysiology, some studies have begun to investigate brain activities associated with IRA based on EEG signals. For instance, using event-related potential source analysis, Yuan et al. (2021) found that the arousal of IRA will significantly enhance the activation intensity of emotional processing-related areas, including the orbital frontal cortex and insula; attention control-related areas, including the frontal eye field and cingulate cortex; visual processing-related areas; and social evaluation-related areas, including the left dorsolateral prefrontal cortex. In another study, Yuan and Liu (2021) used time–frequency (TF) decomposition technology and found that processing of individual face preferences that triggered IRA was associated with a decrease in power in the alpha and lower beta bands over the lateral occipital complex and vertex areas; they hypothesized that changes in alpha and beta power may reflect cortical activation related to emotional stimulus significance (Schubring and Schupp, 2019, 2021). In addition, numerous neuropsychological studies have demonstrated that the asymmetry between the two hemispheres of the frequency band (FB) (especially the alpha and beta bands) was correlated with emotional activities and preferences (Balconi and Mazza, 2009; Liu et al., 2011; Hadjidimitriou and Hadjileontiadis, 2012; Huang et al., 2012; Jatupaiboon et al., 2013; Alarcao and Fonseca, 2019).

In the field of neuroeconomics, although EEG signals have not been used to identify users’ emotional responses and preferences toward potential partners, they have been widely used to identify users’ emotional responses and preferences to other stimuli (Aldayel et al., 2020a,b,c, 2021; Khurana et al., 2021; Naser and Saha, 2021; Zheng et al., 2021). Among previous studies, many researchers have used frequency bands (FBs) as features (Aldayel et al., 2020c, 2021; Khurana et al., 2021; Naser and Saha, 2021; Zheng et al., 2021). For example, Chew et al. (2016) measured the preference of virtual three-dimensional shapes using band power as a feature for two preference categories and obtained accuracies of up to 80%. Aldayel et al. (2020b) measured the preference of consumer using frequency bands features as the feature for two preference categories and obtained accuracies of up to 93%. Meanwhile, several studies on preference also used hemispheric asymmetry scores (ASs) as input features (Aldayel et al., 2020a,b; Naser and Saha, 2021). For instance, Hadjidimitriou and Hadjileontiadis (2013) measured the preference of music using band power and hemispheric ASs as features for two preference categories using the k-nearest neighbors to obtain accuracies of up to 86.52%. Moon measured the preference of visual stimuli using band power and hemispheric ASs as features for four preference categories, achieving accuracies of up to 97.39% (Moon, 2013; Chew et al., 2016).

Although EEG signals have been widely used to identify users' emotional responses and preferences to other stimuli, and EEG activities associated with IRA have also been preliminarily understood, whether users' emotional responses and preferences toward potential partners can be identified on the basis of EEG signals remains unclear. To clarify this, we simulated a mate selection platform similar to Tinder. Participants were asked to imagine that they used the platform to select potentially desirable partners. Their EEG signals were recorded when they viewed and rated the photographs of each potential partner according to their preferences. Specifically, during the EEG recording task, the heterosexual participants were asked to rate photos of opposite-sex potential partners on two dimensions: an four-point IRA rating scale (based on the question "How much would you like to date this person?"; response: "not at all," "a little," "somewhat," or "very much") as well as a three-point zero-acquaintance rating scale (based on the question "Have you ever seen the person in the photo before?"; responses: "no," "not sure," or "yes") (Yuan and Liu, 2021; Yuan et al., 2021). The IRA scale was used to assess the romantic interest of participants toward potential romantic partners, because the desire for emotional union with potential partner is one of the main characteristics of initial romantic attraction arousal. The zero acquaintance scale was used to ensure that participants were at the same level of familiarity with the stimulus material. Numerous studies have demonstrated that the random forest classifier (RFC) performs well in preference classification tasks based on EEG signals; therefore, in this study, the RFC was used to classify and detect the users' IRA toward potential partners based on features obtained through TF analyses.

## MATERIALS AND METHODS

Both the auxiliary experiment and the main experiment were approved by the Ethical Review Committee of Southwest University.

### Auxiliary Experiment Participants

Sixty student volunteers participated in the auxiliary experiment (30 women and 30 men; age:  $21.4 \pm 2.6$  years). All participants reported normal or corrected-to-normal visual acuity and had no history of psychiatric or neurological disorders, as confirmed *via* a screening interview.

### Experimental Procedure

The induction rate of IRA has been reported to be quite low (only a few percent) (Zsok et al., 2017), the IRA induction rate should be increased to obtain enough data to train the model (Yuan and Liu, 2021; Yuan et al., 2021). Numerous studies have shown that physical attractiveness is a good predictor of an individual's popularity (i.e., probability of being selected by the opposite sex) with the opposite sex (Asendorpf et al., 2011; Olderbak et al., 2017; Gerlach and Reinhard, 2018; Yuan et al., 2021). Therefore, in this study, we planned to increase the average induction rate

of IRA by increasing the proportion of stimuli with high physical attractiveness (Yuan and Liu, 2021; Yuan et al., 2021).

To achieve this goal, we first assessed the attractiveness level of each stimulus. To assess the attractiveness level, we first focused on downloading thousands of high-resolution personal portrait photographs from a high-definition copyright commercial photograph library (i.e., Hummingbird<sup>1</sup>) and standardized them (face and hair only; size,  $839 \times 1,080$  pixels). To control the interference factors, we then selected 1,600 photographs from the standardized portrait photograph library; the criteria for screening the photographs were similar orientation and expression of the face and comparable background complexity. Thereafter, the physical attractiveness level of each face was assessed using a nine-point Likert scale. Notably, the male participants rated only female faces, while the female participants rated only male faces. We then calculated the average attractiveness level of each face by averaging the ratings of the same face from 30 participants of the opposite sex. Finally, according to the average attractiveness level, these faces were divided into three categories: high attractiveness [mean = 6.9, standard deviation (SD) = 0.33], medium attractiveness (mean = 5.2, SD = 0.25), and low attractiveness (mean = 3.9, SD = 0.31).

In the natural environment, the proportion of individuals with high, medium, and low attractiveness should conform to the normal distribution. However, in this study, we deliberately increased the proportion of individuals with high attractiveness, reduced the proportion of stimuli with low attractiveness, and adjusted the ratio of high, medium, and low attractiveness to 0.25:0.6:0.15 to increase the average induction rate of IRA. The number of times each participant would need to be exposed to different stimuli was determined to be between 300 and 400 after the comprehensive trade-offs of induction efficiency and participant burden. Ultimately, 360 photographs were selected as the stimulus material for the main experiment for each sex from among 800 photographs of women and 800 photographs of men (Yuan and Liu, 2021; Yuan et al., 2021).

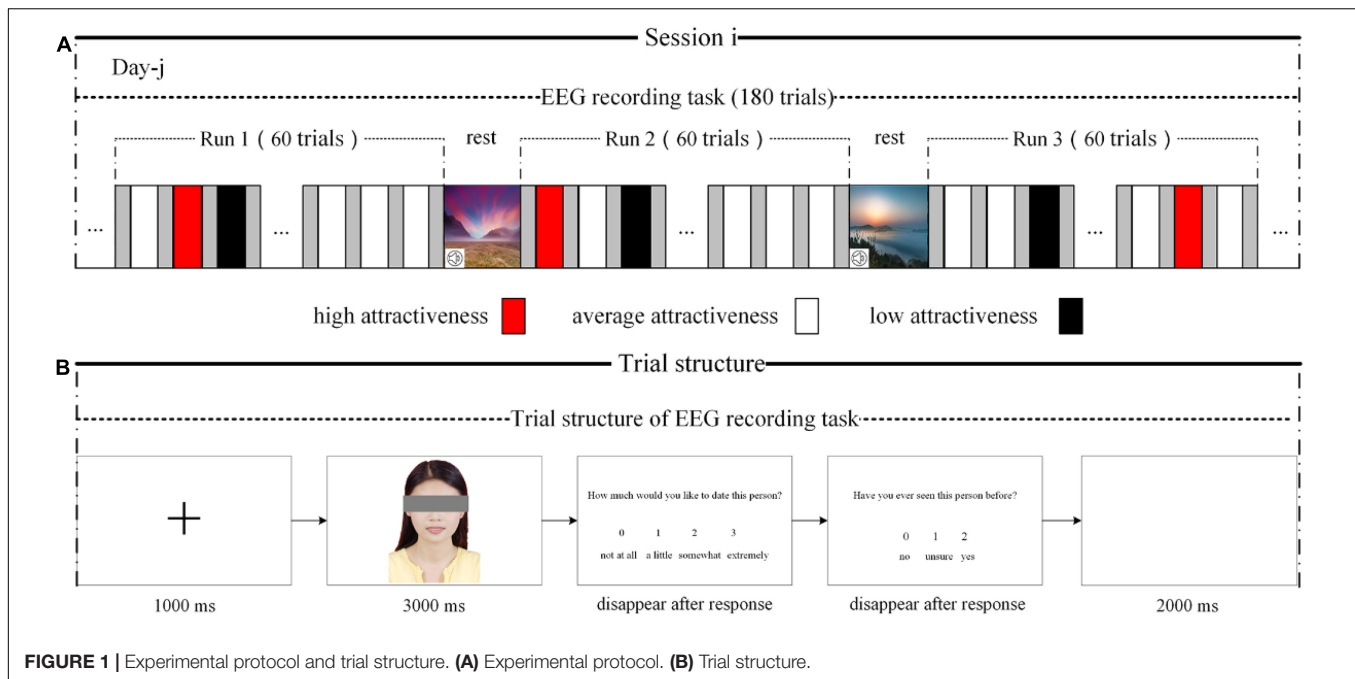
### Main Experiment Participants

Fifty student volunteers participated in the main experiment (all single; 25 women and 25 men; age:  $21.2 \pm 2.4$  years). All participants reported normal or corrected-to-normal visual acuity and had no history of psychiatric or neurological disorders, as confirmed *via* a screening interview.

### Experimental Procedure

The number of stimuli used in the main experiment was significantly reduced by the aforementioned strategy; however, processing of 360 stimuli was still a high-load task for the participants. Specifically, when the participants were asked to complete the task over a short period, they were more likely to experience aesthetic fatigue, which may interfere with the experimental effect. Therefore, to minimize the probability of or delay aesthetic fatigue, we first divided 360 photographs

<sup>1</sup><http://bbs.fengniao.com/>



**FIGURE 1 |** Experimental protocol and trial structure. **(A)** Experimental protocol. **(B)** Trial structure.

of women (or men) equally into two sessions based on their attractiveness level and stipulated that the interval between completing the two parts of the experiment should be at least 1 day (**Figure 1A**). Thereafter, the 180 photographs from each session were divided equally into three runs using the same rules, and a 5–6-min break was provided between every two runs. During the rest period, the participants viewed serene landscapes while listening to soothing music. Notably, the experiment was conducted in a dark and quiet environment to keep the participants focused on the stimulus.

For each session, the trial structure of the EEG recording task is shown in **Figure 1B**. A black fixation cross appeared in the center of a white computer screen for 1,000 ms, followed by a photograph appearing for 3,000 ms. The participants were then asked to assess their romantic interest toward the potential partner based on the question, “How much would you like to date this person?” on a four-point rating scale (0 = not at all; 1 = little; 2 = somewhat; 4 = very much) (Finkel et al., 2007; Cooper et al., 2012; Gerlach and Reinhard, 2018; Yuan and Liu, 2021; Yuan et al., 2021). Thereafter, they were asked, “Have you ever seen the person in the photograph before?” (0 = no; 1 = not sure; 2 = yes). Finally, there was a 2,000-ms blank screen.

### Data Acquisition and Processing

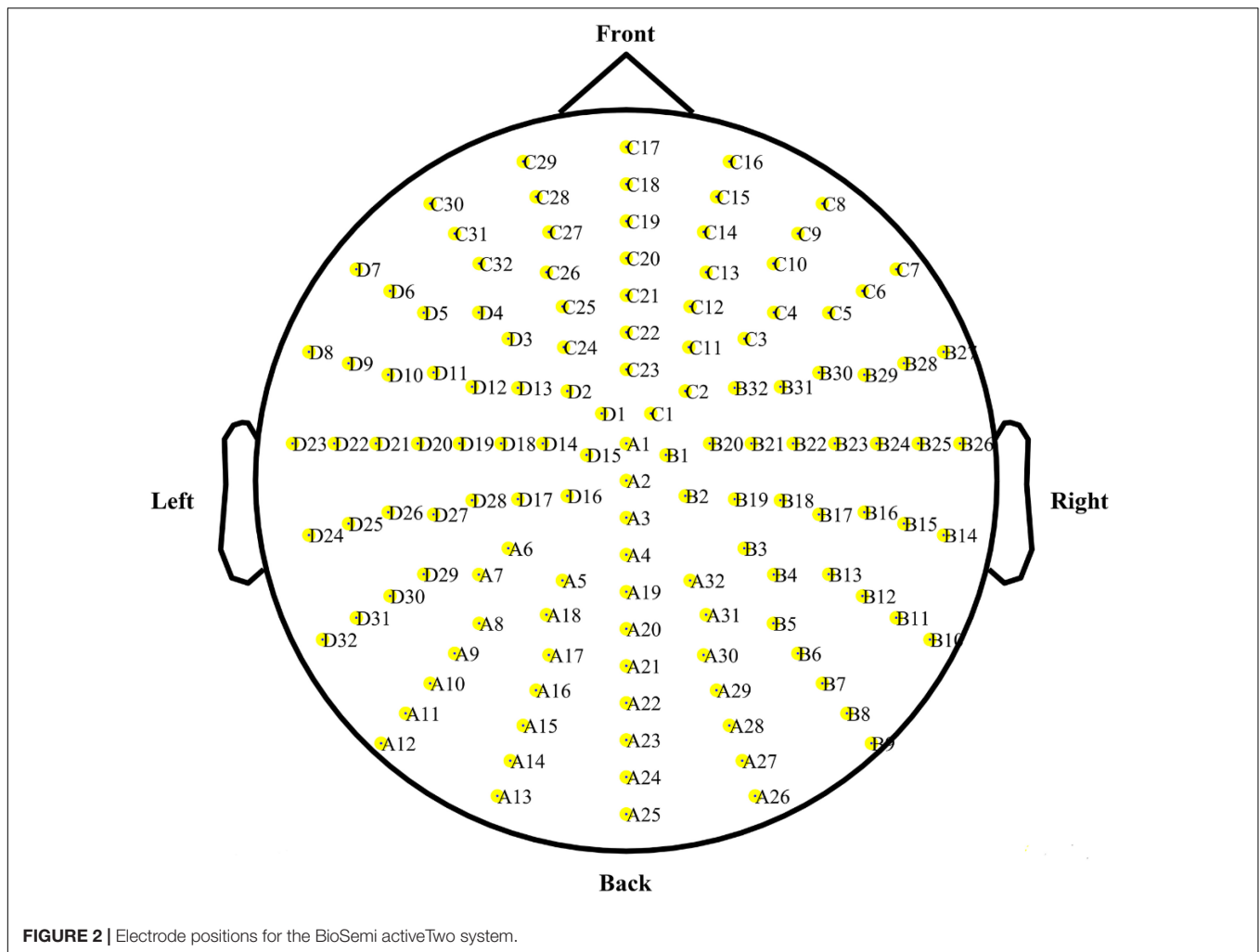
The EEG signals were recorded using the 128-channel BioSemi ActiveTwo system (BioSemi Inc., Heerlen, Netherlands) with a 24-bit analog-to-digital conversion. The 128 electrodes were equally spaced on an electrode cap and customized with an integrated primary amplifier (**Figure 2**). The data were filtered online at a 0.16–100-Hz band-pass filter and sampled at 512 Hz (Yuan and Liu, 2021; Yuan et al., 2021). After the completion of data acquisition, the continuous EEG signals

were re-referenced offline to the average of all channels after rejecting bad segments and interpolating bad traces; the bandpass filter ranged from 0.1 to 50 Hz. An independent component analysis was used to correct electrooculography artifacts from eye movements and blinks. The preprocessed EEG signals were split into epochs from 200 ms before the presentation of the stimulus to 2,000 ms after the onset of the stimulus. EEG data analysis was conducted using the open-source MATLAB signal processing toolbox FieldTrip and in-house functions in MATLAB (Oostenveld et al., 2011).

According to the score for “How much would you like to date this person?,” the EEG epochs were divided into IRA engendered and IRA un-engendered (Fisher et al., 2005; Gerlach and Reinhard, 2018; Yuan and Liu, 2021; Yuan et al., 2021). The IRA engendered category comprised the epochs in which the participants rated their IRA for the potential partners as 3 (very much) or 2 (somewhat). The IRA un-engendered category comprised the epochs in which the participants rated their IRA for the potential partners as 0 (not at all). To minimize ambiguity, we excluded epochs with a rating score of 1. The number of acceptable epochs under the IRA engendered category was 1439, while the number of all acceptable data segments in the IRA un-engendered category was 15298. To solve the problem of serious mismatch in the number of samples between the two preference categories, we randomly selected a number of accepted samples under the IRA un-engendered category to match the number of accepted samples under the IRA engendered category.

### Feature Extraction

To recognize the users’ discrete preferences, we used the wavelet transform (WT) with a sliding time-window approach for TF feature extraction based on the TF analysis (Lindsen et al., 2010;



De Cesarei and Codispoti, 2011; Kang et al., 2015; Yuan and Liu, 2021). Specifically, the time–frequency representation (TFR) was obtained through a five-cycle complex Morlet WT. The sliding windows were advanced in 12-ms and 1-Hz increments to estimate the changes in power over time and frequency in the five FBs: delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–49 Hz). The TF features of the EEG activities were calculated according to event-related oscillations (Pfurtscheller and Lopes da Silva, 1999; Hadjilimitriou and Hadjileontiadis, 2013; Liu et al., 2018; Yuan and Liu, 2021). In this study, two types of TF features were extracted: the power spectral feature (PSF) and the AS (i.e., difference in spectral power between the left and right hemispheres). For each epoch  $j$  and channel  $i$ , each PSF was computed as follows:

$$\text{PSF} = \frac{V - B}{B} \quad (1)$$

where  $V$  represents the quantity estimated during the photograph viewing (PV) period, and  $B$  represents the quantity estimated during the baseline state (BS) period. To obtain the quantity  $V$ , we averaged the TFR during the PV period over the constituent

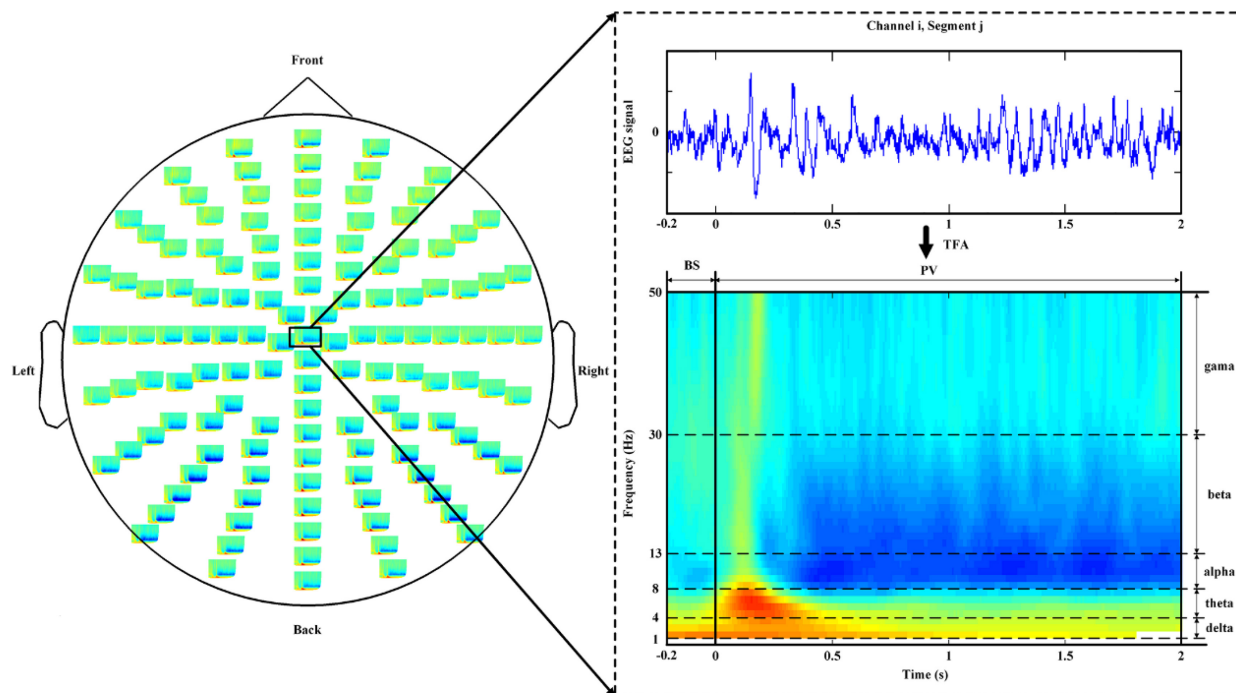
frequencies and time (2). Similarly,  $B$  was computed in the same manner as in PV, as shown in (3).

$$V_w^{\text{fb}}(i, j) = \frac{1}{N_w} \sum_t \left( \frac{1}{N_{\text{fb}}} \sum_f \text{TFR}_{i,j}^{\text{PV}}[t, f] \right) \quad (2)$$

$$B_w^{\text{fb}}(i, j) = \frac{1}{N_w} \sum_t \left( \frac{1}{N_{\text{fb}}} \sum_f \text{TFR}_{i,j}^{\text{BS}}[t, f] \right) \quad (3)$$

where  $[t, f]$  represents the discrete (time and frequency) points in the TF plane;  $\text{TFR}^{\text{PV}}$  represents the obtained TFR during the PV period (Figure 3); and  $N_w$ ,  $N_{\text{fb}}$  denote the number of sample points in the time window of 0–2 s and the number of frequency bins in each FB, respectively (Yuan and Liu, 2021). Similarly,  $\text{TFR}^{\text{BS}}$  represents the obtained TFR during the BS period (Figure 3), and  $N_w$ ,  $N_{\text{fb}}$  denote the number of sample points in the time window of –0.2–0 s and the number of frequency bins in each FB, respectively. Herein, a  $\text{TFR}^{\text{BS}}$  was used to correct the emotional baseline of the  $\text{TFR}^{\text{PV}}$  to exclude the confounding effects of other factors.



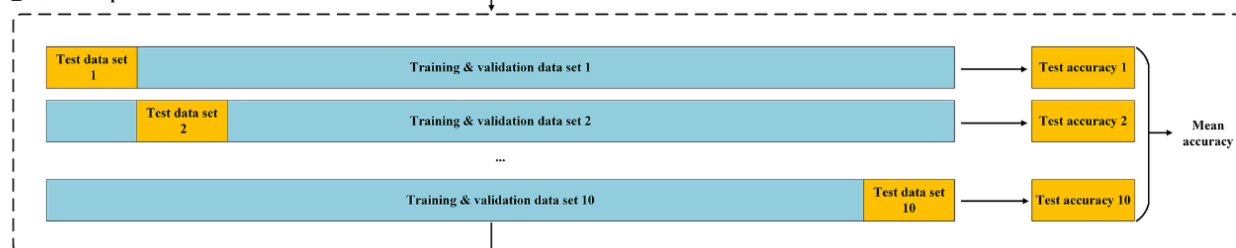


**FIGURE 3 |** Time–frequency plane segmentation for the quantity estimation of B and V, from the TFR of the EEG signal corresponding to channel  $i$  and segment  $j$ . EEG, electroencephalogram. TFR, time–frequency representation.

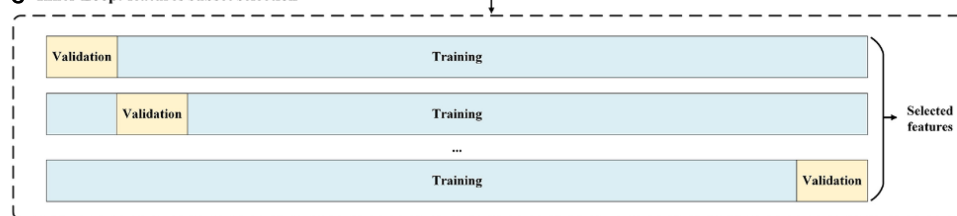
**A Dataset Segmentation: 10 segmentation schemes**



**B Outer Loop: Model assessment**



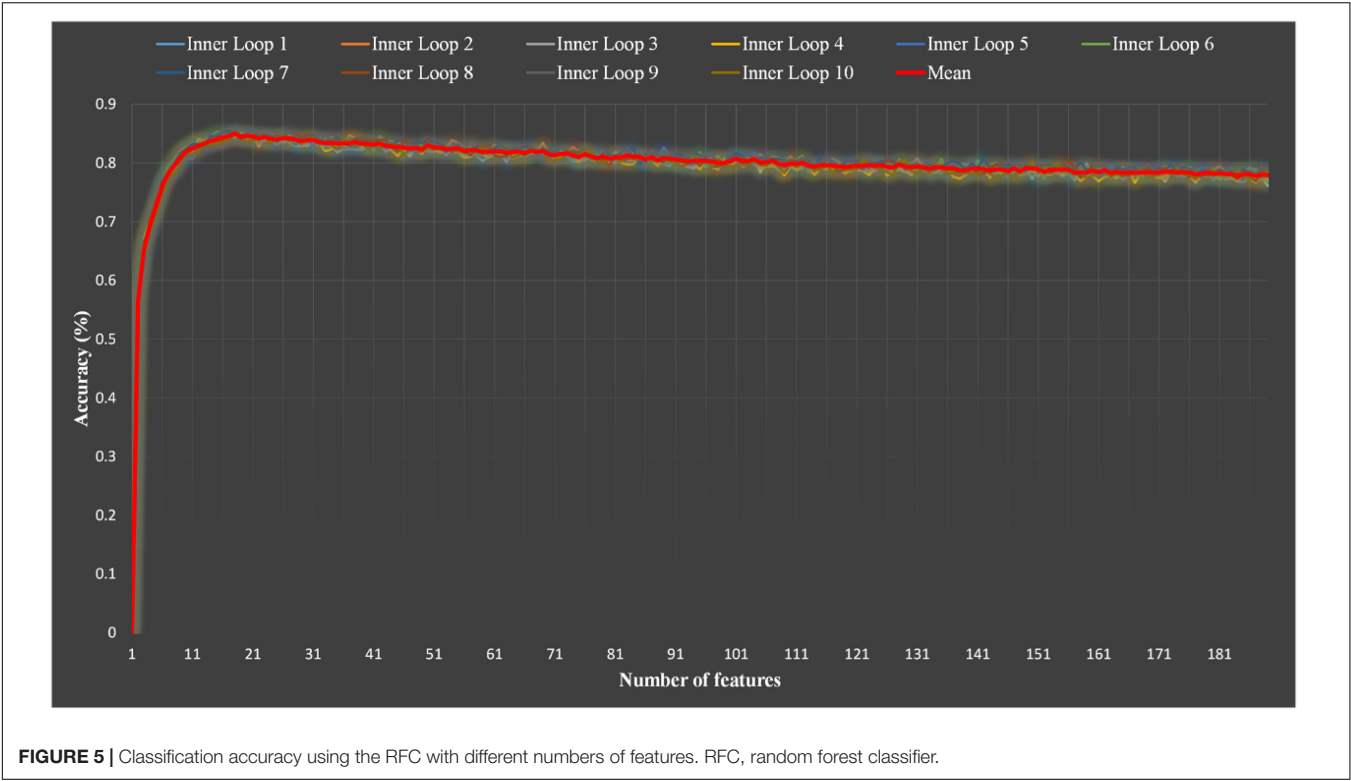
**C Inner Loop: features subset selection**



**FIGURE 4 |** Nested cross-validation architecture used for feature selection and model assessment.

In addition to the PSFs, the ASs of all 55 symmetrical pairs of electrodes on the left and right hemispheres in the five FBs were extracted to measure the possible lateralization of brain activity

that might be caused by emotional stimuli (Liu et al., 2018; Yuan and Liu, 2021). In general, a total of 915 (640 PSFs and 275 ASs) EEG features were extracted.



Preference Recognition With Feature Selection

Nine hundred and fifteen features were extracted from the EEG signals on 128 electrodes, which is undoubtedly a high-dimensional dataset. To effectively analyze the data and

save computational resources, we conducted necessary feature selection before classification (Lu et al., 2020; Zhang et al., 2021). The paired sample *t*-test was used to screen out the feature subsets with significant differences between the IRA engendered and IRA un-engendered categories. A total of 188 features with significant differences ( $p < 0.05$ ) were identified. On this basis, the recursive feature elimination with cross-validation sequential forward feature selector (RFECV) was applied to conduct further feature selection.

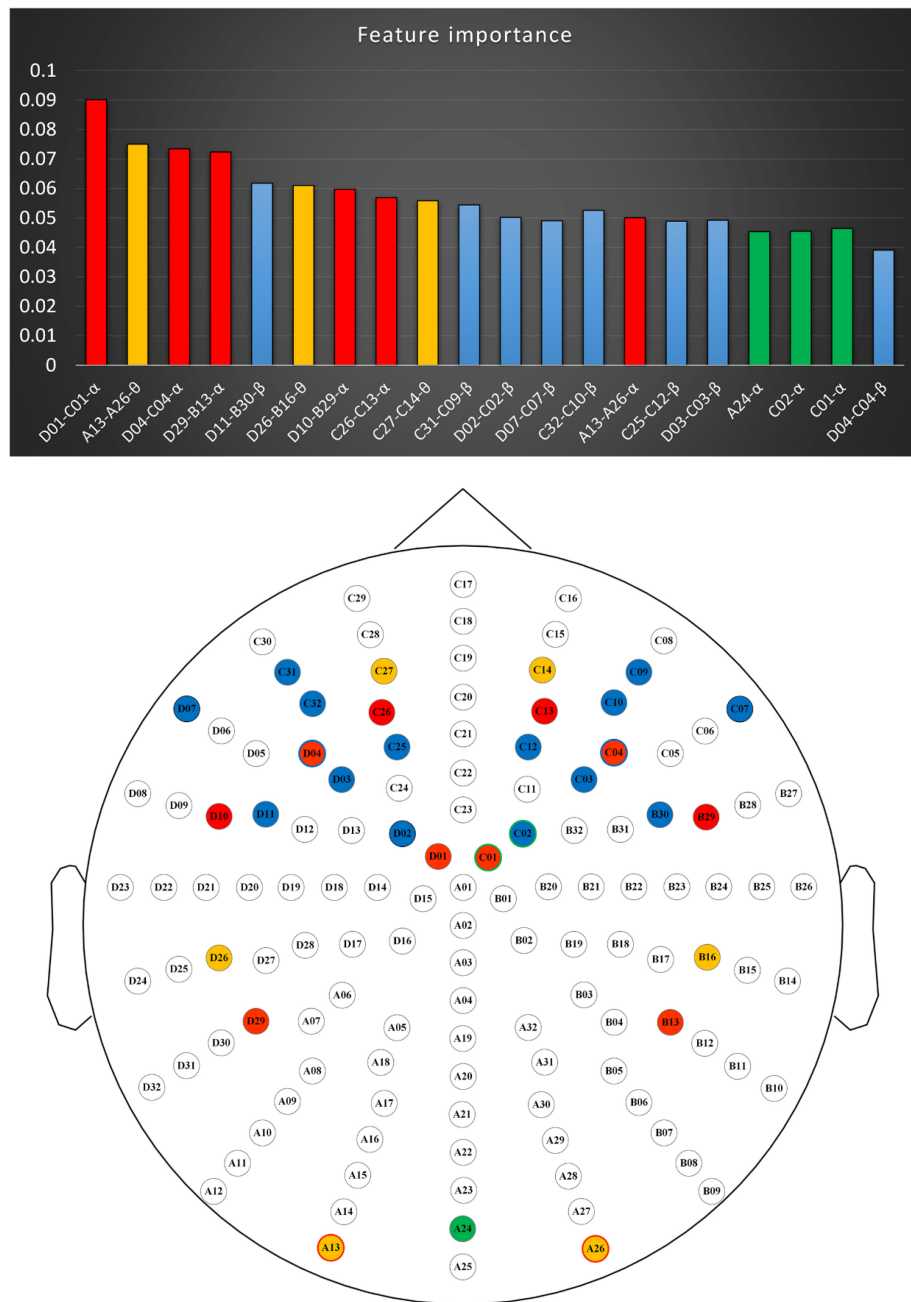
To use the entire dataset to train and test the classifier, we used a nested 10-fold cross-validations to obtain reliable model estimates for feature selection and model training (Pourmohammadi and Maleki, 2020; Zhang et al., 2021).

TABLE 1 | Confusion matrix of each test data set.

	True label Predicted label	IRA engendered	IRA un-engendered
Test data set 1	IRA engendered	144	27
	IRA un-engendered	20	179
Test data set 2	IRA engendered	163	23
	IRA un-engendered	41	141
Test data set 3	IRA engendered	123	19
	IRA un-engendered	20	130
Test data set 4	IRA engendered	139	25
	IRA un-engendered	25	125
Test data set 5	IRA engendered	117	16
	IRA un-engendered	17	116
Test data set 6	IRA engendered	118	23
	IRA un-engendered	18	100
Test data set 7	IRA engendered	137	25
	IRA un-engendered	26	107
Test data set 8	IRA engendered	110	22
	IRA un-engendered	14	105
Test data set 9	IRA engendered	98	21
	IRA un-engendered	17	101
Test data set 10	IRA engendered	80	14
	IRA un-engendered	13	119

TABLE 2 | The results of each test data set.

Metrics	PAM	CA	SE	SP	AUC	JI	FM
Test data set 1	0.72	0.8730	0.8780	0.8689	0.87	0.7539	0.8597
Test data set 2	0.66	0.8261	0.7990	0.8598	0.83	0.7181	0.8539
Test data set 3	0.72	0.8664	0.8601	0.8725	0.87	0.7593	0.8632
Test data set 4	0.68	0.8408	0.8476	0.8333	0.84	0.7354	0.8476
Test data set 5	0.74	0.8579	0.8731	0.8788	0.88	0.7800	0.8764
Test data set 6	0.68	0.8417	0.8676	0.8130	0.84	0.7241	0.8520
Test data set 7	0.74	0.8271	0.8405	0.8106	0.83	0.7287	0.8431
Test data set 8	0.71	0.8566	0.8871	0.8268	0.86	0.7534	0.8594
Test data set 9	0.67	0.8397	0.8522	0.8279	0.84	0.7206	0.8376
Test data set 10	0.73	0.8805	0.8602	0.8947	0.88	0.7477	0.8556
Mean	0.6990	0.8528	0.8566	0.8486	0.8540	0.7439	0.8530



**FIGURE 6 |** Optimal feature subsets of the RFC. Red, blue, and orange present the asymmetry features of the alpha, beta, and theta bands, respectively. Green represents the alpha band PSF. RFC, random forest classifier.

Specifically, the inner loop was responsible for selecting the optimal subset of features (**Figure 4**). In the outer loop and using the selected subset of features, the RFC was evaluated by unseen test data set *via* a subject-wise 10-fold cross-validation (Saeb et al., 2017). Thereafter, the confusion matrix was formed based on the true and predicted labels of sample in the each unseen test data set. Then, based on the confusion matrix, common metrics are calculated to assess performance of machine learning system, including classification accuracy (CA), sensitivity (SE), specificity

(SP), area under curve (AUC), Jaccard index (JI), F-measure (FM), and polygon area metric (PAM) (Aydemir, 2020). The mathematical definitions are, respectively, given as follows:

$$CA = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$SE = \frac{TP}{TP + FN} \quad (5)$$

$$SP = \frac{TN}{TN + FP} \quad (6)$$

$$JI = \frac{TP}{TP + FP + FN} \quad (7)$$

$$FM = \frac{2TP}{2TP + FP + FN} \quad (8)$$

$$AUC = \int_0^1 f(x) dx \quad (9)$$

Where TP is the number of actual positive samples that were predicted to be positive, FN is the number of actual positive samples that were predicted to be negative, TN is the number of actual negative samples that were predicted to be negative, and FP is the number of actual negative samples that were predicted to be positive (Aydemir, 2020). The classifier selected in this study is a widely used classifier with good performance, namely, the RFC (Aldayel et al., 2020a,b). For the RFC, the Gini impurity was used as a function to measure the quality of a split; the maximum depth of the tree was set to 30; and the other super parameters were set to default.

## RESULTS AND DISCUSSION

The classification performance of the proposed EEG-based mate preference recognition algorithm was verified using a total of 2878 EEG samples (including 1439 samples of the IRA engendered category and 1439 samples of the IRA un-engendered category) collected from 50 participants. To obtain an optimal feature subset from 188 features with significant differences ( $p < 0.05$ ) between the two categories, we used a nested 10-fold cross-validation scheme based on the RFECV-RFC algorithm for feature selection. The number of features varied from 1 to 188, and the best feature subset was selected in each step. **Figure 5** displays the mean classification accuracies on the validation sets of each inner loop when selecting different numbers of features. As can be seen from **Figure 5**, the number of features of the optimal feature subset selected by each internal cycle is roughly the same (about 17,  $SD = 0.57$ ). The performance of the model is evaluated on the corresponding test set based on the optimal feature subset selected in each inner loop. The results are shown in **Tables 1, 2**. It can be seen from **Table 2** that the best mean CA value, mean PAM value, mean se value, mean SP value, mean AUC value, mean Ji value and mean FM value are 0.8528, 0.6990, 0.8566, 0.8486, 0.8540, 0.7439, and 0.8530, respectively.

**Figure 6** shows the union of the optimal feature subsets selected by each inner loop and the distribution of each feature. Based on the results shown in **Figure 6**, we found that the asymmetric features over the frontal and parietal lobes play an extremely important role in recognizing initial romantic interest because 15 of the 20 most discriminating features originated from these two regions. Moreover, 14 of these 15 features belonged to the alpha and beta bands. Previous studies have demonstrated that the frontal and parietal lobes are the most informative

regions of emotional states, while the alpha and beta waves appear to be the most discriminative (Alarcao and Fonseca, 2019; Zheng et al., 2020). Yuan and Liu (2021) found that the changes in alpha and beta power on the sensors over the anterior regions play an important role in the generation and evaluation of IRA. In addition, numerous studies have demonstrated that frontal and parietal asymmetries in the alpha and beta FBs are observable for valence and arousal recognition (Cacioppo, 2004; Huang et al., 2012; Alarcao and Fonseca, 2019). In particular, Aldayel et al. (2020a,b) showed that the asymmetric features in alpha and beta frequencies over the frontal and parietal regions can effectively identify users' emotional responses and preferences to market stimuli (Touchette and Lee, 2016; Liu et al., 2018; Ramsay et al., 2018). In addition, Naser et al. showed that asymmetric features of alpha frequency on the frontal and parietal lobe regions could effectively identify users' preference for music (Naser and Saha, 2021). We also found that the asymmetric features in the alpha, beta, and theta bands over the lateral occipital complex and the asymmetric features in the theta bands over the frontal and parietotemporal regions were sensitive in recognizing IRA. Previous studies have observed that the generation of IRA leads to desynchronization of alpha and beta bands in the lateral occipital complex region (Yuan and Liu, 2021). The theta FB over the frontal and parietotemporal areas was also considered to be an effective feature for identifying emotional states (Aftanas et al., 2001; Cartier et al., 2012).

Taken together, these findings demonstrate that users' preferences for potential romantic partners can be determined on the basis of EEG signals. Furthermore, the TF features from channels over the frontal, parietal, and occipital regions are informative and suitable for the identification of IRA toward potential partners.

## CONCLUSION

The purpose of this experiment was to determine the possibility of using EEG signals to identify users' aesthetic preferences for potential romantic partners. In this study, our system achieved a best accuracy of 85.2% ( $SD = 0.03$ ) in recognizing the initial romantic interest. This result demonstrated that based on the information provided by users' EEG signals, we can determine whether they are romantically interested in a potential partner. In addition, the best accuracy 85.2% ( $SD = 0.03$ ) in this study was obtained mainly using the asymmetry features of the alpha, beta, and theta FBs on the sensors over the frontal lobe, parietal lobe, and lateral occipital complex. These results suggest that the TF features from channels over the frontal, parietal, and occipital regions are suitable for identifying human preferences for potential romantic partners. Therefore, in future work, we plan to extract features from different dimensions, such as the time domain and source domain, and explore how to use the minimum channels to optimize the classification accuracy through multi-dimensional feature integration.

In addition, as an exploratory study, this study used portrait photos rather than real people as stimuli to induce IRA based on feasibility considerations. The advantage of this approach is



that by increasing the amount of stimulus, it can effectively solve the problem of insufficient trials in which IRA was successfully induced due to the low average induction rate. However, in real social scenes, initial romantic interest usually occurs in the environment that allows some meaningful interaction, but the types of stimuli used in the present study did not allow participants to interact effectively with potential partners in the photos (Yuan and Liu, 2021; Yuan et al., 2021). This is a problem that needs to be paid attention to and solved in the follow-up research. It is believed that in the near future, mate preference recognition and matching systems based on EEG signals will be applied to online or offline dating scenarios to assist individuals in finding their ideal partners.

## DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethical Review Committee of Southwest University. The patients/participants provided their written informed consent to participate in this study. Written informed consent

was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

## AUTHOR CONTRIBUTIONS

GY and GL conceived the study. GY designed and programmed the tasks, collected the data, analyzed the composite behavioral and EEG data, and wrote the manuscript. GY and WH interpreted the results. GL and WH revised the manuscript. All authors approved the final manuscript.

## FUNDING

This work was supported in part by the National Natural Science Foundation of China (Nos. 61472330 and 61872301). This work was supported in part by the Natural Science Foundation of Shandong Province (No. ZR2019MC048). This work was supported in part by the Natural Science Foundation of Chongqing (No. cstc2021jcyj-msxmX0041).

## ACKNOWLEDGMENTS

The authors thank all the researchers involved in the experiment.

## REFERENCES

- Aftanas, L. I., Varlamov, A. A., Pavlov, S. V., Makhnev, V. P., and Reva, N. V. (2001). Affective picture processing: event-related synchronization within individually defined human theta band is modulated by valence dimension. *Neurosci. Lett.* 303, 115–118. doi: 10.1016/S0304-3940(01)01703-7
- Alarcao, S. M., and Fonseca, M. J. (2019). Emotions recognition using EEG signals: a survey. *IEEE Trans. Affect. Comput.* 10, 374–393.
- Aldayel, M., Ykhlef, M., and Al-Nafjan, A. (2020a). Deep learning for EEG-based preference classification in neuromarketing. *Appl. Sci.* 10:1525. doi: 10.3389/fnhum.2020.604639
- Aldayel, M., Ykhlef, M., and Al-Nafjan, A. (2020b). Recognition of consumer preference by analysis and classification EEG signals. *Front. Hum. Neurosci.* 14:604639.
- Aldayel, M. S., Ykhlef, M., and Al-Nafjan, A. N. (2020c). Electroencephalogram-based preference prediction using deep transfer learning. *IEEE Access* 8, 176818–176829.
- Aldayel, M., Ykhlef, M., and Al-Nafjan, A. (2021). Consumers' preference recognition based on brain-computer interfaces: advances, trends, and applications. *Arab. J. Sci. Eng.* 46, 8983–8997.
- Asendorpf, J. B., Penke, L., and Back, M. D. (2011). From dating to mating and relating: predictors of initial and long-term outcomes of speed-dating in a community sample. *Eur. J. Pers.* 25, 16–30.
- Aydemir, O. (2020). A new performance evaluation metric for classifiers: polygon area metric. *J. Classif.* 38, 16–26. doi: 10.1007/s00357-020-09362-5
- Balconi, M., and Mazza, G. (2009). Brain oscillations and BIS/BAS (behavioral inhibition/activation system) effects on processing masked emotional cues. ERS/ERD and coherence measures of alpha band. *Int. J. Psychophysiol.* 74, 158–165. doi: 10.1016/j.ijpsycho.2009.08.006
- Buston, P. M., and Emlen, S. T. (2003). Cognitive processes underlying human mate choice: the relationship between self-perception and mate preference in Western society. *Proc. Natl. Acad. Sci. U.S.A.* 100, 8805–8810. doi: 10.1073/pnas.1533220100
- Cacioppo, J. T. (2004). Feelings and emotions: roles for electrophysiological markers. *Biol. Psychol.* 67, 235–243. doi: 10.1016/j.biopsycho.2004.03.009
- Cartier, C., Bittencourt, J., Peressutti, C., Machado, S., Paes, F., Sack, A. T., et al. (2012). Premotor and occipital theta asymmetries as discriminators of memory- and stimulus-guided tasks. *Brain Res. Bull.* 87, 103–108. doi: 10.1016/j.brainresbull.2011.10.013
- Chew, L. H., Teo, J., and Mountstephens, J. (2016). Aesthetic preference recognition of 3D shapes using EEG. *Cogn. Neurodyn.* 10, 165–173. doi: 10.1007/s11571-015-9363-z
- Cooper, J. C., Dunne, S., Furey, T., and O'Doherty, J. P. (2012). Dorsomedial prefrontal cortex mediates rapid evaluations predicting the outcome of romantic interactions. *J. Neurosci.* 32, 15647–15656. doi: 10.1523/JNEUROSCI.2558-12.2012
- De Cesarei, A., and Codispoti, M. (2011). Affective modulation of the LPP and alpha-ERD during picture viewing. *Psychophysiology* 48, 1397–1404. doi: 10.1111/j.1469-8986.2011.01204.x
- Finkel, E. J., Eastwick, P. W., and Matthews, J. (2007). Speed-dating as an invaluable tool for studying romantic attraction: a methodological primer. *Pers. Relationsh.* 14, 149–166.
- Fisher, A. A. H., Mashek, D. J., Li, H., and Brown, L. (2002). Defining the brain systems of lust, romantic attraction, and attachment. *Arch. Sex. Behav.* 31, 413–419. doi: 10.1023/a:1019888024255
- Fisher, H. E. (1998). Lust, attraction, and attachment in mammalian reproduction. *Hum. Nat.* 9, 23–52. doi: 10.1007/s12110-998-1010-5
- Fisher, H., Aron, A., and Brown, L. L. (2005). Romantic love: an fMRI study of a neural mechanism for mate choice. *J. Comp. Neurol.* 493, 58–62. doi: 10.1002/cne.20772
- Gerlach, T. M., and Reinhard, S. K. (2018). "Personality and romantic attraction," in *Encyclopedia of Personality and Individual Differences*, eds V. Zeigler-Hill and T. K. Shackelford (Cham: Springer), 1–6. doi: 10.1007/978-3-319-28099-8\_717-2
- Gunes, H., Schuller, B., Pantic, M., and Cowie, R. (2011). "Emotion representation, analysis and synthesis in continuous space: a survey," in *Proceedings of the*

- IEEE International Conference on Automatic Face & Gesture Recognition, Santa Barbara, CA, 827–834.
- Hadjidimitriou, S. K., and Hadjileontiadis, L. J. (2012). Toward an EEG-based recognition of music liking using time-frequency analysis. *IEEE Trans. Biomed. Eng.* 59, 3498–3510. doi: 10.1109/TBME.2012.2217495
- Hadjidimitriou, S. K., and Hadjileontiadis, L. J. (2013). EEG-based classification of music appraisal responses using time-frequency analysis and familiarity ratings. *IEEE Trans. Affect. Comput.* 4, 161–172.
- Huang, D., Guan, C., Kai Keng, A., Haihong, Z., and Yaozhang, P. (2012). “Asymmetric spatial pattern for EEG-based emotion detection,” in *Proceedings of the 2012 International Joint Conference on Neural Networks (IJCNN)*, Brisbane, QLD, 1–7. doi: 10.1186/s12868-016-0283-6
- Jatupaiboon, N., Pan-ngum, S., and Israsena, P. (2013). “Emotion classification using minimal EEG channels and frequency bands,” in *Proceedings of the 2013 10th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, Khon Kaen, 21–24.
- Joel, S., Eastwick, P. W., and Finkel, E. J. (2017). Is romantic desire predictable? Machine learning applied to initial romantic attraction. *Psychol. Sci.* 28, 1478–1489. doi: 10.1177/0956797617714580
- Kang, J. H., Kim, S. J., Cho, Y. S., and Kim, S. P. (2015). Modulation of alpha oscillations in the human EEG with facial preference. *PLoS One* 10:e0138153. doi: 10.1371/journal.pone.0138153
- Khurana, V., Gahalawat, M., Kumar, P., Roy, P. P., Dogra, D. P., Scheme, E., et al. (2021). A survey on neuromarketing using EEG signals. *IEEE Trans. Cogn. Dev. Syst.* 13, 732–749.
- Lin, Y. P., Wang, C. H., Jung, T. P., Wu, T. L., Jeng, S. K., Duann, J. R., et al. (2010). EEG-based emotion recognition in music listening. *IEEE Trans. Biomed. Eng.* 57, 1798–1806. doi: 10.1109/TBME.2010.2048568
- Lindsen, J. P., Jones, R., Shimojo, S., and Bhattacharya, J. (2010). Neural components underlying subjective preferential decision making. *Neuroimage* 50, 1626–1632. doi: 10.1016/j.neuroimage.2010.01.079
- Liu, Y., Sourina, O., and Nguyen, M. K. (2011). “Real-time EEG-based emotion recognition and its applications,” in *Transactions on Computational Science XII: Special Issue on Cyberworlds*, eds M. L. Gavrilova, C. J. K. Tan, A. Sourin, and O. Sourina (Berlin: Springer), 256–277. doi: 10.1007/978-3-642-22336-5\_13
- Liu, Y.-J., Yu, M., Zhao, G., Song, J., Ge, Y., and Shi, Y. (2018). Real-time movie-induced discrete emotion recognition from EEG signals. *IEEE Trans. Affect. Comput.* 9, 550–562. doi: 10.1109/taffc.2017.2660485
- Lu, H., Yuan, G., Zhang, J., and Liu, G. (2020). Recognition of impulse of love at first sight based on photoplethysmography signal. *Sensors* 20:6572. doi: 10.3390/s20226572
- Moon, J. (2013). Extraction of user preference for video stimuli using EEG-based user responses. *ETRI J.* 35, 1105–1114. doi: 10.4218/etrij.13.0113.0194
- Naser, D. S., and Saha, G. (2021). Influence of music liking on EEG based emotion recognition. *Biomed. Signal Process. Control* 64:102251. doi: 10.1016/j.bspc.2020.102251
- Niemic, C. (2004). Studies of emotion: a theoretical and empirical review of psychophysiological studies of emotion. *J. Undergrad. Res.* 1, 15–18.
- Olderbak, S. G., Malter, F., Wolf, P. S. A., Jones, D. N., and Figueredo, A. J. (2017). Predicting romantic interest at zero acquaintance: evidence of sex differences in trait perception but not in predictors of interest. *Eur. J. Pers.* 31, 42–62. doi: 10.1002/per.2087
- Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011:156869. doi: 10.1155/2011/156869
- Pfurtscheller, G., and Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. *Clin. Neurophysiol.* 110, 1842–1857. doi: 10.1016/s1388-2457(99)00141-8
- Pourmohammadi, S., and Maleki, A. (2020). Stress detection using ECG and EMG signals: a comprehensive study. *Comput. Methods Programs Biomed.* 193:105482. doi: 10.1016/j.cmpb.2020.105482
- Ramsey, T. Z., Skov, M., Christensen, M. K., and Stahlhut, C. (2018). Frontal brain asymmetry and willingness to pay. *Front. Neurosci.* 12:138. doi: 10.3389/fnins.2018.00138
- Saeb, S., Lonini, L., Jayaraman, A., Mohr, D. C., and Kording, K. P. (2017). The need to approximate the use-case in clinical machine learning. *Gigascience* 6, 1–9.
- Schubring, D., and Schupp, H. T. (2019). Affective picture processing: alpha- and lower beta-band desynchronization reflects emotional arousal. *Psychophysiology* 56:e13386. doi: 10.1111/psyp.13386
- Schubring, D., and Schupp, H. T. (2021). Emotion and brain oscillations: high arousal is associated with decreases in alpha- and lower beta-band power. *Cereb. Cortex* 31, 1597–1608. doi: 10.1093/cercor/bhaa312
- Spielmann, S. S., MacDonald, G., Maxwell, J. A., Joel, S., Peragine, D., Muise, A., et al. (2013). Settling for less out of fear of being single. *J. Pers. Soc. Psychol.* 105, 1049–1073. doi: 10.1037/a0034628
- Thomas, A. G., Jonason, P. K., Blackburn, J. D., Kennair, L. E. O., Lowe, R., Malouff, J., et al. (2020). Mate preference priorities in the East and West: a cross-cultural test of the mate preference priority model. *J. Pers.* 88, 606–620. doi: 10.1111/jopy.12514
- Touchette, B., and Lee, S.-E. (2016). Measuring neural responses to apparel product attractiveness. *Cloth. Text. Res. J.* 35, 3–15. doi: 10.1177/0887302x16673157
- Yuan, G., and Liu, G. (2021). Mate preference and brain oscillations: initial romantic attraction is associated with decreases in alpha- and lower beta-band power. *Hum. Brain Mapp.* 43, 721–732. doi: 10.1002/hbm.25681
- Yuan, G., Liu, G., and Wei, D. (2021). Roles of P300 and late positive potential in initial romantic attraction. *Front. Neurosci.* 15:718847. doi: 10.3389/fnins.2021.718847
- Zhang, J., Yuan, G., Lu, H., and Liu, G. (2021). Recognition of the impulse of love at first sight based on electrocardiograph signal. *Comput. Intell. Neurosci.* 2021:6631616. doi: 10.1155/2021/6631616
- Zheng, X., Liu, X., Zhang, Y., Cui, L., and Yu, X. (2020). A portable HCI system-oriented EEG feature extraction and channel selection for emotion recognition. *Int. J. Intell. Syst.* 36, 152–176.
- Zheng, X., Yu, X., Yin, Y., Li, T., and Yan, X. (2021). Three-dimensional feature maps and convolutional neural network-based emotion recognition. *Int. J. Intell. Syst.* 36, 6312–6336.
- Zsok, F., Haucke, M., De Wit, C. Y., and Barelds, D. P. H. (2017). What kind of love is love at first sight? An empirical investigation. *Pers. Relationsh.* 24, 869–885. doi: 10.1111/per.12218

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships 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.

Copyright © 2022 Yuan, He and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.



# Closed-Loop Tracking and Regulation of Emotional Valence State From Facial Electromyogram Measurements

Luciano R. F. Branco<sup>1</sup>, Arian Ehteshami<sup>1</sup>, Hamid Fekri Azgomi<sup>1,2</sup> and Rose T. Faghih<sup>1,3\*</sup>

<sup>1</sup> Department of Electrical and Computer Engineering, University of Houston, Houston, TX, United States, <sup>2</sup> Department of Neurological Surgery, University of California, San Francisco, San Francisco, CA, United States, <sup>3</sup> Department of Biomedical Engineering, New York University, New York, NY, United States

Affective studies provide essential insights to address emotion recognition and tracking. In traditional open-loop structures, a lack of knowledge about the internal emotional state makes the system incapable of adjusting stimuli parameters and automatically responding to changes in the brain. To address this issue, we propose to use facial electromyogram measurements as biomarkers to infer the internal hidden brain state as feedback to close the loop. In this research, we develop a systematic way to track and control emotional valence, which codes emotions as being pleasant or obstructive. Hence, we conduct a simulation study by modeling and tracking the subject's emotional valence dynamics using state-space approaches. We employ Bayesian filtering to estimate the person-specific model parameters along with the hidden valence state, using continuous and binary features extracted from experimental electromyogram measurements. Moreover, we utilize a mixed-filter estimator to infer the secluded brain state in a real-time simulation environment. We close the loop with a fuzzy logic controller in two categories of regulation: inhibition and excitation. By designing a control action, we aim to automatically reflect any required adjustments within the simulation and reach the desired emotional state levels. Final results demonstrate that, by making use of physiological data, the proposed controller could effectively regulate the estimated valence state. Ultimately, we envision future outcomes of this research to support alternative forms of self-therapy by using wearable machine interface architectures capable of mitigating periods of pervasive emotions and maintaining daily well-being and welfare.

**Keywords:** closed-loop, control, brain, emotion, valence, electromyogram (EMG), wearable, state-space

## OPEN ACCESS

### Edited by:

Jane Zhen Liang,  
Shenzhen University, China

### Reviewed by:

Qing Yun Wang,  
Beihang University, China  
Johanna J. O'Day,  
Stanford University, United States

### \*Correspondence:

Rose T. Faghih  
rfaghih@nyu.edu

**Received:** 27 July 2022

**Accepted:** 21 February 2022

**Published:** 25 March 2022

### Citation:

Branco LRF, Ehteshami A, Azgomi HF  
and Faghih RT (2022) Closed-Loop  
Tracking and Regulation of Emotional  
Valence State From Facial  
Electromyogram Measurements.  
*Front. Comput. Neurosci.* 16:747735.  
doi: 10.3389/fncom.2022.747735

## 1. INTRODUCTION

Emotions directly influence the way we think and interact with others in different situations, especially when it interferes with rationality in our decision-making or perception (Dolan, 2002). Thus, having a solid grasp of the dynamics of emotions is critical to provide any therapeutic solutions to maintain welfare (Couette et al., 2020). Moreover, deciphering emotions has been an ongoing task among researchers, dictating joint efforts from behavioral, physiological, and computational angles (Scherer, 2005). According to the James A. Russell's circumplex model of affect, emotion can be divided into two perpendicular axes, viz. valence—reflecting the spectrum

of negative to positive emotions—and arousal, accounting for the intensity characteristics (Russell, 1980). In this study, we focus on improving comprehension of emotional valence regulation by proposing an architecture to track and regulate the internal hidden valence state using physiological signals collected via wearable devices. The use of wearable devices to gain insight to the internal brain state provides a good alternative to study the brain dynamics, as usually the procedures either rely on invasive techniques, e.g., extracting bloodstream samples, performing surgery, or require large and expensive equipment for imaging purposes (Villanueva-Meyer et al., 2017; Wickramasuriya et al., 2019a,b).

Affective computing is defined by an interdisciplinary field of research that incorporates both sentiment analysis and emotion recognition (Poria et al., 2017). Scholars have posited the importance of affective computing to endow machines with the means to recognize, interpret or convey emotions and sentiments (Poria et al., 2017; Burzagli and Nalini, 2020). These capabilities allow the development and enhancement of personal care systems that interact better with humans, potentially improving a personal health and daily well-being (Burzagli and Nalini, 2020). Previous attempts in the development of affective computing have focused on emotion feature extraction and classification through human-robot interactions (Azuar et al., 2019; Rudovic et al., 2019; Yu and Tapus, 2019; Filippini et al., 2020; Rosula Reyes et al., 2020; Val-Calvo et al., 2020), facial expressions (Chronaki et al., 2015; rong Mao et al., 2015; Yang et al., 2018a; Zeng et al., 2018), and vocal responses (Wang et al., 2015; Fayek et al., 2017; Noroozi et al., 2017; Anuja and Sanjeev, 2020). The objective of this research is to take this one step further and introduce a tracking and closed-loop control framework to regulate specific emotions.

Within a closed-loop approach, biomarkers are collected in real-time as feedback, which grants the possibility of automatically adjusting brain stimulation levels according to the current emotional state (Wickramasuriya et al., 2019a,b; Thenaisie et al., 2021). Previous studies have shown that this strategy can increase treatment efficacy and decrease the extent of stimulation side-effects, compared to just employing an open-loop stimulation (Price et al., 2020). The benefits of closed-loop neurostimulation have been well reported in addressing conventional-therapy-resistant patients with Parkinson's disease (Little et al., 2016; Weiss and Massano, 2018). However, fewer studies have explored closed-loop therapies for non-motor neuropathologies such as post-traumatic stress disorder or depression (Tegeler et al., 2017; Mertens et al., 2018), even though there is already relevant evidence of improvements with open-loop therapies (Conway et al., 2018; Starnes et al., 2019; Freire et al., 2020). Conversely with the conventional open-loop approach, brain stimulation is manually tuned during in-clinic visits, delivering pre-determined quantities and incurring over or under stimulation of the brain (Wickramasuriya et al., 2019a,b; Price et al., 2020).

To properly regulate the emotional brain state in a closed-loop manner, a suitable biomarker that relates to the internal emotional valence needs to be identified. Prior research has based emotion classification on facial or voice expressions, which not

only requires heavy data acquisition, but also runs into ambiguity issues (Tan et al., 2012). Facial and vocal expressions can vary significantly between person to person, making it difficult to draw any accurate inference about the person's emotional state. Moreover, facial and vocal expressions (e.g., smiling) can be seen as externalized emotions and can be altered at will, confounding the accuracy of such classification approaches, and thus hindering any tracking and control efforts as the true emotional state would not be clear (Cai et al., 2018). In response, our proposed strategy aims to remove this ambiguity by using a more reliable metric: physiological signals (Cacioppo et al., 2000). Physiological signals or biomarkers are involuntary responses initiated by the human's central and autonomic nervous systems, whereas facial and vocal lineaments can voluntarily be hidden to reject certain emotional displays (Cannon, 1927; Cacioppo et al., 2000; Lin et al., 2018; Amin and Faghih, 2020; Wilson et al., 2020). Although overall facial expression can be made to mask certain emotions, several studies have linked electromyogram activity of specific facial muscles to states of affection in varying valence levels, such as happiness, stress and anger (Nakasone et al., 2005; Kulic and Croft, 2007; Gruebler and Suzuki, 2010; Tan et al., 2012; Amin et al., 2016; Cai et al., 2018). Cacioppo et al. described that the somatic effectors of the face are tied to changes in connective tissue rather than skeletal complexes (Cacioppo et al., 1986). Researchers in Cacioppo et al. (1986) posited facial electromyogram could provide insight into valence state recognition even when there is no apparent change in facial expressions. Moreso, the work of Ekman et al. (1980) and Brown and Schwartz (Brown and Schwartz, 1980) are two of the few who showed that using facial electromyogram measurements of the zygomaticus muscle (zEMG) gave the most distinct indicator of valence compared to other facial muscles involved in the act of smiling. Multiple studies have suggested the relation between emotional states and facial electromyogram activity (Van Bostel, 2010; Tan et al., 2011; Koelstra et al., 2012; Künecke et al., 2014; Kordsachia et al., 2018; Kayser et al., 2021; Shiva et al., 2021). Golland et al. (2018) also showcased a consistent relationship between the emotional media viewed and the changes seen in the components of the facial electromyogram signal. We focus on zEMG to build our model and track the hidden valence state. Then, we design a control strategy to automatically regulate the internal emotional valence state in real-time.

It should be noted that electromyogram is not the only physiological metric that has shown promise for valence recognition. Emotional valence can also be represented by many different physiological signals or a combination of them (Egger et al., 2019), such as using electroencephalography (Bozhkov et al., 2017; Wu et al., 2017; Soroush et al., 2019; Feradov et al., 2020), respiration (Zhang et al., 2017; Wickramasuriya et al., 2019a,b), electrocardiography (ECG) (Das et al., 2016; Goshvarpour et al., 2017; Harper and Southern, 2020), blood volume pulse (Das et al., 2016) or heart rate variability (Ravindran et al., 2019). Egger et al. investigated the accuracy of different physiological signals in classifying emotive states such as stress periods, calmness, despair, discontent, erotica, interest, boredom, or elation (Egger et al., 2019). Naji and collaborators displayed the disparity between multimodal and individual signal



measurements regarding emotion classification via ECG and forehead biosignals (Naji et al., 2014).

Previous studies have also investigated different ways of estimating and tracking internal brain states (Sakkalis, 2011). Brain dynamics during resting states have been studied with measurements from functional magnetic resonance imaging (fMRI), using linear and non-linear models, and more recently, employing a tensor based approach (Honey et al., 2009; Abdelnour et al., 2014; Al-Sharoa et al., 2018). The transition of brain states has been examined with machine learning methods and eigenvalue decomposition, by using data from fMRI, electroencephalogram (EEG) or magnetoencephalography (Pfurtscheller et al., 1998; Guimaraes et al., 2007; LaConte et al., 2007; Maheshwari et al., 2020). Moreover, EEG measurements were also employed with machine learning techniques to estimate stress levels (Al-Shargie et al., 2015), and affection (Nie et al., 2011). The method introduced by Yadav et al. uses a state-space formulation to track and classify emotional valence based on two simultaneous assessments of brain activity (Yadav et al., 2019). In the present work, we use a similar approach to estimate and track the hidden valence state, with the help of Bayesian filtering as a powerful statistical tool to improve state estimation under measurement uncertainties (Prerau et al., 2009; Ahmadi et al., 2019; Wickramasuriya and Faghih, 2020). Another contribution of the present work is the use of real measurements from wearable devices to develop a virtual subject environment as a simulation framework for concealed emotional levels. This is the first step to empower the implementation and testing of closed-loop controllers that could track and regulate the internal valence state. In a similar fashion to other control studies, providing a reliable closed-loop simulation framework can pave the way for safe experimentation of brain-related control algorithms here and in future studies (Santaniello et al., 2010; Dunn and Lowery, 2013; Yang et al., 2018b; Wei et al., 2020; Ionescu et al., 2021).

To investigate the validity of regulating emotions through a closed-loop control architecture, we design a simulation system using experimental data. Specifically, in this *in silico* study, we employ features extracted from zEMG data and design a fuzzy logic controller to regulate the emotional valence state in a closed-loop manner. We propose to implement fuzzy logic as this knowledge-based controller works with a set of predetermined fuzzy rules and weights responsible to gauge the degree in which the input variables are classified into output membership functions (Klir and Yuan, 1995; Qi et al., 2019). This process is particularly useful for controlling complex biological systems, as it provides a simple yet effective way of interacting with the uncertainties and impreciseness of these challenging systems (Lilly, 2011). In the literature, previous research have explored the use of a fuzzy logic controller in a simulation environment to control cognitive stress or regulate energy levels of patients with clinical hypercortisolism (Azgomi and Faghih, 2019; Azgomi et al., 2019). A fuzzy controller was also combined with a classical Proportional Integral Derivative (PID) controller to aid the movement of a knee prosthesis leg (Wiem et al., 2018), and to regulate movement of the elbow joint of an exoskeleton during post-stroke rehabilitation (Tageldeen et al., 2016). Scholars have shown fuzzy logic controllers to outperform PID controllers in

the regulation of mean arterial pressure (Sharma et al., 2020), and to improve the anesthetic levels of patients undergoing general anesthesia (Mendez et al., 2018). In light of what is presented, in this *in silico* study, we develop a virtual subject environment to evaluate the efficiency of our proposed architecture.

The remainder of this research is organized as follows. In Section 2 we describe the methods used in this research. Specifically, in Section 2.1 we describe the virtual subject environment and the steps taken toward its development (i.e., the models used, the features extracted, the valence state estimation and the modeling of the environmental stimuli). Next, in Section 2.2 we explain the controller design and the steps taken during implementation. Then, we present our results in Section 3, followed by a discussion of those in Section 4.

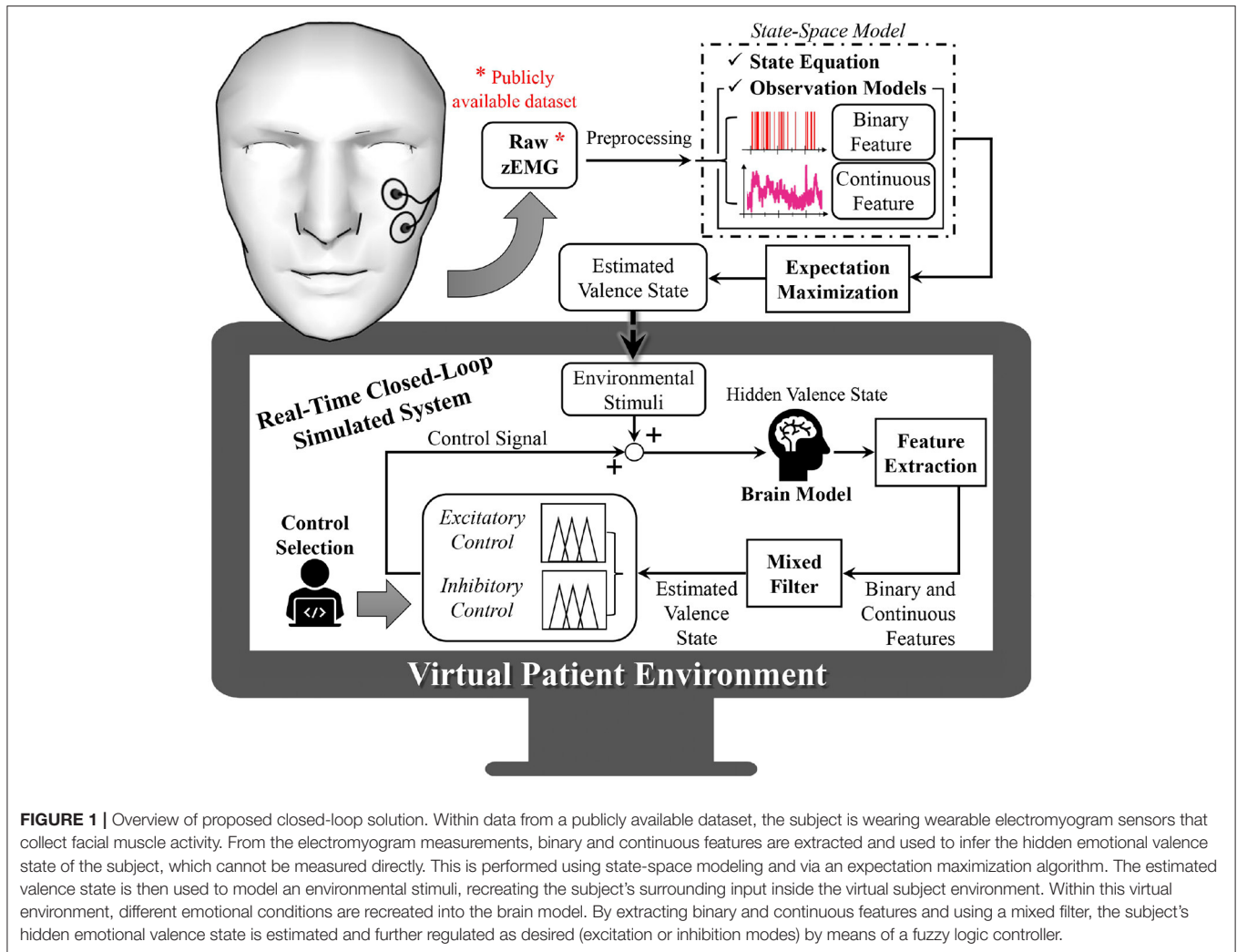
## 2. METHODS

### 2.1. Virtual Subject Environment

An overview of the proposed system is presented in **Figure 1**. As depicted in **Figure 1**, to construct the virtual subject environment we first take the zEMG measurements and preprocess the collected data for our further analysis. From the zEMG data we extract binary and continuous features that will be used both to build the state-space model and to estimate hidden emotional levels. This is possible after the establishment of the continuous and binary observation models associated with the state-space representation of emotional valence. Since the emotional valence progression of the subject is not measurable directly, we use the two simultaneous features and an expectation maximization (EM) algorithm, to model and drive the environmental stimuli within the virtual subject environment. The environmental stimuli are used to recreate, in real-time, different subject-specific emotional valence state-related responses into the simulated brain model. Similarly to the non-real-time case, output from the brain model will then have binary and continuous features extracted before reaching the mixed-filter. The mixed-filter estimates the hidden valence state to supply it to the control method selected of either excitatory or inhibitory control. With these two classes of closed-loop regulation we can analyze the performance of the proposed approach. Finally, the control algorithm determines the control effort necessary and provides it to the brain model, closing the loop. All the simulations of this research were performed using SIMULINK from MATLAB (The Math Works, Inc., Natick, MA) version 2020b.

#### 2.1.1. Dataset

In this research, we develop human brain models using the publicly available Database for Emotion Analysis using Physiological Signals (DEAP) (Koelstra et al., 2012), in which the authors investigated the connection between physiological signals and an associated emotional tag, based on a valence scale. In the DEAP dataset, 32 subjects (16-females and 16-males, mean age 26.9) were asked to watch 1 min segments of 40 different music videos. These videos were selected so that they would capture every aspect of both arousal and valence levels. At the end of each video trial, the researchers gathered each subject's self-assessment regarding emotional valence, on a 1–9



scale. During the experiment, various physiological signals were collected, such as the facial zEMG response at 512 Hz. For our study, the self-assessed emotional valence information is taken as ground truth.

## 2.1.2. State-Space Model

We model the valence state progression by forming stochastic state-space models.

### 2.1.2.1. State Equation

Similar to Prerau et al. (2009), we use a first order autoregressive state-space model,

$$x_{k+1} = x_k + \epsilon_k + s_k + u_k, \quad (1)$$

where  $x_k$  is the hidden valence state at time step  $k$  for  $k = 1, \dots, K$  and  $K$  is the entire experiment duration. The model also includes the process noise as a Gaussian zero-mean random variable  $\epsilon_k \sim \mathcal{N}(0, \sigma_\epsilon^2)$ ,  $s_k$  as a surrogate for any environmental stimuli that influenced the brain at the time of data collection, and  $u_k$  as the input from the controller.

### 2.1.2.2. Observation Models

We include two observation models that capture the evolution of the zEMG signals binary and continuous features so that we can observe the valence state progression in Equation (1). By using two features simultaneously in the model, we achieve a more accurate (i.e., narrower confidence intervals) and more precise emotional state estimation (Prerau et al., 2008). The binary observations  $n_k = \{0, 1\}$ , are modeled as a Bernoulli distribution (McCullagh and Nelder, 1989; Wickramasuriya et al., 2019a,b),

$$P(n_k | x_k) = p_k^{n_k} (1 - p_k)^{1-n_k}, \quad (2)$$

$$p_k = \frac{e^{\gamma + x_k}}{1 + e^{\gamma + x_k}}, \quad (3)$$

where  $p_k$  is the probability of observing a spike given the current valence state amplitude via sigmoidal link function (Equation 3), which has shown to depict frequency or counting datasets well (Wickramasuriya et al., 2019a,b). The continuous observations

$z_k \in \mathbb{R}$  are modeled as,

$$z_k = \alpha + \beta x_k + \omega_k, \quad (4)$$

where  $\alpha$  is a coefficient representing the baseline power of the continuous feature,  $\beta$  is the rate of change in the continuous feature's power, and  $\omega_k$  is a normally distributed zero mean Gaussian random variable  $\omega_k \sim \mathcal{N}(0, \sigma_\omega^2)$ . Both the continuous and binary observations are stated as functions of the valence state  $x_k$ .

### 2.1.3. zEMG Feature Extraction

To perform the estimation process and obtain the hidden valence state, we utilize the zEMG data and extract the binary and continuous features presented in the observation models.

#### 2.1.3.1. Data Preprocessing

We use a third order butterworth bandpass filter between 10 and 250 Hz to remove motion artifacts and other unwanted high frequency noise. Additionally, we use notch filters at 50 Hz and next four harmonics to remove any electrical line interference. Finally, the filtered zEMG signal,  $y_k$ , is segmented into 0.5 second bins with no overlapping.

#### 2.1.3.2. Binary Feature Extraction of Filtered zEMG

As suggested by previous scholars the binary features extracted from the zEMG signal may be associated with the underlying neural spiking activity (Prerau et al., 2008; Amin and Faghih, 2020; Azgomi et al., 2021a). Thus, we estimate the neural spiking pertinent to emotional valence by extracting binary features from the zEMG data. Firstly, the bins of the filtered zEMG signal  $y_k$  are rectified by taking their absolute values and then smoothed with a Gaussian kernel. Similarly to Azgomi et al. (2019) and Yadav et al. (2019), the binary features  $n_k$  are obtained with the Bernoulli distribution,

$$P(n_k|y_k) = q_k^{n_k} (1 - q_k)^{1-n_k}, \quad (5)$$

$$q_k = a y_k, \quad (6)$$

where  $a$  is a scaling coefficient, chosen heuristically to be 0.5, and  $q_k$  is a zEMG amplitude dependent probability function of observing a spike in bin  $k$ , given  $y_k$ .

### 2.1.4. Continuous Feature Extraction of Filtered EMG

Using the filtered zEMG signal  $y_k$ , we also extract the continuous features employing the Welch power spectral density (PSD) of each 0.5 s bin, with a 75% window overlap. Afterwards, for each bin, we compute the bandpower of the PSD result from 10 to 250 Hz, before taking the logarithm. Finally, we normalize the entire signal on a 0–1 scale, to provide insight of the relative band power of the zygomaticus major muscle activity, across all 40 1-min trials.

### 2.1.5. Hidden Valence State Estimation

To estimate the emotional valence fluctuations within the experimental data, we employ the state-space representation shown in Equation (1) without the control effort and

environmental stimuli, since at this time, there is no control signal and the stimuli is inherent in the data. The hidden valence state process is defined by

$$x_{k+1} = x_k + \epsilon_k. \quad (7)$$

Given the complete values for both extracted binary  $N_{1:K} = \{n_1, \dots, n_K\}$  and continuous  $Z_{1:K} = \{z_1, \dots, z_K\}$  features, we use the EM algorithm to estimate the model parameters  $\theta = [\alpha, \beta, \sigma_\epsilon, \sigma_\omega]$  and the hidden valence state  $x_k$ . The EM algorithm provides a way to jointly estimate the latent state and parameters of the state-space models. Composed of two steps, namely, Expectation step (E-step) and Maximization step (M-step), the EM algorithm: (1) finds the expected value of the complete data log-likelihood, and (2) maximizes the parameters corresponding to this data log-likelihood. The algorithm iterates between these two steps until convergence (Wickramasuriya et al., 2019a,b; Yadav et al., 2019). The following equations show how at iteration  $(i + 1)$  values are recursively predicted with estimates and parameters from iteration  $i$  (e.g.,  $x_0^{(i)}, \sigma_\epsilon^{2(i)}$ ).

#### 2.1.5.1. E-Step

##### 2.1.5.1.1. Kalman-Based Mixed-Filter (Forward-Filter).

$$x_{k|k-1} = x_{k-1|k-1} \quad (8)$$

$$\sigma_{k|k-1}^2 = \sigma_{k-1|k-1}^2 + \sigma_\epsilon^{2(i)} \quad (9)$$

$$C_k = \left( \beta^{(i)2} \sigma_{k|k-1}^2 + \sigma_\omega^{2(i)} \right)^{-1} \sigma_{k|k-1}^2 \quad (10)$$

$$\hat{x}_k = x_{k|k} = x_{k|k-1} + C_k \left[ \beta^{(i)} (z_k - \alpha^{(i)} - \beta^{(i)} x_{k|k-1}) + \sigma_\omega^{2(i)} (n_k - p_{k|k}) \right] \quad (11)$$

$$\hat{\sigma}_k^2 = \sigma_{k|k}^2 = \left[ (\sigma_{k|k-1}^2)^{-1} + p_{k|k} (1 - p_{k|k}) + (\sigma_\omega^{2(i)})^{-1} \beta^{(i)2} \right]^{-1} \quad (12)$$

where  $k = 1, \dots, K$ ;  $\hat{x}_k$  is the estimated valence state; and  $\hat{\sigma}_k^2$  constitute the corresponding standard deviation.

##### 2.1.5.1.2. Fixed-Interval Smoothing Algorithm (Backward-Filter).

$$A_k = \sigma_{k|k}^2 (\sigma_{k+1|k}^2)^{-1} \quad (13)$$

$$x_{k|K} = x_{k|k} + A_k (x_{k+1|K} - x_{k+1|k}) \quad (14)$$

$$\sigma_{k|K}^2 = \sigma_{k|k}^2 + A_k^2 (\sigma_{k+1|k}^2 - \sigma_{k+1|K}^2) \quad (15)$$

### 2.1.5.1.3. State-Space Covariance Algorithm.

$$\sigma_{k,u|k} = A_k \sigma_{k+1,u|k} \quad (16)$$

$$W_{k|K} = \sigma_{k|K}^2 + x_{k|K}^2 \quad (17)$$

$$W_{k-1,k|K} = \sigma_{k-1|K} + x_{k-1|K} x_{k|K} \quad (18)$$

for  $1 \leq k \leq u \leq K$ .

### 2.1.5.2. M-Step

$$x_0^{(i+1)} = x_{1|k} \quad (19)$$

$$\begin{aligned} \sigma_\omega^{2(i+1)} = & K^{-1} \sum_{k=1}^K z_k^2 + K \alpha^{2(i+1)} \\ & + \beta^{2(i+1)} \sum_{k=1}^K W_{k|K} - 2\alpha^{(i+1)} \sum_{k=1}^K z_k \\ & - 2\beta^{(i+1)} \sum_{k=1}^K x_{k|K} z_k \\ & + 2\alpha^{(i+1)} \beta^{(i+1)} \sum_{k=1}^K x_{k|K} \end{aligned} \quad (20)$$

$$\begin{aligned} \begin{bmatrix} \alpha^{(i+1)} \\ \beta^{(i+1)} \end{bmatrix} = & \begin{bmatrix} K & \sum_{k=1}^K x_{k|K} \\ \sum_{k=1}^K x_{k|K} & \sum_{k=1}^K W_{k|K} \end{bmatrix}^{-1} \\ & \times \begin{bmatrix} \sum_{k=1}^K z_k \\ \sum_{k=1}^K x_{k|K} z_k \end{bmatrix} \end{aligned} \quad (21)$$

$$\sigma_\epsilon^{2(i+1)} = K^{-1} \sum_{k=1}^K [W_{k|K} - 2W_{k-1,k|K} + W_{k-1|K}] \quad (22)$$

### 2.1.6. Environmental Stimuli Model

We model the environmental stimuli referred to in Equation (1) as a way to capture and recreate the subject's response to high or low valence trials. This allows for the simulation of subject-specific HV and LV conditions. The environmental stimuli are calculated by finding the difference between adjacent elements of the estimated valence state  $\hat{x}_k$ , as in

$$s_k = \hat{x}_{k+1} - \hat{x}_k \quad (23)$$

for  $k = 1, \dots, K-1$ . Then, we assume a sinusoidal harmonic formulation to model the environmental stimuli in either HV or LV trials,

$$s_k = \sum_{j=1}^{100} \rho_j \sin(\zeta_j k + \phi_j) \quad (24)$$

Through inspection across all subjects, we notice that HV trials tend to have a higher mean and standard deviation compared

to LV ones. Thus, to avoid fitting outliers to the harmonic model depicted in Equation (24), we select the six trials with highest mean and standard deviation of estimated valence levels for fitting  $s_k$  to HV, and the six trials with the lowest mean and standard deviation to model LV periods. Additionally, we consider a transition period between each different valence state, as approximated by a linear relationship of 0.5 s in duration. This is done separately for each subject to ensure personalized models. Data from an exemplary subject is depicted in **Figure 2**, in which every step of the process is illustrated separately, i.e., raw zEMG to extracted features and valence state and finally obtaining a corresponding environmental stimuli. In addition, in **Figure 3**, the estimated emotional valence state for the same exemplary subject is presented with 95% confidence intervals. Of the 23 subjects available in the dataset, we excluded five participants due to a lack of emotional response found when comparing between LV and HV periods, that is, both emotional periods have shown equivalent outcomes regarding both features and estimated valence state.

## 2.2. Closed-Loop Control Design

With the virtual subject environment in place, we explore the regulation of emotional valence. Similar to the feature extraction process, we simulate the binary and continuous responses simultaneously from the internal brain state. In other words, we use Equations (2)–(4) to recreate within the virtual subject environment what would be inherent to the zEMG data in the real world. Then, these two features are fed to a Kalman-based mixed-filter to estimate the hidden valence state in an online fashion. The estimated state is averaged out in a 10-s window to smooth any abrupt changes before reaching the fuzzy controller, which then derives the control effort  $u_k$  in real-time. A diagram of the closed-loop is depicted in **Figure 4**. As the hidden valence state cannot be measured directly, we use the recursive, Kalman-based mixed-filter to estimate the latent valence state inherent to the brain model as detailed in Equations (8)–(12). As shown in **Figure 4**, this filter takes in both binary and continuous observations to compute the prior distribution using a Chapman-Kolmogorov equation, then finds the measurement likelihood via Bayes theorem, which can be summarized with, respectively,

$$p(x_k | n_{k-1}, z_{k-1}), \quad (25)$$

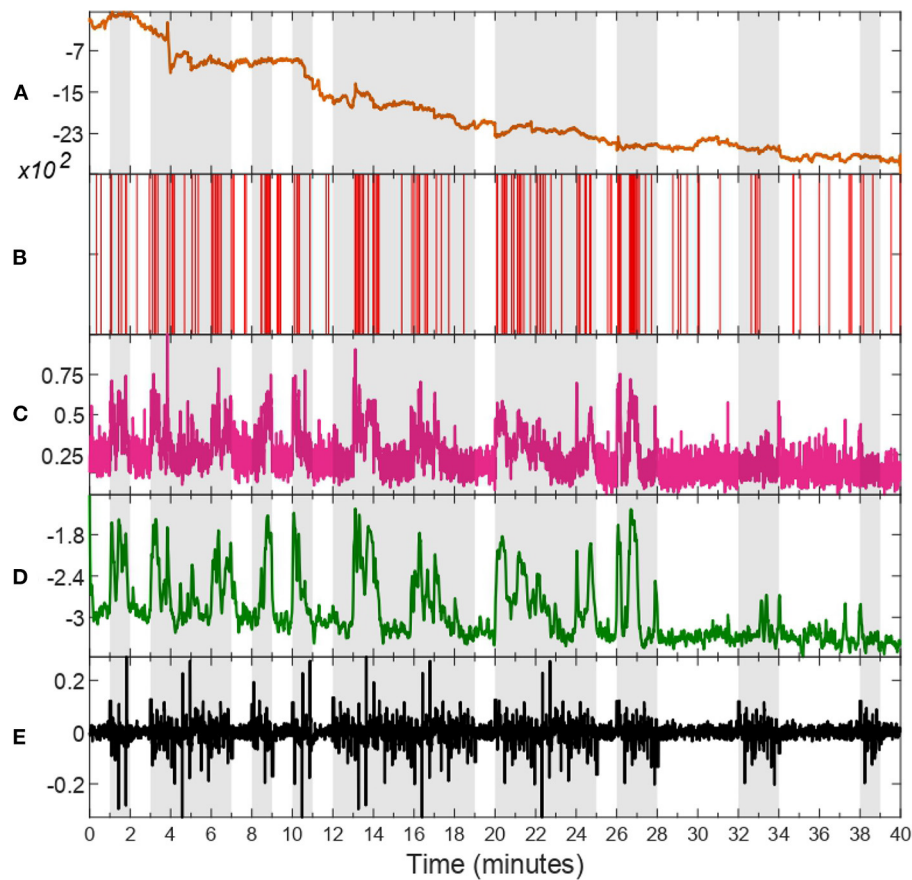
and

$$p(x_k | n_k, z_k). \quad (26)$$

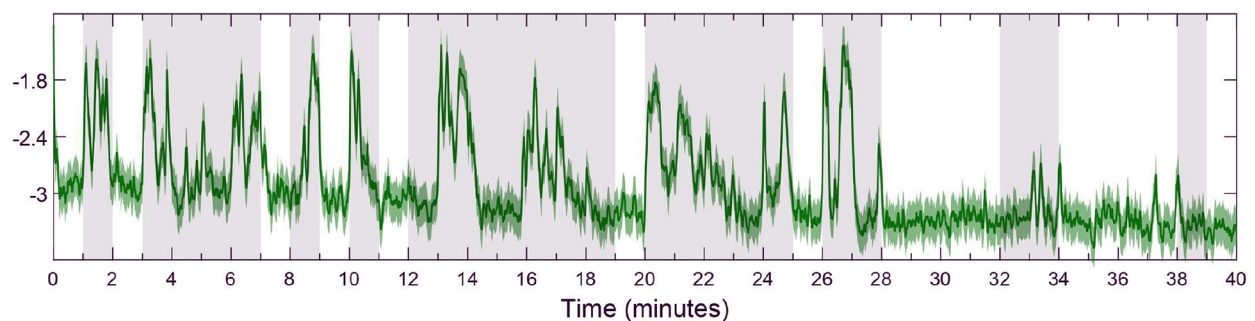
### 2.2.1. Fuzzy Control

We use a Mamdani-type fuzzy logic controller with the fuzzy rules shown in **Table 1** to regulate the subject's emotional valence to a more desired level, i.e., during inhibitory mode of control action, the goal is to achieve and remain in the same valence level characterized by the LV period—and vice-versa for the excitatory controller. As it can be observed in **Figures 1, 4** and **Table 1**, the input signal for the controller is the estimated valence state and not a prediction error as it is more common in control studies. After analyzing the open-loop response of





**FIGURE 2 |** zEMG data, corresponding features, estimated valence state, and environmental stimuli of subject 18. Trials characterized as high valence (HV) are shaded in gray, whilst unshaded ones as representative of low valence (LV). The raw zEMG collected is presented in **(A)** in orange, while **(B,C)** show the extracted features, binary (red) and continuous (pink), respectively. **(D)** illustrates the hidden valence state (green) attained with the EM algorithm by employing both features shown in **(B,C)**. The last **(E)** shows the environmental stimuli (black) obtained from the valence state progression in **(D)**.

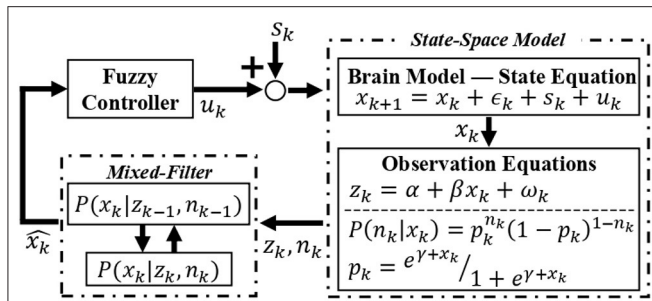


**FIGURE 3 |** Detail of estimated emotional valence state for subject 18 with 95% confidence intervals. The white background depicts LV periods while the gray-shaded areas show HV results. The solid green line shows the estimated valence state while the green region around it is a 95% confidence interval.

all subjects we designed a set of membership functions capable of directly regulating the emotional valence without subtracting it from a target reference. With this, we could employ more intuitive membership functions as depicted in **Figure 5**. Similarly to previous authors (Azgomi et al., 2021b), the fuzzy output can

be obtained with,

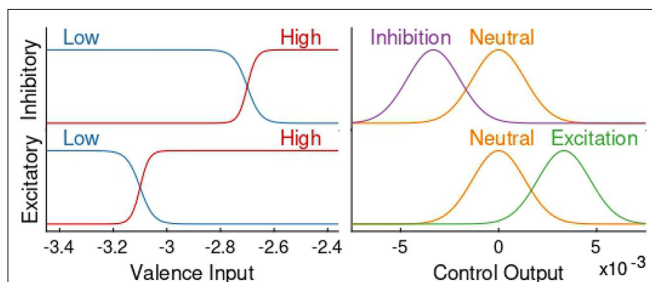
$$\mu_{mamdani}(k) = \mu_m(k) = \max_j(\min(\mu_{valence}(v))) \quad (27)$$



**FIGURE 4 |** Overview of the closed-loop solution. The environmental stimuli  $S_k$  is added to the control signal  $u_k$  to form an input of the state-space brain model. The internal emotional valence state  $x_k$  is governed by the state equation and by employing the observation equations, binary and continuous features are extracted and taken in by the recursive mixed-filter. The filter estimates and tracks the hidden brain state  $\hat{x}_k$ , supplying this signal for the controller. Finally, the controller takes the current estimated valence state and generates a control signal  $u_k$  back to the brain model, thus closing the loop. This control signal is responsible for changing the valence state in the desired direction, i.e., increasing if excitatory or decreasing if inhibitory action.

**TABLE 1 |** Fuzzy controller rule base.

Input (IF): valence levels	Inhibitory Output (THEN): control action	Excitatory Output (THEN): control action
Low valence	Neutral	Excitation
High valence	Inhibition	Neutral



**FIGURE 5 |** Excitatory and inhibitory fuzzy membership functions. The left side shows membership functions of the controller's input, whilst the right side display the ones for the output. The top and bottom row depict, respectively, membership functions of the inhibitory and excitatory controllers. In all four graphs the y axis depicts the degree of membership for every case, in which the lowest value is zero association with that function and the highest value is total association.

where  $j$  designate the active rule at each time step  $k$  and  $\mu_{valence}$  is the fuzzified valence input  $v$ . The crisp output of the fuzzy controller, i.e., the control signal  $u_k$ , is attained using the *centroid* method as follows,

$$u_k = \frac{\int \mu_m(k) \cdot k \, dk}{\int \mu_m(k) \, dk} \quad (28)$$

With a fuzzy logic controller, crisp input values are transformed to degrees of membership of certain functions called membership

functions in the *fuzzification* process. Then, using the pre-determined fuzzy rules the *fuzzy inference* process takes place, in which a connection between all *fuzzified* inputs is made. This results in degrees of membership of a set of output membership functions, which are then *defuzzified* to produce a final representative crisp value (Qi et al., 2019). This fuzzy logic process is convenient when dealing with complex systems, such as those biological in nature, since it allows for the emergence of complex control behaviors using relatively simple constructions (Lilly, 2011).

### 3. RESULTS

In this section, we present the results obtained for subject 7 in three different simulation scenarios: open-loop, inhibitory closed-loop, and excitatory closed-loop. The results associated with other subjects are also available in the **Supplementary Material**. We simulate with an environmental stimulus that is either half LV then half HV or vice versa. During the first minute, the controller is suspended to let the mixed-filter converge. The results are presented in **Figure 6**. As depicted in sub-panel (a) of I and II in **Figure 6**, all three scenarios for one particular subject have the same environmental stimuli in common, either starting with LV or with HV.

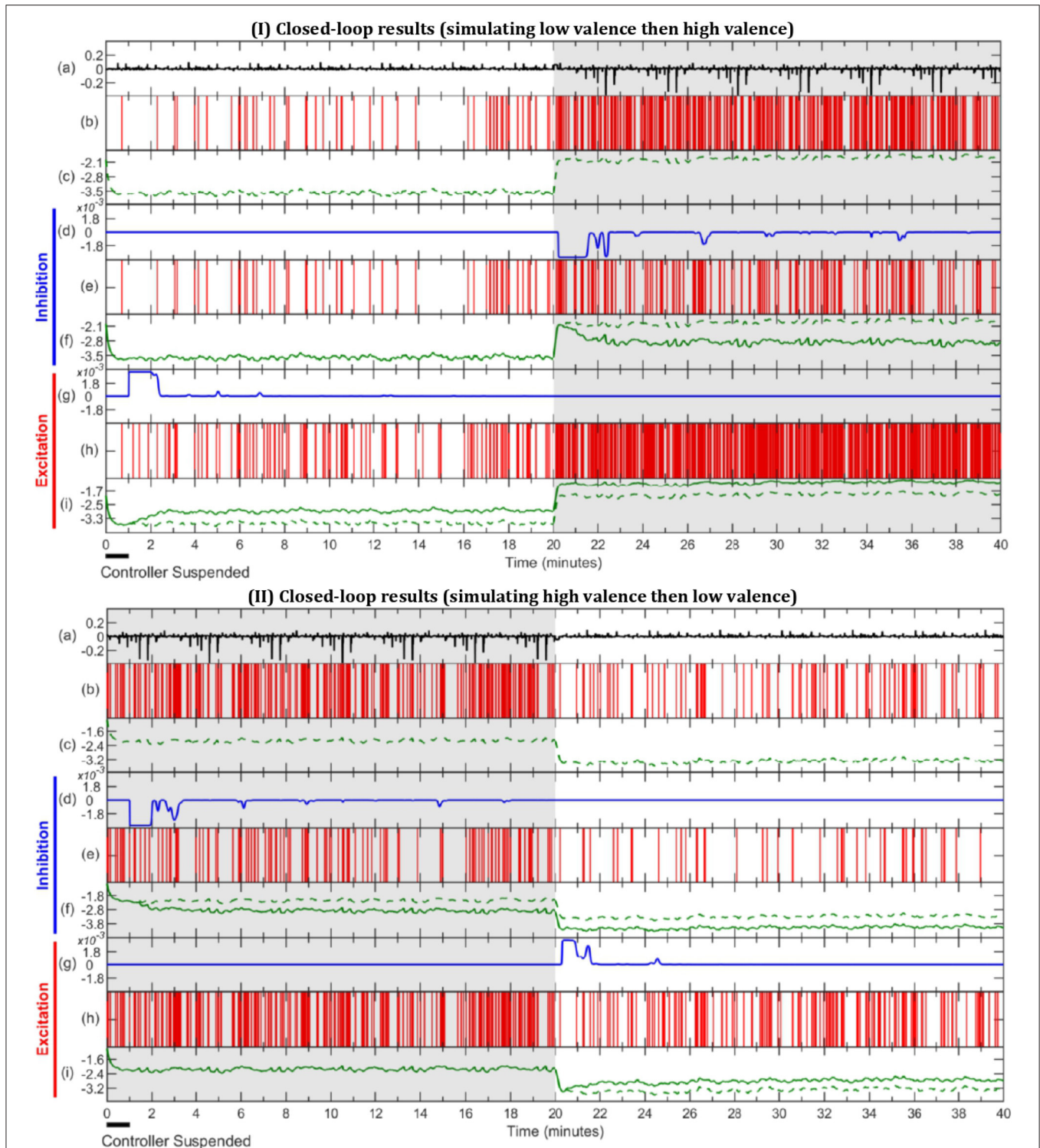
**Scenario 1 - Open-Loop:** Since in the open-loop scenario there is no control effort ( $u_k = 0$ ), it can be omitted and the results are shown within the spike activity, depicted in sub-panel (b), and the corresponding estimated internal state depicted in sub-panel (c) and in dashed lines in both (f) and (i) sub-panels. It is observed, in sub-figure I of **Figure 6**, that the estimated valence state increases from the period of LV in the first half to HV in the second half, and so does the frequency of spikes. In contrast, sub-figure II of **Figure 6** shows valence levels and number of spikes declining from the first half (HV) to the second half (LV).

**Scenario 2 - Inhibitory Closed-Loop:** The inhibitory results are observed in sub-panels (d, e, f) of both I and II in **Figure 6**. The control signal is zero during the LV periods of the simulations (i.e., during the first half in I and for the second half in II). It is not until the controller detects a HV period that the control effort takes a negative value ( $u_k < 0$ ) to inhibit the emotional valence, effectively lowering the number of spikes shown in sub-panel (e) and the estimated valence state depicted in sub-panel (f), as compared to the open-loop case.

**Scenario 3 - Excitatory Closed-Loop:** The last 3 sub-panels (g, h, i) from both I and II of **Figure 6** depict the results of the excitatory controller. From sub-panel (g), we can see there is no control effort in periods of HV; both in the second half of I and first half of II. Once the controller detects a low valence state, it outputs a positive control effort ( $u_k > 0$ ), which increases the number of spikes and estimated valence level in sub-panels (h) and (i), as compared to the open-loop.

### 4. DISCUSSION AND CONCLUSIONS

In this study, we use experimental data to build a virtual subject environment, allowing us to simulate and regulate emotional



**FIGURE 6 |** Simulation results of open-loop, inhibitory closed-loop and excitatory closed-loop scenarios for subject 7. In sub-figure I the external stimulus is comprised of half LV, then half HV, with sub-figure II being the opposite. In both **(I,II)**, LV, and HV periods are represented with unshaded and gray-shaded areas, respectively. (a) depicts environmental stimulus (black) used in all three simulation scenarios. The (b,c) show spike activity (red) and estimated valence state (green, dashed) during the open-loop, respectively. (d–f) display inhibitory closed-loop results, with (d) showing control effort (blue), (e) the corresponding binary signal (red) and (f) the comparison between open-loop (green, dashed) and closed-loop (green, solid) valence state. In a similar fashion, (g–i) exhibit the excitatory closed-loop outcome.

valence levels using a state-space brain model and a fuzzy logic feedback controller. To the best of our knowledge, in this *in silico* feasibility study, we present the first closed-loop control framework for emotional valence state using biofeedback from facial muscles. We use two simultaneous observation models, one binary and one continuous, to relate zEMG measurements to the hidden emotional valence state. The valence state is assumed to be governed by a state-space formulation and is converted from a 1 to 9 valence scale obtained from the self-assessment of subjects from the dataset, to the high (above 5) or low valence level used in this study. These valence labels were previously used by scholars as ground truth and were also employed here to determine subject-specific simulation parameters (Yadav et al., 2019). This was done by selecting specific LV and HV trials for modeling based on a trend in the mean and standard deviation of the estimated valence state between the two categories. To capture the surrounding stimuli influencing the affective levels of the subject and incorporate them into simulation, we use the estimated emotional valence progression and a high-order harmonic formulation. This modeling and simulation of the environmental stimuli is currently necessary to evoke representative subject-specific emotional valence responses within the simulated brain model. Thus, modern control techniques can be systematically investigated *in silico*, allowing for the development of this research field without risking harm to any patients.

In the current stage of this research on closed-loop emotional valence regulation, we focus our contributions on developing the closed-loop simulated framework and opted for using a fuzzy logic controller to regulate the estimated valence state in simulated profiles. While the accuracy of the classification method is paramount for the success of our method, we employed the same methodology for classifying between low valence and high valence states which reported a 89% accuracy in previous works (Yadav et al., 2019). This value is on par with other state-of-the-art methods however, relying on physiological measurements and estimation of the brain state, instead of externalized facial or vocal expressions.

Using the proposed knowledge-based controller we successfully verify the *in silico* feasibility of the presented methods. By employing a set of simple logic rules the fuzzy system is capable of producing complex regulating behaviors (Lilly, 2011). This is extremely valuable since insight about the system can come in many ways, such as from doctors, other researchers, or the individual itself. Moreover, the fuzzy structure allows for an uncomplicated expandability feature which means other physiological signals could be simply incorporated while designing the control systems (Azgomi et al., 2021a,b). This could further enhance the approach for valence regulation.

In previous research for closed-loop regulation of human-related dynamics, scholars have developed simulators to explore controller designs for Parkinson's disease, cognitive stress, depression and other neurological and neuropsychiatric disorders, as well as for anesthetic delivery, hemodynamic stability, and mechanical ventilation (Boayue et al., 2018; Yang et al., 2018b; Azgomi et al., 2019; Parvinian et al., 2019; Fleming et al., 2020; Ionescu et al., 2021). Here, the proposed architectures

set initial steps for a future wearable machine interface (WMI) implementation, as we achieved simulation of emotional valence controllers for both inhibitory and excitatory goals, demonstrating great potential in helping individuals maintain daily mental well-being (Azgomi and Faghih, 2019). While no commercial wearable solution for facial EMG measurement is available yet, the potential for this non-invasive procedure to regulate mental states encourages future efforts.

During excitatory action, we observe an increase in number of spikes and overall emotional valence state when needed and, for inhibition, our approach obtained less spikes and a lower valence level as the need arose. However, the amount of response varied with each subject due to a few reasons. One factor can be attributed to the use of a single mono-objective fuzzy controller design, in which the controller can act locally in the first half of the experiment, correctly adjusting the mental state, without considering that the environmental stimuli are going to further push the subject's valence level in the second half. This architecture also does not account for each individual peculiarities, i.e., lack or abundance of emotional engagement throughout the experiment. Further research needs to explore the optimization of fuzzy membership functions, to adapt for different persons and variations in time. Because the performance of fuzzy logic controllers are highly dependent on their parameters and structure, optimization algorithms could also improve the overall results as the parameters would not rely on pre-determined knowledge of the system (Qi et al., 2019).

In the exemplary subject depicted in **Figure 6** we can observe an inhibitory action taking place in the HV periods of inhibition simulation and lowering of the number of spikes and estimated valence level as compared to the open-loop. Similarly, we can observe the excitatory controller acting in LV periods and increasing the spike frequency and valence levels, accordingly. Overall, subjects 3–5, 8, 11, and 17 (**Supplementary Figures S4–S6, S9, S12, S18**) showed similar results to the exemplary subject depicted in **Figure 6**, accomplishing reasonable regulation across all scenarios. Of the remaining 10 subjects, 7 had good performance in all inhibitory scenarios (**Supplementary Figures S2, S3, S10, S11, S13, S16, S17**) while 4 out of 10 had good performance in at least one excitatory scenario (**Supplementary Figures S10, S14–S16**). This could suggest that HV regulation is more challenging possibly due to the high variability nature of this mental state.

In addition to the subject exemplified in **Figure 6**, *t*-test analysis between the open- and closed-loop simulations with 17 out of 23 subjects was performed, as detailed in **Table 2**. Additionally, **Figure 7** displays the distribution of data used during the *t*-test for the case of LV then HV order of environmental stimuli. The HV then LV order is also included in the **Supplementary Material** and presents a similar analysis. As seen both in **Figure 7** and **Table 2**, the results show LV periods to be significantly different during excitatory action and HV trials to be significantly different throughout inhibition, regarding both the average valence level and number of spikes. This can be an indicative that the proposed controller was able to perform as desired and alter the emotional state of various



**TABLE 2 |** Statistical analysis-*p*-values.

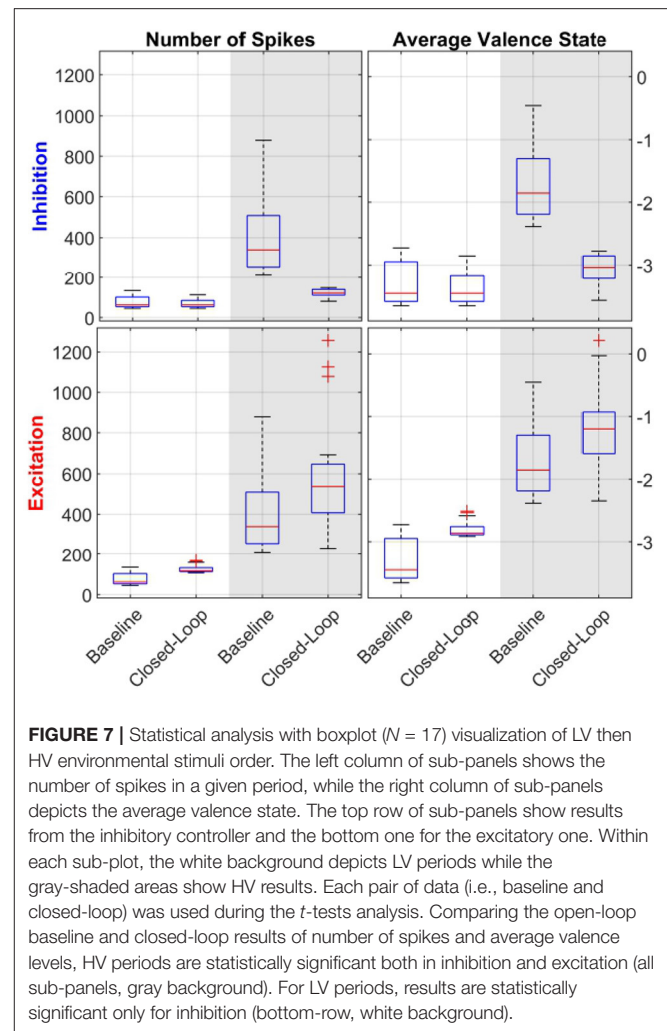
		LV then HV		HV then LV	
		Number of spikes	Average valence	Number of spikes	Average valence
Inhib.	LV	0.1689	0.1947	$2 \times 10^{-6}$	$6 \times 10^{-6}$
	HV	$2 \times 10^{-5}$	$5 \times 10^{-7}$	$3 \times 10^{-4}$	$1 \times 10^{-5}$
Excit.	LV	$7 \times 10^{-6}$	$8 \times 10^{-6}$	$3 \times 10^{-5}$	$3 \times 10^{-5}$
	HV	$4 \times 10^{-5}$	$5 \times 10^{-6}$	0.0554	0.0398

Bold values are significant *p*-values.

subjects when required. In a similar manner, LV periods were not significantly different during inhibitory regulation if the LV was at the beginning of simulation (both in spike count, and mean valence levels). Comparing HV periods throughout excitation, the number of spikes was not significantly different when the HV period happen before LV. These results are indicative that the controller is able to detect when changes to the brain state are not required. The reason the affective state is significantly different in the second half of the experiment in cases it was not necessary (LV inhibition and HV excitation) is due to the fact that the proposed controller is not multi-objective and a regulation goal is selected beforehand | either to excite or to inhibit. Thus, after properly adjusting the brain state in the first half of the simulation, the second half will be different in comparison to the open-loop baseline and the mono-objective nature of this approach is incapable of addressing the matter. Further research is still required.

A few subjects (4, 11, 12, 20, 21, and 23) had poor emotional valence state estimation and were discarded from the statistical analysis which also show directions for improving the proposed approach. These participants showed similar number of spikes and valence levels, during both LV and HV periods, within the open-loop scenario. Thus, when taken to a closed-loop solution, the fuzzy controller is impaired from distinguishing high and low valence levels and leads to unsatisfactory results. However, this poor valence estimation could be due to many factors such as the person not being emotionally engaged during the original data collection or distracted during the experiment (Chaouachi and Frasson, 2010). Similarly to previous scholars (Yadav et al., 2019), we investigate the performance of the emotional valence estimation with a 95% confidence intervals metric, as depicted in **Figure 3**. As it can be observed in **Figure 3**, the confidence intervals reside close to the actual recovered state and further validate the proposed state-space estimation procedure. Moreover, it is possible that these discarded subjects required additional physiological measurements (e.g., electrocardiogram, skin conductance, pupil size) to improve the estimation of the internal brain state. As mentioned above, the flexibility of the proposed state-space and fuzzy logic controller framework could easily incorporate additional physiological signals.

The present study has a few limitations. The dataset used had conflicting metadata on 9 of the 32 subjects, resulting in an impossibility of recovering the position of all 40 trials and



**FIGURE 7 |** Statistical analysis with boxplot ( $N = 17$ ) visualization of LV then HV environmental stimuli order. The left column of sub-panels shows the number of spikes in a given period, while the right column of sub-panels depicts the average valence state. The top row of sub-panels show results from the inhibitory controller and the bottom one for the excitatory one. Within each sub-plot, the white background depicts LV periods while the gray-shaded areas show HV results. Each pair of data (i.e., baseline and closed-loop) was used during the *t*-tests analysis. Comparing the open-loop baseline and closed-loop results of number of spikes and average valence levels, HV periods are statistically significant both in inhibition and excitation (all sub-panels, gray background). For LV periods, results are statistically significant only for inhibition (bottom-row, white background).

thus, these subjects had to be discarded. Additionally, in real-world scenarios as in the dataset used, emotional valence has a spectrum of levels, but we assume only two possible states of high and low valence. This decision also reflects in the controller design in which we experiment with only two classes of closed-loop regulation, i.e., excitation and inhibition. Even with this limitation, it should be noted that both the mixed-filter and designed control provide continuous estimation and control objectives allowing for a finer regulation within this spectrum of emotions. This can be addressed in future research. Moreover, this simulation study does not incorporate the controller dynamics and real-world actuators. To implement the proposed architectures in real-world scenarios, it is paramount to consider how valence needs to be modulated, not only in terms of which actuators to use but also how frequent should interventions take place. These are challenging to address, especially when dealing with such a complex organ as the human brain, and require further investigation. In that sense, future human subject experiments shall be designed to explore the dynamics of possible actuation methods to regulate valence states. Previous scholars

have observed emotional brain responses from changes in lighting or music (Schubert, 2007; Vandewalle et al., 2010; Droit-Volet et al., 2013). These would be interesting to investigate since they are also non-invasive procedures and could be incorporated in a practical system. Future research into using adaptive and predictive control strategies would also be beneficial to address some of the biological intrinsic variations of an individual. Similarly, the applicability of the proposed approach in the real world depends on the real-time estimation of mental states. At this time, we illustrate the feasibility of the approach by incorporating a simulation of brain responses on a per individual basis. Once implemented, this simulation is no longer required. However, a “training” session might be necessary to calibrate the system for each subject’s peculiarities. In addition, robust state estimation or robust control design can be of tremendous importance for a real-world application. Lastly, we extracted features from LV and HV trials from EMG signal of the Zygomaticus major facial muscle, which has been depicted as a good indicator of valence (Brown and Schwartz, 1980; Ekman et al., 1980; Tan et al., 2012). As a future direction of this research, an investigation to quantify the performance in detecting fake emotional expressions via the zEMG signal would be beneficial to further enhance the proposed approach to be implemented in real life.

Using the proposed architecture, we were able to regulate one’s emotional state, specifically emotional valence levels, by implementing a fuzzy controller that acted on a state-space model of the human brain. With a similar approach, a WMI could, in the future, be used to recommend a specific music track for a person feeling down, advise a change in lighting for someone in a bad mental state, or even offer a cup of green tea if the user wants to maintain a desired level of well-being (Athavale and Krishnan, 2017; Cannard et al., 2020). While we used experimental data to design a closed-loop system for regulating an internal valence state in a simulation study, a future direction of this research would be designing human subject experiments

to close the loop in real-world settings. In our future work, we plan to validate the valence state estimator in real-time and close the loop accordingly. For example, we plan to incorporate safe actuators such as music or visual stimulation to close the loop. More research is needed but this suggests an important new step toward new clinical applications and the self-management of mental health.

## DATA AVAILABILITY STATEMENT

The publicly available dataset used in this study can be found in <http://www.eecs.qmul.ac.uk/mmv/datasets/deap/>.

## AUTHOR CONTRIBUTIONS

RF conceived and designed the study. LB, AE, HA, and RF developed the algorithms, analysis tools, and revised the manuscript. LB and AE performed research, analyzed data, and wrote the manuscript. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported in part by NSF CAREER Award 1942585–MINDWATCH: Multimodal Intelligent Noninvasive brain state Decoder for Wearable Adaptive Closed-loop architectures and NSF grant 1755780–CRII: CPS: Wearable-Machine Interface Architectures, and NYU start-up funds.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fncom.2022.747735/full#supplementary-material>

## REFERENCES

- Abdelnour, F., Voss, H. U., and Raj, A. (2014). Network diffusion accurately models the relationship between structural and functional brain connectivity networks. *Neuroimage* 90, 335–347. doi: 10.1016/j.neuroimage.2013.12.039
- Ahmadi, M. B., Craik, A., Azgomi, H. F., Francis, J. T., Contreras-Vidal, J. L., and Faghih, R. T. (2019). “Real-time seizure state tracking using two channels: a mixed-filter approach,” in *2019 53rd Asilomar Conference on Signals, Systems, and Computers* (Pacific Grove, CA: IEEE), 2033–2039.
- Al-Shargie, F., Tang, T. B., Badruddin, N., and Kiguchi, M. (2015). “Mental stress quantification using eeg signals,” in *International Conference for Innovation in Biomedical Engineering and Life Sciences* (Springer), 15–19.
- Al-Sharora, E., Al-Khassawneh, M., and Aviyente, S. (2018). Tensor based temporal and multilayer community detection for studying brain dynamics during resting state fmri. *IEEE Trans. Biomed. Eng.* 66, 695–709. doi: 10.1109/TBME.2018.2854676
- Amin, M. R., and Faghih, R. T. (2020). Identification of sympathetic nervous system activation from skin conductance: a sparse decomposition approach with physiological priors. *IEEE Trans. Biomed. Eng.* 68, 1726–1736. doi: 10.1109/TBME.2020.3034632
- Amin, R., Shams, A. F., Rahman, S. M., and Hatzinakos, D. (2016). “Evaluation of discrimination power of facial parts from 3d point cloud data,” in *2016 9th International Conference on Electrical and Computer Engineering (ICECE)* (Dhaka: IEEE), 602–605.
- Anuja, T., and Sanjeev, D. (2020). “Speech emotion recognition: a review,” in *Advances in Communication and Computational Technology* eds G. Hura, A. Singh, and I. Siong Hoe L (Singapore: Springer).
- Athavale, Y., and Krishnan, S. (2017). Biosignal monitoring using wearables: Observations and opportunities. *Biomed. Signal Process. Control* 38, 22–33. doi: 10.1016/j.bspc.2017.03.011
- Azgomi, H. F., Cajigas, I., and Faghih, R. T. (2021a). Closed-loop cognitive stress regulation using fuzzy control in wearable-machine interface architectures. *IEEE Access*. 9, 106202–106219. doi: 10.1109/ACCESS.2021.3099027
- Azgomi, H. F., and Faghih, R. T. (2019). “A wearable brain machine interface architecture for regulation of energy in hypercortisolism,” in *2019 53rd Asilomar Conference on Signals, Systems, and Computers* (Pacific Grove, CA: IEEE), 254–258.
- Azgomi, H. F., Hahn, J.-O., and Faghih, R. T. (2021b). Closed-loop fuzzy energy regulation in patients with hypercortisolism via inhibitory and excitatory intermittent actuation. *Front. Neurosci.* 15, 695975. doi: 10.3389/fnins.2021.695975
- Azgomi, H. F., Wickramasuriya, D. S., and Faghih, R. T. (2019). “State-space modeling and fuzzy feedback control of cognitive stress,” in *2019 41st Annual*

- International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (Berlin: IEEE), 6327–6330.
- Azuar, D., Gallud, G., Escalona, F., Gomez-Donoso, F., and Cazorla, M. (2019). “A story-telling social robot with emotion recognition capabilities for the intellectually challenged,” in *Iberian Robotics Conference* (Porto: Springer), 599–609.
- Boayue, N. M., Csifcsák, G., Puonti, O., Thielscher, A., and Mittner, M. (2018). Head models of healthy and depressed adults for simulating the electric fields of non-invasive electric brain stimulation. *F1000Research* 7, 704. doi: 10.12688/f1000research.15125.2
- Bozhkov, L., Koprinkova-Hristova, P., and Georgieva, P. (2017). Reservoir computing for emotion valence discrimination from eeg signals. *Neurocomputing* 231, 28–40. doi: 10.1016/j.neucom.2016.03.108
- Brown, S.-L., and Schwartz, G. E. (1980). Relationships between facial electromyography and subjective experience during affective imagery. *Biol. Psychol.* 11, 49–62. doi: 10.1016/0301-0511(80)90026-5
- Burzagli, L., and Naldini, S. (2020). “Affective computing and loneliness: how this approach could improve a support system,” in *International Conference on Human-Computer Interaction* (Copenhagen: Springer), 493–503.
- Cacioppo, J., Berntson, G., Larsen, J., Poehlmann, K., and Ito, T. (2000). “The psychophysiology of emotion,” in *Handbook of Emotions* (The Guilford Press), 173–191.
- Cacioppo, J., Petty, R., Losch, M., and Kim, H. (1986). Electromyographic activity over facial muscle regions can differentiate the valence and intensity of affective reactions. *J. Pers. Soc. Psychol.* 50, 260–268. doi: 10.1037/0022-3514.50.2.260
- Cai, Y., Guo, Y., Jiang, H., and Huang, M.-C. (2018). Machine-learning approaches for recognizing muscle activities involved in facial expressions captured by multi-channels surface electromyogram. *Smart Health* 5, 15–25. doi: 10.1016/j.smhl.2017.11.002
- Cannard, C., Brandmeyer, T., Wabbeh, H., and Delorme, A. (2020). Self-health monitoring and wearable neurotechnologies. *Handbook Clin. Neurol.* 168, 207–232. doi: 10.1016/B978-0-444-63934-9.00016-0
- Cannon, W. B. (1927). The james-lange theory of emotions: a critical examination and an alternative theory. *Am. J. Psychol.* 39, 106–124. doi: 10.2307/1415404
- Chaouachi, M., and Frasson, C. (2010). “Exploring the relationship between learner eeg mental engagement and affect,” in *International Conference on Intelligent Tutoring Systems* (Berlin: Springer), 291–293.
- Chronaki, G., Hadwin, J. A., Garner, M., Maurage, P., and Sonuga-Barke, E. J. S. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. *Br. J. Dev. Psychol.* 33, 218–236. doi: 10.1111/bjdp.12075
- Conway, C. R., Udayar, A., and Schachter, S. C. (2018). Neurostimulation for depression in epilepsy. *Epilepsy Behav.* 88, 25–32. doi: 10.1016/j.yebeh.2018.06.007
- Couette, M., Mouchabac, S., Bourla, A., Nuss, P., and Ferreri, F. (2020). Social cognition in post-traumatic stress disorder: a systematic review. *Br. J. Clin. Psychol.* 59, 117–138. doi: 10.1111/bjc.12238
- Das, P., Khasnobish, A., and Tibarewala, D. N. (2016). “Emotion recognition employing ECG and gsr signals as markers of ans,” in *2016 Conference on Advances in Signal Processing (CASP)* (Pune), 37–42.
- Dolan, R. J. (2002). Emotion, cognition, and behavior. *Science* 298, 1191–1194. doi: 10.1126/science.1076358
- Droit-Volet, S., Bueno, L. J., Bigand, E., et al. (2013). Music, emotion, and time perception: the influence of subjective emotional valence and arousal? *Front. Psychol.* 4, 417. doi: 10.3389/fpsyg.2013.00417
- Dunn, E. M., and Lowery, M. M. (2013). “Simulation of pid control schemes for closed-loop deep brain stimulation,” in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)* (San Diego, CA: IEEE), 1182–1185.
- Egger, M., Ley, M., and Hanke, S. (2019). Emotion recognition from physiological signal analysis: a review. *Electron. Notes Theor. Comput. Sci.* 343, 35–55. doi: 10.1016/j.entcs.2019.04.009
- Ekman, P., Friesen, W. V., and Ancoli, S. (1980). Facial signs of emotional experience. *J. Pers. Soc. Psychol.* 39, 1125–1134. doi: 10.1037/h0077722
- Fayek, H. M., Lech, M., and Cavedon, L. (2017). Evaluating deep learning architectures for speech emotion recognition. *Neural Netw.* 92, 60–68. doi: 10.1016/j.neunet.2017.02.013
- Ferado, F., Mporas, I., and Ganchev, T. (2020). Evaluation of features in detection of dislike responses to audio-visual stimuli from EEG signals. *Computers* 9, 33. doi: 10.3390/computers9020033
- Filippini, C., Spadolini, E., Cardone, D., Bianchi, D., and Preziuso, M. (2020). Facilitating the child-robot interaction by endowing the robot with the capability of understanding the child engagement: the case of mio amico robot. *Int. J. Soc. Rob.* 13, 677–689. doi: 10.1007/s12369-020-00661-w
- Fleming, J. E., Dunn, E., and Lowery, M. M. (2020). Simulation of closed-loop deep brain stimulation control schemes for suppression of pathological beta oscillations in Parkinson's disease. *Front. Neurosci.* 14, 166. doi: 10.3389/fnins.2020.00166
- Freire, R. C., Cabrera-Abreu, C., and Milev, R. (2020). Neurostimulation in anxiety disorders, post-traumatic stress disorder, and obsessive-compulsive disorder. *Anxiety Disord.* 191, 331–346. doi: 10.1007/978-981-32-9705-0\_18
- Golland, Y., Hakim, A., Aloni, T., Schaefer, S. M., and Levit-Binnun, N. (2018). Affect dynamics of facial emg during continuous emotional experiences. *Biol. Psychol.* 139, 47–58. doi: 10.1016/j.biopsycho.2018.10.003
- Goshvarpour, A., Abbasi, A., and Goshvarpour, A. (2017). An accurate emotion recognition system using eeg and gsr signals and matching pursuit method. *Sci. Direct Biomed. J.* 40, 355–368. doi: 10.1016/j.bj.2017.11.001
- Gruebler, A., and Suzuki, K. (2010). Measurement of distal emg signals using a wearable device for reading facial expressions. *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.* 2010, 4594–4597. doi: 10.1109/IEMBS.2010.5626504
- Guimaraes, M. P., Wong, D. K., Uy, E. T., Grosenick, L., and Suppes, P. (2007). Single-trial classification of meg recordings. *IEEE Trans. Biomed. Eng.* 54, 436–443. doi: 10.1109/TBME.2006.888824
- Harper, R., and Southern, J. (2020). A bayesian deep learning framework for end-to-end prediction of emotion from heartbeat. *IEEE Trans. Affect. Comput.* 1:1. doi: 10.1109/TAFFC.2020.2981610
- Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.-P., Meuli, R., et al. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 106, 2035–2040. doi: 10.1073/pnas.081168106
- Ionescu, C. M., Neckebroek, M., Ghita, M., and Copot, D. (2021). An open source patient simulator for design and evaluation of computer based multiple drug dosing control for anesthetic and hemodynamic variables. *IEEE Access.* 9, 8680–8694. doi: 10.1109/ACCESS.2021.3049880
- Kayser, D., Egermann, H., and Barraclough, N. E. (2021). Audience facial expressions detected by automated face analysis software reflect emotions in music. *Behav. Res. Methods.* 1–15. doi: 10.3758/s13428-021-01678-3
- Klir, G., and Yuan, B. (1995). *Fuzzy Sets and Fuzzy Logic, Vol. 4*. Hoboken, NJ: Prentice Hall.
- Koelstra, S., Mühl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., et al. (2012). Deap: a database for emotion analysis using physiological signals. *IEEE Trans. Affect. Comput.* 3, 19–31. doi: 10.1109/T-AFFC.2011.15
- Kordsachia, C. C., Labuschagne, I., Andrews, S. C., and Stout, J. C. (2018). Diminished facial emg responses to disgusting scenes and happy and fearful faces in huntington's disease. *Cortex* 106, 185–199. doi: 10.1016/j.cortex.2018.05.019
- Kulic, D., and Croft, E. A. (2007). Affective state estimation for human-robot interaction. *IEEE Trans. Rob.* 23, 991–1000. doi: 10.1109/TRO.2007.904899
- Künecke, J., Hildebrandt, A., Recio, G., Sommer, W., and Wilhelm, O. (2014). Facial emg responses to emotional expressions are related to emotion perception ability. *PLoS ONE* 9, e84053. doi: 10.1371/journal.pone.0084053
- LaConte, S. M., Peltier, S. J., and Hu, X. P. (2007). Real-time fmri using brain-state classification. *Hum. Brain Mapp.* 28, 1033–1044. doi: 10.1002/hbm.20326
- Lilly, J. H. (2011). *Fuzzy Control and Identification*. Hoboken, NJ: John Wiley & Sons.
- Lin, S., Jinyan, X., Mingyue, Y., Ziyi, L., Zhenqi, L., Dan, L., et al. (2018). A review of emotion recognition using physiological signals. *New Trends Psychophysiol. Mental Health.* 18:2074. doi: 10.3390/s18072074
- Little, S., Beudel, M., Zrinzo, L., Foltynie, T., Limousin, P., Hariz, M., et al. (2016). Bilateral adaptive deep brain stimulation is effective in parkinson's disease. *J. Neurol. Neurosurg. Psychiatry* 87, 717–721. doi: 10.1136/jnnp-2015-310972
- Maheshwari, J., Joshi, S. D., and Gandhi, T. K. (2020). Tracking the transitions of brain states: an analytical approach using EEG. *IEEE Trans. Neural Syst. Rehabil. Eng.* 28, 1742–1749. doi: 10.1109/TNSRE.2020.3005950



- McCullagh, P., and Nelder, J. A. (1989). *Generalized Linear Models*. London: Chapman & Hall; CRC.
- Mendez, J. A., Leon, A., Marrero, A., Gonzalez-Cava, J. M., Reboso, J. A., Estevez, J. I., et al. (2018). Improving the anesthetic process by a fuzzy rule based medical decision system. *Artif. Intell. Med.* 84, 159–170. doi: 10.1016/j.artmed.2017.12.005
- Mertens, A., Raedt, R., Gadeyne, S., Carrette, E., Boon, P., and Vonck, K. (2018). Recent advances in devices for vagus nerve stimulation. *Expert. Rev. Med. Devices* 15, 527–539. doi: 10.1080/17434440.2018.1507732
- Naji, M., Firoozabadi, M., and Azadfallah, P. (2014). Classification of music-induced emotions based on information fusion of forehead biosignals and electrocardiogram. *Cognit. Comput.* 6, 241–252. doi: 10.1007/s12559-013-9239-7
- Nakasono, A., Prendinger, H., and Ishizuka, M. (2005). “Emotion recognition from electromyography and skin conductance,” in *Proceedings of the 5th International Workshop on Biosignal Interpretation* (Tokyo), 219–222.
- Nie, D., Wang, X.-W., Shi, L.-C., and Lu, B.-L. (2011). “Eeg-based emotion recognition during watching movies,” in *2011 5th International IEEE/EMBS Conference on Neural Engineering* (Cancun: IEEE), 667–670.
- Noroozi, F., Sapinski, T., Kaminska, D., and Anbarjafari, G. (2017). Vocal-based emotion recognition using random forests and decision tree. *Int. J. Speech Technol.* 20, 239–246. doi: 10.1007/s10772-017-9396-2
- Parvinian, B., Pathmanathan, P., Daluwatte, C., Yaghoubi, F., Gray, R. A., Weininger, S., et al. (2019). Credibility evidence for computational patient models used in the development of physiological closed-loop controlled devices for critical care medicine. *Front. Physiol.* 10, 220. doi: 10.3389/fphys.2019.00220
- Pfurtscheller, G., Neuper, C., Schlogl, A., and Lugger, K. (1998). Separability of EEG signals recorded during right and left motor imagery using adaptive autoregressive parameters. *IEEE Trans. Rehabil. Eng.* 6, 316–325. doi: 10.1109/86.712230
- Poria, S., Cambria, E., Bajpai, R., and Hussain, A. (2017). A review of affective computing: from unimodal analysis to multimodal fusion. *Inf. Fusion* 37, 98–125. doi: 10.1016/j.inffus.2017.02.003
- Prerau, M. J., Smith, A. C., Eden, U. T., Kubota, Y., Yanike, M., Suzuki, W., et al. (2009). Characterizing learning by simultaneous analysis of continuous and binary measures of performance. *J. Neurophysiol.* 102, 3060–3072. doi: 10.1152/jn.91251.2008
- Prerau, M. J., Smith, A. C., Eden, U. T., Yanike, M., Suzuki, W. A., and Brown, E. N. (2008). A mixed filter algorithm for cognitive state estimation from simultaneously recorded continuous and binary measures of performance. *Biol. Cybern.* 99, 1–14. doi: 10.1007/s00422-008-0227-z
- Price, J. B., Rusheen, A. E., Barath, A. S., Cabrera, J. M. R., Shin, H., Chang, S.-Y., et al. (2020). Clinical applications of neurochemical and electrophysiological measurements for closed-loop neurostimulation. *Neurosurg. Focus* 49, E6. doi: 10.3171/2020.4.FOCUS20167
- Qi, R., Tao, G., and Jiang, B. (2019). *Fuzzy system identification and adaptive control*. Springer.
- Ravindran, A. S., Nakagome, S., Wickramasuriya, D. S., Contreras-Vidal, J. L., and Faghih, R. T. (2019). “Emotion recognition by point process characterization of heartbeat dynamics,” in *2019 IEEE Healthcare Innovations and Point of Care Technologies, (HI-POCT)* (Bethesda, MD: IEEE), 13–16.
- rong Mao, Q., yu Pan, X., zhao Zhan, Y., and jun Shen, X. (2015). Using kinect for real-time emotion recognition via facial expressions. *Front. Inf. Technol. Electron. Eng.* 16, 272–282. doi: 10.1631/FITEE.1400209
- Rosula Reyes, S. J., Depano, K. M., Velasco, A. M. A., Kwong, J. C. T., and Oppus, C. M. (2020). “Face detection and recognition of the seven emotions via facial expression: Integration of machine learning algorithm into the nao robot,” in *2020 5th International Conference on Control and Robotics Engineering (ICCRE)* (Osaka: IEEE), 25–29.
- Rudovic, O. O., Park, H. W., Busche, J., Schuller, B., Breazeal, C., and Picard, R. W. (2019). “Personalized estimation of engagement from videos using active learning with deep reinforcement learning,” in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (Long Beach, CA: IEEE).
- Russell, J. A. (1980). A circumplex model of affect. *J. Pers. Soc. Psychol.* 39, 1161–1178. doi: 10.1037/h0077714
- Sakkalis, V. (2011). Review of advanced techniques for the estimation of brain connectivity measured with eeg/meg. *Comput. Biol. Med.* 41, 1110–1117. doi: 10.1016/j.compbiomed.2011.06.020
- Santaniello, S., Fiengo, G., Glielmo, L., and Grill, W. M. (2010). Closed-loop control of deep brain stimulation: a simulation study. *IEEE Trans. Neural Syst. Rehabil. Eng.* 19, 15–24. doi: 10.1109/TNSRE.2010.2081377
- Scherer, K. R. (2005). What are emotions? and how can they be measured? *Soc. Sci. Inf.* 44, 695–729. doi: 10.1177/0539018405058216
- Schubert, E. (2007). Locus of emotion: The effect of task order and age on emotion perceived and emotion felt in response to music. *J. Music Ther.* 44, 344–368. doi: 10.1093/jmt/44.4.344
- Sharma, R., Deepak, K., Gaur, P., and Joshi, D. (2020). An optimal interval type-2 fuzzy logic control based closed-loop drug administration to regulate the mean arterial blood pressure. *Comput. Methods Programs Biomed.* 185, 105167. doi: 10.1016/j.cmpb.2019.105167
- Shiva, J., Makaram, N., Karthick, P., and Swaminathan, R. (2021). Emotion recognition using spectral feature from facial electromyography signals for human-machine interface. *Stud. Health Technol. Inform.* 281, 486–487. doi: 10.3233/SHTI210207
- Sorosh, M. S., Maghooli, K., Setarehdan, S. K., and Nasrabadi, A. M. (2019). *A Novel EEG-Based Approach to Classify Emotions Through Phase Space Dynamics*. Springer, NM: Signal, Image, and Video Processing.
- Starnes, K., Miller, K., Wong-Kiesel, L., and Lundstrom, B. N. (2019). A review of neurostimulation for epilepsy in pediatrics. *Brain Sci.* 9, 283. doi: 10.3390/brainsci9100283
- Tageldeen, M. K., Elamvazuthi, I., and Perumal, N. (2016). “Motion control for a multiple input rehabilitation wearable exoskeleton using fuzzy logic and pid,” in *2016 IEEE 14th International Workshop on Advanced Motion Control (AMC)* (Auckland: IEEE), 473–478.
- Tan, J.-W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., et al. (2011). “Facial electromyography (fEMG) activities in response to affective visual stimulation,” in *2011 IEEE Workshop on Affective Computational Intelligence (WACI)* (Paris: IEEE), 1–5.
- Tan, J.-W., Walter, S., Scheck, A., Hrabal, D., Hoffmann, H., Kessler, H., et al. (2012). Repeatability of facial electromyography (emg) activity over corrugator supercilii and zygomaticus major on differentiating various emotions. *J. Ambient Intell. Humaniz Comput.* 3, 3–10. doi: 10.1007/s12652-011-0084-9
- Tegeler, C. H., Cook, J. F., Tegeler, C. L., Hirsch, J. R., Shaltout, H. A., Simpson, S. L., et al. (2017). Clinical, hemispheric, and autonomic changes associated with use of closed-loop, allostatic neurotechnology by a case series of individuals with self-reported symptoms of post-traumatic stress. *BMC Psychiatry* 17, 141. doi: 10.1186/s12888-017-1299-x
- Thenaisie, Y., Palmisano, C., Canessa, A., Keulen, B. J., Capetian, P., Jimenez, M. C., et al. (2021). Towards adaptive deep brain stimulation: clinical and technical notes on a novel commercial device for chronic brain sensing. *medRxiv*. doi: 10.1088/1741-2552/ac1d5b
- Val-Calvo, M., Álvarez-Sánchez, J. R., Ferrández-Vicente, J. M., and Fernández, E. (2020). Affective robot story-telling human-robot interaction: exploratory real-time emotion estimation analysis using facial expressions and physiological signals. *IEEE Access* 8, 134051–134066. doi: 10.1109/ACCESS.2020.3007109
- Van Bostel, A. (2010). Facial emg as a tool for inferring affective states. *Proc. Meas. Behav.* 7, 104–108. doi: 10.1371/journal.pone.0226328
- Vandewalle, G., Schwartz, S., Grandjean, D., Vuilleumier, C., Balteau, E., Degueldre, C., et al. (2010). Spectral quality of light modulates emotional brain responses in humans. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19549–19554. doi: 10.1073/pnas.1010180107
- Villanueva-Meyer, J. E., Mabray, M. C., and Cha, S. (2017). Current clinical brain tumor imaging. *Neurosurgery* 81, 397–415. doi: 10.1093/neuros/nyx103
- Wang, K., An, N., Li, B. N., Zhang, Y., and Li, L. (2015). Speech emotion recognition using fourier parameters. *IEEE Trans. Affect. Comput.* 6, 69–75. doi: 10.1109/TAFFC.2015.2392101
- Wei, X., Huang, M., Lu, M., Chang, S., Wang, J., and Deng, B. (2020). “Personalized closed-loop brain stimulation system based on linear state space model identification,” in *2020 39th Chinese Control Conference (CCC)* (Shenyang: IEEE), 1046–1051.
- Weiss, D., and Massano, J. (2018). Approaching adaptive control in neurostimulation for parkinson disease: autopilot on. *Neurology* 90, 497–498. doi: 10.1212/WNL.00000000000005111



- Wickramasuriya, D. S., Amin, M. R., and Faghih, R. T. (2019a). Skin conductance as a viable alternative for closing the deep brain stimulation loop in neuropsychiatric disorders. *Front. Neurosci.* 13, 780. doi: 10.3389/fnins.2019.00780
- Wickramasuriya, D. S., and Faghih, R. T. (2020). A mixed filter algorithm for sympathetic arousal tracking from skin conductance and heart rate measurements in pavlovian fear conditioning. *PLoS ONE* 15, e0231659. doi: 10.1371/journal.pone.0231659
- Wickramasuriya, D. S., Tessmer, M. K., and Faghih, R. T. (2019b). "Facial expression-based emotion classification using electrocardiogram and respiration signals," in *2019 IEEE Healthcare Innovations and Point of Care Technologies, (HI-POCT)* (Bethesda, MD: IEEE), 9–12.
- Wiem, A., Rahma, B., and Naoui, S. B. H. A. (2018). "An adaptive ts fuzzy-pid controller applied to an active knee prosthesis based on human walk," in *2018 International Conference on Advanced Systems and Electric Technologies (IC\_ASET)* (Hammamet: IEEE), 22–27.
- Wilson, K., James, G., and Kilts, C. (2020). Combining physiological and neuroimaging measures to predict affect processing induced by affectively valent image stimuli. *Sci. Rep.* 10, 9298. doi: 10.1038/s41598-020-66109-3
- Wu, S., Xu, X., Shu, L., and Hu, B. (2017). "Estimation of valence of emotion using two frontal EEG channels," in *2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (Kansas City, MO: IEEE), 1127–1130.
- Yadav, T., Atique, M. M. U., Azgomi, H. F., Francis, J. T., and Faghih, R. T. (2019). "Emotional valence tracking and classification via state-space analysis of facial electromyography," in *2019 53rd Asilomar Conference on Signals, Systems, and Computers* (Pacific Grove, CA: IEEE), 2116–2120.
- Yang, D., Alsadoon, A., Prasad, P. C., Singh, A. K., and Elchouemi, A. (2018a). An emotion recognition model based on facial recognition in virtual learning environment. *Procedia Comput. Sci.* 125, 2–10. doi: 10.1016/j.procs.2017.12.003
- Yang, Y., Connolly, A. T., and Shanechi, M. M. (2018b). A control-theoretic system identification framework and a real-time closed-loop clinical simulation testbed for electrical brain stimulation. *J. Neural Eng.* 15, 066007. doi: 10.1088/1741-2552/aad1a8
- Yu, C., and Tapus, A. (2019). Interactive robot learning for multimodal emotion recognition. *Int. J. Soc. Rob.* 633–642. doi: 10.1007/978-3-030-35888-4\_59
- Zeng, N., Zhang, H., Song, B., Liu, W., Li, Y., and Dobaie, A. M. (2018). Facial expression recognition via learning deep sparse autoencoders. *Neurocomputing* 273, 643–649. doi: 10.1016/j.neucom.2017.08.043
- Zhang, Q., Chen, X., Zhan, Q., Yang, T., and Xia, S. (2017). Respiration-based emotion recognition with deep learning. *Comput. Ind.* 92, 84–90. doi: 10.1016/j.compind.2017.04.005

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships 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.

Copyright © 2022 Branco, Ehteshami, Azgomi and Faghih. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.

# Advantages of publishing in Frontiers



## OPEN ACCESS

Articles are free to read  
for greatest visibility  
and readership



## FAST PUBLICATION

Around 90 days  
from submission  
to decision



## HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,  
and constructive  
peer-review



## TRANSPARENT PEER-REVIEW

Editors and reviewers  
acknowledged by name  
on published articles

## Frontiers

Avenue du Tribunal-Fédéral 34  
1005 Lausanne | Switzerland

Visit us: [www.frontiersin.org](http://www.frontiersin.org)

Contact us: [frontiersin.org/about/contact](http://frontiersin.org/about/contact)



## REPRODUCIBILITY OF RESEARCH

Support open data  
and methods to enhance  
research reproducibility



## DIGITAL PUBLISHING

Articles designed  
for optimal readership  
across devices



## FOLLOW US

@frontiersin



## IMPACT METRICS

Advanced article metrics  
track visibility across  
digital media



## EXTENSIVE PROMOTION

Marketing  
and promotion  
of impactful research



## LOOP RESEARCH NETWORK

Our network  
increases your  
article's readership