Analysis of EEG features and study of automatic classification in first-episode and drug-naïve patients with major depressive disorder

Yuanyuan Huang[1,2†], Yun Yi[2,3†], Qiang Chen[3], Hehua Li[2], Shixuan Feng[2], Sumiao Zhou[2], Ziyun Zhang[2], Chenyu Liu[2], Junhao Li[2], Qiuling Lu[2], Lida Zhang[2], Wei Han[2], Fengchun Wu[2,4,5*] and Yuping Ning[1,2,4,5*]

Abstract

Background Major depressive disorder (MDD) has a high incidence and an unknown mechanism. There are no objective and sensitive indicators for clinical diagnosis.

Objective This study explored specific electrophysiological indicators and their role in the clinical diagnosis of MDD using machine learning.

Methods Forty first-episode and drug-naïve patients with MDD and forty healthy controls (HCs) were recruited. EEG data were collected from all subjects in the resting state with eyes closed for 10 min. The severity of MDD was assessed by the Hamilton Depression Rating Scale (HAMD-17). Machine learning analysis was used to identify the patients with MDD.

Results Compared to the HC group, the relative power of the low delta and theta bands was significantly higher in the right occipital region, and the relative power of the alpha band in the entire posterior occipital region was significantly lower in the MDD group. In the MDD group, the alpha band scalp functional connectivity was overall lower, while the scalp functional connectivity in the gamma band was significantly higher than that in the HC group. In the feature set of the relative power of the ROI in each band, the highest accuracy of 88.2% was achieved using the KNN classifier while using PCA feature selection. In the explanatory model using SHAP values, the top-ranking influence feature is the relative power of the alpha band in the left parietal region.

Conclusions Our findings reveal that the abnormal EEG neural oscillations may reflect an imbalance of excitation, inhibition and hyperactivity in the cerebral cortex in first-episode and drug-naïve patients with MDD. The relative

---

†Yuanyuan Huang and Yun Yi contributed equally to this paper. They should be considered as Joint First-authors.

*Correspondence:
Fengchun Wu
13580380071@163.com
Yuping Ning
ningjeny@126.com
Full list of author information is available at the end of the article

power of the alpha band in the left parietal region is expected to be an objective electrophysiological indicator of MDD.

Keywords   Major depressive disorder, Electroencephalogram, Machine learning, Automated classification

Background

Major depressive disorder (MDD) is a debilitating disease that is characterized by noticeable alterations in mood, interest, pleasure, cognition, psychomotor activity and vegetative symptoms [1]. The disease can have a negative impact on people's daily lives, and the number of people suffering from the disease is increasing yearly. The current diagnosis of MDD is dependent on clinical interviewing with patients, which has obvious disadvantages, including poor sensitivity, patient denial, subjective biases and inaccuracy. Detecting the declines in brain physiology prior to the onset of subjective symptoms plays a significant role in the recognition of MDD at an early stage, thus allowing early treatment that is beneficial for MDD patients.

Electroencephalography (EEG) has been characterized as having high temporal resolution and low cost, which makes it more feasible for clinical application [2]. It contributes to the categorization of the electrophysiology of the brain in symptomatic and asymptomatic individuals. Pizzagalli et al. showed that depressive subjects exhibited more β3 in the pronounced right inferior and superior frontal regions than healthy control subjects and less β3 in the posteromedial cluster, including the posterior cingulate cortex and precuneus cortex [3]. Arikan et al. found that the group with suicidal ideation showed significantly higher high-gamma power (40–50 Hz) than other groups through rest-state EEG [4]. A recent study recruiting 44 late-life depression (LLD) patients and 41 healthy controls (HCs) showed that LLD patients had higher beta frequency activity and increased alpha activity than the HC group, and there were no correlations between beta power and the severity of MDD [5]. Because of the heterogeneity of patients and the influence of drug medication, the characteristics of EEG spectrum power in first-episode MDD remain unclear.

From a neuropsychiatric viewpoint, there is a growing awareness that MDD is characterized by disrupted connectivity between cortical and subcortical brain regions. There is evidence that synchronous oscillations can modulate functional connectivity between different brain regions. Liu et al. demonstrated that the dysfunction of oscillatory networks may be a promising indication of the pathoconnectomics of MDD [6]. A study of 16 depressive patients and 14 healthy controls found that depressive patients had abnormally enhanced brain functional connectivity in the gamma band by computing EEG coherence and demonstrated that the gamma band was sensitive to emotion processing [7]. Li et al. demonstrated

that alpha, theta and delta phase synchronization were decreased, but beta phase synchronization was increased in depressive patients, which indicated that underphased synchronization and asynchrony during working memory processing reveal the impairment of attention efficiency, memory and cortical inhibition in patients with MDD [8]. Previous studies focused on cognitive tasks and rest-state EEG can provide more informative insights into the pathophysiology of MDD. Therefore, it is essential to investigate the changes in scalp functional connectivity in the resting state in MDD patients without any drug medication, and the correlation with psychiatric symptoms needs to be further explored.

Recently, EEG-based machine learning (ML) techniques have attracted a considerable amount of attention due to their ability to noninvasively explore neuroimaging data to build computer-aided diagnosis solutions to advance the diagnosis of MDD. Bachmann et al. recruited 13 drug-free outpatients with MDD using linear methods and nonlinear methods to discriminate depressive patients and healthy controls. The results showed that combinations of these features had a maximal classification accuracy of 92% [9]. One study utilized 6 channels (FT7, FT8, T7, T8, TP7, TP8) with the Gaussian kernel of SVM and achieved an accuracy of 96.02% for screening depression [10]. However, no study has used machine learning to combine spectral power and scalp functional connectivity to identify MDD in the early stage.

To date, most studies focusing on the brain spectral signature of MDD have shown inconsistent results, and no study has examined first-episode and drug-naïve MDD with machine learning approaches combining spectral power and scalp functional connectivity. To address the heterogeneity of the results of electroencephalography in MDD studies, we collected detailed data on whole-brain neural activity and functional connectivity to (1) determine the specific electrophysiological indicators in first-episode and drug-naïve patients with MDD and (2) explore the role of these electrophysiological indicators in the clinical diagnosis of MDD using machine learning.

Methods and materials

Participants

Forty right-handed patients with first-episode and drug-naïve MDD and forty right-handed HCs were recruited to participate in this study. All MDD patients were outpatients or inpatients at the Affiliated Brain Hospital of Guangzhou Medical University from June 2021 to

November 2022. All HCs participants were recruited from the community through posters or forums.

The inclusion criteria for patients were as follows: (1) met the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) criteria for MDD; (2) age 18–45 years; (3) the duration of the illness was less than 2 years; (4) no history of psychotic medication or any other somatic therapy (e.g., modified electric convulsive treatment); and 4) the value of the HAMD-17 scale≥17. The inclusion criteria for HCs were as follows: (1) no psychiatric disorders meeting the DSM-V criteria; and (2) no family history of mental disease. The exclusion criteria for all participants were as follows: (1) had serious physical and neurological disease; (2) had a history of epilepsy, febrile convulsions or coma; (3) had history of substances and drug abuse; and (4) could not cooperate with EEG examination.

The study was approved by the Ethics Committee of the Affiliated Brain Hospital of Guangzhou Medical University. The study was conducted according to the Declaration of Helsinki. All participants signed an informed consent form ahead of enrollment.

Clinical measurement
All participants were interviewed and underwent a structured clinical interview by two experienced psychiatrists using the Mini-International Neuropsychiatric Interview (M.I.N.I.). The severity of MDD was assessed using the Hamilton Depression Rating Scale (HAMD-17).

EEG acquisition and processing
EEG Recording
Resting-state EEG data were recorded for each participant with their eyes closed for ten minutes by a trained technician. EEG data were obtained from 32 channels of Ag/AgCl electrodes using the Neuroscan system with a sampling rate of 1000 Hz and paced on the scalp in accordance with the International 10–20 system: C3, C4, F3, F4, F7, F8, Fp1, Fp2, FC1, FC2, FC5, FC6, Fz,T7, T8, P3, P4, Pz, CP1, CP2, CP5, CP6, Oz, P7, P8, O1, O2, PO3, PO4. During EEG recording, the impedance of each electrode was maintained below 10 KΩ, and Pz was set as the reference electrode.

EEG Preprocessing
All EEG data were imported into the EEGLAB toolbox in MATLAB R2013a software. The bandpass filter band was 0.1–80 Hz, and the notch filter band range was 49–51 Hz. The continuous EEG was downsampled to 500 Hz. A 2-s epoch segment was set for bad epoch data rejection and further analysis [11]. Bad channels were interpolated using spherical splines. Independent component analysis was used to remove eye movement artifacts. Segments with voltage values exceeding±80 were manually

rejected. EEG data were rereferenced to the common average reference. Spectral power and functional connectivity in the 2-s segments were measured. Finally, we excluded four EEGs from MDD patients and nine EEGs from HCs for their poor quality.

Power Spectrum
The power spectrum belonged to quantitative EEG, which simply reflected the changes of activity of the brain. Moreover, the power spectrum was regarded as linear parameters which were used in many existed studies. Power spectrum was calculated by performing a Fast Fourier Transform (FFT) algorithm to extract the features in the frequency domain, including: alpha (8–13 Hz), beta (13–30 Hz), theta (4–8 Hz), gamma (30–80 Hz), delta (1–4 Hz). The scalp electrodes were clustered into eight groups according to the regions of interest (ROIs), including the right parietal region (C4, CP6, CP2, P4), right frontal region (FC2, FC6, F4, FP2), right occipital region (PO4, O2), right temporal region (F8, T8, P8), left parietal region (P3, CP1, CP5, C3), left frontal region (FP1, F3, FC5, FC1), left temporal region (P7, T7, FT7), left occipital region (O1, PO3), and left parietal region (C3, CP5, CP1, P3). The relative power in each frequency band was derived by dividing the absolute power of each frequency band by the total broadband of the absolute power, which included the delta band, theta band, alpha band, beta band and gamma band. The formulas were as follows:

The relative of alpha power = the absolute power of alpha / (the absolute power of delta+the absolute power of theta+the absolute power of alpha+the absolute power of beta+absolute power of gamma).

Functional connectivity
The weighted phase lag index (wPLI) was used to compute the functional connectivity between all pairwise combinations of 29 channels. The connectivity between the two regions was calculated by the average wPLI value of all electrode pairs of recording sites between the two regions.

Machine learning analysis
Python was used to conduct the data analysis of the five classification algorithms, including decision tree (DT), gradient boosting decision tree (GBDT), support vector machine (SVM), naïve Bayesian (NB) and K-nearest neighbor (KNN). We used three methods to select features, including recursive feature elimination based on SVM (SVM-RFE), selection operator and least absolute shrinkage based on L1 (LASSO-L1) and principal component analysis (PCA).

The relative power of each band of each scalp area was regarded as set (1) The functional connectivity of any two

scalp areas was regarded as set (2). The combinations of set 1 and set 2 were regarded as set (3). Permutation testing was utilized to evaluate the statistical significance of the classifier performance.

The importance of each feature was interpreted by the SHAP values. SHAP is an acronym for Shapley Additive exPlanations, which was a unified framework introduced for interpreting machine learning predictions. The SHAP technique calculates the Sharpley value by evaluating the contribution and influence of each feature to model prediction [12]. The importance ranking of the top 20 factors with stability and interpretation using optimal model. The red part in feature value represents higher value. The higher SHAP value of a feature is given, the more important feature in predicting depression.

As shown in Fig. 1, we used 1000 sampling points as a sample. Since the data sampling rate was down sampled to 500 Hz, a sample should take 2 s. After the data cutting, the total sample size of the two groups was 19,500. Then, 90% of the data was used for training and 10% for
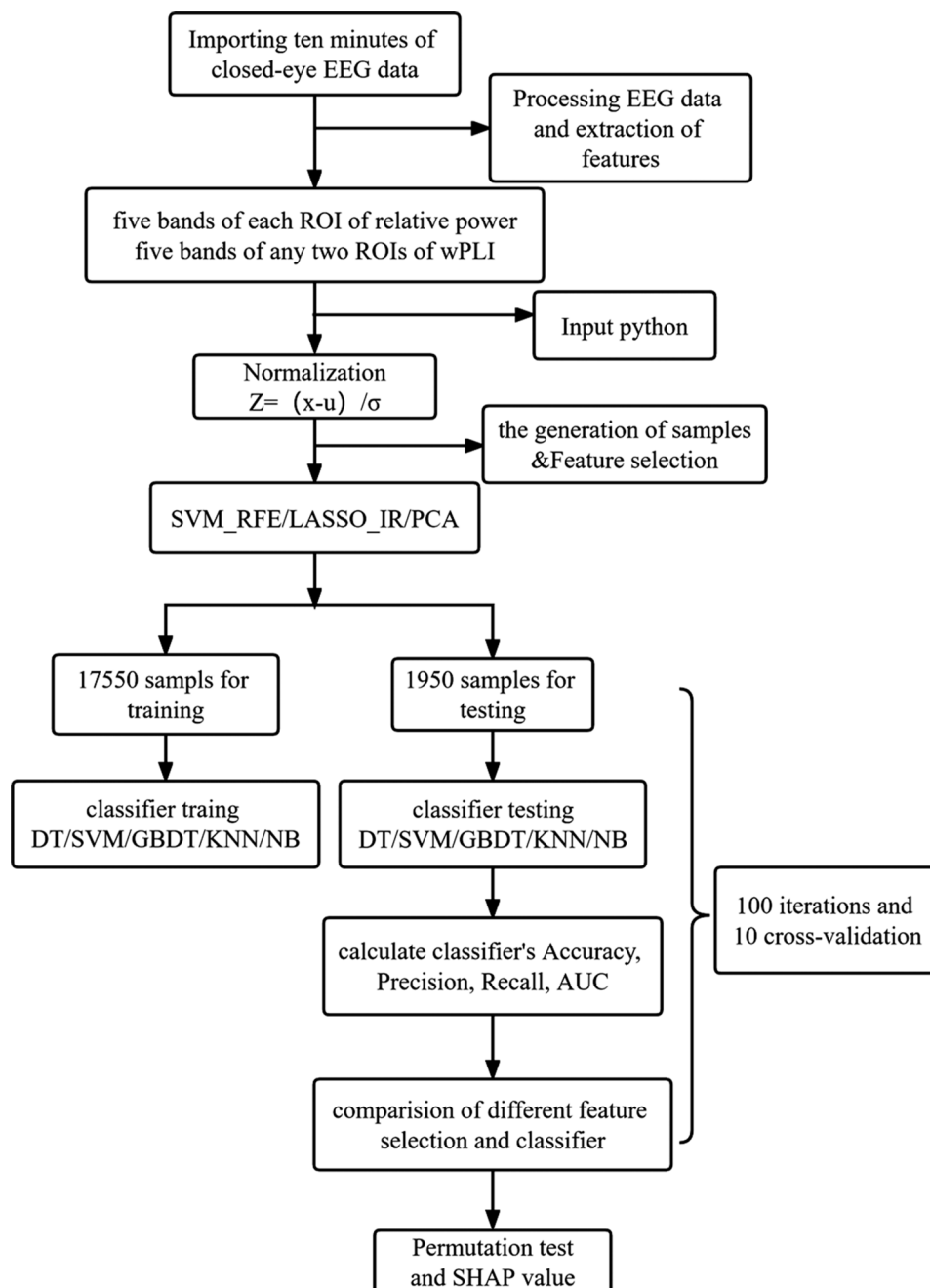

Fig. 1 The flowchart of machine learning

Table 1 Demographic characteristics of two groups

	MDD group (n = 36)	HC group (n = 31)	P-value
Gender(male/female)	14/22	14/17	0.590
Age(years)	22.7 ± 2.1	25.2 ± 6.9	0.054
Education (years)	15.0 ± 1.7	14.6 ± 2.5	0.500
Age of onset (years)	24.5 ± 7.6	—	—
Course of disease (month)	6(1,48)	—	—
HAMD-17	21.9 ± 6.9	—	—

MDD: major depression disorder; HC: Healthy control

testing. The stratification of the dataset was based on the participants.

Statistical analysis

Statistical analyses were carried out using SPSS (version 23.0). For continuous variables, an independent samples t-test was conducted to compare intergroup differences. The comparison of two groups of relative power was computed using the independent sample two-tailed t-test across ROIs in five frequency bands. Pearson's χ^2 test was used to test sex differences between two groups. The

false-discovery rate (FDR) correction was employed to reduce the type I error in multiple comparisons. Pearson correlation analysis was performed to explore the correlations between EEG features and clinical factors. Results were defined as significant if the two-tailed p value was lower than 0.05.

Results

### Demographic characteristics

Table 1 includes demographic information of the subjects in the MDD group and HC group. There was no significant difference between the two groups in terms of sex and age.

EEG feature analysis

The relative power

As shown in Fig. 2, in the right posterior occipital region, the relative power of the delta and theta bands was found to be significantly different. The relative power of the alpha band was also significantly different in the whole posterior occipital region. There were no differences in beta and gamma bands. Furthermore, the relevant data
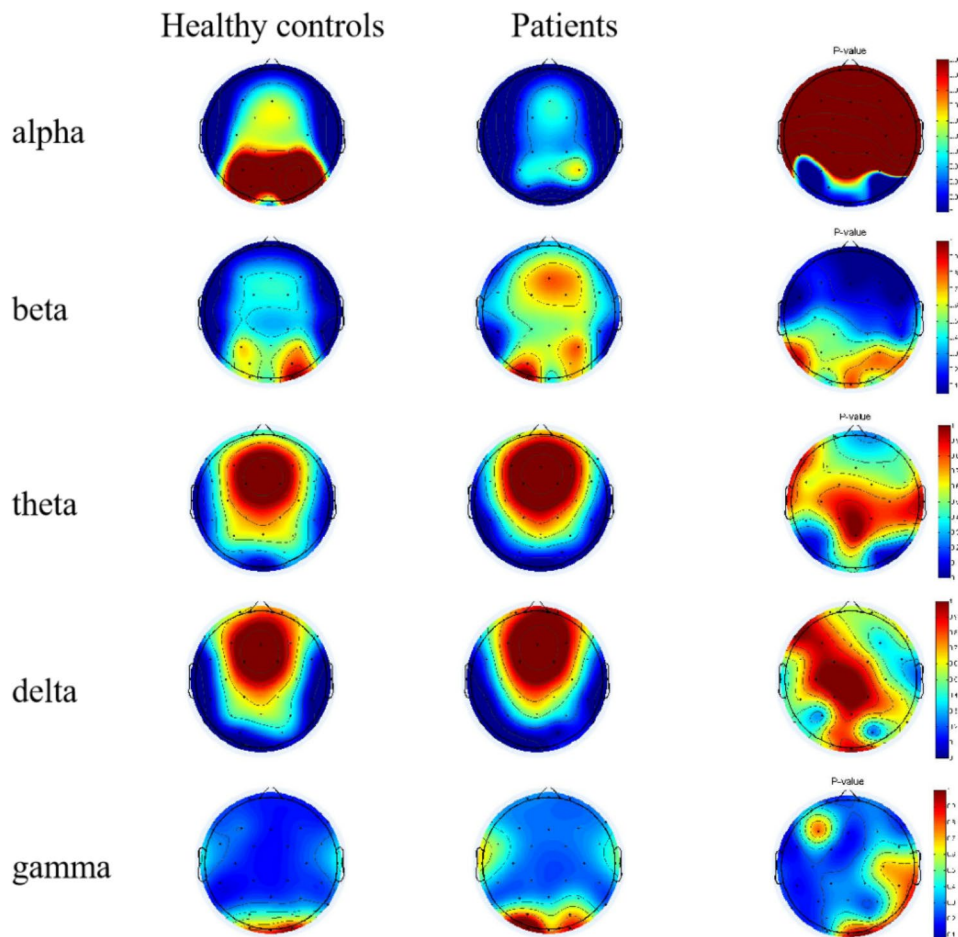
Fig. 2 Comparison of the relative power between two groups in each ROI area in each frequency band

were extracted. In the left occipital region, the relative power of the alpha band in the MDD group was significantly lower than that in the HC group *(t=4.829, FDRp=0.01)*, but the relative power of the delta band was significantly higher than that in the HC group *(t = -3.357, FDRp=0.01).* In the right occipital region, the relative power of delta *(t = -4.934, FDRp=0.01)* and theta *(t = -4.007, FDRp=0.01)* bands in the MDD group was significantly higher than that in the HC group. The relative power of the alpha band in the MDD group was significantly lower than that in the HC group *(t=6.142, FDRp=0.01)*. In the right parietal region, the relative power of the alpha band in the MDD group was significantly lower than that in the HC group *(t=2.886, FDRp=0.03)* (Fig. 3).

Functional connectivity

As shown in Fig. 4, there were no differences in wPLIs in the delta and theta bands. However, wPLIs were significantly weakened in the alpha band and enhanced in the gamma band across all regions of the brain in the MDD group (all p<0.001).

Correlations between EEG features and clinical symptoms

There were no correlations between the relative power of alpha and delta and HAMD-17 scores in the left occipital region (r=0.201, p=0.379; r = -0.233, p=0.171) and right occipital region (r = -0.182, p=0.289; r = -0.110, p=0.521).

Classification of MDD based on EEG features

The classification results of each feature set based on different classifiers with different feature selection methods were shown as Accuracy, Precision, Recall and AUC. They were expressed as mean±sd (Table 2).

The highest classification accuracy was achieved when only feature set 1 was used. That is, when we selected the relative power of each band of the ROI, we achieved the highest accuracy (88.2%) using PCA feature selection and the KNN classifier. The result of permutation test was less than 0.01. This result indicated that the prediction accuracy of the KNN classifier was significantly higher than that of the random prediction case, which is of practical significance.

Figure 5 shows the SHAP values of each feature for each sample. One dot represents a sample, and the color indicates the feature value (the blue color indicates a low value, and the red color indicates a high value). In the model, the top-ranking influence feature is the relative power of the alpha band in the left parietal (LP-alpha).

(LF/RF/LP/RP/LT/RT/LO/RO-delta/theta/alpha/beta/gamma: Left Frontal/ Right Frontal/ Left Parietal/ Right Parietal/ Left Temporal/ Right Temporal/ Left Occipital/ Right Occipital-delta/ theta/ alpha/ beta/ gamma).

Discussion

This was the first study to use machine learning combined with spectral power and scalp functional connectivity to identify first-episode and drug-naïve MDD. The findings of our study included the following: (1) compared to the HC group, the relative power of the low delta and theta bands was significantly higher in the right occipital region, and the relative power of the alpha band in the entire posterior occipital region was significantly lower in the MDD group; (2) in the MDD group, the alpha band scalp functional connectivity was overall lower, while the scalp functional connectivity in the gamma band was significantly higher than that in the HC group; and (3) the relative power of each frequency band in each scalp brain region used as the input feature
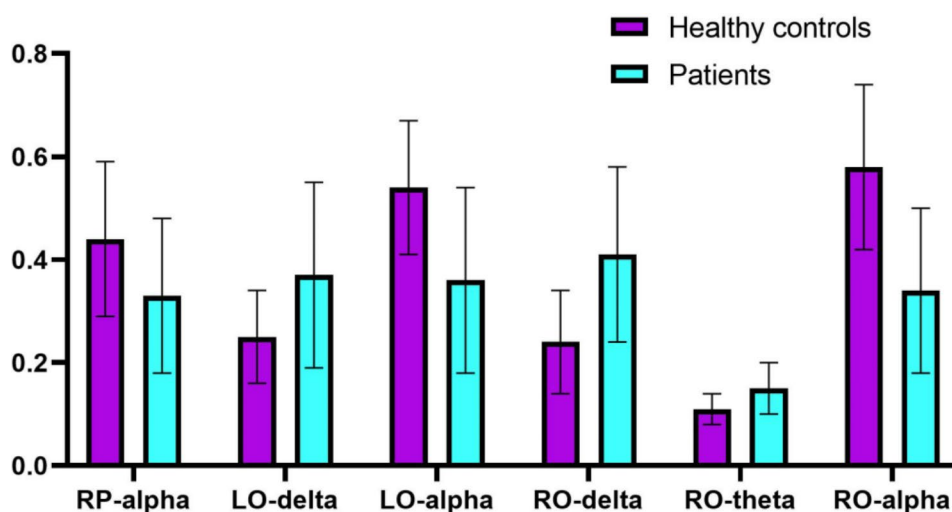
Fig. 3 The differences in the relative power in frequency bands between the two groups (RP-alpha: Right Parietal-alpha; LO-delta: Left Occipital-delta; LO-alpha: Left Occipital-alpha; RO-delta: Right Occipital-delta; RO-theta: Right Occipital-theta; RO-alpha: Right Occipital-alpha)
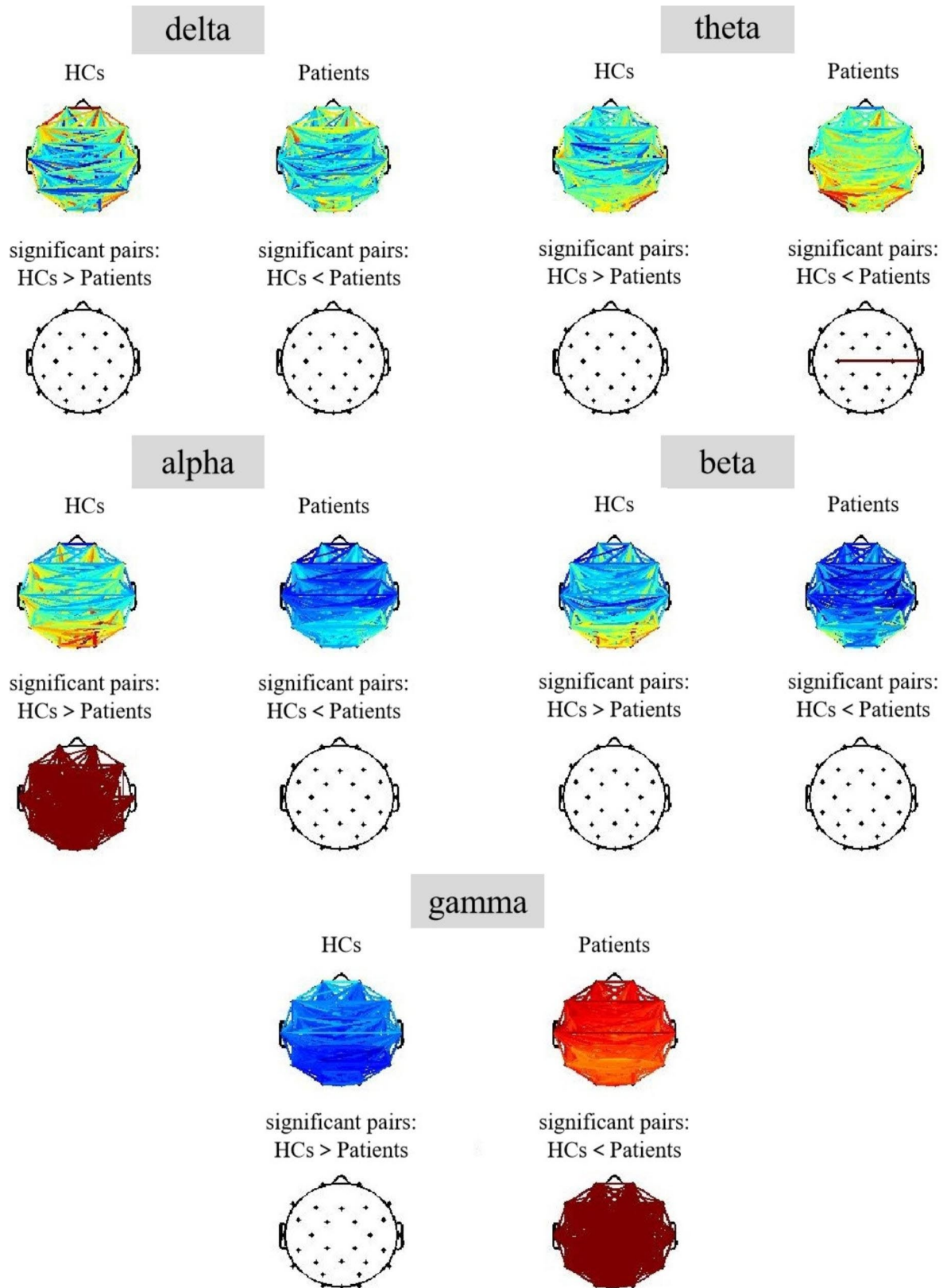
Fig. 4 The functional connectivity of scalp brain regions between the two groups in each frequency band (significant pairs: Functional connectivity pairs representing two groups whose functional connectivity values are still statistically different after FDR correction; HCs > Patients: The functional connectivity was greater in HCs than in patients; HCs < Patients: The functional connectivity was weaker in HCs than in patients)

Table 2 The classification results of different sets

Feature selection	Model	Accuracy	Precision	Recall	AUC	*p*-value
Set1						
SVM-RFE	DT	0.751±0.017	0.774±0.016	0.773±0.013	0.749±0.017	0.0099
	SVM	0.776±0.010	0.787±0.013	0.809±0.012	0.774±0.010	0.0099
	GBDT	0.803±0.013	0.817±0.014	0.817±0.013	0.800±0.012	0.0099
	NB	0.679±0.014	0.741±0.019	0.627±0.013	0.681±0.014	0.0099
	KNN	0.772±0.014	0.731±0.014	0.913±0.012	0.755±0.015	0.0099
LASSO-LR	DT	0.830±0.012	0.843±0.013	0.846±0.013	0.828±0.012	0.0099
	SVM	0.859±0.012	0.896±0.011	0.837±0.022	0.861±0.010	0.0099
	GBDT	0.826±0.008	0.845±0.013	0.831±0.013	0.827±0.008	0.0099
	NB	0.661±0.013	0.778±0.018	0.525±0.016	0.675±0.012	0.0099
	KNN	0.851±0.012	0.811±0.013	0.946±0.010	0.842±0.013	0.0099
PCA	DT	0.792±0.014	0.808±0.015	0.809±0.020	0.791±0.015	0.0099
	SVM	0.870±0.009	0.891±0.011	0.867±0.015	0.871±0.010	0.0099
	GBDT	0.813±0.013	0.832±0.017	0.823±0.015	0.813±0.013	0.0099
	NB	0.687±0.013	0.768±0.013	0.608±0.027	0.694±0.013	0.0099
	KNN	0.882±0.006	0.847±0.008	0.955±0.011	0.877±0.008	0.0099
Set2						
SVM-RFE	DT	0.543±0.034	0.596±0.034	0.577±0.047	0.541±0.035	0.0099
	SVM	0.590±0.037	0.597±0.022	0.845±0.042	0.560±0.037	0.0099
	GBDT	0.598±0.027	0.615±0.019	0.764±0.035	0.580±0.025	0.0099
	NB	0.513±0.029	0.609±0.041	0.353±0.064	0.532±0.026	0.0099
	KNN	0.561±0.017	0.576±0.012	0.817±0.033	0.527±0.019	0.0099
LASSO-LR	DT	0.553±0.024	0.599±0.026	0.602±0.033	0.544±0.027	0.0099
	SVM	0.618±0.032	0.622±0.021	0.822±0.028	0.594±0.033	0.0099
	GBDT	0.585±0.020	0.606±0.013	0.748±0.044	0.579±0.046	0.0099
	NB	0.503±0.026	0.607±0.028	0.296±0.073	0.524±0.021	0.0099
	KNN	0.564±0.019	0.573±0.013	0.854±0.036	0.521±0.022	0.0099
PCA	DT	0.536±0.030	0.585±0.028	0.576±0.049	0.524±0.032	0.0099
	SVM	0.617±0.039	0.615±0.028	0.837±0.025	0.579±0.042	0.0099
	GBDT	0.575±0.032	0.601±0.019	0.730±0.035	0.560±0.027	0.0099
	NB	0.582±0.020	0.623±0.018	0.650±0.048	0.575±0.021	0.0099
	KNN	0.553±0.025	0.571±0.018	0.825±0.039	0.517±0.028	0.0099
Set3						
SVM-RFE	DT	0.721±0.015	0.787±0.019	0.723±0.020	0.719±0.011	0.0099
	SVM	0.755±0.012	0.753±0.010	0.768±0.005	0.719±0.012	0.0099
	GBDT	0.786±0.011	0.768±0.12	0.808±0.011	0.778±0.014	0.0099
	NB	0.639±0.013	0.732±0.021	0.590±0.012	0.655±0.016	0.0099
	KNN	0.758±0.012	0.708±0.019	0.865±0.020	0.734±0.017	0.0099
LASSO_LR	DT	0.828±0.011	0.835±0.018	0.833±0.012	0.802±0.012	0.0099
	SVM	0.837±0.013	0.876±0.008	0.822±0.003	0.842±0.015	0.0099
	GBDT	0.811±0.012	0.832±0.011	0.822±0.022	0.838±0.022	0.0099
	NB	0.661±0.009	0.789±0.012	0.588±0.011	0.655±0.008	0.0099
	KNN	0.844±0.013	0.798±0.018	0.900±0.011	0.802±0.021	0.0099
PCA	DT	0.786±0.016	0.819±0.012	0.758±0.019	0.788±0.016	0.0099
	SVM	0.845±0.011	0.887±0.014	0.856±0.011	0.886±0.012	0.0099
	GBDT	0.822±0.012	0.844±0.017	0.818±0.011	0.823±0.019	0.0099
	NB	0.667±0.011	0.768±0.023	0.668±0.015	0.688±0.011	0.0099
	KNN	0.876±0.018	0.833±0.012	0.923±0.015	0.857±0.012	0.0099

AUC: Area Under the Cure; *p*-value: Permutation test p value; SVM-RFE: Support Vector Machines- Recursive Feature Elimination; LASSO_LR: Least Absolute Shrinkage and Selection Operator- Logistic Regression; PCA: Principal Component Analysis; DT: Decision Tree; GBDT: Gradient Boosting Decision Tree; NB: Naïve Bayesian; KNN: K-Nearest Neighbor
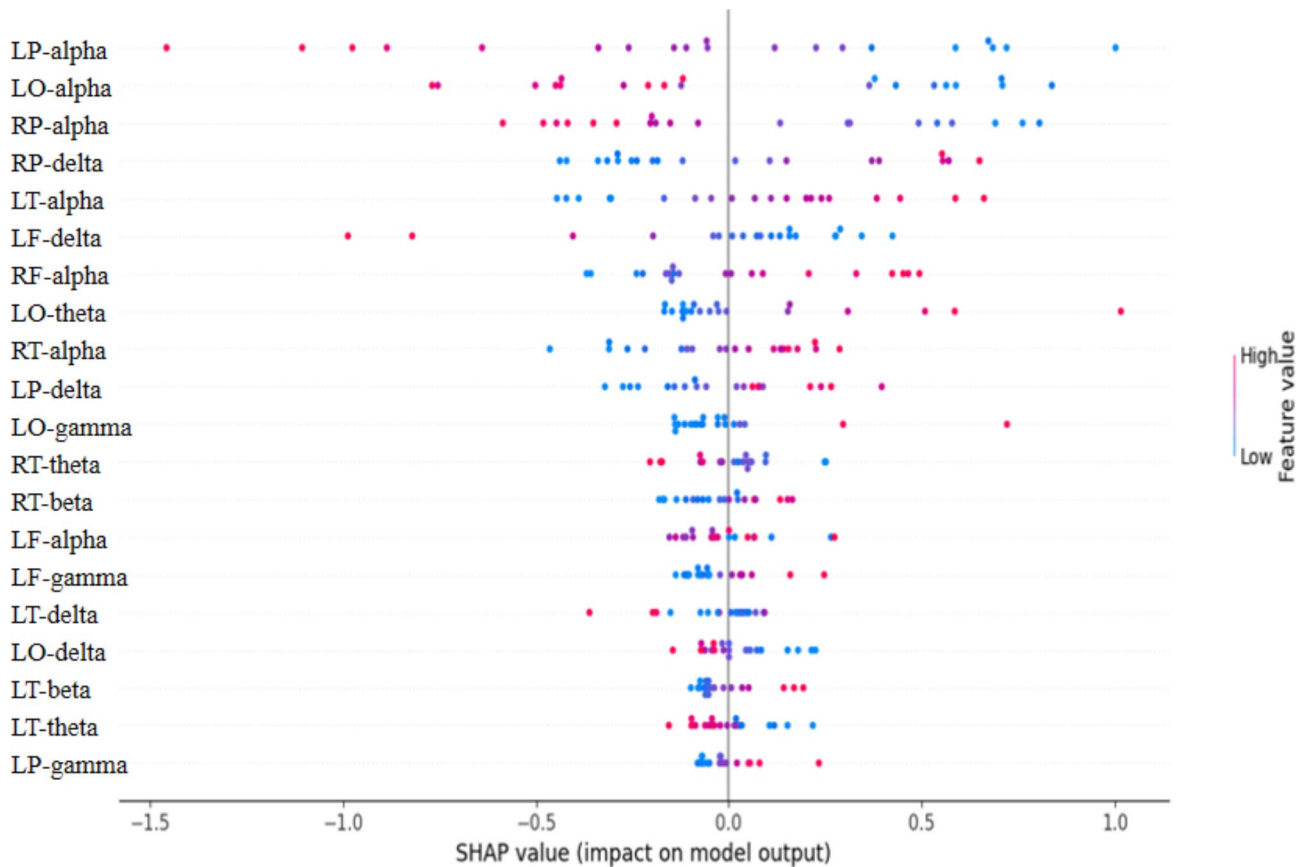
Fig. 5 SHAP model explanation

set achieved the highest accuracy rate (88.2%), and the top-ranking influence feature is the relative power of the alpha band in the left parietal region.

Our results showed that the relative power of the theta and delta bands in the MDD group was significantly higher than that in the HC group. Previous studies have also confirmed that the severity of MDD symptoms is positively proportional to theta power. The activity of the theta band is associated with emotional stimuli and that the activity of theta rhythms can be enhanced by high arousal from negative pictures [13], which may explain that the theta rhythm is associated with the electrical activity of the medial prefrontal cortex [14]. Lee et al. found that compared to the group with lower suicidal ideation scores, the relative power of theta in the central frontal regions (CZ, FZ, FCz, and F3) was significantly higher in the group with high suicidal ideation scores [15]. Delta waves affected the incentive and reward areas of the cerebral cortex, which suggested that they were significantly associated with emotion production [16]. It has been suggested that the high activity of the delta band in MDD patients is a compensatory mechanism for the deficits in cortical function caused by MDD, which is caused by MDD may be achieved through the modulation of cognitive function. An increase in the power of the delta band may inhibit interference affecting cognitive tasks in the cerebral cortex, which increases the attention and concentration required to perform certain cognitive tasks in patients with MDD [17]. However, our findings focused on the right occipital area, while there is no basis to support that the abnormal activation in this brain region is associated with the relative power enhancement of the delta and theta bands of the brain region. Further research is needed to investigate this topic.

Our findings also suggested that patients with MDD have lower relative alpha power in the occipital region than the HC group, which was consistent with previous studies [18–20]. Alpha rhythms are thought to be generated by cortico-thalamic brain interactions, which involve a large number of cognitive operations, particularly sensory system control, working memory and attention [21]. The decline in the relative power of the EEG alpha band in MDD reflected increasing negative emotions and significant activation of the cerebral cortex [22, 23], which may be related to a reduced thalamocortical synchronization system. Additionally, decreased alpha energy in the occipital region of depressed patients may be related to increased arousal and cortical excitability in the posterior occipital region of the brain, which revealed a chronic stress response in the development of

Huang *et al. BMC Psychiatry*   (2023) 23:832

Page 10 of 13

MDD. Moreover, it may also be related to the genotype of the individual. Compared to gene Val carriers, MDD gene Met carriers had lower overall absolute alpha power under closed-eye conditions [20]. This may be because brain-derived neurotrophic factor (BDNF) protein secretion was limited in Met allele carriers [24] and the effect on functional connectivity. However, some studies have shown the opposite results [25, 26]. One reason may be related to the age of onset of the subjects. Alpha oscillations are prone to change during critical neural development, such as puberty [27]. Another explanation can be found in the heterogeneity of subjects, for example, the differences in drugs, sex, and subtypes of MDD. These factors should be more clearly controlled in the future. Significant changes in alpha rhythm can help explain the neural mechanisms of MDD, and our study – which is the first to examined untreated subjects – provides theoretical support for subsequent studies on neural marker identification and detection methods.

MDD is characterized by imbalances in communication between large-scale functional networks, including resting-specific hyper and hypoconnectivity within and between brain networks [28]. Our study demonstrated functional connectivity disorder between cortical regions in patients with MDD, which was consistent with previous studies [29, 30]. Previous studies have shown that MDD exhibits significantly disrupted network characteristics, which are closely related to the emotional processing disorder of MDD. That is, changes in the MDD network in the alpha band may lead to changes in emotional response or emotional arousal [30]. Alpha rhythm synchronization was associated with the activity of the default mode network (DMN) [31]. Thus, any disruption of the DMN in MDD may manifest as changes in EEG alpha oscillations and desynchronization of alpha band connections. However, some studies with other indicators of functional connectivity(e.g., coherence) have produced inconsistent results.A study of 12 outpatients with drug-free MDD and 10 HCs showed that MDD had significantly increased functional connectivity in the alpha and theta bands by using the signal synchronization called the "structural synchronization index" [32]. Leuchter et al. found that patients with MDD had higher connectivity in beta, alpha, theta and delta bands than HCs [33]. Therefore, there is a need to use a uniform functional connectivity evaluation index to explore phase-synchronous changes in the alpha band of MDD with a large sample.

Our study showed that there existed a significant increase in functional connectivity between overall scalp-brain regions in the gamma band, which was in line with previous studies [7, 32, 34]. Previous studies confirmed that gamma oscillations are sensitive to emotional processing [7]. When patients were depressed,

their brains became information activated or overloaded. It may be attributed to the fact that increased gamma functional connectivity leads to abnormal activation of the depressed brain or overload of communication. The increasing connectivity in the gamma band may reflect increased attentional function [35, 36], which reflect the neural mechanisms of information connectivity. Moreover, the functional connectivity enhancement of the gamma band in patients with MDD may be related to the relationship between serotonin and γ-aminobutyric acid (GABA) signaling. Serotonin is a widely distributed neurotransmitter in the central nervous system [37]. Previous studies have demonstrated that activation of 5-hydroxytryptamine signaling has been shown to inhibit the inhibitory function of GABA-energic neurons [38]. In 5-hydroxytryptamine deficiency (e.g., MDD), GABA-energic signaling in the cortex may be enhanced, and GABA plays a role in the production and regulation of endogenous motor cortical rhythm beta and gamma activity [39]. However, there are other results that may differ from our study. An EEG study showed that depressed patients in the beta band of the DMN showed greater functional connectivity [40]. Kim et al. suggested that the pathophysiological mechanisms of depressive states may be related to excessive neural processing in the beta band [41]. Therefore, further longitudinal studies with large samples are needed to investigate gamma band alterations in MDD.

It is worth mentioning that the relative power of the alpha band in the left parietal region was lower in the MDD group than in the HC group, which strongly contributed to the automatic classification of MDD. Previous studies suggest that dysregulation of the left subparietal cortex in depressed patients may not only be associated with deficits in audiovisual integration but also with impaired memory and emotional processing deficits in patients with MDD [42]. Therefore, the above results of this machine learning are reasonable. The results further confirmed the dysfunction of the left parietal lobe in patients with MDD [42]. Moreover, the finding was in agreement with the statistically calculated differences in EEG features. In the two independent sample t-tests, there were group differences in all the above three indicators. However, the relative power of alpha in the left parietal region was not significantly different after FDR correction. The reason for this may be related to the small number of subjects. Therefore, to some extent, there was a significant difference in the relative power of the alpha band in the left top region between the two groups. Additionally, the top three features belong to the alpha band power. The alpha band power achieves a high classification accuracy, which is consistent with a previous study [43]. Finally, our results showed that feature set 1 alone achieves the highest classification accuracy, which was

higher than that of the combination of two feature sets. We have known that the set 2 achieved the lower accuracy in the machine learning. After the calculations of functional connectivity of any two regions above, we found that the values of wPLIs in depression group were less than 0.5. It indicated that the scalp of functional connectivity of any two regions were relatively lower. Therefore, when we regarded the functional connectivity as feature set 2, we achieved accuracy which varied from 0.50 to 0.618. The causes may be owed to the low resolution of the functional connectivity of any two scalp brain regions in each band. When we combined set1 and set2 together, we didn't achieve the highest accuracy in the machine learning. We considered that combining with the connectivity may cause interference, which affects the classification and thus reduces the accuracy. However, our results achieved an accuracy of 88.2%, which was lower than the results of the studies of Abdolkarim Saeedi [44] and Chien-Te Wu [45]. Abdolkarim Saeedi et al. developed E-KNN rather than KNN to give information on feature importance index and improve the results (an accuracy of 98.44%). Chien Te Wu et al. utilized the combination of the optimal feature subset and CK-SVM achieving an accuracy of 91.07% on the training set. In comparison, we analyzed the spectral power and functional connectivity and combined with using five machine-learning algorithms for classification. We speculated that the feature selection and the method of classifier algorithms were contributed to our lower accuracy.

Due to the impact of medication on the EEG and emotional state of MDD patients, we recruit first-episode MDD patients who did not take medication to rule out the impact of medication as an interfering factor. We explored the abnormalities of EEG features in MDD patients to clarify the EEG mechanism of depression in resting state. Combining with the machine learning, we would provide a theoretical basis for early recognition of depression and the application of neuro electrophysiological markers in clinical practice.

Limitations and future research directions

Several limitations should be considered in future studies. First, the sample size was small, which may affect the stability performance in machine learning. Moreover, there may exist correlations between spectrum power and the severity of MDD. However, the correlations may not be significant as a result of the limited sample size. Therefore, although our results provide a strong reference direction for future studies, they need to be further verified through large samples. Second, although we use many measures to control overfitting, including cutting data to increase the sample size, cross-validation and feature selection techniques, our performance estimates are not completely unbiased. To obtain unbiased estimates of our model prediction performance, future studies should be based on independent datasets that were not used for model construction for validation. Third, because our study belonged to cross section study, we wouldn't investigate the impact of LP-Alpha feature in MDD diagnosis. Therefore, we would explore the value of LP-Alpha in the diagnosis of MDD using longitude research in the future study. At last, in the future work, we should learn more the methods of machine learning like the study of Li et al [46] to improve the accuracy and explore the biomarker of MDD patients better.

Conclusions

In conclusion, we found that there were significant changes in EEG features, which indicated that there were potential abnormalities in EEG neural oscillations in first-episode and drug-naïve patients with MDD. This may reflect cortical excitation, inhibition, and overactivity imbalance in patients with drug-free MDD. After the machine learning algorithm, we hypothesized that the relative power of the alpha band in the left parietal region may be a potential neurophysiological marker of MDD. Moreover, it may be used as an indicator to assist in the clinical diagnosis of MDD.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12888-023-05349-9.

Supplementary Material 1: Detailed process of machine learning

Supplementary Material 2: Comparison of the relative power

Huang *et al. BMC Psychiatry*   (2023) 23:832

Page 12 of 13

Data Availability
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The study was approved by the Ethics Committee of the Affiliated Brain Hospital of Guangzhou Medical University. The study was conducted according to the Declaration of Helsinki. All participants signed an informed consent form ahead of enrollment.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]The First School of Clinical Medicine, Southern Medical University, Guangzhou, China
[2]Department of Psychiatry, The Affiliated Brain Hospital of Guangzhou Medical University, Guangzhou, China
[3]Department of Psychiatry, The Brain Hospital of Guangxi Zhuang Autonomous Region, Guangxi, China
[4]Guangdong Engineering Technology Research Center for Translational Medicine of Mental Disorders, Guangzhou, China
[5]Key Laboratory of Neurogenetics and Channelopathies of Guangdong Province and the Ministry of Education of China, Guangzhou Medical University, Guangzhou, China

References
1. Otte C, Gold SM, Penninx BW, et al. Major depressive disorder. Nat Rev Dis Primers. 2016;2:16065. https://doi.org/10.1038/nrdp.2016.65
2. Bailey NW, Hoy KE, Rogasch NC, et al. Responders to rTMS for depression show increased fronto-midline theta and theta connectivity compared to non-responders. Brain Stimul. 2018;11(1):190–203. https://doi.org/10.1016/j.brs.2017.10.015
3. Pizzagalli DA, Nitschke JB, Oakes TR, et al. Brain electrical tomography in depression: the importance of symptom severity, anxiety, and melancholic features. Biol Psychiatry. 2002;52(2):73–85. https://doi.org/10.1016/s0006-3223(02)01313-6
4. Arikan MK, Gunver MG, Tarhan N, Metin B, High-Gamma. A biological marker for Suicide attempt in patients with depression. J Affect Disord. 2019;254:1–6. https://doi.org/10.1016/j.jad.2019.05.007
5. Wu Z, Zhong X, Lin G, et al. Resting-state electroencephalography of neural oscillation and functional connectivity patterns in late-life depression. J Affect Disord. 2022;316:169–76. https://doi.org/10.1016/j.jad.2022.07.055
6. Liu W, Wang X, Xu J, Chang Y, Hamalainen T, Cong F. Identifying Oscillatory Hyperconnectivity and Hypoconnectivity Networks in Major Depression using coupled Tensor Decomposition. IEEE Trans Neural Syst Rehabil Eng. 2021;29:1895–904. https://doi.org/10.1109/TNSRE.2021.3111564
7. Li Y, Cao D, Wei L, Tang Y, Wang J. Abnormal functional connectivity of EEG gamma band in patients with depression during emotional face processing. Clin Neurophysiol. 2015;126(11):2078–89. https://doi.org/10.1016/j.clinph.2014.12.026
8. Li Y, Kang C, Wei Z, et al. Beta oscillations in major depression - signalling a new cortical circuit for central executive function. Sci Rep. 2017;7(1):18021. https://doi.org/10.1038/s41598-017-18306-w
9. Bachmann M, Päeske L, Kalev K, et al. Methods for classifying depression in single channel EEG using linear and nonlinear signal analysis. Comput Methods Programs Biomed. 2018;155:11–7. https://doi.org/10.1016/j.cmpb.2017.11.023
10. Mahato S, Goyal N, Ram D, Paul S. Detection of Depression and Scaling of Severity using six Channel EEG Data. J Med Syst. 2020;44(7):118. https://doi.org/10.1007/s10916-020-01573-y
11. Pan Z, Xiong D, Xiao H, et al. The effects of Repetitive Transcranial Magnetic Stimulation in patients with chronic Schizophrenia: insights from EEG microstates. Psychiatry Res. 2021;299:113866. https://doi.org/10.1016/j.psychres.2021.113866
12. Joseph A. Parametric inference with universal function approximators. 2019. https://doi.org/10.48550/arXiv.1903.04209
13. Balconi M, Pozzoli U. Arousal effect on emotional face comprehension: frequency band changes in different time intervals. Physiol Behav. 2009;97(3–4):455–62. https://doi.org/10.1016/j.physbeh.2009.03.023
14. Zielinski MC, Tang W, Jadhav SP. The role of replay and theta sequences in mediating hippocampal-prefrontal interactions for memory and cognition. Hippocampus. 2020;30(1):60–72. https://doi.org/10.1002/hipo.22821
15. Lee SM, Jang KI, Chae JH. Electroencephalographic correlates of suicidal ideation in the Theta Band. Clin EEG Neurosci. 2017;48(5):316–21. https://doi.org/10.1177/1550059417692083
16. Knyazev GG. EEG delta oscillations as a correlate of basic homeostatic and motivational processes. Neurosci Biobehav Rev. 2012;36(1):677–. https://doi.org/10.1016/j.neubiorev.2011.10.002. 95.
17. Harmony T. The functional significance of delta oscillations in cognitive processing. Front Integr Neurosci. 2013;7:83. https://doi.org/10.3389/fnint.2013.00083
18. Jiang H, Popov T, Jylänki P, et al. Predictability of depression severity based on posterior alpha oscillations. Clin Neurophysiol. 2016;127(4):2108–14. https://doi.org/10.1016/j.clinph.2015.12.018
19. Umemoto A, Panier L, Cole SL, Kayser J, Pizzagalli DA, Auerbach RP. Resting posterior alpha power and adolescent major depressive disorder. J Psychiatr Res. 2021;141:233–40. https://doi.org/10.1016/j.jpsychires.2021.07.003
20. Zoon HF, Veth CP, Arns M, et al. EEG alpha power as an intermediate measure between brain-derived neurotrophic factor Val66Met and depression severity in patients with major depressive disorder. J Clin Neurophysiol. 2013;30(3):261–7. https://doi.org/10.1097/WNP.0b013e3182933d6e
21. Halgren M, Ulbert I, Bastuji H, et al. The generation and propagation of the human alpha rhythm. Proc Natl Acad Sci U S A. 2019;116(47):23772–82. https://doi.org/10.1073/pnas.1913092116
22. Gordeev SA. Clinical-psychophysiological studies of patients with panic Attacks with and without agoraphobic disorders. Neurosci Behav Physiol. 2008;38(6):633–7. https://doi.org/10.1007/s11055-008-9016-3
23. Kline JP, Blackhart GC, Williams WC. Anterior EEG asymmetries and opponent process theory. Int J Psychophysiol. 2007;63(3):302–7. https://doi.org/10.1016/j.ijpsycho.2006.12.003
24. Ozan E, Okur H, Eker C, Eker OD, Gönül AS, Akarsu N. The effect of depression, BDNF gene val66met polymorphism and gender on serum BDNF levels. Brain Res Bull. 2010;81(1):61–5. https://doi.org/10.1016/j.brainresbull.2009.06.022
25. Grin-Yatsenko VA, Baas I, Ponomarev VA, Kropotov JD. Independent component approach to the analysis of EEG recordings at early stages of depressive disorders. Clin Neurophysiol. 2010;121(3):281–9. https://doi.org/10.1016/j.clinph.2009.11.015
26. Hosseinifard B, Moradi MH, Rostami R. Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal. Comput Methods Programs Biomed. 2013;109(3):339–45. https://doi.org/10.1016/j.cmpb.2012.10.008
27. Segalowitz SJ, Santesso DL, Jetha MK. Electrophysiological changes during adolescence: a review. Brain Cogn. 2010;72(1):86–100. https://doi.org/10.1016/j.bandc.2009.10.003
28. Kaiser RH, Andrews-Hanna JR, Wager TD, Pizzagalli DA. Large-Scale Network Dysfunction in Major Depressive disorder: a Meta-analysis of resting-state functional connectivity. JAMA Psychiatry. 2015;72(6):603–11. https://doi.org/10.1001/jamapsychiatry.2015.0071
29. Mohammadi Y, Moradi MH. Prediction of Depression Severity scores based on functional connectivity and complexity of the EEG Signal. Clin EEG Neurosci. 2021;52(1):52–60. https://doi.org/10.1177/1550059420965431
30. Shim M, Im CH, Kim YW, Lee SH. Altered cortical functional network in major depressive disorder: a resting-state electroencephalogram study. Neuroimage Clin. 2018;19:1000–7. https://doi.org/10.1016/j.nicl.2018.06.012
31. Clancy KJ, Andrzejewski JA, You Y, Rosenberg JT, Ding M, Li W. Transcranial stimulation of alpha oscillations up-regulates the default mode network. Proc Natl Acad Sci U S A. 2022;119(1). https://doi.org/10.1073/pnas.2110868119
32. Fingelkurts AA, Fingelkurts AA, Rytsälä H, Suominen K, Isometsä E, Kähkönen S. Impaired functional connectivity at EEG alpha and theta frequency bands

in major depression. Hum Brain Mapp. 2007;28(3):247–61. https://doi.org/10.1002/hbm.20275

33. Leuchter AF, Cook IA, Hunter AM, Cai C, Horvath S. Resting-state quantitative electroencephalography reveals increased neurophysiologic connectivity in depression. PLoS ONE. 2012;7(2):e32508. https://doi.org/10.1371/journal.pone.0032508

34. Siegle GJ, Condray R, Thase ME, Keshavan M, Steinhauer SR. Sustained gamma-band EEG following negative words in depression and schizophrenia. Int J Psychophysiol. 2010;75(2):107–18. https://doi.org/10.1016/j.ijpsycho.2008.04.008

35. Burgess AP, Ali L. Functional connectivity of gamma EEG activity is modulated at low frequency during conscious recollection. Int J Psychophysiol. 2002;46(2):91–100. https://doi.org/10.1016/s0167-8760(02)00108-3

36. Herrmann CS, Munk MH, Engel AK. Cognitive functions of gamma-band activity: memory match and utilization. Trends Cogn Sci. 2004;8(8):347–55. https://doi.org/10.1016/j.tics.2004.06.006

37. Lee S, Seol HS, Eom S, et al. Hydroxy Pentacyclic Triterpene Acid, Kaempferol, inhibits the human 5-Hydroxytryptamine type 3A receptor activity. Int J Mol Sci. 2022;23(1). https://doi.org/10.3390/ijms23010544

38. Coulston CM, Perdices M, Henderson AF, Malhi GS. Cannabinoids for the treatment of schizophrenia? A balanced neurochemical framework for both adverse and therapeutic effects of cannabis use. Schizophr Res Treatment. 2011;2011:501726. https://doi.org/10.1155/2011/501726

39. Gaetz W, Edgar JC, Wang DJ, Roberts TP. Relating MEG measured motor cortical oscillations to resting γ-aminobutyric acid (GABA) concentration. NeuroImage. 2011;55(2):616–21. https://doi.org/10.1016/j.neuroimage.2010.12.077

40. Whitton AE, Deccy S, Ironside ML, Kumar P, Beltzer M, Pizzagalli DA. Electroencephalography Source Functional Connectivity Reveals Abnormal high-frequency communication among large-scale functional networks in Depression. Biol Psychiatry Cogn Neurosci Neuroimaging. 2018;3(1):50–8. https://doi.org/10.1016/j.bpsc.2017.07.001

41. Kim S, Baek JH, Shim SH, et al. Alteration of cortical functional networks in mood disorders with resting-state electroencephalography. Sci Rep. 2022;12(1):5920. https://doi.org/10.1038/s41598-022-10038-w

42. Müller VI, Cieslik EC, Laird AR, Fox PT, Eickhoff SB. Dysregulated left inferior parietal activity in schizophrenia and depression: functional connectivity and characterization. Front Hum Neurosci. 2013;7:268. https://doi.org/10.3389/fnhum.2013.00268

43. Chang H, Zong Y, Zheng W, et al. EEG-based major depressive disorder recognition by selecting discriminative features via stochastic search. J Neural Eng. 2023;20(2). https://doi.org/10.1088/1741-2552/acbe20

44. Saeedi M, Saeedi A, Maghsoudi A. Major depressive disorder assessment via enhanced k-nearest neighbor method and EEG signals. Phys Eng Sci Med. 2020;43(3):1007–18. https://doi.org/10.1007/s13246-020-00897-w

45. Wu CT, Huang HC, Huang S, et al. Resting-State EEG Signal for Major Depressive Disorder Detection: a systematic validation on a large and diverse dataset. Biosens (Basel). 2021;11(12). https://doi.org/10.3390/bios11120499

46. Li Y, Shen Y, Fan X, et al. A novel EEG-based major depressive disorder detection framework with two-stage feature selection. BMC Med Inform Decis Mak. 2022;22(1):209. https://doi.org/10.1186/s12911-022-01956-w

Publisher's Note