



Effects of Brain Atlases and Machine Learning Methods on the Discrimination of Schizophrenia Patients: A Multimodal MRI Study

Jinyu Zang^{1,2,3}, Yuanyuan Huang^{2,4}, Lingyin Kong^{1,2,3}, Bingye Lei^{1,2,3}, Pengfei Ke^{1,2,3}, Hehua Li⁴, Jing Zhou^{1,2,3}, Dongsheng Xiong^{1,2,3}, Guixiang Li^{5,6}, Jun Chen^{5,6}, Xiaobo Li⁷, Zhiming Xiang^{5,8}, Yuping Ning^{2,4}, Fengchun Wu^{4*} and Kai Wu^{1,2,3,4,5,6,9,10*}

¹ Department of Biomedical Engineering, School of Material Science and Engineering, South China University of Technology, Guangzhou, China, ² Guangdong Engineering Technology Research Center for Translational Medicine of Mental Disorders, Guangzhou, China, ³ National Engineering Research Center for Tissue Restoration and Reconstruction, South China University of Technology, Guangzhou, China, ⁴ The Affiliated Brain Hospital of Guangzhou Medical University, Guangzhou Hui'ai Hospital, Guangzhou, China, ⁵ Guangdong Engineering Technology Research Center for Diagnosis and Rehabilitation of Dementia, Guangzhou, China, ⁶ National Engineering Research Center for Healthcare Devices, Guangzhou, China, ⁷ Department of Biomedical Engineering, New Jersey Institute of Technology, Newark, NJ, United States, ⁸ Department of Radiology, Panyu Central Hospital of Guangzhou, Guangzhou, China, ⁹ Key Laboratory of Biomedical Engineering of Guangdong Province, South China University of Technology, Guangzhou, China, ¹⁰ Department of Nuclear Medicine and Radiology, Institute of Development, Aging and Cancer, Tohoku University, Sendai, Japan

OPEN ACCESS

Edited by:

Ahmed Shalaby,
University of Louisville, United States

Reviewed by:

Stavros I. Dimitriadis,
Greek Association of Alzheimer's
Disease and Related Disorders,
Greece
Lotfi Chaari,
UMR5505 Institut de Recherche en
Informatique de Toulouse (IRIT),
France

*Correspondence:

Fengchun Wu
13580380071@163.com
Kai Wu
kaiwu@scut.edu.cn

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 19 April 2021

Accepted: 07 July 2021

Published: 27 July 2021

Citation:

Zang J, Huang Y, Kong L, Lei B,
Ke P, Li H, Zhou J, Xiong D, Li G,
Chen J, Li X, Xiang Z, Ning Y, Wu F
and Wu K (2021) Effects of Brain
Atlases and Machine Learning
Methods on the Discrimination of
Schizophrenia Patients: A Multimodal
MRI Study.
Front. Neurosci. 15:697168.
doi: 10.3389/fnins.2021.697168

Recently, machine learning techniques have been widely applied in discriminative studies of schizophrenia (SZ) patients with multimodal magnetic resonance imaging (MRI); however, the effects of brain atlases and machine learning methods remain largely unknown. In this study, we collected MRI data for 61 first-episode SZ patients (FESZ), 79 chronic SZ patients (CSZ) and 205 normal controls (NC) and calculated 4 MRI measurements, including regional gray matter volume (GMV), regional homogeneity (ReHo), amplitude of low-frequency fluctuation and degree centrality. We systematically analyzed the performance of two classifications (SZ vs NC; FESZ vs CSZ) based on the combinations of three brain atlases, five classifiers, two cross validation methods and 3 dimensionality reduction algorithms. Our results showed that the groupwise whole-brain atlas with 268 ROIs outperformed the other two brain atlases. In addition, the leave-one-out cross validation was the best cross validation method to select the best hyperparameter set, but the classification performances by different classifiers and dimensionality reduction algorithms were quite similar. Importantly, the contributions of input features to both classifications were higher with the GMV and ReHo features of brain regions in the prefrontal and temporal gyri. Furthermore, an ensemble learning method was performed to establish an integrated model, in which classification performance was improved. Taken together, these findings indicated the effects of these factors in constructing effective classifiers for psychiatric diseases and showed that the integrated model has the potential to improve the clinical diagnosis and treatment evaluation of SZ.

Keywords: multimodal MRI, schizophrenia, brain atlas, machine learning, classification

INTRODUCTION

Schizophrenia (SZ) is a chronic psychiatric disorder, characterized by disabling mental symptoms such as auditory delusions, hallucinations and disrupted higher-order cognitive functions (Austin, 2005; Leucht et al., 2007). With the development of machine learning methods, both structural and functional magnetic resonance imaging (MRI) data have been applied into the discriminative analyses of SZ patients (Kasperek et al., 2011; Deanna et al., 2012; Ota et al., 2012; Liu Y. et al., 2017; Chen et al., 2020). For example, support vector machine (SVM) is the most widely used method to distinguish SZ patients from normal controls (NCs) (Liu Y. et al., 2017; Chen et al., 2020) or to differentiate illness stages of SZ, such as first-episode schizophrenia (FESZ) and chronic schizophrenia (CSZ) (Lu et al., 2018; Wu et al., 2018). Similarly, other classifiers such as random forest (Deanna et al., 2012) and linear discriminant analysis (LDA) (Kasperek et al., 2011; Ota et al., 2012) have also been utilized in discriminative analyses of SZ patients.

Recently, a number of discriminative studies of SZ patients have adopted the strategy of multiple classifiers, including LDA (Junhua et al., 2018), SVM (Watanabe et al., 2014; Raymond et al., 2017) and extreme learning machine (Iqbal et al., 2017), and multiple dimensionality reduction algorithms, such as principle component analysis (PCA) (Raymond et al., 2017) and *t*-test (Junhua et al., 2018). Importantly, different classifiers have been selected as the best classifier in previous studies, which shows diversity among the machine learning methods. Thus, to achieve an optimal performance of a discriminative analysis, a systematic evaluation with multiple machine learning methods is essential and of great importance.

Moreover, previous discriminative analyses using different brain atlases have shown that the choice of brain atlases seems rather arbitrary and could lead to different results (Kalmady et al., 2019). A number of researchers have performed discriminative analyses of SZ patients based on the automated anatomical labeling (AAL) atlas (Tzourio-Mazoyer et al., 2002) with accuracies from 76.3% to 85% (Longfei et al., 2013; Kim et al., 2015; Junhua et al., 2018; Matsubara et al., 2019). A brain atlas with 95 regions of interest (ROIs) has also been utilized in the discriminative analysis of SZ patients to achieve 89.3% sensitivity and 93.6% specificity (Karageorgiou et al., 2011). Additionally, another study used the Desikan atlas for discriminative analysis and obtained an accuracy of 85.0% (Xiao et al., 2017). However, few studies have evaluated the effects of brain atlases on discriminative analyses of SZ patients.

In this study, we collected structural MRI (sMRI) and resting-state functional MRI (rs-fMRI) data from 345 subjects and used three brain atlases to calculate 4 MRI measurements, including regional gray matter volume (GMV), regional homogeneity (ReHo), amplitude of low-frequency fluctuation (ALFF) and degree centrality (DC). We then performed a systematic evaluation of the classification performances in two classifications (NC vs SZ, FESZ vs CSZ) using five classifiers, two cross validation methods, and 3 dimensionality reduction algorithms. Moreover, an ensemble learning method was performed to establish an integrated model to improve the clinical diagnosis.

MATERIALS AND METHODS

Subjects

A total of 61 FESZ patients, 79 CSZ patients and 205 NCs were included (Table 1). The inclusion and exclusion criteria of subjects were the same as those in our previous studies (Lu et al., 2016; Wu et al., 2018). All subjects, aged 18 to 45 and of Han nationality, underwent a clinical assessment with the Positive and Negative Syndrome Scale (PANSS) which contains three subscales (general psychopathology, positive symptoms and negative symptoms) and indicates the severity of the symptoms (Kay et al., 1987; Van Tol et al., 2014). Only the subjects with PANSS scores over 60 and with a period of education of more than 6 years were chosen for the project. Meanwhile, they also had to be diagnosed by experienced clinical psychiatrists to be SZ in accordance with the Diagnostic and Statistical Manual of Mental Disorders-IV-Text Revision (DSM-IV-TR) criteria (First et al., 1997). Among these subjects, first-episode patients with a course of disease under 2 years were categorized as FESZ if they had not taken any antipsychotic drugs. Meanwhile, the patients who had suffered recurrent symptoms and had already undergone drug therapy with a course of disease over 2 years were categorized as CSZ.

If one of the following criteria were met, the SZ patient was excluded: (1) alcohol dependence or other mental disorders, such as depressive disorder, dementia or mental retardation, based on DSM-IV-TR criteria; (2) severe physical disorders potentially derived from substance dependence including definite diabetes, hypertension, heart disease, thyroid diseases or narrow-angle glaucoma; (3) history of epilepsy or febrile convulsions; (4) electroconvulsive therapy within the past six months; (5) serious tardive dyskinesia or drug-induced neuroleptic malignant syndrome; (6) contraindication for MRI; (7) lack of legal guardians or noncompliant with drug administration; (8) an irritative state or a serious suicide attempt; and (9) lactation, pregnancy or anticipated pregnancy. Meanwhile, NCs with pregnancy, contraindications for MRI or relatives diagnosed with psychiatric Axis I disorders based on the DSM-IV-TR criteria were also excluded. All the subjects' data were collected from the Affiliated Brain Hospital of Guangzhou Medical University, and all subjects were informed about the experimental details and signed informed consent before clinical tests. The research was strictly subject to the Declaration of Helsinki and was under approval of the ethics committees of the Affiliated Brain Hospital of Guangzhou Medical University.

MRI Data Acquisition

Magnetic resonance imaging data of all subjects were collected by a Philips 3T MR device system in the Affiliated Brain Hospital of Guangzhou Medical University. The echo-planar imaging (EPI) sequence (repetition time = 2,000 ms, echo time = 30 ms, acquisition time = 2,000 ms, field of view = 210 mm × 210 mm, flip angle = 90°, spatial resolution = 3.4 mm × 3.4 mm × 4 mm, 64 × 64 × 33 matrix) was used to generate the functional MRI data. The gradient-echo T1-weighted sequence (repetition time = 8.2 ms, echo time = 3.7 ms, flip angle = 7°, spatial

TABLE 1 | Demographic and clinical characteristics.

	FESZ patients	CSZ patients	NC	Statistic value	P-value
Gender (M/F)	41/20	54/25	110/95	$\chi^2 = 3.53$	0.03
Age (years)	32.08 ± 7.42	33.21 ± 8.37	32.52 ± 8.40	$F = 5.39$	<0.05 ^{a,b}
Education (years)	10.39 ± 3.25	11.97 ± 3.22	12.84 ± 2.83	$F = 21.33$	<0.05 ^{a,b}
PANSS-PScore	24.02 ± 4.50	22.47 ± 5.70	–	$T = 1.74$	0.083
PANSS-NScore	21.64 ± 7.70	23.22 ± 7.29	–	$T = -1.24$	0.218
PANSS-GScore	40.31 ± 8.85	39.54 ± 10.18	–	$T = 0.47$	0.641
PANSS-TScore	85.97 ± 17.49	85.23 ± 19.44	–	$T = 0.23$	0.816

All data presented above were in format: average ± standard deviation. The factor age (years) and education (years) were analyzed by separate one-way ANOVA and the gender was analyzed by χ^2 test. Post hoc pairwise comparison was utilized to analyze distinguished discrimination with two-sample t-test. P-value < 0.05 was considered significant. ^aPost hoc pairwise comparison showed the significant discrepancy between FESZ and NC. ^bPost hoc pairwise comparison showed the significant discrepancy between CSZ and NC.

CSZ, chronic schizophrenia; F, female; FESZ, first-episode schizophrenia; GScore, general score; M, male; NCs, normal controls; NScore, negative syndrome score; PScore, positive syndrome score; TScore, total syndrome score.

resolution = 1 mm × 1 mm × 1 mm, 256 × 256 × 188 matrix) was used to generate the structural MRI data. All participants were instructed to minimize head movement with the eyes closed in a sober state.

Preprocessing

The sMRI data were preprocessed by the SPM12 software package¹ to calculate GMV. The raw images were first standardized with a customized template provided by the DARTEL template creation tool to eliminate the deviation caused by individual discrepancies. Then they were separated into gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF) by the VBM toolkit embedded in the SPM12 software package. The images were also smoothed by an 8 mm full width at half maximum (FWHM) Gaussian kernel. Finally, the structural brain data from GM images were calculated in each region of three brain atlases, including the AAL atlas with 90 brain regions (Tzourio-Mazoyer et al., 2002), the human brainnetome (HBN) atlas with 246 brain regions (Fan et al., 2016), and the groupwise whole-brain (GWB) atlas with 268 brain regions (Shen et al., 2013; Finn et al., 2015).

The rs-fMRI data were also preprocessed by SPM12 and DPARSF V4.4 software². The image data from the first 10 time series were excluded owing to the instability of the device and fluctuations in the subjects' mental state at the beginning. The noise from the variance in signal acquisition times and from head motion was eliminated to amplify the valid image signal. The images were then normalized by the standard EPI template. Finally, the bandpass filter (0.01–0.08 Hz) was utilized to reduce the high-frequency physiological noise and the low-frequency drift. Three rs-fMRI measurements including ReHo, ALFF and DC were calculated in each region of the three brain atlases.

ReHo is Kendall's coefficient of concordance on the time series of a certain voxel with respect to its 26 adjacent neighbors, suggesting functional synchronization within the voxel and its neighbors (Govindarajulu, 1992). The ReHo values were normalized to lessen the deviation resulting from individual

variance and were averaged in each region of the different brain atlases. ALFF is measured as the averaged square root within the bandpass (0.01–0.08 Hz) after fast Fourier transform (FFT) for each voxel and represents the level of regional spontaneous neuronal activity (Yu-Feng et al., 2007). Similarly, the ALFF values of each voxel were divided by the global average ALFF value for normalization and averaged in each region of the different brain atlases. The DC is described as the average of the Pearson correlation coefficients between the time series of a certain ROI and those of other ROIs, evaluating the connection degree of a certain ROI to other ROIs (Zang et al., 2004). The time series of a certain ROI was calculated as the averaged time series of all voxels in that region.

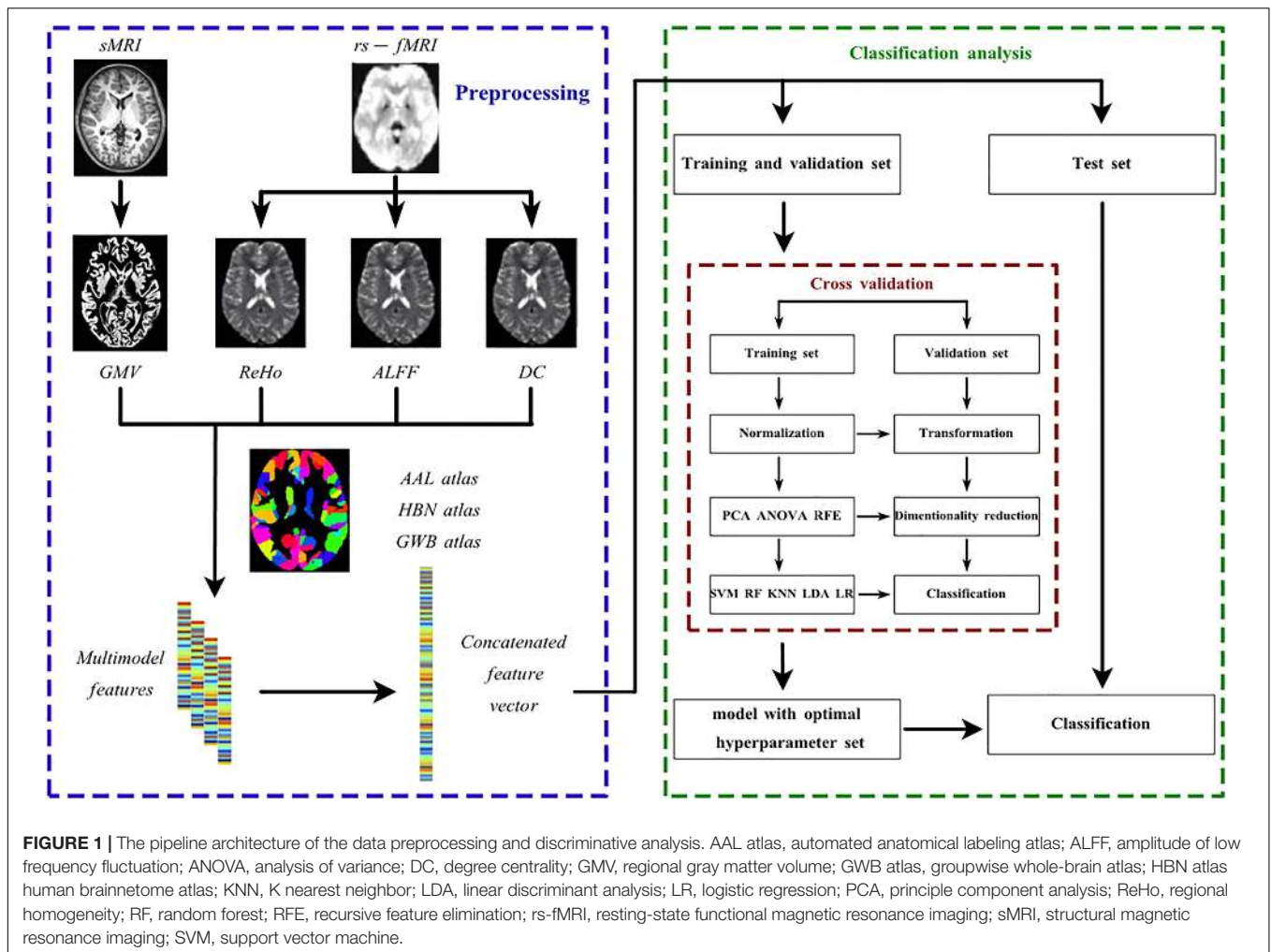
Classification Analysis

After the preprocessing, the multimodal features were combined to form the concatenated feature vector. The concatenated feature vector was composed of 360 measurements (90 GMV measurements+90 ReHo measurements+90 ALFF measurements+90 DC measurements) if AAL atlas was used, of 986 measurements (246 GMV measurements+246 ReHo measurements+246 ALFF measurements+246 DC measurements) if HBN atlas was used, and of 1,072 measurements (268 GMV measurements+268 ReHo measurements+268 ALFF measurements+268 DC measurements) if GWB atlas was used.

With the concatenated feature vector available, the whole pipeline architecture of classification is shown in **Figure 1**. In both classifications, five classifiers were utilized, including SVM, logistic regression (LR), LDA, random forest (RF) and K nearest neighbor (KNN). First developed by Vapnick in 1995, SVM aims to find the optimal hyperplane separating the dataset with different labels into multiple hyperspaces (Sain, 1997). Similar to SVM, LR also generates a hyperplane by a linear transformation function and sigmoid activation function to separate the data and to further provide the probability of unseen data being classified into a certain group (Peng and Ingersoll, 2002). LDA was first suggested by Fisher in 1936 (Fisher, 1936). Its principle is to project the dataset to the 1D dimension where the points representing data within the same group tend to get close

¹<http://www.fil.ion.ucl.ac.uk/spm/>

²<http://rfmri.org/DPARSF>



and points representing data in different groups separate from each other (Fisher, 1936). RF is a bagging algorithm, and it predicts unseen data labels based on votes from all decision trees embedded in RF (Breiman, 1996). KNN simply counts the labels of a datum's K nearest neighbors and predicts this datum's label as the one with the highest frequency (Zhang, 2016). The list of hyperparameters available for all classifiers is shown in **Supplementary Table 1**. The best hyperparameter set was selected by cross validation as mentioned later.

Considering the redundancy or irrelevance of the features which may lead to overfitting when training, dimensionality reduction was performed before classification process. The dimensionality reduction algorithms applied in this study included PCA, analysis of variance (ANOVA) and recursive feature elimination (RFE). PCA has been widely used as a feature selection method in machine learning classification (Cao et al., 2003; Kriti Virmani et al., 2016). It basically projects the data to lower dimensions with the largest variance by a linear function (Jolliffe, 2005). It calculates the eigenvalues and eigenvectors of the covariance matrix in original feature space, and then selects n% percent of eigenvalues to represent the discriminative energy of data on different level. The corresponding n% percent

of eigenvectors would then form the transformation matrix. ANOVA is also a common method for feature selection (Sheikhan et al., 2013; Bejani and Gharavian, 2014; Li et al., 2018; Abdulsalam et al., 2020). It first performs the F test on each separate feature together with data labels. Then it selects features according to the percentile of the highest F scores (Neter et al., 1996). Similarly, RFE also excludes the features with low relevance to label prediction, but the criteria refer to the weights derived from a certain classifier such as SVM (Guyon et al., 2002). Basically, it prunes the least important features iteratively according to the weights derived from the classifier until the desired number of features to select is reached. It has also been widely used in feature selection and achieved good results (Fernandez-Lozano et al., 2014; Xue et al., 2018; Albashish et al., 2021). The list of hyperparameters available for all dimensionality reduction methods is shown in **Supplementary Table 2**. The best hyperparameter set was selected by cross validation as mentioned below.

Twenty percent of the whole dataset was randomly separated to establish a separate test set for classification evaluation after training, and the rest of the dataset was subjected to cross validation to construct the predictive model with the

best hyperparameter set. The application of a separate test set guaranteed the generalization of the classification model. In this study, there were two types of cross validation methods available, namely, 10-fold cross validation (10FCV) and leave-one-out cross validation (LOOCV). The cross validation in the study is mainly used for hyperparameter selection, which is also known as grid search cross validation (Sarah and Media, 2017). In 10FCV, the dataset was split into 10 portions of equal size, where 1 portion was used for validation and the remaining nine portions were used for training, and this occurred in an iterative manner. During the cross validation, the data were standardized by removing the mean and scaling to unit variance before the dimensionality reduction and classification. The normalization, dimensionality reduction and classification constituted the model as a whole. All possible combinations of hyperparameters for the model, as shown in **Supplementary Tables 1, 2**, were validated by the 10 validation sets in the 10FCV. The optimal hyperparameter set for the model was selected based on the average accuracy generated from the 10 iterations and was applied to construct the predictive model trained by 10 portions of data taken together. The performance of the model was assessed using the separate test set. Similar to 10FCV, LOOCV simply selects one portion of the data for validation and all others for training.

The classification performances were systematically analyzed in both classifications for different combinations of brain atlases, classifiers, cross validation methods, and dimensionality reduction algorithms. The receiver operating characteristic (ROC) curve was also plotted to calculate the area under the ROC curve (AUC), which was between 0 and 1. It is believed that the closer the AUC is to 1, the better the classification is. In parallel, the sensitivity and specificity were measured for a further assessment of the performance, and these definitions are shown below.

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (2)$$

$$\text{Accuracy} = \frac{TP + TN}{TN + FN + TN + TP} \quad (3)$$

True positive (TP): the number of positive samples predicted as positive; true negative (TN): the number of negative samples predicted as negative; false positive (FP): the number of negative samples predicted as positive; and false negative (FN): the number of positive samples predicted as negative.

Furthermore, the permutation test, a widely used nonparametric test examining a null hypothesis (Golland and Fischl, 2003; Liu et al., 2015), was performed to analyze the significance of the classification results. In this study, the permutation test was carried out by randomly permutating the labels of all datasets and evaluating the classification performance with these permuted data 1,000 times. If the P -value was small enough ($P < 0.05$ was used in this study), the hypothesis that the classifier had significantly discovered the difference between the two groups with given set of imaging data could be safely

accepted. The P -value was calculated as the percentage of classifications with better performance based on permuted data over all 1,000 trials.

Subsequently, the parameters from the best model (i.e., the combination of a certain dimensionality reduction algorithm and classifier) were analyzed to discover the brain regions with the greatest contribution in both classifications. In detail, the weights extracted from a certain classifier were first transformed to their absolute values. Then these absolute values were further transformed to their original feature space according to the dimensionality reduction algorithm and normalized as brain region contributions for the ranking process. We selected the top 5% features (this involved a different number of features for the different atlases) with the greatest contribution and further calculated the actual percentage of their contribution.

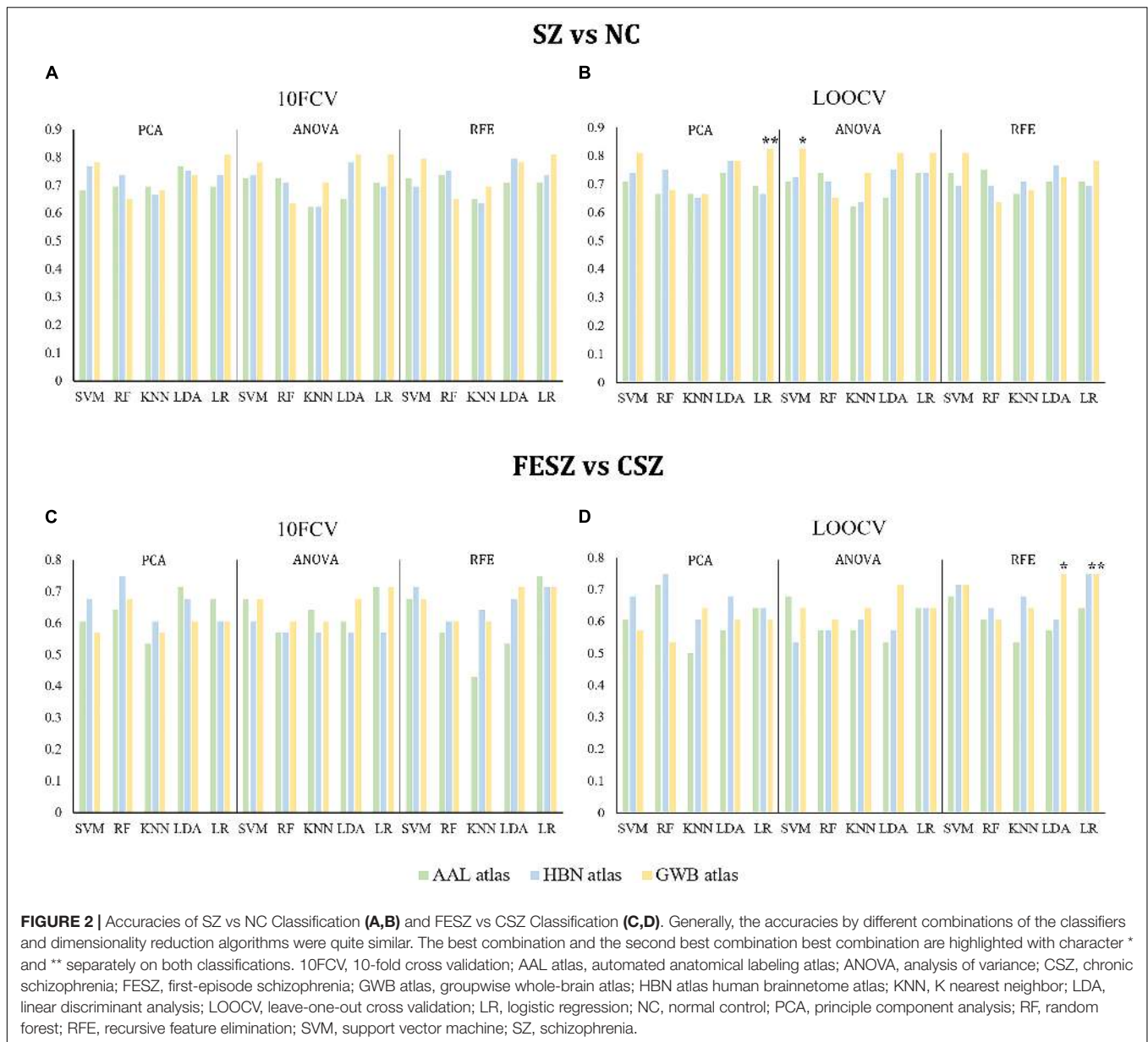
Finally, to improve the clinical diagnosis, we established an integral model for each classification with the stacking technique (Wolpert, 1992; Ting and Witten, 1999). All models (the combinations of three brain atlases, five classifiers, two cross validation methods, and 3 dimensionality reduction algorithms) generated by the pipeline were selected as level-0 generalizers and the gradient boosting algorithm (Friedman, 2001) was selected as the level-1 generalizer. The best hyperparameter set selected by the pipeline structure mentioned above is also applied for each model in level-0 generalizer, and the hyperparameter set for the level-1 generalizer is selected by a fivefold cross validation. Since it is binary classification on both classifications (SZ or NC in SZ vs NC classification; FESZ or CSZ in FESZ vs NC classification), the input data for the level-1 generalizer is generated as the probability to be classified as one class by all level-0 generalizers in both classifications. Therefore, the dimension of the features was identical to the number of the level-0 generalizers (90 for each classification). The train set and separate test set for level-1 generalizer were generated by fivefold cross validation as detailed described in references (Wolpert, 1992; Ting and Witten, 1999). The performance of the final integral model was tested by the separate test set to guarantee generalization. The hyperparameter set available for the gradient boosting algorithm is shown in **Supplementary Table 2**.

The whole classification process was realized by the sklearn software package (<https://scikit-learn.org/stable/>) for machine learning in python code and an in-house software "NEURO-LEARN" (<https://www.github.com/Raniac/NEURO-LEARN>).

RESULTS

Overall Classifier Performance

The results of both classifications are shown in **Figure 2** and **Supplementary Tables 3, 4**, and the optimal hyperparameter sets selected for all models are shown in **Supplementary Table 5**. We selected the models with the highest accuracy and then ranked them by AUC. In the classification between SZ and NC, the best combination was PCA with LR using the GWB atlas with LOOCV (accuracy: 0.83, $P < 0.05$; AUC: 0.89, $P < 0.05$; sensitivity: 0.89; specificity: 0.78; **Figure 2B**, indicated by **). Moreover, the second-best combination was ANOVA with SVM



using the GWB atlas with LOOCV (accuracy: 0.83, $P < 0.05$; AUC: 0.86, $P < 0.05$; sensitivity: 0.71; specificity: 0.90; **Figure 2B**, indicated by *). Similarly, in the classification between FESZ and CSZ, the best combination was RFE with LR using the GWB atlas with LOOCV (accuracy: 0.75, $P < 0.05$; AUC: 0.77, $P < 0.05$; sensitivity: 0.80; specificity: 0.69; **Figure 2D**, indicated by **) and the second-best combination was RFE with LDA using the GWB atlas with LOOCV (accuracy: 0.75, $P < 0.05$; AUC: 0.77, $P < 0.05$; sensitivity: 0.80; specificity: 0.69; **Figure 2D**, indicated by *).

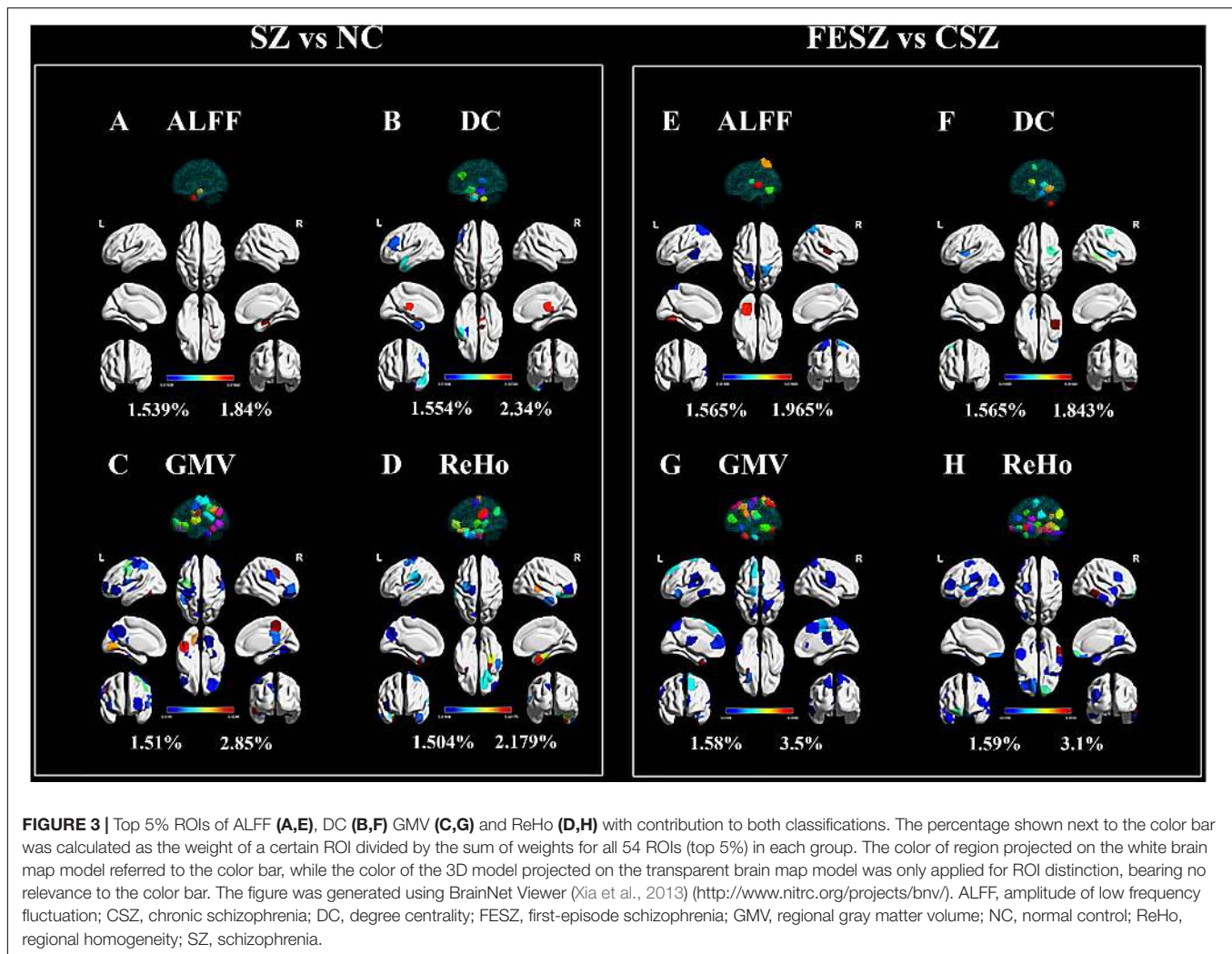
Together, it was discovered that: (1) the GWB atlas was the optimal atlas for both classifications and the best results by the HBN atlas (SZ vs NC: RFE-LDA-10FCV; FESZ vs CSZ: RFE-LR-LOOCV) were also comparable; (2) LR and RFE showed a slight advantage over the others, but generally the results with the various combinations of the classifiers and dimensionality

reduction algorithms were quite similar; and (3) LOOCV was the best method to identify the best hyperparameter set for both classifications.

Feature Importance Analysis

The best combination for two classifications (the combination of the GWB atlas, LR, LOOCV, and PCA for the SZ vs NC classification; the combination of the GWB atlas, LR, LOOCV, and RFE for the FESZ vs CSZ classification) were utilized to generate weights for feature ranking. Based on the methods stated above, the results are shown in **Figure 3** and **Supplementary Tables 6, 7**.

Generally, the GMV and ReHo features equally made the greatest contributions in both classifications that were much more than the contributions of ALFF and DC. In detail, ALFF



from the right limbic hippocampus possessed relatively higher weights in the classification between SZ and NC (Figure 3A). For the classification between FESZ vs CSZ, ALFF from the right parietal primary sensory, left occipital primary sensory and somatosensory association cortex made greater contributions (Figure 3E). DC contributed slightly more to the classification than ALFF. The DC with the highest weight came from the right brainstem and right subcortical thalamus for the SZ vs NC classification (Figure 3B), while the highest weights came from the inferior temporal gyrus, middle temporal gyrus and premotor cortex for the FESZ vs CSZ classification (Figure 3F). Features from GMV and ReHo made up approximately 80% of all 54 features selected. The GMV with the greatest contribution was derived from the right premotor cortex, dorsal posterior cingulate cortex, left temporal fusiform cortex, right prefrontal pars opercularis and left orbitofrontal area in the classification between SZ and NC (Figure 3C), while the highest weights were derived from the left temporal pole, left prefrontal visual field, left motor strip, right motor strip and inferior prefrontal gyrus in the classification between FESZ and CSZ (Figure 3G). Meanwhile, the ReHo features with the greatest contributions

stemmed from the left orbitofrontal area, left temporal pole, right limbic parahippocampus and right middle temporal gyrus in the classification between SZ and NC (Figure 3D) and the highest weights stemmed from right the temporal fusiform region, right orbitofrontal area, left insula, left cerebellum and left orbitofrontal area in the classification between FESZ and CSZ (Figure 3H).

Further measurements were also performed for the brain region contribution in both classifications with all four features (ReHo, ALFF, DC, and GMV) considered (Figure 4A for SZ vs NC classification; Figure 4B for FESZ vs CSZ classification). The contribution of a certain brain region was calculated as the sum of contributions of all four features located in that brain region. It is evident that the brain regions that contributed most to the SZ vs NC classification were the left prefrontal cortex, right prefrontal cortex, right limbic cortex, left temporal cortex and left motor strip. Similarly, the brain regions that contributed most to the FESZ vs CSZ classification were the left prefrontal cortex, right temporal cortex and left temporal cortex.

Besides, the best models by the HBN atlas (the combination of the LDA, 10FCV, and RFE for the SZ vs NC classification;

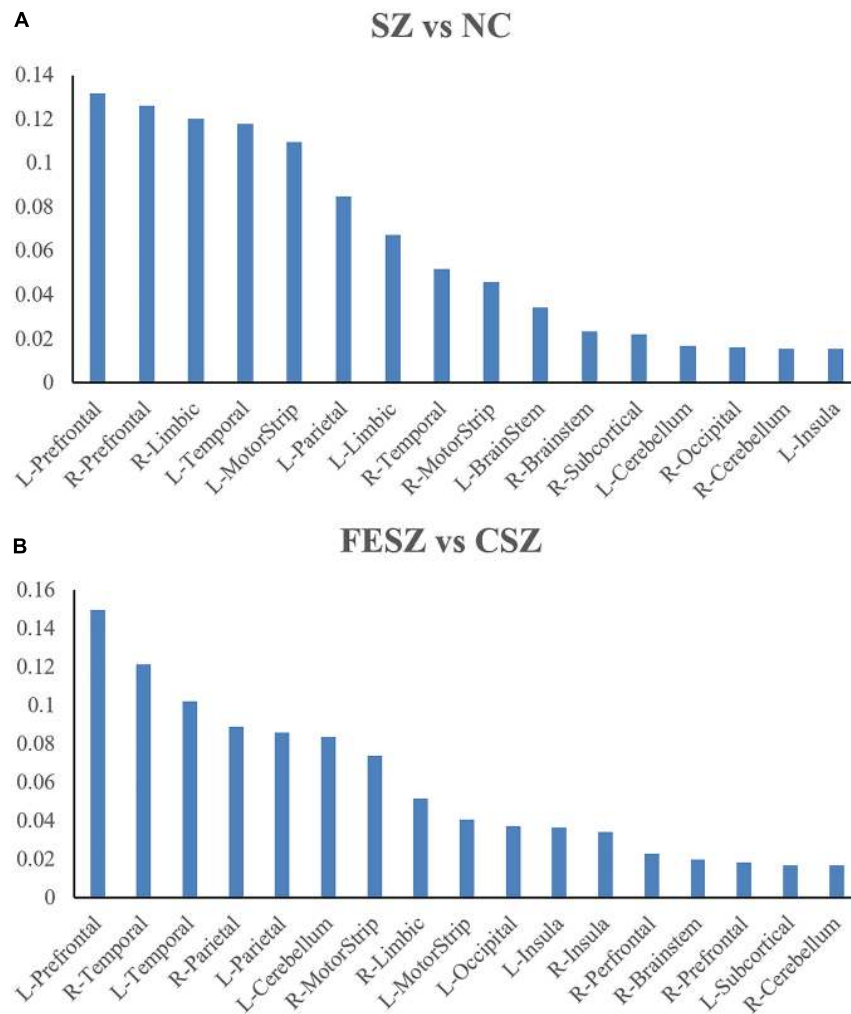


FIGURE 4 | Brain regions with most contribution to SZ vs NC Classification (A) and FESZ vs CSZ Classification (B). The percentage shown as the y axis is calculated as the weight of features from a certain brain region divided by the sum of weights of all top 5% features in each classification. The matching between the ROI and the brain region refers to <https://bioimagesuitweb.github.io/webapp/connviewer.html>. CSZ, chronic schizophrenia; FESZ, first-episode schizophrenia; L, left; NC, normal control; R, right; SZ, schizophrenia.

the combination of LR, LOOCV, and RFE for the FESZ vs CSZ classification) were also utilized to generate weights for feature ranking. The result shows that GMV and ReHo features have made more contribution than ALFF and DC features to both classifications (in **Supplementary Figure 3** and **Supplementary Tables 8, 9**), which is consistent with the result by GWB atlas. Moreover, the contributory features from HBN atlas were also mainly from frontal cortex and temporal cortex as the same of GWB atlas, indicating the commonality of two atlas on extracting features on discriminative analysis for schizophrenia (in **Supplementary Figure 4** and **Supplementary Tables 8, 9**). On the other hand, the contributory ROIs from HBN atlas were not exactly the same as those from GWB atlas as shown in **Supplementary Figure 3**. This shows different brain atlases, although with similar number of ROIs, might still result in the different selection of features for the machine learning models.

Predictive Model Performance

The improvements in the performance with the integral model was evident as shown in **Table 2**. The accuracy and AUC were significantly increased by the stacking algorithm on the separate test set in both classifications (SZ vs NC: accuracy = 0.88, AUC = 0.92; FESZ vs CSZ: accuracy = 0.86, AUC = 0.80).

DISCUSSION

In this study, we systematically analyzed classification performances by using multiple brain atlases and multiple machine learning methods with multimodal MRI data. The main findings are as follows: 1) the GWB parcellation with 268 ROIs outperformed the other two brain atlases; (2) the LOOCV was the best method of cross validation to select the best hyperparameter set, but the results with different classifiers

TABLE 2 | Classification performance improvement for integral model.

Classification group	SZ vs NC		FESZ vs CSZ	
	Before stacking	After stacking	Before stacking	After stacking
Accuracy	0.83	0.88	0.75	0.86
AUC	0.89	0.92	0.77	0.80

The accuracy and AUC before stacking were selected from the optimal model for both classifications, respectively. It is clear from the results that the accuracy and AUC were significantly improved by stacking technique to establish the integral model.

AUC, the area under receiver operating characteristic curve; CSZ, chronic schizophrenia; FESZ, first episode schizophrenia; NC, normal control.

and dimensionality reduction algorithms were quite similar; (3) the GMV and ReHo features in the prefrontal and temporal gyri made the greatest contributions in both classifications; and (4) the ensemble learning method substantially improved classification performance.

Generally, the selection of the brain atlas may result in striking differences in performance of the classification of psychiatric diseases (Koikkalainen et al., 2011; Min et al., 2014; Liu J. et al., 2017; Asim et al., 2018; Kalmady et al., 2019). Kalmady et al. (2019) used 14 brain atlases for discriminative analyses of SZ patients and found that the accuracies of the classifications varied significantly across different brain atlases. Similarly, a number of discriminative analyses have also been performed with patients with Alzheimer's disease (AD) based on multiple brain atlases (Koikkalainen et al., 2011; Min et al., 2014; Asim et al., 2018), in which the features based on all atlases were used to establish the integral model and achieved the best classification performance (Koikkalainen et al., 2011; Min et al., 2014; Asim et al., 2018). Consistent with previous studies, our results also showed discrepancies in the classification performance with different brain atlases. Moreover, our results also suggested the apparent superiority of the GWB atlas, compared with the AAL atlas and the HBN atlas. Previous studies have indicated that the utilization of the GWB atlas has also resulted in satisfactory performances on other discriminative studies, which is consistent with our findings (Valizadeh et al., 2018; de Souza Rodrigues et al., 2019). We speculated that this superiority may derive from the node number and the construction method of the GWB atlas. First, the number of nodes in the GWB atlas is consistent with the range proposed by other studies (Craddock et al., 2012; Van Essen et al., 2012), which enables the brain atlas to provide a more fine-grained scheme than other brain atlases such as the AAL atlas (Finn et al., 2015). Second, the construction of the GWB atlas is based on a groupwise parcellation method, which guaranteed that each node contains voxels with similar resting-state timecourses (Bianciardi et al., 2009; Finn et al., 2015). This could ensure homogeneity within each node and thus better discriminability of the features from the MRI data (Shen et al., 2013). Therefore, these two traits may serve as important criteria for the selection of brain atlases in discriminative analysis of SZ patients.

The selection of machine learning methods has been consequential with regard to the classification process in recent studies (Watanabe et al., 2014; Iqbal et al., 2017; Raymond et al., 2017; Junhua et al., 2018). In this study, we found that the results with different combinations of classifiers and dimensionality reduction algorithms were quite similar, which

is consistent with previous studies (Khondoker et al., 2013; Raymond et al., 2017). In detail, although the LR and RFE exhibited a slight advantage over the others, PCA, ANOVA, SVM, and LDA could also be acceptable choices for the classification. Importantly, our results were surprising because most of the classifiers and dimensionality reduction algorithms were mathematically distinct. One of the plausible explanations is the inherent similarities within different classifiers (Hastie et al., 2008). Almost all classifiers are able to generate a hyperplane, which is the best geometrical feature to classify the distributions of the data in multidimensional feature space with unstructured noise (Raymond et al., 2017). Thus, if the multimodal MRI data confirmed to a specific distribution, similar performances would be achieved by different classifiers and dimensionality reduction algorithms. Furthermore, compared with 10-fold cross validation, LOOCV was discovered to be the optimal method for the selection of hyperparameter sets. Taking into consideration that more training data could be applied to classifiers for discriminative analyses with the LOOCV method (Efron, 1983), the advantage of LOOCV can be easily comprehended.

We also found that the GMV and ReHo features better represented the major discrepancies in both classifications than the ALFF and DC features. These results suggested the necessity of using both structural and functional MRI data for the discriminative analysis of SZ patients, which is consistent with previous studies (Dyrba et al., 2015; Zhuang et al., 2019). The abnormalities in brain regions between SZ patients and NCs were primarily in the bilateral prefrontal cortex, right limbic system, left temporal cortex and left motor strip. While the findings on the prefrontal cortex (Janousova et al., 2015; Ou et al., 2015; Zhuang et al., 2019; Webler et al., 2020), limbic system (Shon et al., 2018; Abdolalizadeh et al., 2020; Falakshahi et al., 2020), and temporal cortex (Shu et al., 2012; Ehrlich et al., 2014; Schnack et al., 2014; Koch et al., 2018; Chatterjee et al., 2020) are aligned with previous studies, the discovery of differences in the motor strip has rarely been reported. The motor strip is imperative as the neural hub that participates in perception, action and anticipation in relation to the environment (Schroeder et al., 1994). Thus, abnormalities in the motor strip might elucidate the abnormal conduct behaviors of SZ patients. Moreover, abnormalities were also found in the left prefrontal cortex, right temporal cortex and left temporal cortex between FESZ and CSZ patients, which is consistent with previous findings of the influence of antipsychotic therapies on brain structure and function (Chua et al., 2009; Goghari et al., 2013; Ren et al., 2013; Lesh et al., 2015). Therefore, we hypothesized that the abnormalities in these

brain regions might be derived from the side effects of long term antipsychotic drug intake.

In this study, we also established an integrated model with the stacking technique, which remarkably improved the performance of the integral model. Kalmady et al. (2019) applied stacking technique in discriminative analyses of SZ patients and found that the classification performance (accuracy of 87%) outperformed earlier machine learning models. Similarly, Irandoost et al. also found that the stacking technique for classification of individuals with AD and cognitively normal individuals was better than using one classifier and comparable to the state-of-the-art methods (Irandoost and Asadi, 2019). Consistent with previous studies, our results showed apparent improvements in classification performance after the stacking technique was applied (SZ vs NC: accuracy = 0.88, AUC = 0.92; FESZ vs CSZ: accuracy = 0.86, AUC = 0.80). The advantage of the stacking technique may derive from both the diversity of the level-0 generalizers and the diversity of the atlases, which offer the integral model with more information to learn that reduced the variance (Wolpert, 1992; Ting and Witten, 1999).

LIMITATIONS

There were several limitations in this study. First, the cross validation and the separate test set prevented overfitting and guaranteed model performance for generalization to unseen data, but as a consequence, the accuracies in the FESZ vs CSZ classification (86%) and in the SZ vs NC classification (88%) were still not as satisfactory as in other studies (Iqbal et al., 2017; Liu J. et al., 2017). Similarly, the performance of the model was checked with a limited dataset, because only a modicum of examples had been provided for the model to discover the significant discrepancies between the two groups, especially for the FESZ vs CSZ classification with fewer data. Second, all classifiers applied in this study were traditional classifiers. Recent studies have shown satisfactory classification performance by deep learning methods for psychiatric diseases (Zeng et al., 2018; Matsubara et al., 2019). In future studies, we plan to perform systematic estimations using deep learning methods. Third, more brain atlases of different sizes can be included in the studies (Kalmady et al., 2019). Wu et al. (2018) have used a brain atlas with 1,024 ROIs in the discriminative analysis of SZ patients and achieved high classification performance. Thus, we plan to estimate classification performances based on brain atlases with relatively larger sizes in the future. Moreover, numerous researches using other biological data have also found significant discrepancies between schizophrenia patients and normal controls, including gut microbiota data (Li et al., 2020), blood data (Chan et al., 2014), and electroencephalogram data (Alfimova and Uvarova, 2008). Therefore, we also plan to use multi-biological data on the discriminative analysis for further improvement.

CONCLUSION

In this study, a systematic analysis of classifications with different combinations of brain atlases, classifiers, cross

validation methods and dimensionality reduction algorithms was performed in two classifications (NC vs SZ, FESZ vs CSZ). The performances of the models were analyzed and the weights from the best combination model were used for feature ranking. Further estimation was also performed to provide information indicating the most significant abnormalities in different brain regions. Moreover, an integral model with higher accuracy and AUC was generated with an ensemble learning method. Our findings indicated effects of these factors in constructing effective classifiers for psychiatric diseases and showed that the integrated model has the potential to improve the clinical diagnosis and treatment evaluation of SZ.

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 Ethics Committees of the Affiliated Brain Hospital of Guangzhou Medical University. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

All authors contributed toward data analysis, drafting and critically revising the manuscript, gave final approval of the version to be published, and agreed to be accountable for all aspects of the work.

FUNDING

This work was supported by the National Key Research and Development Program of China (2020YFC2004300, 2020YFC2004301, 2019YFC0118800, 2019YFC0118802, 2019YFC0118804, and 2019YFC0118805), the National Natural Science Foundation of China (31771074 and 81802230), the Key Research and Development Program of Guangdong (2018B030335001, 2020B0101130020, and 2020B0404010002), Guangdong Basic and Applied Basic Research Foundation Outstanding Youth Project (2021B1515020064), the Science and Technology Program of Guangzhou (201807010064, 201803010100, 201903010032, and 202103000032), and Key Laboratory Program of Guangdong Provincial Education Department (2020KSYS001).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2021.697168/full#supplementary-material>

REFERENCES

- Abdolzadeh, A. H., Ostadrahimi, H., Mohajer, B., Darvishi, A., and Abbasi, N. (2020). White matter microstructural properties associated with impaired attention in chronic schizophrenia: a multi-center study. *Psychiatry Res. Neuroimaging* 302:111105. doi: 10.1016/j.pscychres.2020.111105
- Abdulsalam, S. O., Mohammed, A. A., Ajao, J. F., Babatunde, R. S., Ogundokun, R. O., Nnodim, C. T., et al. (2020). "Performance evaluation of ANOVA and RFE algorithms for classifying microarray dataset using SVM," in *Information Systems. EMCS 2020. Lecture Notes in Business Information Processing*, eds M. Themistocleous, M. Papadaki, and M. M. Kamal (Cham: Springer International Publishing), 480–492. doi: 10.1007/978-3-030-63396-7_32
- Albhashish, D., Hammouri, A. I., Braik, M., Atwan, J., and Sahran, S. (2021). Binary biogeography-based optimization based SVM-RFE for feature selection. *Appl. Soft Comput.* 101:107026. doi: 10.1016/j.asoc.2020.107026
- Alfimova, M. V., and Uvarova, L. G. (2008). Changes in EEG spectral power on perception of neutral and emotional words in patients with schizophrenia, their relatives, and healthy subjects from the general population. *Neurosci. Behav. Physiol.* 38, 533–540. doi: 10.1007/s11055-008-9013-6
- Asim, Y., Raza, B., Malik, A. K., Rathore, S., Hussain, L., and Iftikhar, M. A. (2018). A multi-modal, multi-atlas-based approach for Alzheimer detection via machine learning. *Int. J. Imag. Syst. Tech.* 28, 113–123. doi: 10.1002/ima.22263
- Austin, J. (2005). Schizophrenia: an update and review. *J. Genet. Couns.* 14, 329–340. doi: 10.1007/s10897-005-1622-4
- Bejani, M., and Gharavian, D. (2014). Audiovisual emotion recognition using ANOVA feature selection method and multi-classifier neural networks. *Neural Comput. Appl.* 24, 399–412. doi: 10.1007/s00521-012-1228-3
- Bianciardi, M., Fukunaga, M., Gelderen, P. V., Horovitz, S. G., Zwart, J. A. D., and Duyn, J. H. (2009). Modulation of spontaneous fMRI activity in human visual cortex by behavioral state. *Neuroimage* 45, 160–168. doi: 10.1016/j.neuroimage.2008.10.034
- Breiman, L. (1996). Bagging predictors. *Mach. Learn.* 24, 123–140. doi: 10.1007/BF00058655
- Cao, L. J., Chua, K. S., Chong, W. K., Lee, H. P., and Gu, Q. M. (2003). A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine. *Neurocomputing* 55, 321–336. doi: 10.1016/S0925-2312(03)00433-8
- Chan, M. K., Gottschalk, M. G., Haenisch, F., Tomasik, J., Ruland, T., Rahmoune, H., et al. (2014). Applications of blood-based protein biomarker strategies in the study of psychiatric disorders. *Prog. Neurobiol.* 122, 45–72. doi: 10.1016/j.pneurobio.2014.08.002
- Chatterjee, I., Kumar, V., Rana, B., Agarwal, M., and Kumar, N. (2020). Identification of changes in grey matter volume using an evolutionary approach: an MRI study of schizophrenia. *Multimed. Syst.* 26, 383–396. doi: 10.1007/s00530-020-00649-6
- Chen, Z., Yan, T., Wang, E., Jiang, H., Tang, Y., Yu, X., et al. (2020). Detecting abnormal brain regions in schizophrenia using structural MRI via machine learning. *Comput. Intell. Neurosci.* 2020:6405930. doi: 10.1155/2020/6405930
- Chua, S. E., Deng, Y., Chen, E. Y. H., Law, C. W., Chiu, C. P. Y., Cheung, C., et al. (2009). Cambridge journals online—psychological medicine—abstract—early striatal hypertrophy in first-episode psychosis within 3 weeks of initiating antipsychotic drug treatment. *Psychol. Med.* 39, 793–800. doi: 10.1017/S0033291708004212
- Craddock, R., James, G., Holtzheimer, P., Hu, X., and Mayberg, H. (2012). A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Mapp.* 33, 1914–1928. doi: 10.1002/hbm.21333
- de Souza Rodrigues, J., Ribeiro, F. L., Sato, J. R., Mesquita, R. C., and Júnior, C. E. B. (2019). Identifying individuals using fNIRS-based cortical connectomes. *Biomed. Opt. Express* 10:2889. doi: 10.1364/BOE.10.002889
- Deanna, G., Malley, J. D., Brian, W., Liv, C., and Nitin, G. (2012). Using multivariate machine learning methods and structural MRI to classify childhood onset schizophrenia and healthy controls. *Front. Psychiatry* 3:53. doi: 10.3389/fpsy.2012.00053
- Dyrba, M., Grothe, M., Kirste, T., and Teipel, S. J. (2015). Multimodal analysis of functional and structural disconnection in Alzheimer's disease using multiple kernel SVM. *Hum. Brain Mapp.* 36, 2118–2131. doi: 10.1002/hbm.22759
- Efron, B. (1983). Estimating the error rate of a prediction rule: improvement on cross-validation. *J. Am. Stat. Assoc.* 78, 316–331. doi: 10.1080/1983.10477973
- Ehrlich, S., Geisler, D., Yendiki, A., Panneck, P., Roessner, V., Calhoun, V. D., et al. (2014). Associations of white matter integrity and cortical thickness in patients with schizophrenia and healthy controls. *Schizophr. Bull.* 40, 665–674. doi: 10.1093/schbul/sbt056
- Falakshahi, H., Vergara, V. M., Liu, J., MATHALON, D. H., and Plis, S. (2020). Meta-modal information flow: a method for capturing multimodal modular disconnectivity in schizophrenia. *IEEE Trans. Biomed. Eng.* 67, 2572–2584. doi: 10.1109/TBME.2020.2964724
- Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., et al. (2016). The human brainnetome atlas: a new brain atlas based on connectural architecture. *Cereb. Cortex* 26, 3508–3526. doi: 10.1093/cercor/bhw157
- Fernandez-Lozano, C., Fernandez-Blanco, E., Dave, K., Pedreira, N., Gestal, M., Dorado, J., et al. (2014). Improving enzyme regulatory protein classification by means of SVM-RFE feature selection. *Mol. Biosyst.* 10, 1063–1071. doi: 10.1039/c3mb70489k
- Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., et al. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* 18, 1664–1671. doi: 10.1038/nn.4135
- First, M., Gibbon, M., Williams, J., and Spitzer, R. (1997). *User's Guide for the Structured Clinical Interview for DSM-IV Axis II Personality Disorders (SCID-II)*. Washington, DC: American Psychiatric Press, Inc.
- Fisher, R. A. (1936). The use of multiple measurements intaxonomic problems. *Ann. Hum. Genet.* 7, 179–188. doi: 10.1111/j.1469-1809.1936.tb02137.x
- Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Ann. Stat.* 29, 1189–1232. doi: 10.2307/2699986
- Goghari, V. M., Smith, G. N., Honer, W. G., Kopala, L. C., Thornton, A. E., Su, W., et al. (2013). Effects of eight weeks of atypical antipsychotic treatment on middle frontal thickness in drug-naïve first-episode psychosis patients. *Schizophr. Res.* 149, 149–155. doi: 10.1016/j.schres.2013.06.025
- Golland, P., and Fischl, B. (2003). Permutation tests for classification: towards statistical significance in image-based studies. *Inf. Process. Med. Imaging* 18, 330–341. doi: 10.1007/978-3-540-45087-0_28
- Govindarajulu, Z. (1992). Rank correlation methods (5th ed.). *Technometrics* 34:108. doi: 10.1080/00401706.1992.10485252
- Guyon, I., Weston, J., Barnhill, S., and Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Mach. Learn.* 46, 389–422. doi: 10.1023/A:1012487302797
- Hastie, T., Tibshirani, R., and Friedman, J. (2008). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, 2nd Edn. NewYork, NY: Springer.
- Iqbal, Q. M. N., Jooyoung, O., Dongrae, C., Joon, J. H., and Boreom, L. (2017). Multimodal discrimination of schizophrenia using hybrid weighted feature concatenation of brain functional connectivity and anatomical features with an extreme learning machine. *Front. Neuroinform.* 11:59. doi: 10.3389/fninf.2017.00059
- Irandoost, S. A., and Asadi, F. (2019). Classification of brain MRI for Alzheimer's disease detection based on ensemble machine learning. *Iran J. Radiol.* 16:e99157. doi: 10.5812/iranradiol.99157
- Janousova, E., Schwarz, D., and Kasperek, T. (2015). Combining various types of classifiers and features extracted from magnetic resonance imaging data in schizophrenia recognition. *Psychiatry Res. Neuroimaging* 232, 237–249. doi: 10.1016/j.pscychres.2015.03.004
- Jolliffe, I. T. (2005). Principal component analysis. 2nd ed. *Weather* 98, 111–143. doi: 10.1002/0470013192.bsa501
- Junhua, L., Yu, S., Yi, H., Anastasios, B., and Rongjun, Y. (2018). Machine learning technique reveals intrinsic characteristics of schizophrenia: an alternative method. *Brain Imaging Behav.* 13, 1386–1396. doi: 10.1007/s11682-018-9947-4
- Kalmady, S. V., Greiner, R., Agrawal, R., Shivakumar, V., Narayanaswamy, J. C., Brown, M. R. G., et al. (2019). Towards artificial intelligence in mental health by improving schizophrenia prediction with multiple brain parcellation ensemble-learning. *NPJ Schizophr.* 5:2. doi: 10.1038/s41537-018-0070-8
- Karageorgiou, E., Schulz, S. C., Gollub, R. L., Andreasen, N. C., Ho, B. C., Lauriello, J., et al. (2011). Neuropsychological testing and structural magnetic resonance imaging as diagnostic biomarkers early in the course of schizophrenia and related psychoses. *Neuroinformatics* 9, 321–333. doi: 10.1007/s12021-010-9094-6

- Kasperek, T., Thomaz, C. E., Sato, J. R., Schwarz, D., Janousova, E., Marecek, R., et al. (2011). Maximum-uncertainty linear discrimination analysis of first-episode schizophrenia subjects. *Psychiatry Res.* 191, 174–181. doi: 10.1016/j.psychres.2010.09.016
- Kay, S. R., Fiszbein, A., and Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophr. Bull.* 13, 261–276. doi: 10.1093/schbul/13.2.261
- Khondoker, M., Dobson, R., Skirrow, C., Simmons, A., and Stahl, D. (2013). A comparison of machine learning methods for classification using simulation with multiple real data examples from mental health studies. *Stat. Methods Med. Res.* 25, 1804–1823. doi: 10.1177/0962280213502437
- Kim, J., Calhoun, V. D., Shim, E., and Lee, J. H. (2015). Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *Neuroimage* 124, 127–146. doi: 10.1016/j.neuroimage.2015.05.018
- Koch, S. P., Claudia, H., John-Dylan, H., Andreas, H., Florian, S., and Philipp, S. (2018). Diagnostic classification of schizophrenia patients on the basis of regional reward-related fMRI signal patterns. *PLoS One* 10:e0119089. doi: 10.1371/journal.pone.0119089
- Koikkalainen, J., Lötjönen, J., Thurfjell, L., Rueckert, D., Waldemar, G., and Soininen, H. (2011). Multi-template tensor-based morphometry: application to analysis of Alzheimer's disease. *Neuroimage* 56, 1134–1144. doi: 10.1016/j.neuroimage.2011.03.029
- Kriti, Virmani, J., Dey, N., and Kumar, V. (2016). "PCA-PNN and PCA-SVM based CAD systems for breast density classification," in *Applications of Intelligent Optimization in Biology and Medicine. Intelligent Systems Reference Library*, Vol. 96, eds A. E. Hassanien, C. Grosan, and M. Fahmy Tolba (Cham: Springer International Publishing), 159–180. doi: 10.1007/978-3-319-21212-8_7
- Lesh, T. A., Tanase, C., Geib, B. R., Niendam, T. A., Yoon, J. H., Minzenberg, M. J., et al. (2015). A multimodal analysis of antipsychotic effects on brain structure and function in first-episode schizophrenia. *JAMA Psychiatry* 72, 226–234. doi: 10.1001/jamapsychiatry.2014.2178
- Leucht, S., Burkard, T., Henderson, J., Maj, M., and Sartorius, N. (2007). Physical illness and schizophrenia: a review of the literature. *Acta Psychiatr. Scand.* 116, 317–333. doi: 10.1111/j.1600-0447.2007.01095.x
- Li, S., Zhuo, M., Huang, X., Huang, Y., and Wu, K. (2020). Altered gut microbiota associated with symptom severity in schizophrenia. *PeerJ* 8:e9574. doi: 10.7717/peerj.9574
- Li, X., Song, D., Zhang, P., Zhang, Y., Hou, Y., and Hu, B. (2018). Exploring EEG features in cross-subject emotion recognition. *Front. Neurosci. Switz.* 12:162. doi: 10.3389/fnins.2018.00162
- Liu, F., Guo, W., Fouché, J. P., Wang, Y., Wang, W., Ding, J., et al. (2015). Multivariate classification of social anxiety disorder using whole brain functional connectivity. *Brain Struct. Funct.* 220, 101–115. doi: 10.1007/s00429-013-0641-4
- Liu, J., Wang, X., Zhang, X., Pan, Y., Wang, X., and Wang, J. (2017). MMM: classification of schizophrenia using multi-modality multi-atlas feature representation and multi-kernel learning. *Multimed. Tools Appl.* 77, 29651–29667. doi: 10.1007/s11042-017-5470-7
- Liu, Y., Zhang, Y., Lv, L., Wu, R., Zhao, J., and Guo, W. (2017). Abnormal neural activity as a potential biomarker for drug-naïve first-episode adolescent-onset schizophrenia with coherence regional homogeneity and support vector machine analyses. *Schizophr. Res.* 192, 408–415. doi: 10.1016/j.schres.2017.04.028
- Longfei, S., Lubin, W., Hui, S., Guiyu, F., and Dewen, H. (2013). Discriminative analysis of non-linear brain connectivity in schizophrenia: an fMRI Study. *Front. Hum. Neurosci.* 7:702. doi: 10.3389/fnhum.2013.00702
- Lu, X., Yang, Y., Wu, F., Gao, M., Xu, Y., Zhang, Y., et al. (2016). Discriminative analysis of schizophrenia using support vector machine and recursive feature elimination on structural MRI images. *Medicine* 95:e6669. doi: 10.1097/01.md.0000504794.22466.69
- Lu, X., Zhang, Y., Yang, D., Yang, Y., Wu, F., Ning, Y., et al. (2018). Analysis of first-episode and chronic schizophrenia using multi-modal magnetic resonance imaging. *Eur. Rev. Med. Pharmacol. Sci.* 22, 6422–6435. doi: 10.26355/eurrev_201810_16055
- Matsubara, T., Tashiro, T., and Uehara, K. (2019). Deep neural generative model of functional MRI images for psychiatric disorder diagnosis. *IEEE Trans. Biomed. Eng.* 66, 2768–2779. doi: 10.1109/TBME.2019.2895663
- Min, R., Wu, G., Cheng, J., Wang, Q., Shen, D., and The Alzheimer's Disease Neuroimaging Initiative (2014). Multi-atlas based representations for Alzheimer's disease diagnosis. *Hum Brain Mapp* 35, 5052–5070. doi: 10.1002/hbm.22531
- Neter, J., Wasserman, W., and Kutner, M. H. (1996). Applied linear statistical models. *Technometrics* 39, 880–880. doi: 10.2307/1271154
- Ota, M., Sato, N., Ishikawa, M., Hori, H., Sasayama, D., Hattori, K., et al. (2012). Discrimination of female schizophrenia patients from healthy women using multiple structural brain measures obtained with voxel-based morphometry. *Psychiatry Clin. Neurosci.* 66, 611–617. doi: 10.1111/j.1440-1819.2012.02397.x
- Ou, J., Xie, L., Li, X., Zhu, D., Terry, D. P., Puente, A. N., et al. (2015). Atomic connectomics signatures for characterization and differentiation of mild cognitive impairment. *Brain Imaging Behav.* 9, 663–677. doi: 10.1007/s11682-014-9320-1
- Peng, C. Y. J., and Ingersoll, L. G. M. (2002). An introduction to logistic regression analysis and reporting. *J. Educ. Res.* 96, 3–14. doi: 10.2307/27542407
- Raymond, S., Joaquim, R., Canales-Rodríguez, E. J., Aleix, S., Salvador, S., Goikolea, J. M., et al. (2017). Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. *PLoS One* 12:e0175683. doi: 10.1371/journal.pone.0175683
- Ren, W., Lui, S., Deng, W., Li, F., Li, M., Huang, X., et al. (2013). Anatomical and functional brain abnormalities in drug-naïve first-episode schizophrenia. *Am. J. Psychiatry* 170, 1308–1316. doi: 10.1176/appi.ajp.2013.12091148
- Sain, R. S. (1997). The nature of statistical learning theory. *Technometrics* 38, 409–409. doi: 10.1080/00401706.1996.10484565
- Sarah, G., and Media, O. (2017). *Introduction to Machine Learning with Python*. Sebastopol, CA: O'Reilly Media, Inc.
- Schnack, H. G., Nieuwenhuis, M., Van Haren, N. E. M., Abramovic, L., Scheewe, T. W., Brouwer, R. M., et al. (2014). Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *Neuroimage* 84, 299–306. doi: 10.1016/j.neuroimage.2013.08.053
- Schroeder, J., Buchsbaum, M. S., Siegel, B. V., Geider, F. J., Haier, R. J., Lohr, J., et al. (1994). Patterns of cortical activity in schizophrenia. *Psychol. Med.* 24, 947–955. doi: 10.1017/S0033291700029032
- Sheikhan, M., Bejani, M., and Gharavian, D. (2013). Modular neural-SVM scheme for speech emotion recognition using ANOVA feature selection method. *Neural Comput. Appl.* 23, 215–227. doi: 10.1007/s00521-012-0814-8
- Shen, X., Tokoglu, F., Papademetris, X., and Constable, R. T. (2013). Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage* 82, 403–415. doi: 10.1016/j.neuroimage.2013.05.081
- Shon, S., Yoon, W., Kim, H., Joo, S. W., Kim, Y., and Lee, J. (2018). Deterioration in global organization of structural brain networks in schizophrenia: a diffusion MRI tractography study. *Front. Psychiatry* 9:272. doi: 10.3389/fpsy.2018.00272
- Shu, N., Liang, Y., Li, H., Zhang, J., Li, X., Wang, L., et al. (2012). Disrupted topological organization in white matter structural networks in amnesic mild cognitive impairment: relationship to subtype. *Radiology* 265, 518–527. doi: 10.1148/radiol.12112361
- Ting, K. M., and Witten, I. H. (1999). Issues in stacked generalization. *J. Artif. Intell. Res.* 10, 271–289. doi: 10.1613/jair.594
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Valizadeh, S. A., Liem, F., Mérillat, S., Hänggi, J., and Jäncke, L. (2018). Identification of individual subjects on the basis of their brain anatomical features. *Sci. Rep. U. K.* 8:5611. doi: 10.1038/s41598-018-23696-6
- Van Essen, D. C., Glasser, M. F., Dierker, D. L., Harwell, J., and Coalson, T. (2012). Parcellations and hemispheric asymmetries of human cerebral cortex analyzed on surface-based atlases. *Cereb. Cortex* 22, 2241–2262. doi: 10.1093/cercor/bhr291
- Van Tol, M., Van Der Meer, L., Bruggeman, R., Modinos, G., Knegtering, H., and Aleman, A. (2014). Voxel-based gray and white matter morphometry correlates

- of hallucinations in schizophrenia: the superior temporal gyrus does not stand alone. *Neuroimage* 4, 249–257. doi: 10.1016/j.neuroimage.2003.12.030
- Watanabe, T., Kessler, D., Scott, C., Angstadt, M., and Sripada, C. (2014). Disease prediction based on functional connectomes using a scalable and spatially-informed support vector machine. *Neuroimage* 96, 183–202. doi: 10.1016/j.neuroimage.2014.03.067
- Webler, R. D., Hamady, C., Molnar, C., Johnson, K., Bonilha, L., Anderson, B. S., et al. (2020). Decreased interhemispheric connectivity and increased cortical excitability in unmedicated schizophrenia: a prefrontal interleaved TMS fMRI study. *Brain Stimul.* 13, 1467–1475. doi: 10.1016/j.brs.2020.06.017
- Wolpert, D. H. (1992). Stacked generalization. *Neural Netw.* 5, 241–259. doi: 10.1016/S0893-6080(05)80023-1
- Wu, F., Zhang, Y., Yang, Y., Lu, X., Fang, Z., Huang, J., et al. (2018). Structural and functional brain abnormalities in drug-naïve, first-episode, and chronic patients with schizophrenia: a multimodal MRI study. *Neuropsychiatr. Dis. Treat.* 14, 2889–2904. doi: 10.2147/ndt.s174356
- Xia, M., Wang, J., and He, Y. (2013). BrainNet viewer: a network visualization tool for human brain connectomics. *PLoS One* 8:e68910. doi: 10.1371/journal.pone.0068910
- Xiao, Y., Yan, Z., Zhao, Y., Tao, B., and Lui, S. (2017). Support vector machine-based classification of first episode drug-naïve schizophrenia patients and healthy controls using structural MRI. *Schizophr. Res.* 214, 11–17. doi: 10.1016/j.schres.2017.11.037
- Xue, Y., Zhang, L., Wang, B., Zhang, Z., and Li, F. (2018). Nonlinear feature selection using Gaussian kernel SVM-RFE for fault diagnosis. *Appl. Intell.* 48, 3306–3331. doi: 10.1007/s10489-018-1140-3
- Yu-Feng, Z., Yong, H., Chao-Zhe, Z., Qing-Jiu, C., Man-Qiu, S., Meng, L., et al. (2007). Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain Dev.* 29, 83–91. doi: 10.1016/j.braindev.2006.07.002
- 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
- Zeng, L., Wang, H., Hu, P., Yang, B., Pu, W., Shen, H., et al. (2018). Multi-site diagnostic classification of schizophrenia using discriminant deep learning with functional connectivity MRI. *Ebiomedicine* 30, 74–85. doi: 10.1016/j.ebiom.2018.03.017
- Zhang, Z. (2016). Introduction to machine learning: k-nearest neighbors. *Ann. Transl. Med.* 4:218. doi: 10.21037/atm.2016.03.37
- Zhuang, H., Liu, R., Wu, C., Meng, Z., and Wang, D. (2019). Multimodal classification of drug-naïve first-episode schizophrenia combining anatomical, diffusion and resting state functional resonance imaging. *Neurosci. Lett.* 705, 87–93. doi: 10.1016/j.neulet.2019.04.039

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 Zang, Huang, Kong, Lei, Ke, Li, Zhou, Xiong, Li, Chen, Li, Xiang, Ning, Wu and Wu. 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.