frontiers in Bioengineering and Biotechnology

ORIGINAL RESEARCH
published: 15 January 2020
doi: 10.3389/fbioe.2019.00479

Schizophrenia Identification Using Multi-View Graph Measures of Functional Brain Networks

Yizhen Xiang[1], *Jianxin Wang*[1,2], *Guanxin Tan*[1], *Fang-Xiang Wu*[3] *and Jin Liu*[1]*

[1] School of Computer Science and Engineering, Central South University, Changsha, China, [2] Hunan Provincial Key Lab on Bioinformatics, Central South University, Changsha, China, [3] Division of Biomedical Engineering and Department of Mechanical Engineering, University of Saskatchewan, Saskatoon, SK, Canada

Schizophrenia (SZ) is a functional mental disorder that seriously affects the social life of patients. Therefore, accurate diagnosis of SZ has raised extensive attention of researchers. At present, study of brain network based on resting-state functional magnetic resonance imaging (rs-fMRI) has provided promising results for SZ identification by studying functional network alteration. However, previous studies based on brain network analysis are not very effective for SZ identification. Therefore, we propose an improved SZ identification method using multi-view graph measures of functional brain networks. Firstly, we construct an individual functional connectivity network based on Brainnetome atlas for each subject. Then, multi-view graph measures are calculated by the brain network analysis method as feature representations. Next, in order to consider the relationships between measures within the same brain region in feature selection, multi-view measures are grouped according to the corresponding regions and Sparse Group Lasso is applied to identify discriminative features based on this feature grouping structure. Finally, a support vector machine (SVM) classifier is employed to perform SZ identification task. To evaluate our proposed method, computational experiments are conducted on 145 subjects (71 schizophrenic patients and 74 healthy controls) using a leave-one-out cross-validation (LOOCV) scheme. The results show that our proposed method can obtain an accuracy of 93.10% for SZ identification. By comparison, our method is more effective for SZ identification than some existing methods.

Keywords: Schizophrenia identification, fMRI, functional brain networks, multi-view graph measures, SVM

OPEN ACCESS

Edited by:
Fa Zhang,
Institute of Computing Technology (CAS), China

Reviewed by:
Leyi Wei,
Tianjin University, China
Renmin Han,
Shandong University, China

***Correspondence:**
Jin Liu
liujin06@csu.edu.cn

Specialty section:
This article was submitted to Bioinformatics and Computational Biology, a section of the journal Frontiers in Bioengineering and Biotechnology

Received: *05 September 2019*
Accepted: *23 December 2019*
Published: *15 January 2020*

Citation:
Xiang Y, Wang J, Tan G, Wu F-X and Liu J (2020) Schizophrenia Identification Using Multi-View Graph Measures of Functional Brain Networks. Front. Bioeng. Biotechnol. 7:479. doi: 10.3389/fbioe.2019.00479

1. INTRODUCTION

Schizophrenia (SZ) is a functional mental disorder which caused by genetic factors and environmental effects. Patients with SZ (SZs) share some common symptoms which include depression, hallucinations, cognitive dysfunction and disorganized thinking (Marín, 2012). Impairments of this disorder cover multiple cognitive areas, including memory (He et al., 2012), attention and executive function (Heinrichs and Zakzanis, 1998). One percent of the population is affected by the serious psychiatric disease worldwide (Ripke et al., 2013). The clinical diagnosis of SZ relies mainly on mental state examination rather than any biomarker (Arbabshirani et al., 2013; Liu et al., 2017d) since the cause and mechanism of the disease are not clearly revealed. However, this diagnosis method is usually subjective and not completely effective. Therefore, it is urgent to find an objective method to realize the automatic diagnosis of SZ and improve the accuracy of recognition.

Nowadays, Magnetic resonance imaging technology has been widely used in various studies related to brain disease diagnosis (Nieuwenhuis et al., 2012; Liu et al., 2016, 2017b,c, 2018a; Yang and Wang, 2018). Since SZ is reported to be a functional disease, functional magnetic resonance imaging (fMRI) is increasingly used to study brain dysfunction in patients with mental illness (Castro et al., 2011; Huang et al., 2018; Liu et al., 2018b; Moghimi et al., 2018; Chen et al., 2019). In addition, fMRI provides a database for functional analysis of these brain diseases owing to it's massive spatial and temporal information.

In recent years, the number of neurobiological literatures using fMRI to study SZ disease has increased significantly. fMRI is usually applied to discover anomalous patterns present in activation maps [i.e., Regional Homogeneity (REHO), Amplitude of Low Frequency Fluctuations (ALFF), fractional Amplitude of Low Frequency Fluctuations (FALFF)] (Guo et al., 2014; Chyzhyk et al., 2015; Huang et al., 2018) of SZ. These activation maps are widely used as potential clinical biomarkers for the diagnosis of SZ. For example, Huang et al. (2018) used tree-guided group sparse learning method to perform feature selection on fALFF data in multi-frequency bands, and then used multi-kernel learning (MKL) method to achieve an accuracy of 91.10% on 34 subjects. Chyzhyk et al. (2015) combined these activation maps by using extreme learning machines and successfully distinguished SZs from healthy controls (HCs). However, these methods focus on the voxel-wise information in these maps rather than the connectivity between regions of interest (ROIs).

Functional connectivity has been reported to analyze the differences in the functional organization of brain networks between patients and HCs (Lynall et al., 2010; Pettersson-Yeo et al., 2011). Functional connectivity networks are usually derived from fMRI data (Van Den Heuvel and Pol, 2010; Craddock et al., 2013). Nodes of a functional brain network could be the voxels of fMRI data, ROIs defined by brain atlas or the discrete regions with similar size by randomly parcellating the brain (Fornito et al., 2013). Links of a functional brain network could be determined by the correlations estimated from time courses between pairs of nodes (Liu et al., 2017a). For example, Yu et al. (2015) created functional brain network using group ICA and Pearson correlation coefficient, and they found the new evidence about altered dynamic brain graphs in SZ. Abraham et al. (2017) investigated the most predictive biomarkers for Autism spectrum disorders (ASD) by building participant-specific connectomes from functionally-defined brain areas. For these methods, the connections between all pairs of nodes in a brain network are employed as features, but the topological measures of connectivity networks are not considered.

To quantitatively analyze functional brain networks, graph theoretical analysis is employed for investigating the topological organization of functional connectivity (Anderson and Cohen, 2013; Brier et al., 2014). The most commonly used graph measures include betweenness centrality, degree, local efficiency, participation coefficient, average clustering coefficient, average path length, global efficiency, and small-worldness (Liu et al., 2017a). These topological measures have been applied in the brain disease classifications (Cheng et al., 2015; Khazaee et al., 2015, 2017; Moghimi et al., 2018). For example, Moghimi et al. (2018) calculated a set of 25 graph measures including global and local measures for each subject and obtained a classification accuracy of 80% with a double-cross validation scheme. Cheng et al. (2015) achieved an accuracy of 79% by using betweenness centrality measure in SZ identification, and they found that changes in functional hubs were associated with SZ. Overall, these methods using graph measures for SZ identification have not achieved a good classification performance.

In this paper, we propose an improved method based on multi-view graph measures to identify SZs from HCs. Functional brain networks are constructed based on fMRI scans. Nodes of functional brain network are brain regions parcellated with the Brainnetome atlas (Fan et al., 2016), and edges of functional brain networks are determined by Pearson's correlation coefficients. Five local graph measures are calculated from functional brain networks by graph theoretical approach as features. The five local graph measures include betweenness centrality, nodal clustering coefficient, local efficiency, degree and participation coefficient. In order to consider the relationship of features within the same region, firstly we need to group graph measures according to brain regions defined by Brainnetome atlas. Then Sparse Group Lasso feature selection method is employed to select the most important regions as well as discriminative features within the selected regions. Finally, support vector machine (SVM) is trained for SZ identification. Our experiments are conducted on 145 samples with fMRI data, including 74 HCs and 71 SZs. Our proposed method achieves a mean classification accuracy of 93.10% using a leave-one-out cross-validation (LOOCV) scheme. The overall framework of our proposed method is shown in **Figure 1**, which consists of four main components include image preprocessing, feature representation, feature selection, and classification with SVM classifier. The code for this classification framework is available for download at https://github.com/xyzxzj/SZClassification.

2. MATERIALS AND METHODS

2.1. Subject Description and Image Preprocessing

The data involved in this study is collected by the Center for Biomedical Research Excellence (COBRE). COBRE[1] dataset consists of 148 subjects with functional and anatomical MRI data. 74 HCs and 71 SZs of the dataset are employed for our subsequent experiments owing to the class labels of the other three subjects are not given. During the scan, all participants are asked to remain relaxed and keep their eyes open. A brief summary of demographic information of subjects is listed in **Table 1**.

All of the fMRI images are preprocessed by using Data Processing & Analysis for Brain Imaging (DPABI) (Yan et al., 2016). The preprocessing procedure is as follows: the first 10 volumes of functional runs are removed owing to the fMRI signal instability. Then, the rest volumes are performed slice time correction, head-motion correction, and co-registration of T1-weighted MRI images and fMRI images. After that, the fMRI images are normalized to Montreal Neurological Institute (MNI)

[1]http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html

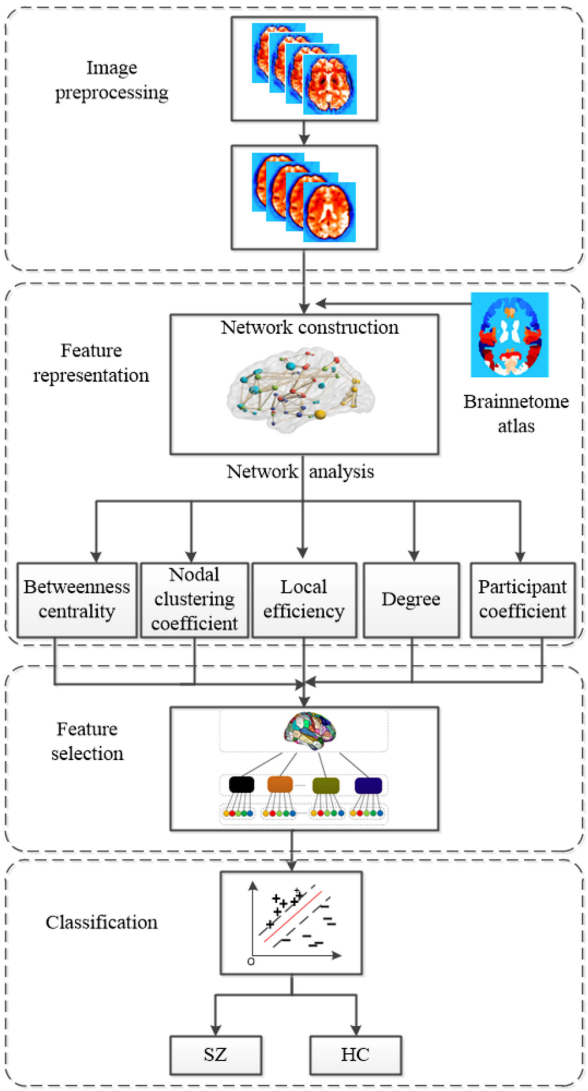

FIGURE 1 | The overall framework of our proposed method using multi-view graph measures of functional brain network for SZ/HC classification.

space and resampled to $3 \times 3 \times 3\,mm^3$ voxels. Smooth (4-mm FWHM) and band-pass filter (0.01–0.1Hz) are applied to the images which are transformed to MNI space.

In order to construct time series matrices for all subjects, first all brain images are parcellated into 246 regions by registering images to the Brainnetome atlas after fMRI data preprocessing. Then we extract the averaged time series for each of 246 brain regions for each subject. The time series of each brain region is derived from averaging fMRI signals of all voxels within the region. Finally, a time series matrix consists of 246 regional time series.

2.2. Feature Representation

2.2.1. Brain Network Construction

A network is composed of a collection of nodes and links. It can be described as a graph $G = (V, E)$, where V denotes the set of nodes and E is the set of links. There are four types of network topology, including weighted undirected, weighted directed, binary undirected and binary directed. In this study, the functional connectivity network is represented by an weighted undirected graph. The nodes in functional connectivity network usually are defined by brain regions, and links can represent temporal correlation in activity between pairs of nodes. Given a time series matrix, we can construct a functional connectivity network by calculating Pearson correlation coefficients (Pedersen et al., 2018) between signals of all pairs of regions. The generated functional brain network has $246 \times (246 - 1)/2 = 30,315$ weighted edges under the condition of 246 brain regions and the strength of each edge is the Pearson correlation coefficient between a pair of connected nodes.

TABLE 1 | Demographic information of 145 subjects from COBRE dataset.

Type	Number	Age	Gender (M/F)
SZ	71	38.1 ± 13.9	57/14
HC	74	35.8 ± 11.5	51/23

2.2.2. Brain Network Analysis

A great deal of functional connections in the network may lead to feature redundancy. A threshold t is employed in the dense network to keep a certain proportion of edges with the highest correlation. Graph-theoretic measures can quantify topological organization of network. Thus, we can extract some measures which can characterize the global or local functional connectivity from the threshold network. We compute 5 local graph measures using brain network analysis as feature representations, including degree, betweenness centrality, nodal clustering coefficient, local efficiency, and participation coefficient.

Degree is the most fundamental and important measure to characterize the centrality of nodes. In general, nodes with a higher degree are more important in networks. Betweenness centrality can also reflect the centrality of nodes. The betweenness centrality of a brain region can measure its ability on information transmission. Nodal clustering coefficient represents the possibility that any two neighbors of a given node are also neighbors of each other. It measures the ability of the node on functional segregation. Local efficiency measures the efficiency of a subnetwork formed by a given node and all its direct neighbors to transfer information. Local efficiency is related to the shortest path length of the node, the shorter the shortest path length, the greater the local efficiency of the node, the faster the information transmission in the subnetwork. Participation coefficient of a node measures its diversity of intermodular interconnections. The nodes with low participation coefficient but high degree in the module are regarded as provincial hubs, it indicates that the hubs are likely to have a great impact on the modular segregation. These five local measures play an important role in information exchange of functional networks. They can be calculated as follow:

$$K(i) = \sum_{j \in N} a_{ij} \tag{1}$$

$$B(i) = \frac{1}{(N-1)(N-1)} \sum_{m \neq j \neq i} \frac{n_{mj}(i)}{n_{mj}} \quad (2)$$

$$C(i) = \frac{2sw_i}{(K_i(K_i-1))} \quad (3)$$

$$E_{loc}(i) = \frac{1}{N_{G_i}(N_{G_i}-1)} \sum_{j \neq h \neq G_i} \frac{1}{l_{jh}} \quad (4)$$

$$PC(i) = 1 - \sum_{m \in M} \left(\frac{k_i(m)}{k_i}\right)^2 \quad (5)$$

where $K(i)$, $B(i)$, $C(i)$, $E_{loc}(i)$, and $PC(i)$ are the degree, betweenness centrality, clustering coefficient, local efficiency, and participation coefficient of node i, respectively. N is the number of nodes in a network, $a_{ij} = 1$ if node i and node j are connected, $a_{ij} = 0$ otherwise; $n_{mj}(i)$ is the number of shortest paths between m and j that pass through node i, and n_{mj} is the number of shortest paths between m and j; sw_i is the sum of the weights of all the connected edges between the neighbors of node i; G_i is the subnetwork that contains node i and its all direct neighbors, N_{G_i} is the number of nodes in the subnetwork G_i, l_{jh} is the length of shortest path between node j and node h in the subgraph; M denotes the set of modules, k_i is determined as the number of links between i and the nodes within module m.

In this study, we adopt the Brain Connectivity Toolbox (http://www.brain-connectivity-toolbox.net) (Rubinov and Sporns, 2010) to calculate these five local graph measures. For each local graph measure (gm), we compute 246 values corresponding to the 246 brain regions. Therefore, the dimension of the final feature vector for each subject is 1,230.

2.3. Feature Selection

The raw feature matrices have high dimension, multiple redundancy and multi-noise characteristics. Thus, applying a suitable feature selection algorithm to identify features related to SZ/HC identification and remove unnecessary information appears especially important. Least absolute shrinkage and selection operator (Lasso) (Chan et al., 2015) is widely used in various areas due to the very low data requirements. In addition, lasso can filter variables and reduce the complexity of the model. It aims to select the most important features from dense data matrix by using l_1 norm constraint. The optimization problem can be formulated as follow:

$$\min_{\alpha} ||y - X\alpha||^2 + \lambda_1 ||\alpha||_1 \quad (6)$$

where X denotes an $n \times p$ feature matrix, and n is the number of subjects, p represents the dimension of a feature vector. y is defined as a class label, α is a coefficient vector, and λ_1 is a regularization parameter.

Graph measures within the same region usually have a certain correlation. However, Lasso has not consider the relationship between graph measures derived in the same brain region. Hence we use the priori information of brain regions to group measures and then perform feature selection based on this feature grouping structure. Group Lasso (GLasso) (Yuan and Lin, 2006), a group variable selection method, is the extension of Lasso. It can select

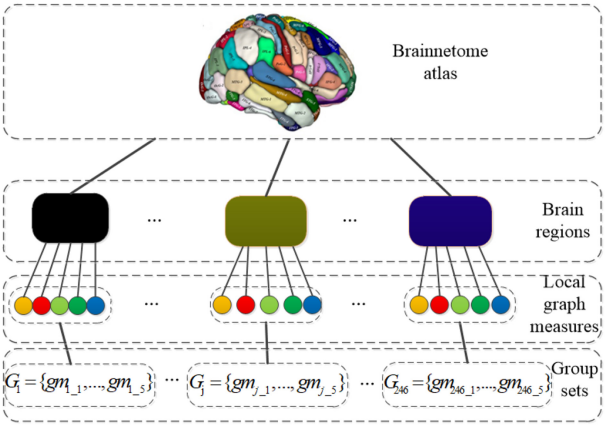

FIGURE 2 | The grouping structure: the nodes in the third layer represent local graph measures and the blocks in the second layer represent brain regions; $G_j = \{gm_{j\_1}, ..., gm_{j\_5}\}$ is a group set which consists of 5 local graph measures calculated for j_th region.

the most important groups by grouping all the variables and penalizing the l_2 norm of each group. The effect is that we can eliminate the entire set of coefficients into zero at the same time and then this set of features are excluded. The objective function of GLasso is as follow:

$$\min_{\alpha} ||y - X\alpha||^2 + \lambda_2 \sum_{j=1}^{M} w_j ||\alpha_{G_j}||_2 \quad (7)$$

where α_{G_j} denotes the set of coefficients of all features in the group G_j, w_j is a weight for group G_j.

Actually, there are also many redundant features in the important groups selected by GLasso. It is necessary to perform another feature selection to choose the most important features from these selected groups. Sparse Group Lasso (SGLasso) (Liu et al., 2009) is introduced to select the most significant groups as well as the discriminative features within the selected groups by adding l_1 and l_2 penalties. The objective function of the SGLasso can be written as:

$$\min_{\alpha} ||y - X\alpha||^2 + \lambda_1 ||\alpha||_1 + \lambda_2 \sum_{j=1}^{M} w_j ||\alpha_{G_j}||_2 \quad (8)$$

Before performing SGLasso, 1230-dimensional feature vector for each subject is grouped as $G = \{G_1, ..., G_j, ..., G_M\}$ according the brain regions defined by Brainnetome atlas. M is the number of groups. $G_j = \{gm_{j\_1}, gm_{j\_2}, gm_{j\_3}, gm_{j\_4}, gm_{j\_5}\}$ is a group consists of 5 local graph measures calculated for j_th region. The grouping structure is shown in **Figure 2**. In addition, z-score transformation is used to normalize the feature matrix before feature selection. It is worth noting that, after feature selection, those features are kept with corresponding regression coefficients greater than the mean value of absolute values of all elements in coefficient vectors.

2.4. Classification

SVM (Chang and Lin, 2011) is widely applied in various fields such as natural language processing, target detection, pattern classification due to its good performance as a supervised machine learning approach. The choice of SVM kernel functions is critical to their performance. In this study, we choose the linear kernel SVM (LSVM) to identify SZs from HCs. Linear kernel is mainly used in linear separability cases, and the dimension of the feature space and input space is the same. It performs good classification in most linear separable problems owing to the less parameters and fast calculation. The formulation of SVM model and linear kernel function are as follows:

$$\begin{aligned} &\max_{\lambda} -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\lambda_i\lambda_j y_i y_j K(x_i,x_j)+\sum_{i=1}^{N}\lambda_i \\ &s.t.\sum_{i=1}^{N}\lambda_i y_i = 0 \\ &0 \le \lambda_i \le C, i = 1,2,...,N \end{aligned} \tag{9}$$

$$K(x_i,x_j) =< x_i,x_j > \tag{10}$$

where λ is the Lagrange multiplier, N is the number of samples, x_i represents the feature vector of the i-th sample, and y_i is the label corresponding to x_i, $K(.,.)$ denotes the kernel function, C is determined as the soft margin parameter.

After feature selection, the optimal feature set $X = \{x_1,...,x_i,...,x_n\}$ is used as the input to SVM classifier, i = 1,..., n. Giving a test subject x, the trained SVM will predict its label based on a decision function P(x) as follows:

$$P(x) = sign(\sum_{i=1}^{N}\lambda_i y_i K(x_i,x)) \tag{11}$$

3. EXPERIMENTS AND RESULTS

3.1. Experiment Settings

In our study, the classification performance of our proposed method is estimated by adopting LOOCV scheme. LOOCV scheme is not affected by the random sample partitioning because n samples are only divided into n subsets in a unique way, each subset contains one sample. Each subset will be tested as a test data in turn while remaining subjects as the training data. In addition, we usually adopt the LIBSVM library (Chang and Lin, 2011) to solve SVM classification. We further calculate classification accuracy (ACC), sensitivity (SEN), specificity (SPE) to measure the performance of the method. These three metrics can be written as follows:

$$ACC = \frac{TP+TN}{TP+FP+FN+TN} \tag{12}$$

$$SPE = \frac{TN}{TN+FP} \tag{13}$$

$$SEN = \frac{TP}{TP+FN} \tag{14}$$

where true positive (TP), true negative (TN), false negative (FN), and false positive (FP) are defined as the number of correctly classified SZs, HCs and misidentified SZs, HCs, respectively.

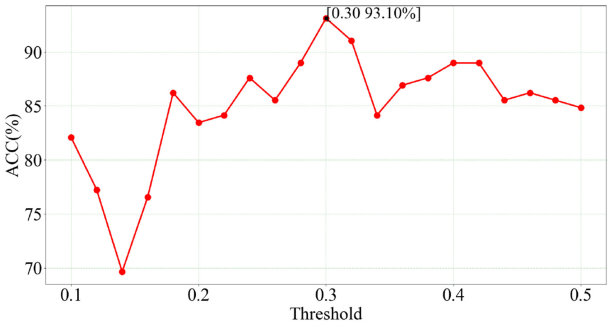

FIGURE 3 | Classification accuracies for SZ identification based on different network thresholds.

In addition, the area under receiver operating characteristic (ROC) curve (AUC) is also used to evaluate overall classification performance of our method.

At the stage of feature representation, we set $t = [0.1, 0.12, ..., 0.48, 0.5]$ to represent a collection of threshold values from 0.1 to 0.5 by the step of 0.02, and then calculate the 5 local graph measures at these 21 thresholds. The two regularization parameters for SGLasso are set as $\lambda_1 = [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]$ and $\lambda_2 = [0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1.0]$, which are optimized with the grid search algorithm.

3.2. Identification Performance for SZ

We use LSVM to perform SZ/HC classification on the optimal feature set obtained from feature selection of SGLasso at each of 21 thresholds. The classification results corresponding to 21 thresholds are showed in **Figure 3**.

According to **Figure 3**, we can see that the best accuracy (93.10%) is achieved at $t = 0.30$. Furthermore, the classification accuracies at these 21 thresholds are all higher than 70%. In addition, the number of selected features is 55 and SEN, SPE, AUC values are 92.96%, 93.24%, 0.950, respectively. The experimental results indicate that the feature combination of five local measures extracted at $t = 0.30$ has a relatively strong correlation with SZ identification.

4. DISCUSSION

4.1. Comparison With Different Feature Selection Methods

In order to demonstrate the SGLasso method is more effective than the common feature selection methods based on these five local measures for SZ classification, we compare four feature selection methods. The first one is t-test which is the one of the most basic feature selection method and the most critical part of this method is selecting features based on the p-value (i.e., 0.05). The rest methods are Lasso, GLasso and Elastic Net (Enet). These three methods are based on linear sparse models. GLasso and Enet are the extension of Lasso. GLasso is used to solve l_1/l_q-norm regularized problem. Enet is used for the situations

TABLE 2 | Classification with different feature selection methods.

Methods	Number of selected features	ACC (%)	SEN (%)	SPE (%)
t-test	153	78.62	80.28	77.03
Lasso	123	83.45	88.73	78.38
GLasso	225	86.21	85.92	86.49
ENet	64	85.52	84.51	86.19
SGLasso	55	93.10	92.96	93.24

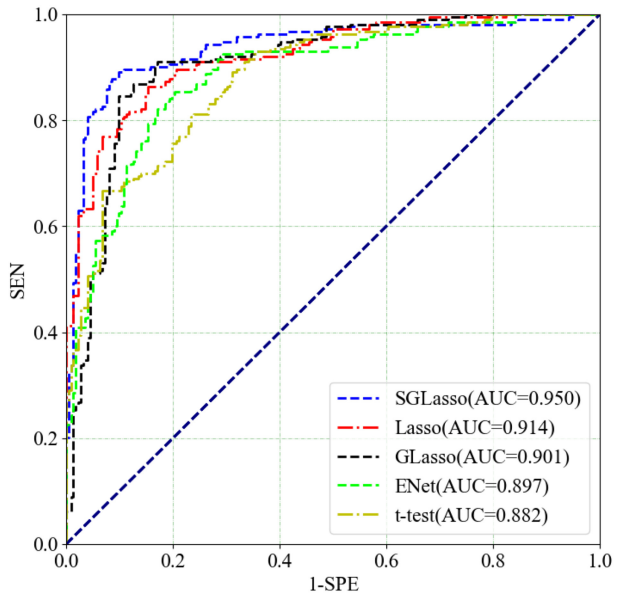

FIGURE 4 | ROC curves for SZ/HC classification for different feature selection methods.

where features are related to each other and always produce valid solution.

These four feature selection methods perform the same experimental procedure as SGLasso for the sake of fairness. It's worth noting that the five local graph measures are extracted at the threshold of 0.30. **Table 2** shows the experimental results of the above mentioned four methods and SGLasso feature selection method. As we can see that SGLasso method selects the least features (55) but achieves the best ACC (93.10%), SEN (92.96%), SPE (93.24%). The ROC curves for SZ/HC classification for different feature selection methods as shown in **Figure 4**. We notice that SGLasso achieves the highest AUC (0.950) than other four feature selection methods. Experimental result shows that considering within- and between- group sparsity is likely helpful for selecting significant features that are effective for SZ identification.

4.2. Comparison With Different Classifiers

In order to prove that LSVM is optimal to conduct classification in this context, a series of comparative experiments using several SVMs with different kernels including Radial Basis Function kernel (RBF), Ploynomial kernel (Poly), Sigmoid kernel (Sigm) under the same condition as the LSVM have been performed. These SVMs are denoted as RBF-SVM, Poly-SVM, Sigm-SVM, respectively. The experimental results of SVMs with different kernels are shown in **Table 3**. It is worth mentioning that bold text indicates that the best result is obtained on a certain evaluation metric.

TABLE 3 | Comparison with other SVMs using different kernels.

Methods	ACC (%)	SEN (%)	SPE (%)	AUC
RBF-SVM	80.00	76.06	83.78	0.8601
Poly-SVM	82.07	77.46	86.49	0.8506
Sigm-SVM	87.59	83.10	91.89	0.9393
LSVM	**93.10**	**92.96**	**93.24**	**0.950**

Bold text indicates that the best result is obtained on a certain evaluation metric.

TABLE 4 | Comparison with other commonly used classifiers.

Methods	ACC (%)	SEN (%)	SPE (%)	AUC
KNN	82.07	74.65	89.19	0.7912
RForest	77.93	74.65	81.08	0.8378
NBayes	84.83	83.10	86.49	0.9069
LDA	90.34	87.32	93.24	0.9418
LSVM	**93.10**	**92.96**	**93.24**	**0.950**

Bold text indicates that the best result is obtained on a certain evaluation metric.

In addition, we also compare four commonly used classifiers, such as k-nearest neighbors (KNN), Random Forest (RForest), NaiveBayes (NBayes), and Linear Discriminant Analysis (LDA). These classifiers are all implemented on the platform of Matlab2016a. We evaluate the performance of the above four classifiers under the same conditions as LSVM. The experimental results of these five classifiers are shown in **Table 4**. As can be seen from **Tables 3**, **4**, LSVM can achieve the best classification performance than other classifiers.

4.3. Regularization Parameter Selection

The regularization parameters of SGLasso have a great influence on the results of feature selection. Using different regularization parameters, the selected features are also different. It affects not only the feature dimension, but also the final classification performance. Therefore, selecting the appropriate regularization parameters can improve the efficiency of SGLasso method and obtain more effective features associated with the labels.

The two regularization parameters of SGLasso are λ_1 and λ_2. λ_1 is used to control the model sparseness, and λ_2 can control the sparse constraint of each feature group. We use the grid search algorithm to find the optimal combination of regularization parameters. **Figure 5** shows the classification results using different combination of λ_1, λ_2. According to **Figure 5**, when the parameter combination is (λ_1=9, λ_2=0.1), the features obtained from SGLasso feature selection method are the most effective for SZ/HC classification.

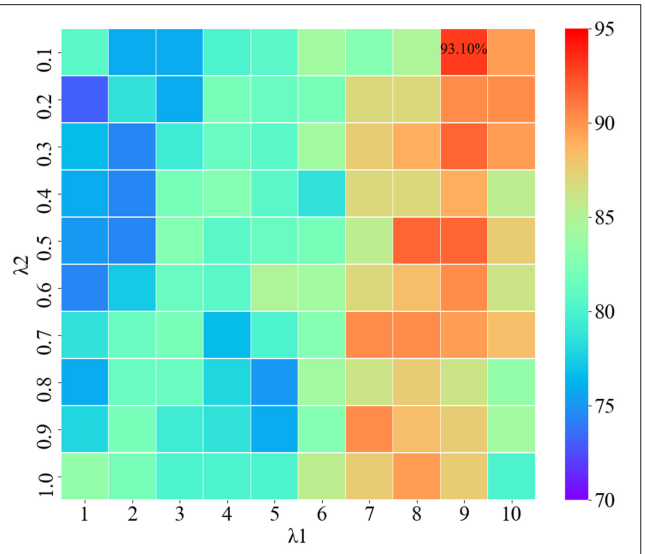

FIGURE 5 | Classification results using different combination of λ_1,λ_2.

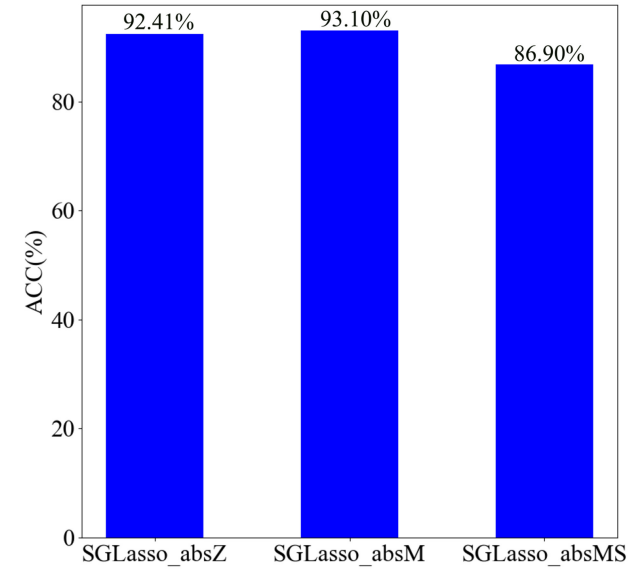

FIGURE 6 | Classification results using three different regression coefficient selection strategies.

4.4. Regression Coefficient Selection

In general, the non-zero elements in the coefficient vector α generated from the SGLasso feature selection algorithm indicate that the corresponding features are selected. In order to retain the least but most informative features according to α, we test the impact of the three coefficient selection strategies on classification performance. We named these three strategies as SGLasso_absZ, SGLasso_absM, and SGLasso_absMS. The description of these three strategies is as follows:

- SGLasso_absZ is a common strategy to retain non-zero coefficients of α.
- SGLasso_absM strategy is to retain those coefficients which are greater than the mean value of absolute values of all elements in α.
- SGLasso_absMS strategy is more strict for selecting coefficients, since it retains the coefficients which are larger than the mean value of absolute values of all non-zero coefficients in α.

We apply the above mentioned three strategies to feature selection, and then select the corresponding features according to the retained coefficients in α. SVM performs SZ identification using these selected features. The classification results using three different regression coefficient selection strategies are shown in **Figure 6**. According to **Figure 6**, the classification accuracy is the best when using SGLasso_absM strategy. Experimental result indicates that using SGLasso_absM strategy in feature selection can select the most effective features for SZ/HC classification. Therefore, we finally choose the SGLasso_absM strategy to select the regression coefficients.

4.5. Classification Comparison Using Different Feature Combinations

In order to explore the impact of different feature combinations on SZ/HC identification, we combine these five local measures extracted at the threshold of 0.30 in $C_5^2 + C_5^3 + C_5^4 + C_5^5 = 26$ ways. Furthermore, we don't consider individual graph measure because we only investigate multiple measures in this study. We evaluate these 26 feature sets under the same experimental settings. The classification results are shown in **Figure 7**.

As can be seen from **Figure 7**, the combination of 5 local graph measures achieves the best classification performance compared to other feature sets. At the same time, we also find that the classification accuracies obtained by using feature sets including two measures are lower than the classification accuracies obtained by using feature sets including three measures, four measures and five measures. It indicates that using fewer measures may not be enough to characterize brain network alternation, and we find that the combination of five local measures can provide more useful information for SZ identification.

4.6. Comparison With Existing Classification Methods

To verify the effectiveness of our proposed classification method, we compare some recently proposed methods for SZ classification using fMRI in the literature. Huang et al. (2018) proposed a tree-guided group sparse learning method to select the most important information from FALFF data in four frequency bands and get a classification accuracy of 91.1% by using multi-kernel SVM. Cheng et al. (2015) calculated only betweenness centrality measure to characterize the network. They used the rank of betweenness centrality of all nodes as feature representations and used SVM to classify SZs from HCs.

The two above mentioned methods are performed on the COBRE dataset. The classification results and ROC curves for SZ/HC classification of the two methods and our proposed method are shown in **Table 5** and in **Figure 8**, respectively.

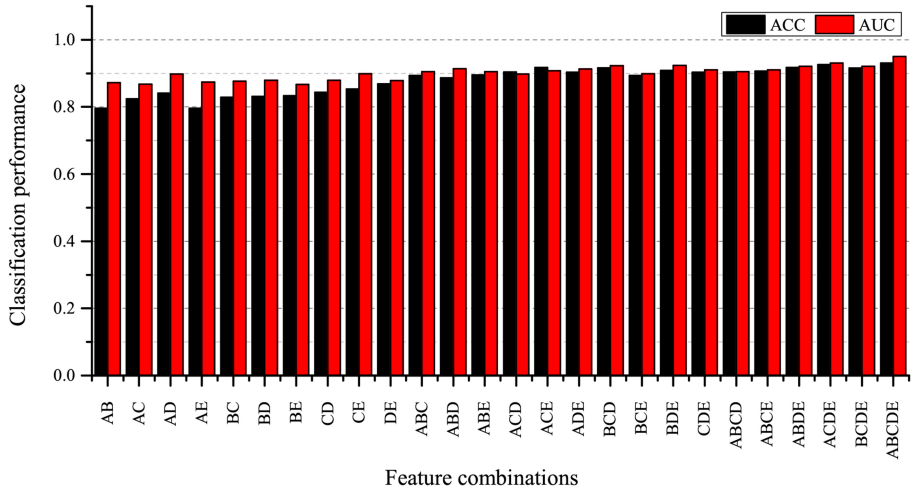

FIGURE 7 | Classification result for different feature combinations. A: betweenness centrality, B: nodal clustering coefficient, C: local efficiency, D: degree, E: participation coefficient.

TABLE 5 | Comparison with some existing methods for SZ/HC classification.

Methods	ACC (%)	SEN (%)	SPE (%)	AUC
Huang et al. (2018)	77.24	77.46	76.58	0.815
Cheng et al. (2015)	74.48	73.53	69.12	0.792
Proposed	**93.10**	**92.96**	**93.24**	**0.950**

Bold text indicates that the best result is obtained on a certain evaluation metric.

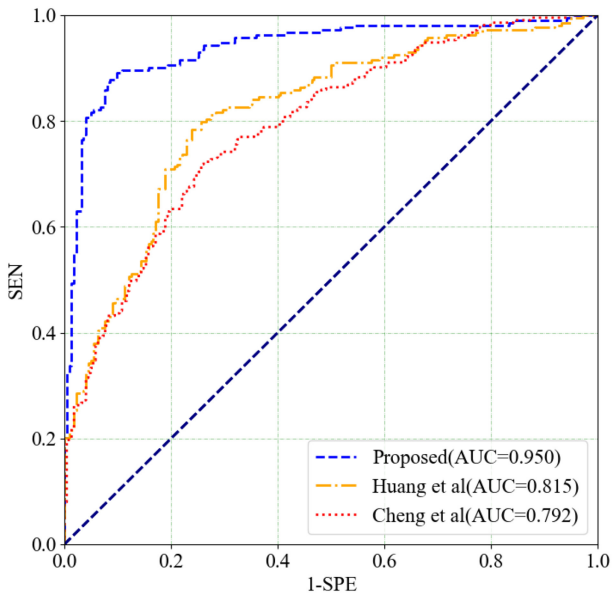

FIGURE 8 | ROC curves for SZ/HC classification for different classification methods.

According to **Table 5** and **Figure 8**, Our proposed method gets the best ACC (93.10%), SEN (92.96%), SPE (93.24%), and AUC (0.950) values. The experimental result illustrates that our proposed method has made a significant improvement in classification performance on the COBRE dataset.

4.7. Analysis of Discriminative Graph Measures and Corresponding Regions

The graph measures selected in the feature selection stage are considered to be related to their corresponding brain regions. Our method can select the most discriminative brain regions as the biomarkers to guide the disease-induced interpretation. There is a total of 145 experiments in the LOOCV scheme due to 145 subjects. And the number of feature occurrence in 145 experiments is introduced to indicate the contribution of the feature to classification. We assume that if the occurrence number of a local graph measure extracted from a certain brain region is greater than 140 in a total of 145 experiments, the brain region is considered to have the most discriminative power to distinguish between SZs and HCs. Based on this hypothesis, 21 salient brain regions have been found. These significant brain regions are shown in **Table 6**. Five brain regions include left superior frontal gyrus (SFG_L_7_2), right inferior temporal gyrus (ITG_R_7_7), right inferior parietal lobule (IPL_R_6_4), right postcentral gyrus (PoG_R_4_1), and right thalamus (Tha_R_8_7) are related to more than one local graph measure.

These findings on discriminative brain regions are in agreement with the following studies: superior frontal gyrus,cingulate gyrus, postcentral gyrus (Szeszko et al., 1999; Gur et al., 2000; Arbabshirani et al., 2013; Chyzhyk et al., 2015), parahippocampal gyrus (Shenton et al., 1992; Chyzhyk et al., 2015), middle temporal gyrus, fusiform gyrus and thalamus (Chyzhyk et al., 2015; Li et al., 2019), inferior parietal lobule, inferior temporal gyrus (Peng et al., 1994; Goldstein et al., 1999; Li et al., 2019). However, we cannot report agreement with these

TABLE 6 | The most discriminative graph measures and corresponding Brainnetome regions.

Graph measures	Hemisphere	Brainnetome regions	Occurrence number
Nodal clustering coefficient	SFG_L_7_2	Superior Frontal Gyrus	144
Degree	SFG_L_7_2	Superior Frontal Gyrus	145
Nodal clustering coefficient	SFG_R_7_2	Superior Frontal Gyrus	140
Participation coefficient	SFG_R_7_7	Superior Frontal Gyrus	144
Betweenness centrality	IFG_L_6_3	Inferior Frontal Gyrus	143
Betweenness centrality	OrG_L_6_2	Orbital Gyrus	143
Betweenness centrality	OrG_R_6_6	Orbital Gyrus	145
Betweenness centrality	PrG_L_6_3	Precentral Gyrus	142
Degree	MTG_L_4_4	Middle Temporal Gyrus	145
Betweenness centrality	MTG_L_4_1	Middle Temporal Gyrus	141
Participation coefficient	ITG_R_7_7	Inferior Temporal Gyrus	145
Betweenness centrality	ITG_R_7_7	Inferior Temporal Gyrus	145
Betweenness centrality	FuG_R_3_3	Fusiform Gyrus	145
Betweenness centrality	PhG_L_6_3	Parahippocampal Gyrus	144
Degree	PhG_R_6_5	Parahippocampal Gyrus	145
Local efficiency	IPL_R_6_4	Inferior Parietal Lobule	145
Participation coefficient	IPL_R_6_4	Inferior Parietal Lobule	145
Degree	IPL_R_6_2	Inferior Parietal Lobule	145
Degree	PCun_L_4_3	Precuneus	145
Nodal clustering coefficient	PoG_R_4_1	Postcentral Gyrus	145
Betweenness centrality	PoG_R_4_1	Postcentral Gyrus	145
Local efficiency	PoG_R_4_1	Postcentral Gyrus	143
Degree	PoG_R_4_1	Postcentral Gyrus	145
Participation coefficient	CG_L_7_4	Cingulate Gyrus	145
Betweenness centrality	CG_R_7_3	Cingulate Gyrus	145
Participation coefficient	LOcC_L_4_3	lateral Occipital Cortex	145
Degree	BG_R_6_1	Basal Ganglia	145
Betweenness centrality	BG_R_6_4	Basal Ganglia	145
Participation coefficient	Tha_L_8_8	Thalamus	145
Degree	Tha_L_8_5	Thalamus	145
Degree	Tha_R_8_8	Thalamus	145
Nodal clustering coefficient	Tha_R_8_7	Thalamus	140
Local efficiency	Tha_R_8_7	Thalamus	141

regions:inferior frontal gyrus, orbital gyrus, precentral gyrus, precuneus, lateral occipital cortex and basal ganglia.

5. CONCLUSION

In this paper, we propose a method to classify SZs from HCs using multi-view graph measures of functional brain networks. We get five local network measures using graph theoretical approach from multiple views. These measures paly an important role in the information exchange of brain networks. Our proposed method achieves a good classification performance on the COBRE dataset. Experimental results demonstrate that this approach is efficient for the clinical diagnosis of SZ. Furthermore, multiple measures have the potential to be used as clinical biomarkers to differentiate SZs from HCs.

DATA AVAILABILITY STATEMENT

The imaging data and phenotypic information was collected and shared by the Mind Research Network and the University of New Mexico funded by a National Institute of Health Center of Biomedical Research Excellence (COBRE) grant 1P20RR021938-01A2. The dataset for this study can be found in this website: http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html.

AUTHOR CONTRIBUTIONS

JW and JL conceived the project. YX, GT, F-XW, and JL designed the experiments. YX and GT performed the experiments. YX and JL wrote the paper. All authors read and approved the final manuscript.

FUNDING

This work was supported in part by the National Natural Science Foundation of China under Grant Nos. 61802442, 61877059, the Natural Science Foundation of Hunan Province under Grant No. 2019JJ50775, the 111 Project (No. B18059), the Hunan Provincial Science and Technology Program (2018WK4001). The funding bodies did not influence the design of the study and collection, analysis, and interpretation of data or writing the manuscript.

ACKNOWLEDGMENTS

A partial version of this work was accepted at the Fourth CCF Bioinformatics Conference (CBC 2019) in Guangzhou, China (Aug 23-25, 2019). We would like to thank the reviewers for their detailed suggestions which greatly improved the quality and readability of this work.

REFERENCES

Abraham, A., Milham, M. P., Di Martino, A., Craddock, R. C., Samaras, D., Thirion, B., et al. (2017). Deriving reproducible biomarkers from multi-site resting-state data: an autism-based example. *Neuroimage* 147, 736–745. doi: 10.1016/j.neuroimage.2016.10.045

Anderson, A., and Cohen, M. S. (2013). Decreased small-world functional network connectivity and clustering across resting state networks in schizophrenia: an fMRI classification tutorial. *Front. Hum. Neurosci.* 7:520. doi: 10.3389/fnhum.2013.00520

Arbabshirani, M. R., Kiehl, K., Pearlson, G., and Calhoun, V. D. (2013). Classification of schizophrenia patients based on resting-state functional network connectivity. *Front. Neurosci.* 7:133. doi: 10.3389/fnins.2013.00133

Brier, M. R., Thomas, J. B., Fagan, A. M., Hassenstab, J., Holtzman, D. M., Benzinger, T. L., et al. (2014). Functional connectivity and graph theory in preclinical Alzheimer's disease. *Neurobiol. Aging* 35, 757–768. doi: 10.1016/j.neurobiolaging.2013.10.081

Castro, E., Martínez-Ramón, M., Pearlson, G., Sui, J., and Calhoun, V. D. (2011). Characterization of groups using composite kernels and multi-source

fmri analysis data: application to schizophrenia. *Neuroimage* 58, 526–536. doi: 10.1016/j.neuroimage.2011.06.044

Chan, M., Krebs, M., Cox, D., Guest, P., Yolken, R. H., Rahmoune, H., et al. (2015). Development of a blood-based molecular biomarker test for identification of schizophrenia before disease onset. *Transl. Psychiatry* 5:e601. doi: 10.1038/tp.2015.91

Chang, C.-C., and Lin, C.-J. (2011). Libsvm: a library for support vector machines. *ACM Trans. Intell. Syst. Technology* 2:27. doi: 10.1145/1961189.1961199

Chen, Q., Lai, D., Lan, W., Wu, X., Chen, B., Chen, Y.-P. P., et al. (2019). ILDMSF: inferring associations between long non-coding RNA and disease based on multi-similarity fusion. *IEEE/ACM Trans. Comput. Biol. Bioinformatics*. doi: 10.1109/TCBB.2019.2936476. [Epub ahead of print].

Cheng, H., Newman, S., Goñi, J., Kent, J. S., Howell, J., Bolbecker, A., et al. (2015). Nodal centrality of functional network in the differentiation of schizophrenia. *Schizophrenia Res.* 168, 345–352. doi: 10.1016/j.schres.2015.08.011

Chyzhyk, D., Savio, A., and Graña, M. (2015). Computer aided diagnosis of schizophrenia on resting state fMRI data by ensembles of ELM. *Neural Netw.* 68, 23–33. doi: 10.1016/j.neunet.2015.04.002

Craddock, R. C., Jbabdi, S., Yan, C.-G., Vogelstein, J. T., Castellanos, F. X., Di Martino, A., et al. (2013). Imaging human connectomes at the macroscale. *Nat. Methods* 10:524. doi: 10.1038/nmeth.2482

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 connectional architecture. *Cereb. Cortex* 26, 3508–3526. doi: 10.1093/cercor/bhw157

Fornito, A., Zalesky, A., and Breakspear, M. (2013). Graph analysis of the human connectome: promise, progress, and pitfalls. *Neuroimage* 80, 426–444. doi: 10.1016/j.neuroimage.2013.04.087

Goldstein, J. M., Goodman, J. M., Seidman, L. J., Kennedy, D. N., Makris, N., Lee, H., et al. (1999). Cortical abnormalities in schizophrenia identified by structural magnetic resonance imaging. *Arch. Gen. Psychiatry* 56, 537–547. doi: 10.1001/archpsyc.56.6.537

Guo, W., Su, Q., Yao, D., Jiang, J., Zhang, J., Zhang, Z., et al. (2014). Decreased regional activity of default-mode network in unaffected siblings of schizophrenia patients at rest. *Eur. Neuropsychopharmacol.* 24, 545–552. doi: 10.1016/j.euroneuro.2014.01.004

Gur, R. E., Cowell, P. E., Latshaw, A., Turetsky, B. I., Grossman, R. I., Arnold, S. E., et al. (2000). Reduced dorsal and orbital prefrontal gray matter volumes in schizophrenia. *Arch. Gen. Psychiatry* 57, 761–768. doi: 10.1001/archpsyc.57.8.761

He, H., Sui, J., Yu, Q., Turner, J. A., Ho, B.-C., Sponheim, S. R., et al. (2012). Altered small-world brain networks in schizophrenia patients during working memory performance. *PLoS ONE* 7:e38195. doi: 10.1371/journal.pone.0038195

Heinrichs, R. W., and Zakzanis, K. K. (1998). Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. *Neuropsychology* 12:426. doi: 10.1037/0894-4105.12.3.426

Huang, J., Zhu, Q., Hao, X., Shi, X., Gao, S., Xu, X., and Zhang, D. (2018). Identifying resting-state multifrequency biomarkers via tree-guided group sparse learning for schizophrenia classification. *IEEE J. Biomed. Health Informatics* 23, 342–350. doi: 10.1109/JBHI.2018.2796588

Khazaee, A., Ebrahimzadeh, A., and Babajani-Feremi, A. (2015). Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory. *Clin. Neurophysiol.* 126, 2132–2141. doi: 10.1016/j.clinph.2015.02.060

Khazaee, A., Ebrahimzadeh, A., Babajani-Feremi, A., for the Alzheimer's Disease Neuroimaging Initiative (2017). Classification of patients with mci and ad from healthy controls using directed graph measures of resting-state fMRI. *Behav. Brain Res.* 322, 339–350. doi: 10.1016/j.bbr.2016.06.043

Li, J., Sun, Y., Huang, Y., Bezerianos, A., and Yu, R. (2019). Machine learning technique reveals intrinsic characteristics of schizophrenia: an alternative method. *Brain Imaging Behav.* 13, 1386–1396. doi: 10.1007/s11682-018-9947-4

Liu, J., Ji, S., and Ye, J. (2009). *Slep: Sparse Learning With Efficient Projections*. Arizona State University, 7.

Liu, J., Li, M., Lan, W., Wu, F.-X., Pan, Y., and Wang, J. (2016). Classification of Alzheimer's disease using whole brain hierarchical network. *IEEE/ACM Trans. Comput. Biol. Bioinformatics* 15, 624–632. doi: 10.1109/TCBB.2016.2635144

Liu, J., Li, M., Pan, Y., Lan, W., Zheng, R., Wu, F.-X., and Wang, J. (2017a). Complex brain network analysis and its applications to brain disorders: a survey. *Complexity* 2017:8362741. doi: 10.1155/2017/8362741

Liu, J., Li, M., Pan, Y., Wu, F.-X., Chen, X., and Wang, J. (2017b). Classification of schizophrenia based on individual hierarchical brain networks constructed from structural MRI images. *IEEE Trans. Nanobiosci.* 16, 600–608. doi: 10.1109/TNB.2017.2751074

Liu, J., Pan, Y., Li, M., Chen, Z., Tang, L., Lu, C., et al. (2018a). Applications of deep learning to MRI images: a survey. *Big Data Mining Analyt.* 1, 1–18. doi: 10.26599/BDMA.2018.9020001

Liu, J., Wang, J., Hu, B., Wu, F.-X., and Pan, Y. (2017c). Alzheimer's disease classification based on individual hierarchical networks constructed with 3-D texture features. *IEEE Trans. Nanobiosci.* 16, 428–437. doi: 10.1109/TNB.2017.2707139

Liu, J., Wang, J., Tang, Z., Hu, B., Wu, F.-X., and Pan, Y. (2017d). Improving Alzheimer's disease classification by combining multiple measures. *IEEE/ACM Trans. Comput. Biol. Bioinformatics* 15, 1649–1659. doi: 10.1109/TCBB.2017.2731849

Liu, J., Wang, X., Zhang, X., Pan, Y., Wang, X., and Wang, J. (2018b). MMM: classification of schizophrenia using multi-modality multi-atlas feature representation and multi-kernel learning. *Multimedia Tools Appl.* 77, 29651–29667. doi: 10.1007/s11042-017-5470-7

Lynall, M.-E., Bassett, D. S., Kerwin, R., McKenna, P. J., Kitzbichler, M., Muller, U., et al. (2010). Functional connectivity and brain networks in schizophrenia. *J. Neurosci.* 30, 9477–9487. doi: 10.1523/JNEUROSCI.0333-10.2010

Marín, O. (2012). Interneuron dysfunction in psychiatric disorders. *Nat. Rev. Neurosci.* 13:107. doi: 10.1038/nrn3155

Moghimi, P., Lim, K. O., and Netoff, T. I. (2018). Data driven classification of fMRI network measures: application to schizophrenia. *Front. Neuroinform.* 12:71. doi: 10.3389/fninf.2018.00071

Nieuwenhuis, M., van Haren, N. E., Pol, H. E. H., Cahn, W., Kahn, R. S., and Schnack, H. G. (2012). Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. *Neuroimage* 61, 606–612. doi: 10.1016/j.neuroimage.2012.03.079

Pedersen, M., Omidvarnia, A., Zalesky, A., and Jackson, G. D. (2018). On the relationship between instantaneous phase synchrony and correlation-based sliding windows for time-resolved fMRI connectivity analysis. *Neuroimage* 181, 85–94. doi: 10.1016/j.neuroimage.2018.06.020

Peng, L. W., Lee, S., Federman, E. B., Chase, G. A., Barta, P. E., and Pearlson, G. D. (1994). Decreased regional cortical gray matter volume in schizophrenia. *Am. J. Psychiatry* 151:843. doi: 10.1176/ajp.151.6.842

Pettersson-Yeo, W., Allen, P., Benetti, S., McGuire, P., and Mechelli, A. (2011). Dysconnectivity in schizophrenia: where are we now? *Neurosci. Biobehav. Rev.* 35, 1110–1124. doi: 10.1016/j.neubiorev.2010.11.004

Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J. L., Kähler, A. K., Akterin, S., et al. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat. Genet.* 45, 1150–1159. doi: 10.1038/ng.2742

Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003

Shenton, M. E., Kikinis, R., Jolesz, F. A., Pollak, S. D., LeMay, M., Wible, C. G., et al. (1992). Abnormalities of the left temporal lobe and thought disorder in schizophrenia: a quantitative magnetic resonance imaging study. *N. Engl. J. Med.* 327, 604–612. doi: 10.1056/NEJM199208273270905

Szeszko, P. R., Bilder, R. M., Lencz, T., Pollack, S., Alvir, J. M. J., Ashtari, M., et al. (1999). Investigation of frontal lobe subregions in first-episode schizophrenia. *Psychiatry Res. Neuroimaging* 90, 1–15. doi: 10.1016/S0925-4927(99)00002-5

Van Den Heuvel, M. P., and Pol, H. E. H. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol.* 20, 519–534. doi: 10.1016/j.euroneuro.2010.03.008

Yan, C.-G., Wang, X.-D., Zuo, X.-N., and Zang, Y.-F. (2016). Dpabi: data processing & analysis for (resting-state) brain

imaging. *Neuroinformatics* 14, 339–351. doi: 10.1007/s12021-016-9299-4

Yang, Y., and Wang, H. (2018). Multi-view clustering: a survey. *Big Data Mining and Analyt.* 1, 83–107. doi: 10.26599/BDMA.2018.9020003

Yu, Q., Erhardt, E. B., Sui, J., Du, Y., He, H., Hjelm, D., et al. (2015). Assessing dynamic brain graphs of time-varying connectivity in fMRI data: application to healthy controls and patients with schizophrenia. *Neuroimage* 107, 345–355. doi: 10.1016/j.neuroimage.2014.12.020

Yuan, M., and Lin, Y. (2006). Model selection and estimation in regression with grouped variables. *J. R. Stat. Soc. Ser. B* 68, 49–67. doi: 10.1111/j.1467-9868.2005.00532.x

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

Copyright © 2020 Xiang, Wang, Tan, Wu and Liu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.