

The Effect Analysis of Atlas and Global Signal Regression in Classification based on Brain Network for Major Depression Disorders

Dan Long

Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital)
Institute of Cancer and Basic Medicine (IBMC), Chinese Academy of Sciences

Yingjun Liu

School of Biomedical Engineering, Southern Medical University

Zengsi Chen

College of Sciences, China Jiliang University

Jingsi Xie

Jiangxi V&T College of Communication

Lei Shi and Cong Luo

Cancer Hospital of the University of Chinese Academy of Sciences (Zhejiang Cancer Hospital)
Institute of Cancer and Basic Medicine (IBMC), Chinese Academy of Sciences
E-mail: shilei@zjcc.org.cn, lw939291@126.com

Abstract. Automatic classification of major depression disorders (MDD) is an arduous task. When constructing the brain network automatic classification model based on functional magnetic resonance imaging (fMRI), the selection of global signal regression (GSR) and brain atlas are two key factors. However, their impact on the classification has not reached a consensus so far. The main reasons include the following two points: first, the sample size of previous studies is small, and different studies lead to inconsistent results; second, there are too many parameters in their models, which could not clearly reveal the effects of the above two factors. Therefore, we believe that only by using the data of multi-center and large samples, it is possible to find out the influence of these two factors on the classification results. To test our hypothesis, data sets (The REST-meta-MDD project) from 17 centers were used in this study. The set was divided into two parts, training set and independent validation set. The training set used 10-fold cross-validation to evaluate the classification performance, and the independent validation set used the features of the first part to classify directly. Feature selection adopted two sample t-test plus least absolute shrinkage and selection operator (LASSO), and the classifier was linear support vector machine (SVM). Finally, the classification effect of factors was confirmed by statistical analysis. The results showed that the impact of GSR on the classification results was related to the selection of brain atlas. In anatomical automatic labeling (AAL)-based networks, GSR would reduce the classification accuracy. But for Dosenbach networks, GSR would improve the classification performance. The classification ability of networks constructed by different brain templates was different, and the AAL was the best. In conclusion, the choice of brain atlas was a

key factor affecting classification performance in MDD classification.
© 2022 Society for Imaging Science and Technology.
[DOI: 10.2352/J.ImagingSci.Technol.2022.66.4.040413]

1. INTRODUCTION

Depression is a highly prevalent mental illness and more than 264 million people are affected worldwide [1]. Its main pathophysiological mechanism is still unclear. Depression not only affects the normal life of patients but may also lead to suicide. According to WHO statistics, depression is the second leading cause of death among young people aged 15–29. At present, the diagnosis of depression mainly relies on neuropsychological scores, the most commonly used are the diagnostic and statistical manual of mental disorders (DSM-5) and Hamilton Rating Scale for Depression (HRSD). However, a study showed that the consistency of the diagnosis of major depression disease (MDD) by clinicians through DSM-5 was only 0.25 [2], which might weaken the effectiveness of MDD diagnosis. Because of unclear classification and confusion of subjective clinical impressions, objective diagnostic tools are urgently needed to assess depression.

In recent years, many scholars hoped to use the state of the art machine learning technology to extract imaging markers to establish MDD diagnosis and efficacy evaluation [3]. The latest research results showed that functional connectivity (FC) based on graph theory could sensitively detect changes in the brain network of patients with

depression [4–7], and these changes could be transformed into biomarkers to identify MDD and its subtypes [8, 9], and even distinguished other diseases with similar clinical symptoms (such as bipolar disorder [10]). An impressive study showed that four different neurophysiological biotypes of MDD could be defined according to FC between the limbic and frontostriatal networks. These subtypes exceeded current diagnostic boundaries and may help determine who was most likely to receive neurostimulation therapy help [9]. In short, FC based on Pearson's correlation (PC) has great potential as a biological marker for effectively distinguishing MDD from normal controls.

Building an automatic classification system based on FC remains a challenging task, in which global signal regression (GSR) and brain template selection were two key steps. Although some studies attempted to find out the impact of the two factors on classification [11, 12], no convincing conclusion has been drawn so far. A very important factor for this lack of consensus was that the influence of imaging conditions in the signal processing was often ignored. Functional magnetic resonance imaging (fMRI) signal was based on blood oxygen level dependence (BOLD) contrast. BOLD signal was the result of the complex interaction of metabolism, blood flow and blood volume. Any factor affecting the balance between these three parameters will lead to BOLD signal change, leading to low signal-to-noise ratio of BOLD signal. Therefore, we believe that using multi-center data could better understand the impact of these two factors on classification.

To test our hypothesis, this study used 1160 subjects' data from 17 research centers in China (REST-meta-MDD project). The data set was divided into two parts, training set and independent validation set. The training set evaluated the classification performance through the cross-validation method. The independent validation set used the features selected from the training set to classify directly, which further strengthens our conclusion. The rest of this paper was organized as follows. The second section reviews the effects of GSR and brain template selection in MDD classification. The third section describes our proposed algorithm. The fourth part reports the implementation process, experimental results and analysis of the model design. The fifth section summarizes our work.

2. RELATED WORK

Although FC has made good achievements in the classification of MDD, there were great disputes on two key factors affecting the classification performance. One is based on GSR used in data preprocessing. Two studies in 2009 gave opposite suggestions on whether GSR should be used in the preprocessing of resting-state FC [13, 14]. Some scholars thought that GSR will introduce “pseudo-anticorrelation” [14]. Li's research showed that GSR can improve FC correlation with behavior [15], but Dadi's research showed that the use of GSR did not improve classification accuracy [16], such contradictory results required us to carefully verify the effectiveness of GSR in classification.

Another key factor is the choice of brain atlas, which is usually ignored. An important step of FC-based pattern recognition is to extract time-series information based on ROI (region of interest) in order to calculate FC. Dadi's study compared the effects of three different brain region definition methods (anatomy based, function based and independent component analysis) on the classification results in detail, and their results showed that the function-based brain template has better classification performance than the anatomy-based brain template [16]. Wang used resting-state fMRI (rsfMRI) to study the influence of different brain maps on the topology of brain function network. The results showed that no matter which template is chosen, the brain function network has strong small world characteristics and truncated power-law degree distribution. However, significant intergroup differences were observed in the local and global characteristics of the network. These results provided quantitative evidence of topological parameters in brain functional networks for the first time [11].

3. PROPOSED MODEL

Generally speaking, the framework of the brain network-based classification model includes four parts: data preprocessing, brain network definition, feature extraction, and classifier selection. There are many choices for each part, among which the most controversial are the choice of GSR and brain template. In order to highlight the impact of these two factors on classification, we used the most common method with the least parameters to construct the classification model. Data preprocessing included slice timing, realignment and spatial normalization. The classifier adopted linear support vector machine (SVM). The whole experimental process is shown in Figure 1.

3.1 Data Preprocessing

Each subject's data was preprocessed by Data Processing Assistant for Resting-State fMRI (DPARSF). The details were divided into the following four steps [17]:

- step 1 Discarded the first 10 volumes, and then used slice timing to reduce the impact of the difference in acquisition time between image layers at the same time point;
- step 2 Adopted a rigid body linear transformation to realign the time series of images for each subject. After realignment, individual T1-weighted images were co-registered to the mean functional image using a 6 degrees-of-freedom linear transformation without resampling and then segmented into gray matter (GM), white matter (WM) and cerebrospinal fluid (CSF);
- step 3 Space standardization, and images were registered to MNI space;
- step 4 Applied Friston-24 parameter model to regress the effect of head movement [18]. As GSR was still a controversial practice in the rsfMRI field, there were two

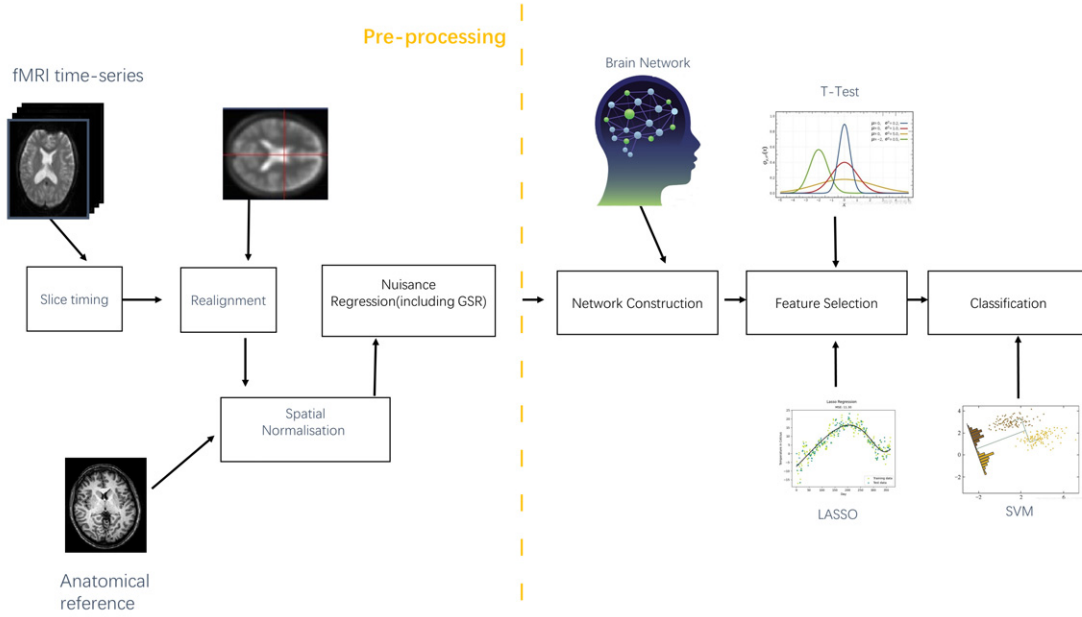


Figure 1. Flow chart of the study.

types data in this study. One did not perform GSR and the other included analyses with GSR. WM and CSF signals were also removed from the data through linear regression to reduce respiratory and cardiac effects. Additionally, the linear trend was included as a regressor to account for drifts in the BOLD signal. Finally, we applied bandpass filtering to all time series (0.01–0.1 Hz) [17].

After the data preprocessing, the time series were extracted from different brain atlases, and then the average time series of each brain area were calculated. The detailed data preprocessing please refer to literature [17].

The method of slice timing of different layers in step 1 was as follows: assuming the time series y of the number of layers n obtained at the time point $t(n)$, the linear interpolation of the layer time of the reference layer $t(r)$ can be formally expressed as:

$$y_n^{(r)} = \frac{(t(r) - t(n-1))y_n - (t(n) - t(r))y_{n-1}}{t(n) - t(n-1)}. \quad (1)$$

The effective implementation of sinc interpolation (see Eq. (1)) requires the application of phase shift (i.e., adding a constant value) in the frequency domain of the signal obtained by fast Fourier transform. Phase shift is used to compensate interpolation and surround effect. The specific calculation is as follow:

$$y_n^{(r)} = \sum_{i=-\infty}^{\infty} y_i \sin c \left(\frac{\pi}{TR} (r - iTR) \right), \quad (2)$$

where TR is the pulse repetition time.

3.2 Global Signal Regression

The specific algorithm of GSR is as follows. Suppose that each $y_i(t)$ is a column vector representing the time series of the i th voxel ($i = 1, \dots, N$). Since the average value of the time series of each voxel can be removed in the regression process, it is assumed that the average value of each voxel was zero. The corresponding time series after global signal regression is expressed as column vector $x_i(t)$. The regression is showed in Eq. (3) and the global signal is calculated according to Eq. (4).

$$y_i(t) = g(t)\beta_i + x_i(t) \quad (3)$$

$$g(t) = \frac{1}{N} \sum_{j=1}^N y_j(t). \quad (4)$$

Where j is same as i , denotes j th voxel. The estimation of regression coefficient β_i in Eq. (3) is obtained from Eq. (5).

$$\beta_i = (g(t)'g(t))^{-1}g(t)'y_i(t), \quad i = 1, 2, \dots, N. \quad (5)$$

Thus, the mean value of the coefficients of all voxels is,

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \beta_i &= (g(t)'g(t))^{-1}g(t)'\frac{1}{N} \sum_{i=1}^N y_i(t) \\ &= (g(t)'g(t))^{-1}g(t)'g(t). \end{aligned} \quad (6)$$

Have reason,

$$\frac{1}{N} \sum_{i=1}^N \beta_i = 1. \quad (7)$$

Substitute Eq. (4) into Eq. (3), we have,

$$y_i(t) = \left(\frac{1}{N} \sum_{j=1}^N y_j(t) \right) \beta_i + x_i(t). \quad (8)$$

Shift the term of Eq. (8) and calculate the sum of time series after global signal regression:

$$\begin{aligned}\sum_{i=1}^N x_i(t) &= \sum_{i=1}^N \left(y_i(t) - \beta_i \frac{1}{N} \sum_{j=1}^N y_j(t) \right) \\ &= \sum_{i=1}^N y_i(t) - \left(\frac{1}{N} \sum_{j=1}^N \beta_j \right) \left(\sum_{j=1}^N y_j(t) \right) \quad (9)\end{aligned}$$

replaced by Eq. (7), we can obtain the result of Eq. (10).

$$\sum_{i=1}^N x_i(t) = 0, \quad \forall t. \quad (10)$$

Let x_1 be the time series of seed voxels in correlation analysis. We apply inner product on Eq. (10):

$$\left\langle \sum_{i=1}^N x_i(t), x_1(t) \right\rangle = \sum_{i=1}^N \langle x_i(t), x_1(t) \rangle = 0. \quad (11)$$

Thus there:

$$\sum_{i=2}^N \langle x_i(t), x_1(t) \rangle = -\langle x_1(t), x_1(t) \rangle \leq 0. \quad (12)$$

Equation (12) shows that the sum of inner products is less than or equal to 0, and the sum of Pearson correlation coefficients is as follow:

$$\sum_{i=2}^N c(x_1(t), x_i(t)) = \sum_{i=2}^N \frac{\langle (x_1(t) - \mu_{x_1}), (x_i(t) - \mu_{x_i}) \rangle}{T \sigma_{x_1} \sigma_{x_i}}. \quad (13)$$

As mentioned earlier, where T is the number of time points, μ_{x_i}, σ_{x_i} are the mean and standard deviation of time series $x_i(t)$. Because we have assumed $\mu_{x_i} = 0$, Eq. (13) can be simplified to Eq. (14):

$$\sum_{i=2}^N c(x_1(t), x_i(t)) = \sum_{i=2}^N \frac{\langle x_1(t), x_i(t) \rangle}{T \sigma_{x_1} \sigma_{x_i}}. \quad (14)$$

If we assume that the standard deviation is unbiased for the positive and negative inner product between time series, we would obtain the result of Eq. (15) from Eq. (12).

$$\sum_{i=2}^N c(x_1(t), x_i(t)) \leq 0. \quad (15)$$

It can be seen from Eq. (15) that the negative correlation does increase after GSR treatment.

3.3 Network Construction

All networks in this study were generated by the Brain-NetClass (Brain Network Construction and Classification) toolbox [19]. We used the Pearson-based FC to construct network. It was defined as follows: the brain was divided into N region of interests (ROIs) according to the atlas.

Table I. Subject information.

Number of subjects	1160	Number of sites	10
Male	434	Female	726
MDD	597	Normal Controls	563

The i th ROI was expressed as a column vector $x_i = [x_{1i}, x_{2i}, \dots, x_{Ti}]' \in R^T$ ('means transpose), the whole brain signal was represented by the matrix $X = [x_1, x_2, \dots, x_N] \in R^{T \times N}$. Brain network can be expressed as a weighted graph $W \in R^{N \times N}$ that each element in the matrix W was the PC coefficient of the average time series of two different brain regions. The FC between two ROIs (x_i and x_j) could be calculated by the following formula (Eq. (16)):

$$FC_{ij} = \frac{\sum_{t=1}^T (x_i(t) - \bar{x}_i)(x_j(t) - \bar{x}_j)}{\sqrt{\sum_{t=1}^T (x_i(t) - \bar{x}_i)^2} \sqrt{\sum_{t=1}^T (x_j(t) - \bar{x}_j)^2}}. \quad (16)$$

3.4 Features Selection

Too many features with no discriminative power might adversely affect the classification results. In this study, FC in the brain network was regarded as the feature and used a unified two-step method for feature selection. First, the two-sample t -test was used to initially select the features. And the features with no significant differences were discarded ($p < 0.01$). Second, the features are further selected through the least absolute shrinkage and selection operator (LASSO) [19], and only the features with a coefficient other than 0 were selected as the last features for model training.

4. EXPERIMENTS

4.1 Dataset

All data came from the REST-meta-MDD project (<http://rfmri.org/REST-meta-MDD>). By combining 25 research groups in 17 hospitals in China, the brain imaging data of 1300 patients with depression and 1128 healthy controls were successfully collected. After screening, the data set contained 1642 subjects from 17 different sites (848 MDD Patients and 794 normal controls) [17]. In order to meet the needs of this research, we further screened the data, the details were as follows:

- removed data which time of repetition was not 2.0.
- removed subjects that contain all 0 time series.
- performed gender and age matching tests on the data of each site, deleted the unmatched sites, and finally a total of 1160 subject data from 10 sites were included in this experiment.

Subject information was shown in Table I and please refer to literature for more details [17]. The data were randomly divided into two parts. Data set 1 contained 70% of the samples, and 10-fold cross-validation (CV) was used to verify the impact of different factors on classification. The

Table II. Average accuracy based on dataset1 10-fold CV (10 repetitions).

	AAL*	Harvard Oxford*	Craddock*	Dosenbach*	Power
GSR	59.64 ± 1.40	59.20 ± 0.88	59.00 ± 1.33	57.32 ± 1.58	58.17 ± 1.15
noGSR	61.12 ± 0.94	54.69 ± 0.98	60.18 ± 1.34	55.37 ± 1.77	56.66 ± 2.32

* Indicated that there was a significant difference between GSR and noGSR for a single atlas ($p < 0.05$).

Table III. Average accuracy based on dataset2 10-fold CV (10 repetitions).

	AAL*	Harvard-Oxford	Craddock	Dosenbach*	Power
GSR	54.25 ± 1.69	55.98 ± 0.96	56.68 ± 2.48	56.07 ± 1.95	55.35 ± 2.38
noGSR	58.32 ± 1.23	56.36 ± 1.23	56.36 ± 2.80	51.53 ± 1.25	52.95 ± 1.27

* Indicated that there was a significant difference between GSR and noGSR for a single atlas ($p < 0.05$).

remaining data was an independent validation set (called data set 2) that can be used to evaluate the generalization performance of the classifier which trained by the most discriminative features selected from data set 1. The gender and age of each site in the two datasets were matched.

There were two options for the GSR: GSR and noGSR (data not processed by GSR). A total of five brain atlases were selected: Anatomical Automatic Labeling (AAL) [20], Harvard-Oxford [21], Craddock [22], Dosenbach [23], Powers [24] (the data set contained data from the Schaefer and Zalesky atlases. These brain atlases with too much ROIs which would lead to too much calculation, so they were not included in this study). These two factors will produce a total of 10 different situations.

4.2 Classification and Effect Analysis

The training model used was linear SVM. We used cross-validation to evaluate the classification performance for data set 1. Due to the relatively large sample size, we used a stratified 10-fold CV method. Briefly, all subjects were randomly divided into 10 parts at random, and the whole process was divided into ten rounds. One of which was taken as the test set in each round, and the remaining parts were used as the training set. The classification accuracy was calculated in each round. The final performance index was the average of the 10 classification results. Since the parameter λ in LASSO needs to be determined before feature selection (the range of this study is [0.02, 0.08]), the parameter can be selected by the nested-CV method [25]. In short, nested-CV was to perform a 10-fold CV for each candidate parameter combination on each round of training set. The parameter combination with the best classification performance was the selected parameter, which can ensure the external CV did not involve parameter selection. In conclusion, a total of two CVs were performed for data set 1. The internal CV determined the parameter λ and the external CV determined the classification performance. The whole process was repeated 10 times.

Data set 2 was an independent validation set and unlike the situation of data set 1, there was no CV. For feature selection, we used the features selected in each round in data set 1 as training features and then trained an SVM classifier to classify data set 2.

To determine whether a certain factor will affect the classification effect, we used two sample t -test for statistical analysis ($p < 0.05$).

4.3 Results

In general, AAL is a very special brain atlas. The AAL-based network has the best classification performance, and GSR would reduce the classification ability.

Different brain atlas might have an impact on classification. Through statistical analysis, we found that AAL+noGSR network had the best classification performance for both dataset 1 and dataset 2. The impact of GSR on classification depends on the selection of atlases. In the two data sets, the networks based on AAL and Dosenbach have statistical differences, reflecting good consistency. The GSR effect of the remaining brain template networks appeared only in one data set. For AAL, GSR would reduce the classification accuracy, which was particularly prominent in dataset 2. But for the Dosenbach network, the result was the opposite. The detailed results are shown in Tables II and III.

For data set 1, there are 10 different combinations in total and 10 λ parameters would be selected for each combination because of the 10-fold CV. This process was repeated 10 times, and a total of 1000 parameters were selected. And 0.04 is the most selected parameter (73.7%). It showed that the random effect of the λ parameter on the classification result was very small. The details are shown in Table IV.

If the data was not processed by GSR, the classification performance of AAL network was better than the others. The classification accuracy of PowersROI was lower than that of Harvard-Oxford and Craddock networks. The classification performance of Dosenbach network was lower than that of Craddock network. If the data was processed by GSR, these

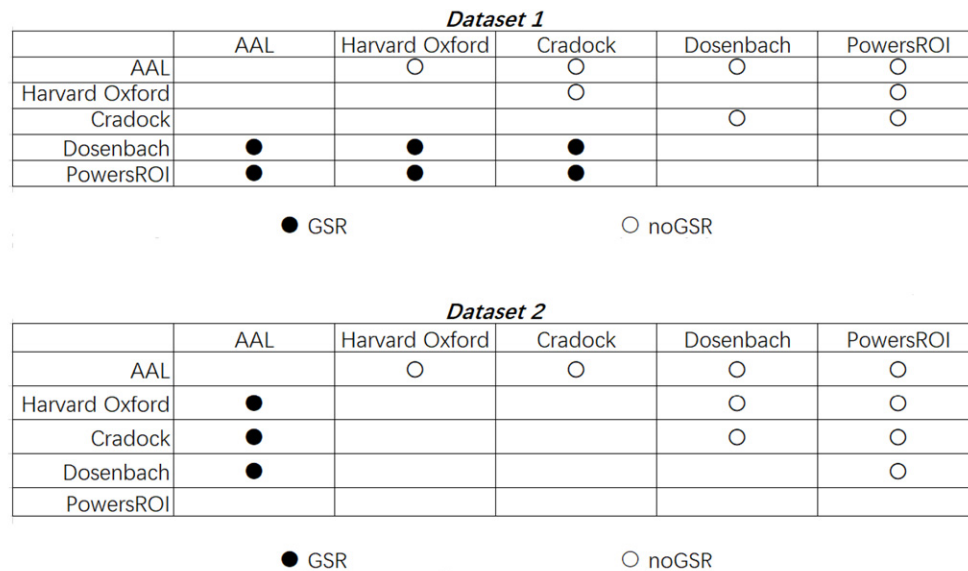


Figure 2. The atlas effect in two datasets. In the above two tables, the upper triangle area represented the comparison results of classification performance between different templates without GSR, and the hollow circle represented a statistical difference ($p < 0.05$); The lower triangle represented the comparison results of classification performance among different templates after processed by GSR, and the solid circle represented the statistical difference ($p < 0.05$).

Table IV. λ number of times selected ($p < 0.01$).

0.02	0.04	0.06	0.08
190 (19%)	737 (73.7%)	58 (5.8%)	15 (1.5%)

statistical differences were inconsistent in the two data sets. The details were shown in Figure 2.

Since the AAL-based network showed the best classification performance, we only presented the features selected from AAL, and the detailed information is shown in Table V and Figure 3. The most discriminating brain regions are located in the frontal lobe, occipital lobe, amygdala and cerebellum. The FC between the amygdala and the frontal-parietal control system (postcentral, supramarginal) and temporal lobe are the most distinguishing which indicates the amygdala play a vital role in the pathological changes of MDD. The cerebellum is also a very discriminative brain area. Changes in the cerebellum, occipital lobe, and frontal lobe could also be used as features to distinguish MDD from normal people. The FC alteration between the cerebellum and the occipital lobe was considered to be a visual compensatory effect [26] which also has an important impact on classification.

4.4 Analysis

To the best of our knowledge, in the MDD classification based on brain network, the influence of two important factors, atlas and GSR has been studied for the first time. The results show that whether GSR would affect the classification performance depends on the selection of brain template. When GSR is not used, the classification ability of AAL network is higher than GSR network, while the Dosenbach

is the opposite. Apart from other factors, the classification efficiency of different brain atlas is different and AAL-based networks have the best classification ability.

Although there are many machine learning methods used for MDD classification, such as Gaussian process classifier, linear discriminant analysis, decision tree [3] and more recently deep learning [27], but SVM is still the most common method. In order to ensure that our results have the greatest universality, we used SVM as a classifier (LIBSVM v3.23 [28]). There are many different kernel functions that can be chosen to construct the SVM classifier. We used linear SVM as the classification model instead of the most commonly used the radial basis function (RBF) kernel function. The main reason is that the RBF kernel function needed to determine two parameters, the random effect may cover the main effect. So, we did not use it as the kernel function.

The brain atlases used in this study were divided into three categories. The first category was based on anatomical structures, including AAL and Harvard-Oxford; the second category was based on FC, including Cradock and PowersROI; the third category define brain atlases based on the previous paper, such as Dosenbach. We can see a significant difference in the classification between these three types, the AAL network is slightly better than others. The possible reason for this result is that the location and size of ROI in different maps are different, and the information extracted from them is different, which would lead to the change of brain network structure [11], and finally affect the classification.

GSR in rsfMRI data preprocessing is a controversial issue. Two studies discussed this issue at the same time but gave opposite suggestions [13, 14]. Fox et al., also found that several characteristics of anti-correlated networks

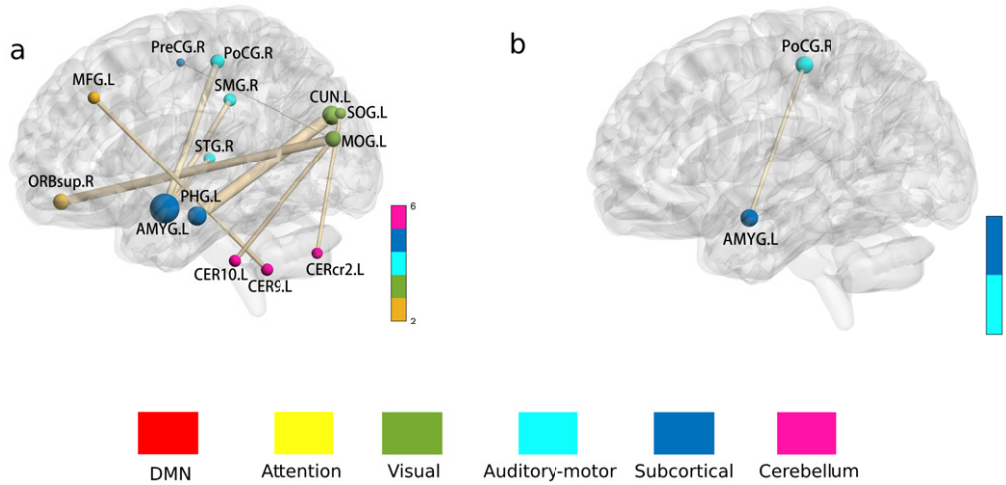


Figure 3. Features selected in each round based on the AAL. The whole brain was divided into six subnets marked with different colors: 1. DMN; 2. attention network; 3. visual network; 4. auditory-motor network; 5. subcortical network; 6. cerebellum. Figure a showed using GSR preprocessing and then building a network; Figure B showed that the network was built directly without GSR. The brain networks were visualized with the BrainNet Viewer (<http://www.nitrc.org/projects/bnv/>) [29].

Table V. Features selected in each round based on the AAL.

AAL	
GSR	(PreCG.R,MOG.L)(ORBsup.R,MOG.L)(MFG.L,CER9.L)(PHG.L,CUN.L) (AMYG.L,SMG.R) (AMYG.L,STG.R)(SOG.L,CERcr2.L)(MOG.L,CER10.L)
noGSR	(AMYG.L, PoCG.R)

PreCG:Precentral gyrus, MOG: middle occipital gyrus; ORBsup: Superior frontal gyrus, orbital part; MFG: Middle frontal gyrus; CER9: Cerebellum_9; PHG: Parahippocampal gyrus; CUN: Cuneus; AMYG: Amygdala; SMG: Supramarginal gyrus; STG: Superior temporal gyrus; SOG: Superior occipital gyrus; CERcr2: Cerebellum_Crus2; CER10: Cerebellum_10; PoCG: Postcentral gyrus; L: left; R: right.

were not the product of GSR. They concluded that GSR can be beneficial because GSR enhanced the detection of system-specific correlations and improved the correspondence between resting-state correlations and anatomy [13]. Subsequent studies have found that GSR can improve the specificity of positive correlation [13, 30] and help remove the effects of breathing [31] and movement [32, 33] in the global signal. However, Murphy first proved that GSR would produce anti-correlation in mathematics which may be the product of processing technology and concluded that GSR should not be used [14]. Subsequent studies have found that GSR reduces the test-retest reliability of the brain network of the elderly [34], but increases the consistency of FC between different scans within the subject [35]; impairs the ability to find relationships between connectivity and behavior in Autism Spectrum Disorder [36], but improves the relationship between resting-state FC and behavior in normal people [15]. Although the authors of these two papers jointly published a paper on Neuroimage in 2017, hoping to reach a consensus, we still cannot give a definite answer on whether to use GSR [12]. In short, the effect of GSR on fMRI signals has always been a controversial issue. Long et al., found that GSR has no effect on the classification results [37], which was consistent with some part of our results. However, it is worth mentioning that from our results, we can see that

GSR would affect the classification results when some certain brain atlases (AAL and Dosenbach) were used. Therefore, in the study of MDD classification using fMRI-FC, the brain atlas should be carefully selected and confirmed whether to use GSR to preprocess the data.

4.5 Limitation

There is no doubt that this study has a few limitations. First, we chose SVM as the only classifier. With the increasing data in future, SVM might not be able to do this job very well. Recently, deep learning has also been used to predict the efficacy of antidepressants based on fMRI with encouraging results [27]. Therefore, we have reasons to believe that in order to improve the accuracy of classification, future research should apply deep learning to MDD classification. Second, our search range for λ is [0.02, 0.08]. Because we found that the value still cannot be selected when λ exceeds this range. We chose this range in order to decrease the calculation cost, which the random effect cannot be completely ignored even if the probability of being selected with $\lambda = 0.04$ exceeds 70%.

5. CONCLUSION

This study systematically analyzes the impact of GSR and atlas on the classification in MDD based on brain network.

We found that the selection of brain atlas is the key factor affecting the classification performance. Whether GSR affect the classification results depends on the selection of brain templates; different brain templates have different classification ability. In this experiment, the AAL network has the best classification performance.

ACKNOWLEDGMENTS

This work was funded in part by the Zhejiang Natural Science Foundation (LSY19A010001, LQ18A010003), in part by general research program of the Health Commission of Zhejiang Province (2015KYB047), in part by Jiangxi Provincial Department of Education (GJJ181252). I would like to thank Dr. Han Zhang from ShanghaiTech University in China for his technical support.

REFERENCES

- GBD 2017 Disease and Injury Incidence and Prevalence Collaborators "Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017," *Lancet* **392**, 1789–1858 (2018).
- D. A. Regier, W. E. Narrow, D. E. Clarke, H. C. Kraemer, S. J. Kuramoto, E. A. Kuhl, and D. J. Kupfer, "DSM-5 field trials in the United States and Canada, Part II: test-retest reliability of selected categorical diagnoses," *Am. J. Psychiatry* **170**, 59–70 (2013).
- S. Gao, V. D. Calhoun, and J. Sui, "Machine learning in major depression: from classification to treatment outcome prediction," *CNS Neurosci. Ther.* **24**, 1037–1052 (2018).
- K. Yoshida, Y. Shimizu, J. Yoshimoto, M. Takamura, G. Okada, Y. Okamoto, S. Yamawaki, and K. Doya, "Prediction of clinical depression scores and detection of changes in whole-brain using resting-state functional MRI data with partial least squares regression," *PLoS One* **12**, e0179638 (2017).
- X. Zhong, H. Shi, Q. Ming, D. Dong, X. Zhang, L. L. Zeng, and S. Yao, "Whole-brain resting-state functional connectivity identified major depressive disorder: a multivariate pattern analysis in two independent samples," *J. Affect. Disord.* **218**, 346–352 (2017).
- X. Wang, Y. Ren, and W. Zhang, "Depression disorder classification of fMRI data using sparse low-rank functional brain network and graph-based features," *Comput. Math. Methods Med.* **2017**, 3609821 (2017).
- P. Fang, L. L. Zeng, H. Shen, L. Wang, B. Li, L. Liu, and D. Hu, "Increased cortical-limbic anatomical network connectivity in major depression revealed by diffusion tensor imaging," *PLoS One* **7**, e45972 (2012).
- L. L. Zeng, H. Shen, L. Liu, L. Wang, B. Li, P. Fang, Z. Zhou, Y. Li, and D. Hu, "Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis," *Brain* **135**, 1498–1507 (2012).
- A. T. Drysdale, L. Grosenick, J. Downar, K. Dunlop, F. Mansouri, Y. Meng, R. N. Fetcho, B. Zebley, D. J. Oathes, A. Etkin, and A. F. Schatzberg, "Resting-state connectivity biomarkers define neurophysiological subtypes of depression," *Nat. Med.* **23**, 28–38 (2017).
- H. He, J. Sui, Y. Du, Q. Yu, D. Lin, W. C. Drevets, J. B. Savitz, J. Yang, T. A. Victor, and V. D. Calhoun, "Co-altered functional networks and brain structure in unmedicated patients with bipolar and major depressive disorders," *Brain Struct. Funct.* **222**, 4051–4064 (2017).
- J. Wang, L. Wang, Y. Zang, H. Yang, H. Tang, Q. Gong, Z. Chen, C. Zhu, and Y. He, "Parcellation-dependent small-world brain functional networks: a resting-state fMRI study," *Hum. Brain Mapp.* **30**, 1511–1523 (2009).
- K. Murphy and M. D. Fox, "Towards a consensus regarding global signal regression for resting state functional connectivity MRI," *Neuroimage* **154**, 169–173 (2017).
- M. D. Fox, D. Zhang, A. Z. Snyder, and M. E. Raichle, "The global signal and observed anticorrelated resting state brain networks," *J. Neurophysiol.* **101**, 3270–3283 (2009).
- K. Murphy, R. M. Birn, D. A. Handwerker, T. B. Jones, and P. A. Bandettini, "The impact of global signal regression on resting state correlations: are anti-correlated networks introduced," *Neuroimage* **44**, 893–905 (2009).
- J. Li, R. Kong, R. Liégeois, C. Orban, Y. Tan, N. Sun, A. J. Holmes, M. R. Sabuncu, T. Ge, and B. T. Yeo, "Global signal regression strengthens association between resting-state functional connectivity and behavior," *Neuroimage* **196**, 126–141 (2019).
- K. Dadi, M. Rahim, A. Abraham, D. Chyzyk, M. Milham, B. Thirion, and G. Varoquaux, "Benchmarking functional connectome-based predictive models for resting-state fMRI," *Neuroimage* **192**, 115–134 (2019).
- C. G. Yan, X. Chen, L. Li, F. X. Castellanos, T. J. Bai, Q. J. Bo, J. Cao, G. M. Chen, N. X. Chen, W. Chen, and C. Cheng, "Reduced default mode network functional connectivity in patients with recurrent major depressive disorder," *Proc. Natl. Acad. Sci. USA* **116**, 9078–9083 (2019).
- K. J. Friston, S. Williams, R. Howard, R. S. Frackowiak, and R. Turner, "Movement-related effects in fMRI time-series," *Magn. Reson. Med.* **35**, 346–355 (1996).
- Z. Zhou, X. Chen, Y. Zhang, D. Hu, L. Qiao, R. Yu, P. T. Yap, G. Pan, H. Zhang, and D. Shen, "A toolbox for brain network construction and classification (BrainNetClass)," *Hum. Brain Mapp.* **41**, 2808–2826 (2020).
- N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot, "Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain," *Neuroimage* **15**, 273–289 (2002).
- D. N. Kennedy, N. Lange, N. Makris, J. Bates, J. Meyer, and V. S. Caviness Jr, "Gyri of the human neocortex: an MRI-based analysis of volume and variance," *Cereb. Cortex* **8**, 372–384 (1998).
- R. C. Craddock, G. A. James, P. E. Holtzheimer 3rd, X. P. Hu, and H. S. Mayberg, "A whole brain fMRI atlas generated via spatially constrained spectral clustering," *Hum. Brain Mapp.* **33**, 1914–1928 (2012).
- N. U. Dosenbach, B. Nardos, A. L. Cohen, D. A. Fair, J. D. Power, J. A. Church, S. M. Nelson, G. S. Wig, A. C. Vogel, C. N. Lessov-Schlaggar, and K. A. Barnes, "Prediction of individual brain maturity using fMRI," *Science* **329**, 1358–1361 (2010).
- J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig, K. A. Barnes, J. A. Church, A. C. Vogel, T. O. Laumann, F. M. Miezin, B. L. Schlaggar, and S. E. Petersen, "Functional network organization of the human brain," *Neuron* **72**, 665–678 (2011).
- X. Chen, H. Zhang, Y. Gao, C. Y. Wee, G. Li, and D. Shen, "High-order resting-state functional connectivity network for MCI classification," *Hum. Brain Mapp.* **37**, 3282–3296 (2016).
- A. Etkin, K. E. Prater, A. F. Schatzberg, V. Menon, and M. D. Greicius, "Disrupted amygdalar subregion functional connectivity and evidence of a compensatory network in generalized anxiety disorder," *Arch. Gen. Psychiatry* **66**, 1361–1372 (2009).
- K. P. Nguyen, C. C. Fatt, A. Treacher, C. Mellema, M. H. Trivedi, and A. Montillo, "Predicting response to the antidepressant bupropion using pretreatment fMRI," *Predict. Intell. Med.* **11843**, 53–62 (2019).
- C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.* **2**, 1–27 (2011).
- M. Xia, J. Wang, and Y. He, "BrainNet Viewer: a network visualization tool for human brain connectomics," *PLoS One* **8**, e68910 (2013).
- A. Weissenbacher, C. Kasess, F. Gerstl, R. Lanzenberger, E. Moser, and C. Windischberger, "Correlations and anticorrelations in resting-state functional connectivity MRI: a quantitative comparison of preprocessing strategies," *Neuroimage* **47**, 1408–1416 (2009).
- R. M. Birn, "The role of physiological noise in resting-state functional connectivity," *Neuroimage* **62**, 864–870 (2012).

- ³² J. D. Power, A. Mitra, T. O. Laumann, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, "Methods to detect, characterize, and remove motion artifact in resting state fMRI," *Neuroimage* **84**, 320–341 (2014).
- ³³ C. G. Yan, B. Cheung, C. Kelly, S. Colcombe, R. C. Craddock, A. Di Martino, Q. Li, X. N. Zuo, F. X. Castellanos, and M. P. Milham, "A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics," *Neuroimage* **76**, 183–201 (2013).
- ³⁴ C. C. Guo, F. Kurth, J. Zhou, E. A. Mayer, S. B. Eickhoff, J. H. Kramer, and W. W. Seeley, "One-year test-retest reliability of intrinsic connectivity network fMRI in older adults," *Neuroimage* **61**, 1471–1483 (2012).
- ³⁵ J. Song, A. S. Desphande, T. B. Meier, D. L. Tudorascu, S. Vergun, V. A. Nair, B. B. Biswal, M. E. Meyerand, R. M. Birn, P. Bellec, and V. Prabhakaran, "Age-related differences in test-retest reliability in resting-state brain functional connectivity," *PLoS One* **7**, e49847 (2012).
- ³⁶ S. J. Gotts, Z. S. Saad, H. J. Jo, G. L. Wallace, R. W. Cox, and A. Martin, "The perils of global signal regression for group comparisons: a case study of Autism Spectrum Disorders," *Front. Hum. Neurosci.* **7**, 356 (2013).
- ³⁷ Z. Long, J. Huang, B. Li, Z. Li, Z. Li, H. Chen, and B. Jing, "A comparative atlas-based recognition of mild cognitive impairment with voxel-based morphometry," *Front Neurosci.* **12**, 916 (2018).