

## Research article

# Magnetic resonance imaging study of gray matter in schizophrenia based on XGBoost

Wang Yu<sup>1,\*</sup>, Zhang Na<sup>1</sup>, Yan Fengxia<sup>2</sup>, Gao Yanping<sup>1</sup>

<sup>1</sup> Beijing Key Laboratory of Big Data Technology for Food Safety, School of Computer and Information Engineering, Beijing Technology and Business University, Beijing, 100048, China

<sup>2</sup> School of Science, National University of Defense Technology, Changsha, 410073, China

\*Correspondence: wangyu@btbu.edu.cn (Wang Yu)

<https://doi.org/10.31083/j.jin.2018.04.0410>

## Abstract

Brain structural abnormalities of schizophrenia subjects are often considered as the main neurobiological basis of this brain disease. Therefore, with the rapid development of artificial intelligence and medical imaging technologies, machine learning and structural magnetic resonance imaging have often been applied to computer-aided diagnosis of brain diseases such as schizophrenia, Alzheimer, glioma segmentation, etc. In this paper, statistical analysis of schizophrenic and normal subjects is initially made. Additionally, a slicing and weighted average method is proposed for gray matter images of the structural magnetic resonance imaging stored as three-dimensional volume data. Grey-level co-occurrence matrix texture features from the previously processed gray matter images of structural magnetic resonance imaging are then extracted and normalized. Finally, an eXtreme Gradient Boosting classifier is used for schizophrenia classification. Experiments employed 100 schizophrenic subjects and 100 normal controls. Results show the proposed method improves the respective classification accuracy of healthy controls and schizophrenic subjects by 8% and 10.6% of the area under the receiver operating characteristic. This suggests that the textural features of gray matter changes may be of diagnostic value in schizophrenia.

## Keywords

Schizophrenia; structural magnetic resonance imaging; feature extraction; classifier

Submitted: October 30, 2017; Accepted: December 18, 2017

## 1. Introduction

Schizophrenia (SCZ) is a chronic mental disorder often characterized by abnormal social behavior and a diverse range of symptoms. Traditional medical diagnosis of SCZ is mostly based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), International Classification of Diseases (ICD-10), and the diagnostic criteria for classification of mental disorders [1, 2]. With the rapid development of technologies such as state-of-the-art brain imaging, especially magnetic resonance imaging (MRI) [3, 4], mass data can be used to assist the diagnosis of SCZ [5–7]. At the same time, neuroimaging studies have demonstrated their clinical value by using machine learning and image processing methods to individually distinguish SCZ subjects from normal controls (NC).

Several meta-analyses on structural magnetic resonance imaging (sMRI) have been conducted to identify brain regions that exhibit pathological changes that could potentially act as disease markers for SCZ. These findings include abnormalities in the brain gray matter, temporal and parietal lobes, etc [8–11]. Much research has been based on the above findings. Wang et al [12] performed feature extraction in the region of interest (ROI) after analyzing brain gray matter images SCZ subjects. Subsequently, a support vector machine (SVM) was used to distinguish and classify SCZ subjects and HCs. In a machine learning for signal processing competition (MLSP) [13], Solin & Särkkä [14] received first place for automatically diagnosing subjects with SCZ based on multimodal features derived from their

brain MRI scans. Sui et al [15] combined functional MRI (fMRI) and sMRI features, and improved the classification accuracy between SCZ subjects and NCs. Lu et al [16] employed a voxel-based morphometry (VBM) method, ROI analysis on gray matter images of SCZ, and the classification of SCZ and NC was completed by a SVM classifier.

Besides the sMRI based SCZ classification, there has been growing interest in the use of machine learning classifiers for analyzing functional magnetic resonance imaging (fMRI) data. A discriminative model of multivariate pattern classification, based on fMRI and an anatomical template is presented in [17]. In [18] SCZ characterization is improved by a multiple kernel learning (MKL) based methodology which uses both magnitude and phase fMRI data. It also detects brain regions that convey most of the discriminative information between subjects and controls. A hybrid machine learning method to classify SCZ subjects and HCs is proposed in [19] which uses fMRI and single nucleotide polymorphism (SNP) data. Experimental results show that combining genetic and fMRI data is an effective way to reassess biological classification of individuals with SCZ.

It can be seen that the feature extraction and classification between SCZ and NC based on MRI scans are well completed with the help of computer-aided diagnosis techniques. The objective of this work is to distinguish SCZ diseases from controls focused on the gray matter of sMRI images. Here, gray matter images were sliced and then weighted and averaged images were obtained. The grey-

level co-occurrence matrix (GLCM) texture features were extracted and normalized. Finally, a classic machine learning classifier such as extreme Gradient Boosting (XGBoost, KNN), SVM, LR, and a GB classifier were applied to the data obtained from the classification experiment. The motivation was to develop a machine classification process for evaluation of the performance of different classifiers for this problem, in terms of a statistical performance measure. Fully-automatic determination of whether subjects are sick can then provide necessary biological reference indicators for doctors.

## 2. Method

### 2.1. Image preprocessing

Brain MRI data is typically stored in a three-dimensional format. In this study, gray matter sMRI images ( $96 \times 113 \times 94$  voxels) were analyzed. If features are extracted from each voxel a dimensional disaster will occur as model performance will be greatly reduced by the extraction of a large number of irrelevant or redundant features. Therefore, here a preprocessing method that includes the slicing and calculation of weighted sums of average gray images is proposed. The detailed steps include:

(1) Slicing original image: For each subject, the size of the gray matter image was  $96 \times 113 \times 94$  voxels. And the volume image was sliced in the Z-axis to give 94 slices.

(2) Selecting and converting to gray image: By removing 10 slices (the first and last five slices in the image stack) which did not include feature information, and converting the remaining slices into gray images, sequentially numbered slices were obtained and denoted as  $a_i$  ( $i = 0, 1, 2, \dots, 83$ ). A part of the sliced and grayed image slices from the first subject numbered NC001 are shown in Fig. 1.

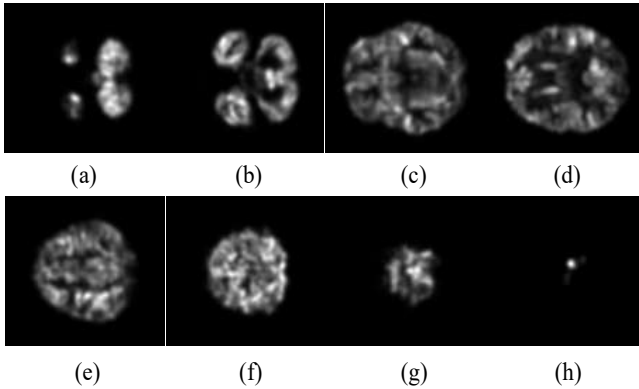


Fig. 1. Examples of sliced gray matter images (from (a) to (h) respectively the 13th, 20th, 34th, 48th, 62nd, 69th, 76th, 83rd slices)

(3) Weighting and averaging the gray image: Different weights are given according to the information contained in each slice. Here, the ordered slices are divided into three groups, and the weighted average gray image (Img1, Img2, and Img3) is calculated for each group. The calculations are:

$$Img1 = \frac{m_0 \times 1 + m_1 \times 2 + m_2 \times 3 + \dots + m_{27} \times 28}{1 + 2 + 3 + \dots + 28}, \quad (1)$$

$$Img2 = \frac{m_{28} \times 1 + m_{29} \times 1 + m_{30} \times 1 + \dots + m_{55} \times 1}{28}, \quad (2)$$

$$Img3 = \frac{m_{29} \times 28 + m_{30} \times 27 + m_{31} \times 26 + \dots + m_{83} \times 1}{28 + 27 + 26 + \dots + 2 + 1}, \quad (3)$$

The sMRI data from each subject gives three image sets (Img1, Img2, and Img3) after preprocessing to give 600 preprocessed images from 200 subjects, which provides a data set suited for subsequent feature extraction.

### 2.2. Feature extraction

Image feature extraction is a fundamental and critical step in medical image processing. The purpose is to obtain the characteristics or attributes of samples as numerical values, symbols, and feature vectors. The resulting feature extraction directly affects classification accuracy. Because texture information in the image is not sensitive to noise, light, and color, it is this feature that is chosen for the analysis.

#### 2.2.1. Texture features based on the gray-level co-occurrence matrix

The gray-level co-occurrence matrix (GLCM) is a statistical matrix that describes the grayscale of adjacent pixels (or within a certain distance) [17, 18]. Assume the gray level of a digital image is  $N$ ,  $p(q, w)$  represents the possibility (or frequency) of the appearance of grayscale  $w$  under the condition that the starting grayscale is  $q$ , where it is assumed that  $w$  is along the direction  $\theta$  of  $q$  and the space distance is  $d$ .  $\theta$  denotes the angle between the position whose grayscale is  $q$  and the position whose grayscale is  $w$ . A GLCM contains statistical information which reflects the gray direction, interval, and amplitude variation of an image. It is calculated as:

(1) Mean:

$$Mean = \bar{x} = \sum_{q=0}^N \sum_{w=0}^N p(q, w) \times q, \quad (4)$$

The mean reflects the regularity of texture. The smaller the mean, the more disorganized the texture.

(2) Variance:

$$Variance = \sum_{q=0}^N \sum_{w=0}^N p(q, w) \times (q - \bar{x})^2, \quad (5)$$

Variance measures deviation of the pixel value from the mean. The larger the variance, the greater the change of gray scale.

(3) Entropy:

$$Entropy = - \sum_{q=0}^N \sum_{w=0}^N p(q, w) \times \ln p(q, w), \quad (6)$$

Entropy gives a measure of the information contained in an image. The greater the entropy, the more complex the texture.

(4) Contrast:

$$Contrast = \sum_{q=0}^N \sum_{w=0}^N p(q, w) \times (q - w)^2, \quad (7)$$

Contrast reflects the total amount of local gray scale changes in an image. The greater the contrast of an image, the clearer the visual effect of an image.

(5) Correlation:

$$Correlation = \sum_{q=0}^N \sum_{w=0}^N \frac{(q - Mean) \times (w - Mean) \times p(q, w)^2}{Variance}, \quad (8)$$

Correlation is a measure of the linear relationship of the gray scale. The longer the extension of the gray value in a certain direction, the greater the correlation.

(6) Homogeneity:

$$\text{Homogeneity} = \sum_{q=0}^N \sum_{w=0}^N p(q, w) \times \frac{1}{1 + (q - w)^2}, \quad (9)$$

Homogeneity is a measure the uniformity of the local gray level of an image. The more homogeneous the local gray scale, the greater the homogeneity value.

(7) Energy:

$$\text{Energy} = \sum_{q=0}^N \sum_{w=0}^N p(q, w)^2, \quad (10)$$

Energy is the measurement of uniformity of gray distribution in the image.

A texture vector representing an image can be obtained by use of the foregoing seven textural features.

### 2.2.2. Normalization

Feature vectors must be normalized so that no particular feature dominates all the others. Feature vectors are normalized to zero mean and unit variance. The normalization is performed as:

$$x'_i = \frac{x_i - \mu}{\sigma_x}, \quad (11)$$

where  $\mu$  and  $\sigma_x$  are respectively the mean and standard deviation of all the features  $x_i$ .

In summary, for each of the three weighted and averaged gray images (Img1, Img2, Img3), the above seven statistics are calculated and normalized. Thus, each subject can be represented by a vector containing 21 features.

### 2.3. Classifier

Machine learning can often train a model by use of the given data and can perform tasks such as classification, recognition, and segmentation. Machine learning algorithms can extract knowledge from data and make predictions using data. Here, k-nearest neighbor (KNN), SVM, logistic regression (LR), gradient boosting (GB), and an improved eXtreme gradient boosting (XGBoost) classifier are adopted as the methods by which to compare for medical image processing techniques.

The boosting method [19, 20], which combines the additive model (linear combination of basis functions) and the forward step-wise algorithm, is a widely used and effective statistical method for optimizing learning. A boosting method based on a decision tree is called a boosting tree. The GB model is established in the gradient descent direction of loss function which is derived from the above model of boosting method.

A GB classifier is used to train and classify the feature vector of the sample, and the complete set of algorithmic steps is:

(1) Determine the training and testing sets. 80% of the 200 samples were used as training sets, i.e.  $T = (x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)$ ,  $x_i \in X \subseteq \mathbb{R}^n, y_i \in \{0, +1\}$ . The remaining samples were used as a testing set. Let  $x_i$  denote the feature vector of each sample,  $y_i$  denotes its class label. When the sample is either SCZ or NC,  $y_i$  is 0 or +1, respectively.

(2) Confirm the loss function:

$$L(y, f(x)) = [y - f(x)]^2, \quad (12)$$

where  $f(x)$  is the fitting function of  $y$ .

(3) Initialize model variables:

$$f_0(x) = \arg \min_c \sum_{i=1}^N L(y_i, c), \quad (13)$$

where  $c$  is the constant that minimizes the loss function and represents a tree with only one root node.

(4) For each model do the following:

(a) For the  $i$ th ( $i = 1, 2, \dots, N$ ) sample, calculate the negative gradient of the loss function in the current model:

$$r_{mi} = - \left[ \frac{\partial L(y, f(x_i))}{\partial f(x_i)} \right]_{f(x)=f_{m-1}(x)}, \quad (14)$$

where  $m$  is the model number,  $m = 1, 2, \dots, M$ .  $M$  denotes the maximum value of  $m$ .  $J$  is the maximum value of  $j$ , and  $j$  is the number of leaf node area.

(b) Fit a regression tree for  $r_{mi}$  and obtain the leaf node area  $R_{mj}$  ( $j = 1, 2, \dots, J$ ) of the  $m$ th tree.

(c) The number of leaf nodes  $j$  ( $j = 1, 2, \dots, J$ ) is calculated from:

$$c_{mj} = \arg \min_c \sum_{x_i \in R_{mj}} L(y_i, f_{m-1}(x_i) + c), \quad (15)$$

In this step, linear search is used to estimate the value of the leaf node region and to minimize the loss function.

(d) Update regression tree:

$$f_m(x) = f_{m-1}(x) + \sum_{j=1}^J c_{mj} I(x \in R_{mj}), \quad (16)$$

where if  $x \in R_{mj}$ ,  $I = 1$ , or  $I = 0$ .

(5) Obtain the final regression tree model:

$$\hat{f}(x) = f_M(x) = \sum_{m=1}^M \sum_{j=1}^J c_{mj} I(x \in R_{mj}), \quad (17)$$

The XGBoost classifier is based on the GB model. The main improvement is that the loss function is constructed as a Taylor expansion and a regular term is added ( $L1$ -norm or  $L2$ -norm) in the objective function:

$$\text{Obj}^t = \sum_{i=1}^n L(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) + a, \quad (18)$$

where  $\sum_{i=1}^n L(y_i, \hat{y}_i^{(t-1)} + f_t(x_i))$  denotes the loss function,  $\Omega(f_t)$  is the regularization term, and  $a$  is a constant.

The complexity of the XGBoost model is controllable and overfitting can effectively be avoided by addition of the regular term. Additionally, this model also has the advantage of parallel processing and fast and high flexibility. Therefore, it is highly suited to MRI data.

### 2.4. Algorithm flow of the proposed method

As described above, the process of classifying SCZ MRI images is given in Table 1.

Table 1. Algorithm for the proposed method

- 1) Input the sMRI image;
- 2) Use formula (1) – (3) to preprocess the image, and to obtain the  $Img1$ ,  $Img2$ ,  $Img3$  of each subject;
- 3) For the image obtained in step 2), use formula (4) – (11) to extract texture features  $T$ ,  $T = (x_{11}, x_{12}, \dots, x_{1n}, y_1), (x_{21}, x_{22}, \dots, x_{2n}, y_2), \dots, (x_{N1}, x_{N2}, \dots, x_{Nn}, y_N), x_{Ni} \in X \subseteq R^n, y_i \in \{0, +1\}$ ;
- 4) Randomly divide the feature vectors into training and test sets, and put into KNN, SVM, LR, GB, XGBoost classifier for testing;
- 5) Tune parameters, and obtain the optimizing classification results.

Table 2. Description of the four classification results

Actual Group	Predicted Group	
	Normal	Abnormal
Normal	TP	FP
Abnormal	FN	TN

### 2.5. Evaluation Criteria

For binary classification, a sample is divided into positive or negative classes. Specifically, four cases will occur and are given in **Table 2**:

True positive (TP) denotes the classification result is positive in the case of clinical normality. False negative (FN) denotes the classification result is negative in the case of clinical normality. False positive (FP) denotes the classification result is positive in the case of clinical abnormality, and true negative (TN) denotes that the classification result is negative in the case of clinical abnormality.

The true positive rate (TPR) indicates the proportion of positive samples predicted by the classifier to all positive samples, and can be described by:

$$TPR = \frac{TP}{(TP + FN)} \times 100\%. \quad (19)$$

Similarly, the false positive rate (FPR) is calculated from:

$$FPR = \frac{FP}{(FP + TN)} \times 100\%. \quad (20)$$

The FPR represents the proportion of negative samples that the classifier mistakes for positive classes. The ROC is obtained by plotting FPR (X-axis) against TPR (Y-axis). Accuracy (ACC) and area under ROC (AUC) are usually used to measure the performance of the classifiers [10], where ACC is defined as:

$$ACC = \frac{(TP + TN)}{(TP + TN + FP + FN)} \times 100\%. \quad (21)$$

## 3. Experimental Results and Analysis

To verify the effectiveness and robustness of the proposed method, comparative evaluations are performed on different state-of-the-art classifiers such as KNN, SVM, LR, GB, and XGBoost. Evaluations were performed on a PC with an Intel Core i5, CPU@2.40 Ghz, speed 800 MHz, and 32 Gb RAM. The compiling environments were Matlab 2013a and Python 2.7. The detailed procedure is given in **Fig. 2**.

### 3.1. Database

MRI data was collected by the Biomedical Image Computing and Analytics Center, Department of Radiology, and Department of Psy-

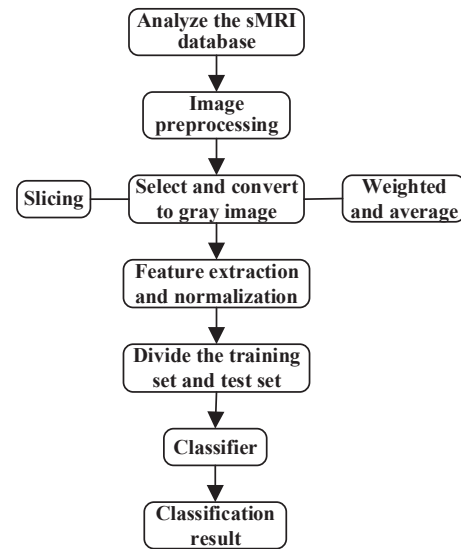


Fig. 2. Procedural flow chart of the evaluation method

chiatry, University of Pennsylvania (USA). 132 NC and 137 SCZ subjects were recruited. All subjects met DSM-IV criteria and were diagnosed as schizophrenic by psychiatrists. All sMRI image scans were acquired using a GE 3-T Signa scanner (GE Medical Systems, Milwaukee WI, USA) with the following protocol: slice thickness = 1 mm, TE = 3.2 ms, TR = 8.2 ms, flip angle = 12°, acquisition matrix = 256 × 256, FOV = 25.6 cm. All subjects remained quiet, without moving, eyes closed, no sleeping, and minimal cognitive activity during fMRI scanning. Subjects had no history of other neurological diseases or serious drug diseases. Written informed consent was obtained from all subjects before fMRI scanning. To obtain a balanced number of controls and subjects, and also simplify the study, 100 NC and 100 SCZ were randomly chosen.

The acquired MRI images were preprocessed using the statistical parametric mapping software package SPM (Wellcome Trust Centre for Neuroimaging, Institute of Neurology, London, UK, <http://www.fil.ion.ucl.ac.uk/spm>) which included the following steps: (1) skull stripping, (2) bias correction, (3) tissue segmentation into gray matter, white matter, cerebrospinal fluid, and lateral ventricles, (4) spatial registration to a Montreal Neurological Institute (MNI) template, (5) generation of the regional analysis of volumes examined in normalized space (RAVENS) [21, 22] maps of gray matter, white matter, and cerebrospinal fluid by the publicly available DRAMMS deformable registration package [23], and (6) the RAVENS maps were then smoothed using a six millimeter full width at half maximum Gaussian filter.

### 3.2. Statistical Analysis

The statistical information of 200 age and gender matched subjects is analyzed. No statistically significant is assumed for  $p < 0.05$ . Please see analytical results in **Table 3**.

### 3.3. Experimental Settings and Results

Two different training and testing sets were analyzed in this study. 80% of the 200 samples are used as training sets, and the remaining samples are used as a test set. The extracted texture features are put into each classifier in which XGBoost, GB, SVM algorithms

Table 3. Characteristics of study participants

Variable	Sample size	Gender (male/female)	Age (years) Mean/SD (range)
SCZ	100	62/38	35.50/37.08(13-60)
NC	100	49/51	34.73/34.86(17-65)
Value	—	0.064 <sup>a</sup>	0.53 <sup>b</sup>

SD = standard deviation <sup>a</sup>Pearson Chi-square test <sup>b</sup>Two-sample t-test

are related to more parameters. Parameter choice directly affected classification accuracy. Therefore, parameter tuning is particularly important. Grid search and cross validation methods were used by the classifiers for parameter tuning in which parameter K = 5 in KNN. A radial basis function (RBF) is selected as the kernel function in SVM. In XGBoost, the overall parameters were divided into three categories by the authors: general parameters that guide the overall functioning, booster parameters that guide the individual booster (tree/regression) at each step, and learning task parameters that guide the optimization performed. Parameters of GB were divided into two categories: tree-specific parameters that affect each individual tree in the model and boosting parameters that affect the boosting operation in the model. The main parameters of these classifiers are given in Table 4.

Table 4. Classifier parameters

Classifiers	Parameters
SVM	C: $\{10^{-3}, 10^{-2}, 10^{-1}, 1, 10, 100, 1000\}$ ; gamma: $\{0.001, 0.0001\}$ ;
G B	Tree min_samples_split: 9; min_samples_leaf: 1; max_depth: 9; max_features: 8;
	Boosting learning_rate: 0.01; n_estimators: 90; subsample: 0.8;
XGBoost	Booster col_sample_bytree: 0.8; min_child_weight: 1; max_depth: 5; learning_rate: 0.01; sub_sample: 0.9; gamma: 0;
	General booster: gbtree;
Learning Task objective: binary-logistic; seed: 1.	

The best classification results obtained by each classifier are given in Table 5. ROC curves of the GB and XGBoost classifier are given in Fig. 3.

Table 5. Experimental result comparison

Classifier	ACC(%)	AUC(%)
KNN	60	60.10
RBF-SVM	54	52.08
LR	52	52.24
GB	64	65.20
XGBoost	<b>72</b>	<b>75.80</b>

It can be seen from Table 4 that on the same data set the XGBoost classifier gives the highest accuracy (up to 72%) and the LR classifier returns the lowest accuracy (52%). GB and KNN classifiers are inferior to XGBoost and classification accuracy is 64% and 60%, respectively. The RBF-SVM classifier is a poor classifier for this SCZ data set. It indicates that XGBoost, which has a better generalization ability and avoids over-fitting effectively, is better than the

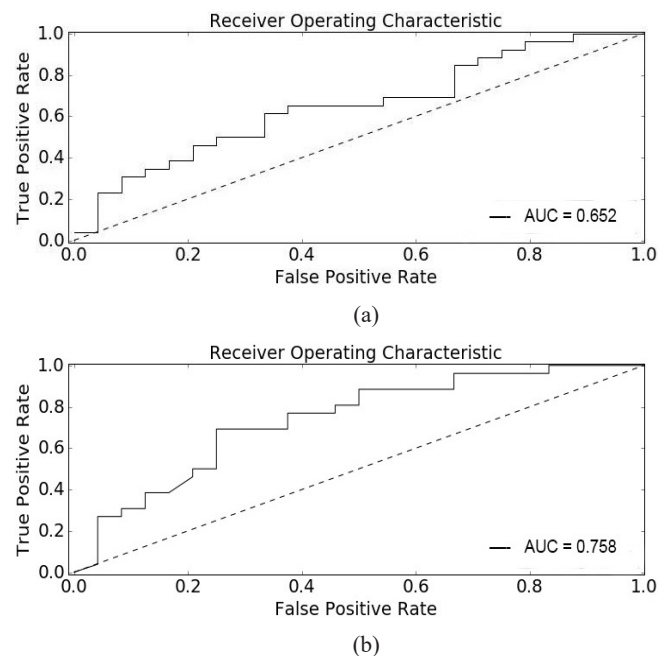


Fig. 3. ROC curve for (a) GB algorithm, and (b) XGBoost algorithm

other classifiers. XGBoost also has an advantage when dealing with irregular data.

It can be seen from Fig. 3 that, compared with the GB algorithm, a higher AUC value is obtained by using the XGBoost classifier, which shows the XGBoost classifier exhibits faster and more efficient processing. Additionally, if better experimental results are required, a more detailed analysis of the features and parameter tuning of the models must be undertaken.

## 4. Conclusion

In this paper, to better discriminate SCZ subjects from HCs, image processing and machine learning based on sMRI are employed to assist diagnosis and analysis of SCZ. Firstly, the gray matter image is preprocessed by being sliced, weighted, and averaged. GLCM texture features are then extracted and normalized. Finally, different machine learning methods are used to train and establish a binary classification model. The different results show that the XGBoost approach has superior performance when compared to KNN, SVM, LR, and GB classifiers. Further work should include two aspects. One is extracting more significant features from the data. The other, the building of more effective models, which may further improve role of computer-aided diagnosis of SCZ.

## Acknowledgments

This work is supported by the Natural Science Foundation of China (No. 61671028), Beijing Natural Science Foundation (No. 4162018), the Beijing Talents Fund (No. 2014000026833ZK14), the Importation and Development of High-Caliber Talents Project of Beijing Municipal Institutions (No. CIT&TCD201504010), and Beijing Natural Science Foundation (No. 4172015).

## Conflict of Interest

All authors declare no conflict of interest.



## References

- [1] Muming P, Xu B, Tan T (2016) Brain Science and Brain-Inspired Intelligence Technology—An Overview. *Bulletin of Chinese Academy of Sciences* **31**(7), 725-736.
- [2] Wheeler AL, Voineskos AN (2014) A review of structural neuroimaging in schizophrenia: from connectivity to connectomics. *Frontiers in Human Neuroscience* **8**(8), 653.
- [3] Tan Z LZ (2015) Advances in magnetic resonance imaging of schizophrenia. *Diagnostic Imaging & Interventional Radiology* **11**(6), 507-511.
- [4] Ferro A, Roiz-Santiañez R, de la Foz VO-G, Tordesillas-Gutiérrez D, Ayesa-Arriola R, de La Fuente-González N, Fañanás L, Brambilla P, Crespo-Facorro B (2015) A cross-sectional and longitudinal structural magnetic resonance imaging study of the post-central gyrus in first-episode schizophrenia patients. *Psychiatry Research: Neuroimaging* **231**(1), 42-49.
- [5] Bi W (2016) A Research of pattern recognition methodology for brain diseases & disorders classification with magnetic resonance imaging. *Sichuan: School of Life Science and Technology*.
- [6] Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, Nelson SM, Wig GS, Vogel AC, Lessov-Schlaggar CN (2010) Prediction of individual brain maturity using fMRI. *Science* **329**(5997), 1358-1361.
- [7] Juneja A, Rana B, Agrawal R (2014) A novel approach for computer aided diagnosis of schizophrenia using auditory oddball functional MRI. in: Proceedings of the 2014 Indian Conference on Computer Vision Graphics and Image Processing.
- [8] Gupta CN, Calhoun VD, Rachakonda S, Chen J, Patel V, Liu J, Segall J, Franke B, Zwiers MP, Arias-Vasquez A (2014) Patterns of gray matter abnormalities in schizophrenia based on an international mega-analysis. *Schizophrenia Bulletin* **41**(5), 1133-1142.
- [9] Li X, Alapati V, Jackson C, Xia S, Bertisch HC, Branch CA, DeLisi LE (2012) Structural abnormalities in language circuits in genetic high-risk subjects and schizophrenia patients. *Psychiatry Research: Neuroimaging* **201**(3), 182-189.
- [10] van Erp TG, Hibar DP, Rasmussen JM, Glahn DC, Pearlson GD, Andreassen OA, Agartz I, Westlye LT, Haukvik UK, Dale AM (2016) Subcortical brain volume abnormalities in 2028 individuals with schizophrenia and 2540 healthy controls via the ENIGMA consortium. *Molecular Psychiatry* **21**(4), 547.
- [11] Andreasen NC, Arndt S, Swayze V, Cizadlo T, Flaum M, O'leary D, Ehrhardt JC, Yuh W (1994) Thalamic abnormalities in schizophrenia visualized through magnetic resonance image averaging. *Science* **266**(5183), 294-298.
- [12] Wang L, Shen H, Li B, Hu D (2011) Classification of schizophrenic patients and healthy controls using multiple spatially independent components of structural MRI data. *Frontiers of Electrical and Electronic Engineering in China* **6**(2), 353.
- [13] Silva RF, Castro E, Gupta CN, Cetin M, Arbabshirani M, Potluru VK, Plis SM, Calhoun VD (2014) The tenth annual MLSP competition: schizophrenia classification challenge. in: Machine Learning for Signal Processing (MLSP), 2014 IEEE International Workshop 1-6.
- [14] Solin A, Särkkä S (2014) The 10th annual MLSP competition: First place. Machine Learning for Signal Processing (MLSP), 2014 IEEE International Workshop 1-3.
- [15] Sui J, Castro E, He H, Bridwell D, Du Y, Pearlson GD, Jiang T, Calhoun VD (2014) Combination of FMRI-SMRI-EEG data improves discrimination of schizophrenia patients by ensemble feature selection, in: Engineering in Medicine and Biology Society (EMBC). 2014 36th Annual International Conference of the IEEE 3889-3892.
- [16] Lu X, Yang Y, Wu F, Gao M, Xu Y, Zhang Y, Yao Y, Du X, Li C, Wu L (2016) Discriminative analysis of schizophrenia using support vector machine and recursive feature elimination on structural MRI images. *Medicine* **95**(30), e3973.
- [17] Shi F, Liu Y, Jiang T, Zhou Y, Zhu W, Jiang J, Liu H, Liu Z (2007) Regional homogeneity and anatomical parcellation for fMRI image classification: application to schizophrenia and normal controls. in: International Conference on Medical Image Computing and Computer-Assisted Intervention, pp.136-143.
- [18] Castro E, Gómez-Verdejo V, Martínez-Ramón M, Kiehl KA, Calhoun VD (2014) A multiple kernel learning approach to perform classification of groups from complex-valued fMRI data analysis: application to schizophrenia. *NeuroImage* **87**(3), 1-17.
- [19] Yang H, Liu J, Sui J, Pearlson G, Calhoun VD (2010) A hybrid machine learning method for fusing fMRI and genetic data: combining both improves classification of schizophrenia. *Frontiers in Human Neuroscience* **4**(192), 192.
- [20] Selvaraj H, Selvi ST, Selvathi D, Gewali L (2007) Brain MRI slices classification using least squares support vector machine. *International Journal of Intelligent Computing in Medical Sciences & Image Processing* **1**(1), 21-33.
- [21] Sulochana S, Vidhya R (2013) Texture based image retrieval using framelet transform-gray level co-occurrence matrix (GLCM). *International Journal of Advanced Research in Artificial Intelligence* **2**(2), 68-73.
- [22] Li H (2012) *Statistical learning Method*. Beijing: Tsinghua University Press.
- [23] Zhou Z (2016) *Machine learning*. Beijing: Tsinghua University Press.
- [24] Goldszal AF, Davatzikos C, Pham DL, Yan MX, Bryan RN, Resnick SM (1998) An image-processing system for qualitative and quantitative volumetric analysis of brain images. *Journal of Computer Assisted Tomography* **22**(5), 827-837.
- [25] Davatzikos C, Genc A, Xu D, Resnick SM (2001) Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. *NeuroImage* **14**(6), 1361-1369.
- [26] Ou Y, Sotiras A, Paragios N, Davatzikos C (2011) DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting. *Medical Image Analysis* **15**(4), 622-639.