Front. Electr. Electron. Eng. China 2011, 6(2): 353–362
DOI 10.1007/s11460-011-0142-2

RESEARCH ARTICLE

Lubin WANG, Hui SHEN, Baojuan LI, Dewen HU

Classification of schizophrenic patients and healthy controls using multiple spatially independent components of structural MRI data

Abstract Several meta-analyses were recently conducted in attempts to identify the core brain regions exhibiting pathological changes in schizophrenia, which could potentially act as disease markers. Based on the findings of these meta-analyses, we developed a multivariate pattern analysis method to classify schizophrenic patients and healthy controls using structural magnetic resonance imaging (sMRI) data. Independent component analysis (ICA) was used to decompose gray matter density images into a set of spatially independent components. Spatial multiple regression of a region of interest (ROI) mask with each of the components was then performed to determine pathological patterns, in which the voxels were taken as features for classification. After dimensionality reduction using principal component analysis (PCA), a nonlinear support vector machine (SVM) classifier was trained to discriminate schizophrenic patients from healthy controls. The performance of the classifier was tested using a 10-fold cross-validation strategy. Experimental results showed that two distinct spatial patterns displayed discriminative power for schizophrenia, which mainly included the prefrontal cortex (PFC) and subcortical regions respectively. It was found that simultaneous usage of these two patterns improved the classification performance compared to using either of them alone. Moreover, the two pathological patterns constitute a prefronto-subcortical network, suggesting that schizophrenia involves abnormalities in networks of brain regions.

Keywords schizophrenia, discriminative analysis, gray matter network, independent component analysis (ICA), support vector machine (SVM)

Received August 10, 2010; accepted January 27, 2011

Lubin WANG, Hui SHEN, Baojuan LI, Dewen HU (✉)
College of Mechatronics and Automation, National University of Defense Technology, Changsha 410073, China
E-mail: dwhu@nudt.edu.cn

1 Introduction

Schizophrenia is a complex mental disorder characterized by a diverse range of symptoms, including altered perception, cognition, language, motor activity, emotion, and so on [1]. Numerous structural magnetic resonance imaging (sMRI) studies have identified subtle but significant brain changes in schizophrenia. Morphological features reported to be associated with the illness include ventricular enlargement, total brain volume deficits, and gray and white matter deficits in temporal, frontal, parietal, and subcortical regions [2]. Recently, some meta-analyses revealed that patients with schizophrenia have structural abnormalities in a distributed network of brain regions [3–5]. These studies not only improve our understanding of the pathology of this mental disorder but also indicate that sMRI may be a useful tool for the clinical diagnosis of schizophrenia.

The most popular methods for detecting structural abnormalities in various mental disorders including schizophrenia are region of interest (ROI) analysis and voxel-based morphometry (VBM) [6]. ROI analysis has been traditionally used to investigate regional differences between groups of subjects. Such analyses rely on a priori defined brain regions of interest. However, structural abnormalities might be part of an ROI, or span over multiple ROIs, thus potentially reducing statistical power of the underlying morphological analysis [7]. VBM involves a voxel-wise comparison of the density of gray or white brain matter between groups. It is an automated, efficient whole-brain analysis to detect structural changes in mental disorders, and it is therefore a useful initial approach to the structural images [3]. However, VBM is a univariate method. It will detect voxels reaching a prespecified criterion and does not utilize any information about the relationships among voxels. Even when voxels in a spatially distributed brain region individually

meet a certain criterion, these spatially distinct regions may not carry similar information. For example, it is possible that the gray matter densities of two voxels are both deficient in schizophrenia, yet they may be uncorrelated with each other. In contrast, multivariate pattern analysis approach provides a way to pool information across different voxels as well as often extracting important novel information from neuroimaging data [8].

Independent component analysis (ICA) is a powerful statistical and computational multivariate pattern analysis technique that attempts to discover hidden factors underlying sets of mixed signals. It has been successfully applied to detect brain activation maps or brain networks in functional magnetic resonance imaging (fMRI) analysis and can separate either spatial or temporal independent sources [9–11]. Xu et al. also utilized ICA to decompose gray matter segmentation images into source images [12]. In their study, using a two-sample t-test, sources that demonstrated significant different loading parameters in patients and controls were determined as significant sources associated with schizophrenia. However, the loading parameters that express the contribution of every source image to the subjects may not obey a Student's t distribution. Thus, some irrelevant sources, such as artifacts, may be identified, and some structural differences that do not reach a pre-specified significance level may be discarded. In addition, ICA is a group-level analysis approach. It cannot distinguish normal from abnormal at individual subject level and therefore cannot be applied directly for clinical diagnose purposes.

In the present study, we integrated ICA and support vector machine (SVM) to distinguish schizophrenic patients from healthy controls at individual subject level. The hypothesis was that brain abnormalities involved in schizophrenia would exhibit specific spatial patterns independent from various artifacts and sources of noise. ICA was used to separate brain abnormalities in schizophrenia from artifacts and noise. The pathological components were identified in an automated manner by spatial multiple regression, which has previously been used to determine brain functional networks in fMRI analysis [13]. Then, a nonlinear SVM classifier was trained to differentiate schizophrenic patient from healthy controls. Because of the limited number of samples in this study, a 10-fold cross-validation strategy was used to test the generalization ability of the classifier.

The remainder of this paper is organized as follows. In the materials section, we introduce the data acquisition as well as the preprocessing. In the methodology section, we detail the three steps in the methodology, i.e., decomposition of sMRI gray matter density images, feature selection and extraction, and classifier design. The experimental results of our method are presented in the results section. The discussions are in the discussion section followed by the conclusion section.

2 Materials and methods

2.1 Subjects

Participants were 32 right-handed patients with schizophrenia and 32 right-handed healthy controls (25 males and 7 females in the patient group, and 21 males and 11 females in the control group). All patients were recruited from outpatient departments and inpatient units at the Department of Psychiatry, Second Xiangya Hospital of Central South University. Diagnoses of schizophrenia were determined by the attending psychiatrists using the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV; American Psychiatric Association, 1994) criteria for schizophrenia. The healthy controls were recruited using advertisements. The symptoms of these patients were assessed with the Positive and Negative Symptom Scale (PANSS) at the time of scanning, and the total PANSS score was 80.06 ± 16.55. The mean duration of illness was 27.2 ± 16.7 months. The two groups exhibit no significant differences in age (24.00 ± 5.66 years for patients and 22.52 ± 4.2 years for controls, $p = 0.84$) and length of education (11.15 ± 2.50 years for patients and 13.32 ± 3.50 years for controls, $p = 0.76$). No participants had suffered major head trauma or had a history of alcohol or drug dependence or neurological disorders. Written informed consent was obtained from all participants.

2.2 Overview of our methods

Our method included the following major steps: 1) All sMRI data were preprocessed to obtain gray matter density images; 2) ICA was then utilized to decompose gray matter density images into a set of spatial components; 3) Spatial multiple regression of an ROI mask with each component was performed to determine pathological patterns; 4) Pathological patterns were concatenated into a single image; 5) Voxels from this concatenated image were selected as features for classification, and principal component analysis (PCA) was used for dimensionality reduction; and 6) An SVM classifier was trained on the dimension-reduced feature space to classify test samples into patient and control groups. A flowchart outlining the analysis stream of our method is presented in Fig. 1.

The classification algorithm for differentiating schizophrenic patients and healthy controls is also presented here.

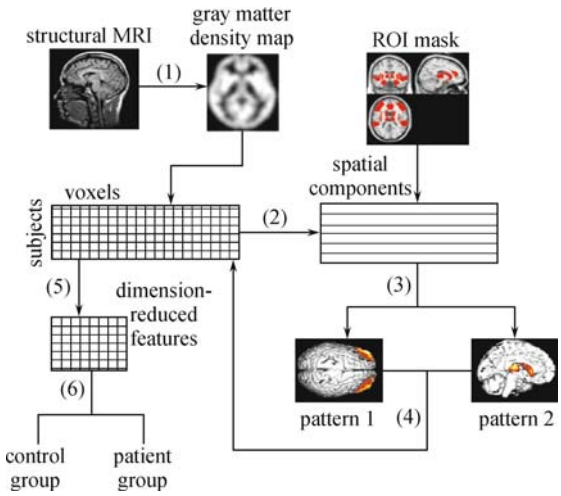

Fig. 1 Flowchart outlining the analysis stream of our method, which involves the following techniques: (1) data preprocessing, (2) ICA, (3) spatial multiple regression, (4) data fusion, (5) feature selection and extraction, and (6) SVM

Classification Algorithm

1: randomly divide n subjects into k groups (e.g., $k = 10$)
2: for $i = 1$ to k do
3: select k–1 groups as training data, the remaining group as test data
4: train a SVM classifier f
5: for each test sample x, if $f(x) > 0$, $x \in$ controls; else, $x \in$ patients
6: repeat each group as test data in turn
7: end
8: output the classification accuracy

2.3 Data acquisition and preprocessing

Magnetic resonance imaging (MRI) data were acquired on a 1.5 T GE Signa System (GE Signa, Milwaukee, Wisconsin, USA). The whole brain was evaluated with T1-weighted spoiled gradient echo pulse sequence in the sagittal plane using the following imaging parameters: repetition time = 12 s, field of view = 24 cm, echo time = 4.2 ms, flip angle =15°, matrix = 256 × 256 × 172, slice thickness = 1.8 mm, no gap.

Images were preprocessed using the statistical parametric mapping software package (SPM2, Wellcome Department of Cognitive Neurology, Institute of Neurology, London, UK, http://www.fil.ion.ucl.ac.uk/spm). All the MRI data were first normalized to the 152 average T1 template in Montreal Neurological Institute space. The normalized images were resliced to a voxel size of 3 mm × 3 mm × 3 mm and then were segmented into gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF) using a modified mixture model cluster analysis technique [14]. The GM images were spatially smoothed with an isotropic Gaussian filter of 12 mm full-width half-maximum kernel to obtain GM density maps. After removing non-brain voxels, these GM density images were flattened to 1-D vectors and placed along the columns and subjects along the rows to construct a 2-D matrix.

2.4 ICA on gray matter density images

ICA is a data-driven method for revealing hidden factors that underlie mixed signals. In our study, it was employed to determine the spatial distribution of structural changes in schizophrenia. We hypothesized that brain abnormalities involved in schizophrenia would exhibit specific spatial patterns independent from various artifacts and sources of noise. The gray matter density images were then modeled as a linear combination of spatial independent source images. Based on such a model, each subject could be considered a mixed signal, with voxels as observations. We used ICA to extract a set of spatial patterns that were maximally independent from each other and expected to find patterns that could represent structural differences between patients and controls. Let $X_1, X_2, \ldots, X_n$ denote n subjects, and let $s_1, s_2, \ldots, s_m$ denote m independent components. The ICA model could be expressed by the following equation:

$$X_i = a_{i1}s_1 + a_{i2}s_2 + \cdots + a_{im}s_m, \quad (1)$$

for all $i = 1, 2, \ldots, n$, where a_{ij} were some real coefficients, $j = 1, 2, \ldots, m$.

The Bayesian information criterion (BIC) was used to estimate the number of components [15]. The formula for the BIC is

$$\mathrm{BIC} = -2 \ln L + k \ln n, \quad (2)$$

where L denotes the maximized value of the likelihood function for the estimated model, k is the number of free parameters to be estimated, and n is the sample size. It is possible to increase the likelihood by adding parameters, which may result in overfitting. The BIC resolves this problem by introducing a penalty term for the number of parameters in the model.

We performed ICA using the infomax algorithm [16]. All independent components were scaled to z-scores. A z-threshold criterion ($z > 2$) was applied to select voxels contributing most strongly to each component [9]. The resulting ICA components were then reconverted into 3-D images. These images could be viewed as a set of independent spatial patterns in GM density maps.

2.5 Feature selection and extraction

An ROI mask of abnormal brain regions in schizophrenia was created based on previous meta-analysis studies. Spatial multiple regression of the ROI mask was performed with each of the components. Components with

relative high regression values were considered pathological patterns. This method has previously been used to select brain functional networks in fMRI analysis [13]. These pathological patterns were concatenated into a single image. Voxels from this image were selected as original features for classification to filter out noise and artifacts.

The number of voxels in the concatenated image was far larger than the number of samples, n. Dimensionality reduction was thus necessary to represent this ill-posed data set in a low-dimensional space, which also decreased the computational complexity of the classifier. PCA is commonly used for data decomposition and dimensionality reduction, attempting to find a set of principle components that capture most of the variance in the original data. In the present work, all n–1 principle components were reserved in the PCA step. That is, the constructed dimension-reduced subspace was only a change of the coordinate system without loss of information.

2.6 Classifier design

Let $\boldsymbol{x}_i$ denote the feature vector of the ith subject after feature selection and extraction ($\boldsymbol{x}_i \in \mathbb{R}^{n-1}, i = 1, 2, \ldots, n$), y_i denote its class label (+1 for control group, −1 for patient group). The main task remained was to train a classifier f from the training samples and predict the class of new samples.

SVM training is to find a decision function $f(\boldsymbol{x}) = \boldsymbol{\omega}'\phi(\boldsymbol{x})$ by solving the following optimization problem, where $\boldsymbol{\omega}$ is the normal of the hyperplane and ϕ is the feature map by some kernel function:

$$\begin{aligned} \min_{\boldsymbol{\omega},\boldsymbol{\xi}} \quad & \frac{1}{2}\|\boldsymbol{\omega}\|^2 - \rho + \frac{C}{2}\sum_{i=1}^{n}\xi_i^2, \\ \text{s.t.} \quad & y_i\boldsymbol{\omega}'\phi(\boldsymbol{x}) \geqslant \rho - \xi_i, \end{aligned} \tag{3}$$

where ξ_i is a measure of the misclassification errors for non-separable cases, and C trades off the empirical risk and model complexity. After mapping, the feature vectors into a high-dimensional feature space via a kernel function, SVM constructs an optimal separating hyperplane in this high-dimensional space, which equips SVM with high generalization ability. In our study, the sigmoid function was used in SVM as the kernel function, which was defined as

$$\phi(\boldsymbol{x}_1, \boldsymbol{x}_2) = \tanh(p_1(\boldsymbol{x}_1 \cdot \boldsymbol{x}_2) - p_2),$$

where $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are two feature vectors, p_1 and p_2 are the scale and offset parameters of the sigmoid kernel.

2.7 Evaluating the performance of our method

Because of the limited number of samples in our study, we used a cross-validation strategy to evaluate the performance of our study. Leave-one-out cross-validation (LOOCV) strategy is often used to estimate the generalization ability of supervised classifiers in cases where sample sizes are small. Suppose there are n samples in total. LOOCV trains classifier n times, each time leaving one sample out, training on the remaining ones, and making a prediction for this sample. However, this procedure can be computationally expensive. In the present study, we used a k-fold cross-validation procedure, where k is the number of parts into which the data set is divided [8]. Here, we specified $k = 10$, corresponding to leaving 10% of the samples out as test data on each fold.

The performance of our classifier was then quantified by sensitivity, specificity, and classification accuracy:

$$\text{sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{4}$$

$$\text{specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \tag{5}$$

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}}, \tag{6}$$

where TP, TN, FP, and FN denote the number of patients correctly predicted, controls correctly predicted, controls classified as patients, and patients classified as controls, respectively.

3 Results

3.1 Creation of ROI

The ROI mask was formed based on three meta-analyses of brain abnormalities in schizophrenia: (A) Honea et al., 2005; (B) Ellison-Wright et al., 2008; and (C) Glahn et al., 2008. They have been conducted for clearly identifying specific structural changes in patients with schizophrenia. Among them, Honea et al. mainly focused on chronic schizophrenia, while Ellison-Wright et al. sought to map gray matter changes in first-episode schizophrenia and to compare these with the changes in chronic schizophrenia. Glahn et al. did not distinguish the two types of schizophrenia. The brain regions that were referred to in the three meta-analyses are presented in Table 1. Most of these brain regions show gray matter deficits. Regions of increased gray matter density were smaller and more discrete than areas of decrease, which might be caused by the antipsychotic medication. Ellison-Wright et al. identified the putamen bilaterally showed gray matter increases in schizophrenia. Glahn et al. also included the right head of the caudate.

Table 1 shows considerable overlap of brain regions in the three studies. It can also be found that gray matter decreases of cortical regions were more extensive in chronic schizophrenia than in first-episode schizophrenia. The schizophrenic patients in our study were not

Table 1 Abnormal brain regions that were referred to in three meta-analysis studies

meta-analysis studies	a	b	c	d	e	f	g	h	i	j	k	l
A	↓	↓	↓	↓	↓	↓	↓	↓	↓			
B1		↓		↓		↓	↓	↓	↓	↑		
B2	↓	↓	↓	↓		↓	↓	↓		↑	↓	
C	↓	↓	↓	↓	↓		↓	↓	↑	↑		↓

1) The three studies are (A) Honea et al., 2005, mainly chronic schizophrenia; (B1) Ellison-Wright et al., 2008, first-episode schizophrenia; (B2) Ellison-Wright et al., 2008, chronic schizophrenia; and (C) Glahn et al., 2008, did not distinguish first-episode from chronic schizophrenia.

2) "↓" represents regions of reduced gray matter density in schizophrenia and "↑" represents regions of increased gray matter density in schizophrenia.

3) (a) superior temporal gyrus, (b) inferior frontal gyrus, (c) middle frontal gyrus, (d) insula, (e) parahippocampal gyrus, (f) hippocampus/amygdala, (g) anterior cingulated gyrus, (h) thalamus, (i) caudate, (j) putamen, (k) fusiform gyrus, and (l) postcentral gyrus.

exactly chronic or first-episode. Thus, the abnormal brains regions of the patients in our study maybe slightly different from the above results. For the purpose of robust, we picked the brain regions that were referred to at least two times. So the ROI was derived by using masks of these brain regions: superior temporal gyrus, inferior frontal gyrus, middle frontal gyrus, insula, parahippocampal gyrus, hippocampus, amygdala, anterior cingulated gyrus, thalamus, caudate, and putamen. Although previous studies have also reported structural changes of cerebellum in schizophrenia, the exact abnormal areas were not identified, so we did not add cerebellum to our ROI mask.

3.2 Pathological patterns associated with schizophrenia

Eighteen independent components were estimated from the data. The multiple regression values of these components are shown in Fig. 2. Components 3 and 6 exhibited higher regression values relative to the others and were considered pathological patterns associated with schizophrenia. For convenience, these are referred to as pattern 1 and pattern 2, respectively. Some other components also comprised part of the ROI but had large areas outside the ROI. Such components exhibited relatively small regression values and were not considered to be reliable pathological patterns in schizophrenia.

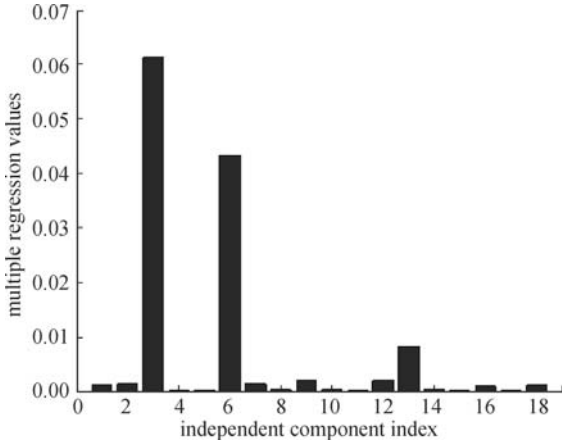

Fig. 2 Multiple regression values of all components estimated by ICA. Components 3 and 6 have high values and are considered pathological patterns

The structural changes associated with pattern 1 were primarily located in the prefrontal cortex (PFC). The corresponding Brodmann areas were mainly BA 9, 10, and 46. This pattern also included the bilateral middle temporal pole and left fusiform gyrus (see Fig. 3). The structural changes associated with pattern 2 were primarily in the subcortical regions. These areas included the bilateral caudate, bilateral thalamus, right putamen, and right hippocampus as well as the right superior occipital gyrus and right cerebellum (see Fig. 3). A detailed list of brain regions in pattern 1 and pattern 2 are presented in Table 2.

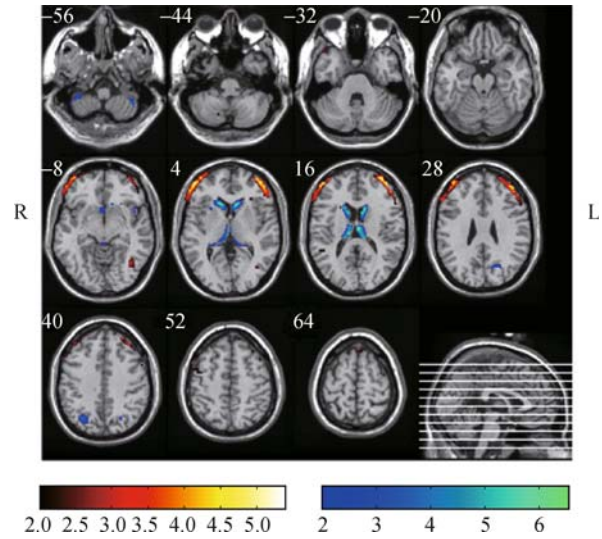

Fig. 3 Spatial distributions of two pathological patterns at a z-threshold of 2. L and R indicate the left and right hemispheres of the brain. Pattern 1 is observed predominantly in the PFC, while pattern 2 is primarily located in the subcortical regions

3.3 Classification performance

Voxels in the two pathological patterns were selected as original features for classification. After dimensionality reduction using PCA, we applied a nonlinear SVM classifier in the dimension-reduced feature space. Sigmoid function was used as the kernel function. In SVM, the regularization parameter C should be specified to trade off the empirical risk and model complexity in non-separable cases. In our study, the C parameter is selected in a range of $\{0.01, 0.1, 1, 10, \text{Inf}\}$. For our data set, SVM performed best when C was set to be 0.1. Increasing C reduced the SVM margin and classification accuracy.

Table 2 Anatomical localization, cluster size, MNI coordinates, and max z-values of pathological pattern 1 and pattern 2 at a cluster threshold of 30 voxels and a z-threshold of 2. L and R indicate the left and right hemispheres of the brain

	L/R size	L/R max $Z(x, y, z)$	Brodmann area
pathological pattern 1:			
middle frontal gyrus	404/419	5.2(−39,57,15)/5.4(42,60,0)	9, 10, 46
inferior frontal gyrus	194/271	4.9(−48,48,9)/5.2(51,48,−3)	45, 46, 47
superior frontal gyrus	72/67	4.1(−33,63,12)/3.9(33,66,3)	9, 10
middle temporal pole	38/35	3.2(−30,18,−42)/3.0(42,24,−36)	38
fusiform gyrus	32/na	3.3(−42,−60,−15)/na	37
pathological pattern 2:			
caudate	204/177	6.3(−12,18,9)/6.5(12,21,6)	—
thalamus	123/130	6.0(−6,−12,18)/5.7(9,−21,18)	—
putamen	na/69	na/2.9(21,21,−3)	—
superior occipital gyrus	na/69	na/3.7(27,−69,39)	19
hippocampus	na/31	na/4.3(15,−33,9)	—
cerebellum	na/64	na/3.6(30,−66,−63)	—

There were totally 64 samples for classification in our study, including 32 schizophrenic patients and 32 healthy controls. A 10-fold cross-validation was conducted to estimate to the prediction ability of the classifier. Each time there were 58 training samples and 6 test ones. Classification result was shown in Fig. 4. From Fig. 4, we found eight controls and three patients were misclassified. So the specificity, sensitivity, and overall classification accuracy of our method were 75.0%, 90.6%, and 82.8%, respectively.

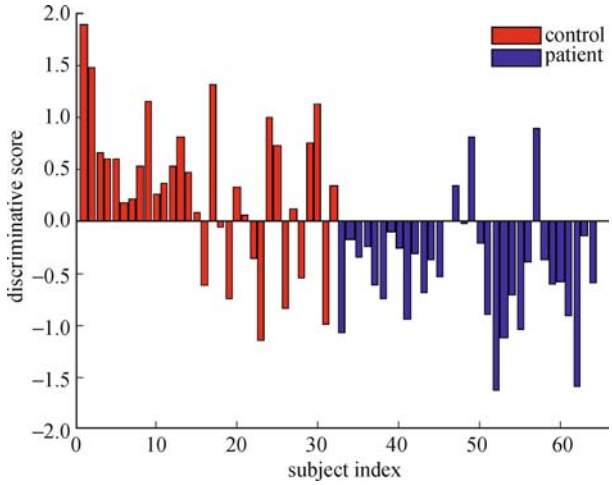

Fig. 4 Discriminative scores of all subjects for pattern 1 + 2. The subjects' permutation is randomly ordered. Controls are denoted by red bars, and patients are denoted by blue bars. Negative scores represent subjects are classified to the patient group, and positive scores represent subjects are classified to the control group

In our study, the z-threshold value is a user-specified parameter, which was used to select voxels contributing most strongly to each component. More voxels will be selected when lowering the z-threshold value. The experiments were repeated for different z-threshold values (z = 0, 1, 2, 3, 4), which are summarized in Table 3. We can see that the classification results are stable with respect to the z-threshold value. Note that the classifier reached the best performance (82.8% classification accuracy) when $z = 2$. With an increasing z-threshold value, some important features were filtered out, and the remainder features could not effectively reveal the character of structural changes in schizophrenia, the classification performance therefore gradually reduced. On the contrary, decreasing the z-threshold value would also reduce the classification performance, since more and more low discriminative features were selected, and the signal-to-noise ratio was decreased.

Table 3 Classification results of SVM on different z-threshold values

z-threshold setup	sensitivity/%	specificity/%	accuracy/%
$z = 0$	68.8	68.8	68.8
$z = 1$	75.0	87.5	81.3
$z = 2$	75.0	90.6	82.8
$z = 3$	71.9	87.5	79.7
$z = 4$	68.8	81.3	75.0

3.4 Discriminative powers of different patterns

For further evaluating the discriminative powers of these two pathological patterns, the voxels in pattern 1 and pattern 2 were selected as features for classification individually. In our study, discriminative scores of subjects were defined as the distance between them and the hyperplane. The discriminative scores of all subjects for pattern 1 and pattern 2 are presented in Fig. 5. For each pattern, controls predominantly exhibited positive scores, while patients tended to show negative values. Line 1 and line 2 were two separating boundaries between patients and controls. Results revealed that 9 controls and 6 patients were wrongly classified in pattern 1, and 13 controls and 5 patients were wrongly classified in pattern 2.

We can see that subjects in quadrant I were all healthy controls, while subjects in quadrant III were mainly

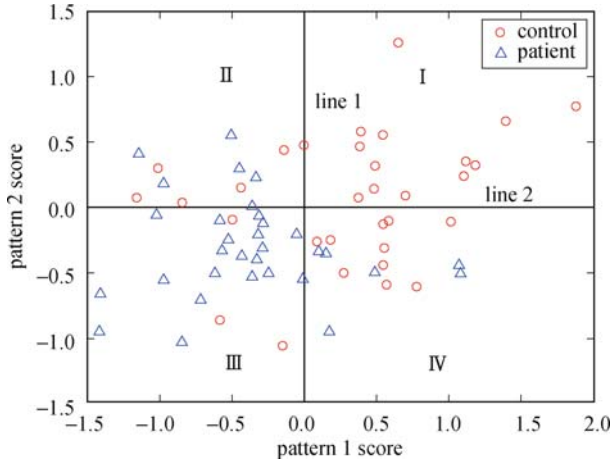

Fig. 5 Distributions of discriminative scores for the two pathological patterns. Line 1 and line 2 are two boundaries separating patients from controls

schizophrenic patients showing gray matter deficits in both the PFC and subcortical regions. There were also five patients in quadrant II, who only showed gray matter deficits in the PFC, and six patients in quadrant IV, who only showed gray matter deficits in subcortical regions. Our results suggest that brain abnormalities in both of the two pathological patterns are associated with schizophrenia. In addition, healthy subjects also located in quadrants II, III, and IV, making it difficult to distinguish them from schizophrenic patients.

When voxels in these two patterns were combined, three patients and eight controls were wrongly classified. It indicates that combing these two patterns achieved better classification performance than using either of them alone. Furthermore, our classifier was also performed using voxels of the whole brain. The results revealed that 13 patients and 13 controls were wrongly classified. Given these findings, it could be inferred that noise and artifacts would decrease the performance of our classifier and that ICA was an effective way to detect the spatial distribution of abnormal brain regions in schizophrenia. Data reflecting the specificity, sensitivity, and classification accuracy of SVM using different spatial patterns are presented in Table 4.

Table 4 Classification results of SVM on different spatial patterns

spatial patterns	sensitivity/%	specificity/%	accuracy/%
pattern 1	81.3	71.9	76.6
pattern 2	84.4	59.4	71.9
pattern 1 + 2	90.6	75.0	82.8
whole brain	59.4	59.4	59.4

4 Discussion and conclusion

Based on prior knowledge of structural abnormalities in schizophrenia, this paper proposed a multivariate pattern analysis method to successfully extract pathological patterns associated with schizophrenia as well as distinguishing schizophrenic patients from healthy controls at individual subject level. We uncovered two disease-correlated spatial patterns (i.e., a PFC pattern and a subcortical region pattern), which were statistically independent from each other. The results revealed that combing these two patterns could achieve better classification performance than using either of them alone. This novel method thus provides potential ability to improve current diagnosis of this mental disorder.

4.1 Advantages of ICA relative to univariate methods

In MRI studies of schizophrenia, ROI and morphological analyses have been intensively employed to study brain anatomical changes in patients with schizophrenia. No single brain region has been consistently reported to be abnormal in schizophrenia perhaps because of variation in demographic characteristics and methodological differences between studies [3]. In recent years, several meta-analyses have been conducted in attempts to identify the core brain regions associated with pathological changes in schizophrenia, which could potentially act as disease markers. For example, taking account of the possible progression of neuropathology during the illness, Ellison-Wright et al. employed an anatomical likelihood estimation method to map gray matter changes in first-episode schizophrenia and to compare these with the changes in chronic schizophrenia [4]. However, these studies quantify brain changes in schizophrenia on a voxel by voxel or ROI by ROI basis, ignoring statistical relationships between variables. By pooling information from many variables, multivariate methods may be more sensitive for detecting group differences in some circumstances [8]. In the present study, a multivariate method, ICA, was employed to decompose gray matter density images into a set of spatial patterns. In particular, two pathological patterns were successfully extracted, both of which have significant discriminative power in distinguishing schizophrenia. Thus, our findings demonstrate that ICA can effectively separate brain abnormalities in schizophrenia from artifacts and various types of noise.

4.2 Neuropathological changes in schizophrenia detected by ICA

As shown in Fig. 3, several localized areas were identified as abnormal in schizophrenia, including the PFC, caudate nucleus, thalamus, putamen, hippocampus, cerebellum, fusiform gyrus, middle temporal pole, and superior occipital gyrus. Evidence from previous sMRI studies has indicated that most of these regions are

implicated in the pathophysiology of schizophrenia [2]. Other parts of our ROI, such as the superior temporal gyrus and insula, have also been reported as abnormal in schizophrenia in many previous studies. However, these brain regions did not show significant structural alterations in our data set. Ellison-Wright et al. reported more extensive cortical changes in chronic schizophrenia compared with first-episode schizophrenia in the PFC, insula, and temporal cortex [4]. Although the patients in our study were not first-episode, they were in the early stages of the illness. The mean duration of illness was 27.2 ± 16.7 months. The patients in our study may have thus exhibited less cortical change than chronic patients because of a possible progression of neuropathology during the illness.

The current experimental results revealed that combining the two pathological patterns could achieve better classification accuracy than using only a single pattern. Figure 5 indicates that 26 out of 32 patients showed gray matter deficits in the PFC, while 27 out of 32 patients showed gray matter deficits in the subcortical regions. No single brain region was sensitive enough to distinguish patients from healthy controls. However, we found that all patients had gray matter deficits at least in one of the pathological patterns. One could speculate that a lesion in any brain region of the prefronto-subcortical network may cause schizophrenia. It is interesting to note that the gray matter density of the PFC pattern and the subcortical pattern are independent from each other. Rüsch et al. also reported that the left dorsolateral PFC had negative volumetric correlations with thalamic, cerebellar, pontine, and right parahippocampal areas in schizophrenia. They attributed this to a compensatory mechanism, i.e., gray matter in the subcortical areas may be increased to make up for reduced dorsolateral PFC gray matter [17]. We propose that the heterogeneity of schizophrenia is important when considering underlying causes. Schizophrenia presents with a diversity of symptoms that involve multiple psychological domains, for example, perception, cognition, language, motor activity, emotion, and so on. Not all patients have symptoms involving all these domains. Some studies have already found that the subgroups or subsyndromes of schizophrenia are associated with different brain structural changes [18]. We suggest that psychopathological variability may result in multiple independent abnormal morphological patterns in schizophrenia.

4.3 Potential ability of ICA + SVM to improve clinical diagnosis of schizophrenia

There has been a growing interest in applying machine learning methods to distinguish patients with schizophrenia from healthy controls using neuroimaging data and have demonstrated high classification performance [7,19]. However, due to limited number of samples, artifacts and noise of various types, high-dimensional measurement space, and the clinical heterogeneity of schizophrenia, applying machine learning methods to clinical diagnosis remains challenging. In the present study, we sought to solve these limitations by two separate steps. First, ICA was employed to determine the spatial distribution of brain changes in schizophrenia to filter out noise and artifacts. Then, SVM was used to deduce a precise model of the relationship between brain morphological characteristics and schizophrenia, which can easily achieve nonlinear classification through a kernel matrix. Experimental results showed that ICA + SVM (82.8%) had much better classification accuracy than performing SVM using voxels of the whole brain (59.4%) (see Table 4). So it could be inferred that noise and artifacts would decrease the performance of our classifier and that ICA was an effective way to detect the spatial distribution of abnormal brain regions in schizophrenia.

4.4 Limitations

Some limitations of the present work also need to be considered. It should be noted that the best classification accuracy of our method was around 80%. Although the results are promising, even better accuracy is need for clinical diagnosis purpose of schizophrenia. It might suggest that brain gray matter abnormality is not enough sufficient for differentiating schizophrenic patients and healthy controls. Other neuroimaging evidences such as functional connectivity abruption would be also important to do a synthesized diagnosis of this complex disorder [19]. In the further, we expect to develop a multimodal analysis combing both structural and functional information for achieving better classification accuracy [20]. The specificity of our method was only 75.0%, indicating that there is a high likelihood of misclassifying a healthy subject as a schizophrenic patient. Previous studies have reported that factors associated with age [21], sex [22], and intelligence quotient [23] can contribute to gray matter changes in some brain regions in healthy subjects. Therefore, possible explanations include the fact that we did not take account of the possible variability of gray matter patterns in our healthy subjects. Finally, generalizing our findings to clinical applications is also problematic due to the limited number of samples and interindividual anatomical variations. As such, it will be necessary to evaluate our method with larger sample size and multi-center imaging data in the future before it can be implemented in clinical practice.

4.5 Conclusion

In conclusion, this study integrated ICA and SVM to classify schizophrenic patients and healthy controls using sMRI data. The proposed method was capable of discovering the pathological changes in schizophrenia as well as demonstrating good classification accuracy. Experimental results indicated that both the PFC and the subcortical regions play important roles in the pathophysiology of schizophrenia, suggesting that schizophrenia involves abnormalities in networks of brain regions.

Acknowledgements This work was supported by the National Natural Science Foundation of China (Grant Nos. 61003202, 90820304, and 60835005) and the National Basic Research Program of China (No. 2011CB707802).

References

1. Andreasen N C, Nopoulos P, O'Leary D S, Miller D D, Wassink T, Flaum M. Defining the phenotype of schizophrenia: cognitive dysmetria and its neural mechanisms. Biological Psychiatry, 1999, 46(7): 908–920
2. Shenton M E, Dickey C C, Frumin M, McCarley R W. A review of MRI findings in schizophrenia. Schizophrenia Research, 2001, 49(1–2): 1–52
3. Honea R, Crow T J, Passingham D, Mackay C E. Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. American Journal of Psychiatry, 2005, 162(12): 2233–2245
4. Ellison-Wright I, Glahn D C, Laird A R, Thelen S M, Bullmore E. The anatomy of first-episode and chronic schizophrenia: an anatomical likelihood estimation meta-analysis. American Journal of Psychiatry, 2008, 165(8): 1015–1023
5. Glahn D C, Laird A R, Ellison-Wright I, Thelen S M, Robinson J L, Lancaster J L, Bullmore E, Fox P T. Meta-analysis of gray matter anomalies in schizophrenia: application of anatomic likelihood estimation and network analysis. Biological Psychiatry, 2008, 64(9): 774–781
6. Giuliani N R, Calhoun V D, Pearlson G D, Francis A, Buchanan R W. Voxel-based morphometry versus region of interest: a comparison of two methods for analyzing gray matter differences in schizophrenia. Schizophrenia Research, 2005, 74(2–3): 135–147
7. Fan Y, Shen D G, Gur R C, Gur R E, Davatzikos C. COMPARE: classification of morphological patterns using adaptive regional elements. IEEE Transactions on Medical Imaging, 2007, 26(1): 93–105
8. Pereira F, Mitchell T, Botvinick M. Machine learning classifiers and fMRI: a tutorial overview. NeuroImage, 2009, 45(1): S199–S209
9. McKeown M J, Jung T P, Makeig S, Brown G, Kindermann S S, Lee T W, Sejnowski T J. Spatially independent activity patterns in functional MRI data during the Stroop color-naming task. Proceedings of the National Academy of Sciences of the United States of America, 1998, 95(3): 803–810
10. Calhoun V D, Adali T, Pearlson G D, Pekar J J. Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. Human Brain Mapping, 2001, 13(1): 43–53
11. Hu D W, Yan L, Liu Y, Zhou Z, Friston K J, Tan C, Wu D. Unified SPM-ICA for fMRI analysis. NeuroImage, 2005, 25(3): 746–755
12. Xu L, Groth K M, Pearlson G, Schretlen D J, Calhoun V D. Source-based morphometry: the use of independent component analysis to identify gray matter differences with application to schizophrenia. Human Brain Mapping, 2009, 30(3): 711–724
13. Calhoun V D, Maciejewski P K, Pearlson G D, Kiehl K A. Temporal lobe and "default" hemodynamic brain modes discriminate between schizophrenia and bipolar disorder. Human Brain Mapping, 2008, 29(11): 1265–1275
14. Ashburner J, Friston K J. Voxel-based morphometry — the methods. NeuroImage, 2000, 11(6): 805–821
15. Stoica P, Selen Y. Model-order selection: a review of information criterion rules. IEEE Signal Processing Magazine, 2004, 21(4): 36–47
16. Bell A J, Sejnowski T J. An information-maximization approach to blind separation and blind deconvolution. Neural Computation, 1995, 7(6): 1129–1159
17. Rüsch N, Spoletini I, Wilke M, Bria P, Di Paola M, Di Iulio F, Martinotti G, Caltagirone C, Spalletta G. Prefrontal-thalamic-cerebellar gray matter networks and executive functioning in schizophrenia. Schizophrenia Research, 2007, 93(1–3): 79–89
18. Nenadic I, Sauer H, Gaser C. Distinct pattern of brain structural deficits in subsyndromes of schizophrenia delineated by psychopathology. NeuroImage, 2010, 49(2): 1153–1160
19. Shen H, Wang L B, Liu Y D, Hu D W. Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI. NeuroImage, 2010, 49(4): 3110–3121
20. Fan Y, Resnick S M, Wu X Y, Davatzikos C. Structural and functional biomarkers of prodromal Alzheimer's disease: a high-dimensional pattern classification study. NeuroImage, 2008, 41(2): 277–285
21. Raz N, Lindenberger U, Rodrigue K M, Kennedy K M, Head D, Williamson A, Dahle C, Gerstorf D, Acker J D. Regional brain changes in aging healthy adults: general trends, individual differences and modifiers. Cerebral Cortex, 2005, 15(11): 1676–1689
22. Witte A V, Savli M, Holik A, Kasper S, Lanzenberger R. Regional sex differences in grey matter volume are associated with sex hormones in the young adult human brain. NeuroImage, 2010, 49(2): 1205–1212
23. Frangou S, Chitins X, Williams S C R. Mapping IQ and gray matter density in healthy young people. NeuroImage, 2004, 23(3): 800–805

Lubin WANG was born in Hebei, China, in 1983. He received the B.Sc. and M.Sc. degrees both from the National University of Defense Technology in 2005 and 2007, respectively. He is currently pursuing his Ph.D degree in the field of cognitive science from February 2008 to now. His research interests include brain image/signal processing, neuroscience, and machine learning.

Hui SHEN was born in Jiangxi, China, in 1975. He received the B.Sc., M.Sc., and Ph.D degrees from the National University of Defense Technology in 1997, 2000 and 2003, respectively. He was promoted associate Professor in 2005. His research interests include neural image/signal processing, computer/biological vision, cognitive neuroscience. He has over 40 papers and 1 academic prize in the areas of his interests.

Baojuan LI was born in Shanxi, China, in 1983. She received the B.Sc. degree from the National University of Defense Technology in 2006. She is currently pursuing her Ph.D degree in the field of cognitive science from February 2008 to now. Her research interests include neuroscience, dynamic causal modeling of fMRI data, and pattern recognition.

Dewen HU was born in Hunan, China, in 1963. He received the B.Sc. and M.Sc. degrees from Xi'an Jiaotong University, in 1983 and 1986, respectively. From 1986, he was with the National University of Defense Technology. From October 1995 to October 1996, he was a Visiting Scholar with the University of Sheffield, UK. He got his Ph.D degree from the National University of Defense Technology in 1999. He was promoted Professor in 1996. His research interests include image processing, system identification and control, neural networks, and cognitive science. He is an action editor of Neural Networks.