



Feature selection using genetic algorithm for classification of schizophrenia using fMRI data

H. Shahamat* and A. A. Pouyan

Computer Engineering and Information Technology Department, University of Shahrood, Shahrood, Iran.

Received 17 December 2013; Accepted 30 April 2014

*Corresponding author: Shahamat@shahroodut.ac.ir (H. Shahamat).

Abstract

In this paper we propose a new method for classification of subjects into schizophrenia and control groups using functional magnetic resonance imaging (fMRI) data. In the preprocessing step, the number of fMRI time points is reduced using principal component analysis (PCA). Then, independent component analysis (ICA) is used for further data analysis. It estimates independent components (ICs) of PCA results. For feature extraction, local binary patterns (LBP) technique is used for the ICs. It transforms the ICs into spatial histograms of LBP values. For feature selection, the genetic algorithm (GA) is used to obtain a set of features with large discrimination power. In the next step of feature selection, linear discriminant analysis (LDA) is used for further extract features that maximize the ratio of between-class and within-class variability. Finally, a test subject is classified into schizophrenia or control group using a Euclidean distance based classifier and a majority vote method. In this paper, a leave-one-out cross validation method is used for performance evaluation. Experimental results prove that the proposed method has an acceptable accuracy.

Keywords: Schizophrenia, ICA, Feature Extraction, Local Binary Patterns, LDA.

1. Introduction

Schizophrenia is a common, chronic and debilitating psychiatric disorder. It affects about 1% of the global population, and another 3% has Schizophrenia-type personality disorders [1]. Schizophrenia is the fourth leading cause of disability in the developed countries [2]. In the last years, researchers have tried to propose methods for classification of patients with severe mental illness. It was done to exam differences between patient and controls groups, based on neuroscientific measures [3]. In this regard, researchers have used event-related potentials (ERP) derived from the electroencephalogram (EEG) for finding abnormalities in schizophrenia patients for many years. ERP waveforms obtained through AOD stimuli show good results in separating schizophrenia from normal controls [4, 5]. However, the studies based on ERP have not proven to be sensitive enough to be used in diagnostic purposes.

On the other hand, functional magnetic resonance imaging (fMRI) data have potential to classify different brain disorders including schizophrenia

with a higher accuracy than other neuroimaging techniques such as ERPs [6–8]. Since, there exist many challenges in the accurate analysis of fMRI data (such as high dimensionality and noisy nature), many algorithms should be employed for preprocessing, statistical analysis, feature selection, and classification. Many algorithms for dimensionality reduction have been developed. Principal component analysis (PCA) [9] is one of the most popular techniques for dimensionality reduction. PCA constructs a low-dimensional representation of data that describes as much of variance in the data as possible. For fMRI analysis, independent component analysis (ICA) is a useful method, which extracts powerful multivariate features for classification [10,11]. ICA decomposes fMRI data into a product of a set of time courses and independent components (ICs).

These ICs show different activation levels in the normal and schizophrenia groups. Finding an optimal feature selection and extraction method is very important for removing the redundancy and

preserving the most discriminative activation patterns from the ICs [3].

Several studies have used fMRI activation levels to discriminate schizophrenia and normal controls. Shinkareva et al. [12] identified groups of voxels showing between-group temporal dissimilarity and worked directly with fMRI time series for classification purposes. In this method, the task-associated stimulus was used to calculate the temporal dissimilarity matrix. However, the rest of data have no such stimulus presented nor are the data task-related. Thus, this method is not applicable for such cases. Ford et al. [13] combined structural and functional MRI data for classification purposes. They used PCA to project the high dimensional data onto a lower dimensional space for the training set. Du et al. [3] proposed a new method to extract classification features from fMRI data collected at rest or during the performance of a task. They proposed a combination of kernel PCA and Fisher's linear discriminant analysis (FLD) for feature identification. Then, a majority vote method was used for classification of subjects into predefined groups.

In this paper, local binary patterns (LBP) [14] is used for feature extraction. Since, after performing this step the data still have a very high dimension, genetic algorithm (GA) is used for feature selection. GA is a search procedure based on the mechanism of natural selection and natural genetics. The first GA was developed by John H. Holland in the 1960s to allow computers to evolve solutions to difficult search and combinatorial problems, such as function optimization and machine learning [15].

GAs offer a particularly attractive approach for problems like feature subset selection since they are generally quite effective for rapid global search of large, non-linear and poorly understood spaces. GAs are based on an imitation of the biological process in which new and better populations among different species are developed during evolution. Thus, unlike most standard heuristics, GA uses information of a population (individuals) of solutions when they search for better solutions.

In this paper, a new approach to discriminate the normal controls and schizophrenia patients is proposed. First, fMRI scans are preprocessed using statistical parametric mapping software version 8 (SPM8) [16], and PCA is used for dimension reduction. Then, independent components of the new data (given by PCA) are estimated using ICA method. For feature extraction, LBP histogram extraction technique is

used for all estimated components. Genetic Algorithm is used for selection of the most significant histogram bins, in next step. Then, linear discriminant analysis (LDA) is performed to further extract features that maximize the ratio of between-class and within-class variability. Finally, a classifier based on Euclidean distance is used for classification. We evaluate the classification performance using a leave-one-out cross-validation method. Figure 1 shows the overall procedure of the proposed method.

The rest of the paper is organized as follows: Section 2 introduces brain fMRI database. In this section the preprocessing steps including preprocessing using SPM8, PCA, and ICA are briefly described. Section 3 explains details of feature extraction using LBP method, and feature selection using GA. Also, in this section, details of GA operators are described. In section 4, we explain the classification process and evaluation of performance of the proposed method. Finally, sections 5 and 6 show experimental results and conclusion.

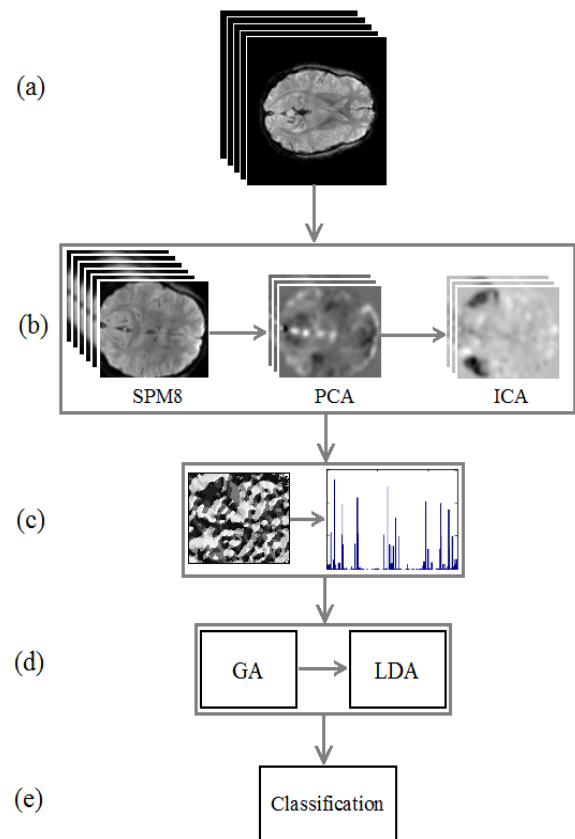


Figure 1. Overall procedure of proposed method: (a) original data, (b) preprocessing using SPM8, PCA, and ICA methods, (c) feature selection using LBP method and its histogram, (d) feature selection using GA and LDA methods, (e) classification using the Euclidean-based classifier.

2. Data and preprocessing

2.1. Database

Multimodal T1 structural MRI, DTI and Resting State FMRI (R-FMRI) datasets of 10 schizophrenia patients (SZ) and 10 (NC) were downloaded from the publicly available NA-MIC dataset [17], but the FMRI scans for case01017 and case01073 do not exist. In this paper, only the fMRI scans are used for further processes. Hence, 18 subjects including 10 NC and 8 SZ are remained for classification. Preprocessing including realignment, normalization, and smoothing, was performed in the statistical parametric mapping software (SPM8) [16]. An example of preprocessing using SPM8 is shown in figure 2.

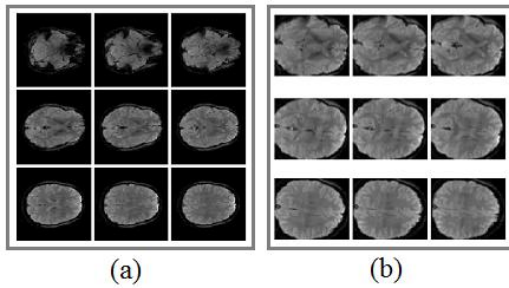


Figure 2. Preprocessing using SPM8: (a) original fMRI scan, (b) preprocessed data.

2.2. PCA and ICA

Dimension reduction is one of the key challenges in most FMRI studies. Principle component analysis (PCA) [9] is a mathematical procedure for solving this problem. PCA transforms the original data onto a smaller number of principal components [18]. It is done by finding a linear basis of reduced dimensionality for the data, which the amount of variance in the data is maximal. In this paper, PCA is used for FMRI time point reduction. For FMRI scans, a data matrix $X = [x_1, \dots, x_T]$ is constructed. Where, X is a V -by- T matrix, V is the number of voxels, and T is the number of FMRI time points. Finally, PCA is applied to the data matrix X using MATLAB toolbox for dimensionality reduction proposed in [19]. After dimension reduction, ICA method is used for further data analysis. It decomposes data into a set of independent components (ICs), which have very high discrimination power. The ICA analysis of FMRI data is started with $X=AS$ model [3]. Where, $S = [s_1, \dots, s_N]^T$ is an N -by- V source matrix, N is the number of sources (the principal components in PCA), V is the number of voxels and s_i is the i th spatial component. The mixing matrix A is an M -by- N matrix where each column a_i represents the time course for the i th source. The goal of the ICA algorithm is to

determine a demixing matrix W such that the sources are estimated using $\hat{S}=WX$ under the assumption of statistical independence of spatial components. Several algorithms for ICA were proposed, and FastICA is one of the most popular of them. FastICA provides a simple way for independent components extraction. It does not depend on any user-defined parameters, and is fast to converge to the most accurate solution allowed by the data [20]. In this paper, ICA is applied to the FMRI scans using FastICA MATLAB toolbox proposed by Hyvarinen [21].

3. Feature extraction and selection

3.1. Local binary patterns

Local binary patterns (LBP) [14] is a simple and efficient image texture operator. Texture analysis based on LBP has excellent discriminative power for many applications in the domain of computer vision. Therefore, it can be used to extract features from medical images [22]. In this paper, LBP technique operates on the ICs, which are estimated by ICA algorithm in the preprocessing step. The LBP operator can be defined as:

$$LBP(x) = \sum_{p=0}^7 C(v_{x_p}, v_x) 2^p \quad (1)$$

where, for labeling voxels of ICs using the original LBP, the voxel value v_x at position x is compared to the voxel values v_{x_p} of the eight neighbors of the center position x , as follows:

$$C(v_{x_p}, v_x) = \begin{cases} 1, & v_{x_p} \geq v_x \\ 0, & v_{x_p} < v_x \end{cases} \quad (2)$$

where, $p = 0, 1, \dots, 7$.

The LBP codes for all voxels in the ICs are calculated, and these coded ICs are transformed into a histogram of LBP values. This paper uses the LBP technique in a 3×3 neighborhood mode (Figure 3). Thus, there will exist $2^8 = 256$ possible texture units (histogram bins) for one IC.

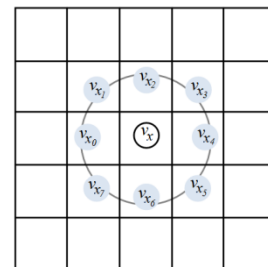


Figure 3. The 3×3 neighborhood in LBP method.

3.2. Feature selection using genetic algorithm

All LBP histograms have 256 bins. Each histogram is considered as a feature vector and

genetic algorithm (GA) is used for feature selection. For 256 bins, there exists 2^{256} subset of bins. Finding a subset of features with sufficiently large discrimination power requires a very large search space. GA is very effective in solving large-scale problems, and can be used to find an optimal or near optimal feature subset [23]. In GA, the individuals are typically represented by n -bit binary vectors. In feature selection problem, each individual would represent a feature subset. It is assumed that the quality of each candidate solution (or fitness of the individual in the population) can be evaluated using a fitness function, with respect to some criteria of interest. GA components are adjusted as follows:

3.2.1. Encoding

Each chromosome in the population represents a candidate solution for feature selection problem. If m is total number of features (here, $m = 256$), each chromosome is represented by a binary vector of dimension m . If a bit is equal to 0 it means that the corresponding feature is not selected, and if the bit is equal to 1 means the feature is selected [24]. This is the simplest and most straightforward representation scheme.

3.2.2. Initial population

The initial population is generated randomly. A random binary vector creates each chromosome. The number of chromosomes in the initial population is an important issue for GA performance. A large population causes more genetic diversity, but it suffers from slower convergence. A very small population explores only a reduced part of the search space and it may converge to a local extreme.

3.2.3. Fitness function

The fitness function gives the quality of the produced member of the population. In this paper, the quality is measured with the Fisher criterion [3] and GA is used for finding a feature subset (corresponding chromosome), which has maximum or near-maximum amount of Fisher criterion in training data.

3.2.4. Genetic operators

(a) *Selection*: Roulette wheel selection is used to probabilistically select individuals from a population for later breeding.

(b) *Crossover*: Single-point crossover operator is used in this paper. The crossover point i is chosen randomly. The new solutions (offspring) will be created using first i bits of one parent and the remaining bits of the other parent.

(c) *Mutation*: Each individual has a probability P_m to mutate. We randomly choose 10% of the total bits of each selected individual, which should be flipped in the mutation stage.

3.2.5. Genetic algorithm parameters

Finally, GA parameters are adjusted as follows:

- 1) *Population size*: 100
- 2) *Number of generation*: 50
- 3) *Probability of crossover*: 0.7
- 4) *Probability of mutation*: 0.4
- 5) *Crossover strategy*: Random single point
- 6) *The bits of selected chromosomes that will be mutated*: 0.1

3.3. Linear discriminant analysis

Linear Discriminant Analysis (LDA) [25] attempts to maximize the linear separability between data points belonging to different classes. In contrast to most other dimensionality reduction techniques, LDA is a supervised technique. LDA finds a linear mapping M that maximizes the linear class separability in the low-dimensional representation of the data. The criteria that are used to formulate linear class separability in LDA are between-class scatter and within-class scatter. LDA optimizes the ratio between these scatters by finding a linear mapping M that maximizes the Fisher criterion [3]. LDA maps data points onto a d -dimensional space. Where, $d < C$, and C is the number of classes. In this paper, we deal with a two-class problem. Therefore, d is equal to 1. In general, the projection onto one dimension leads to a considerable loss of information. However, by using LDA, we can achieve a projection that maximizes the class separation and also does not lose within-class compactness. In this paper, MATLAB toolbox for dimensionality reduction [19] is used for applying LDA technique.

4. Classification process and performance evaluation

The classification procedure uses a leave-one-out cross-validation method to evaluate performance of the proposed method. It involves using a single subject for validation data and the remaining subjects as the training set. This is repeated such that each subject is used once as validation data. In this paper, for each left-out test subject, the remaining 17 subjects (including controls and patients) comprise the training set. Our feature extraction method consists of three steps: LBP, GA and LDA. First, histogram of each independent component is extracted using LBP technique, which provides significant features

based on texture information. Second, GA is performed to select the best subset of the LBP histogram. Finally, LDA is used for project-selected features onto one-dimensional space that maximizes the ratio of between- and within-class variability. It should be noted, GA is an optimization method based on stochastic optimization that generates and uses random variables. Thus, to deal with randomization issues, GA will run three times to prove the robustness of the proposed method. For each run, we show the accuracy, sensitivity, and specificity of the obtained classification result. Accuracy is calculated as the ratio between the number of correctly classified subjects and the total number of subjects. Sensitivity and Specificity [3] are defined and calculated as follows:

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

and,

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (4)$$

where, TP (True positive) is correctly diagnosed patients, FP (False positive) incorrectly identified patients, TN (True negative) correctly diagnosed controls and FN (False negative) incorrectly identified controls. Du et al. [3] proposed a classification algorithm based on Euclidean distance, which shows good results for one-dimensional data. Therefore, we used this algorithm for classification of our data. After obtaining significant features by GA and LDA, the Euclidean distances between the test feature and all training features should be calculated, such that $d_1^{[c]}, \dots, d_{n_1}^{[c]}, d_1^{[p]}, \dots, d_{n_2}^{[p]}$, where c and p denote the healthy control and the patient group, respectively. By comparing the mean distances between the test data and each training group, the test data will be assigned to closest group. The classification process is used for all slices in all time points for all FMRI scans. Finally, using a majority vote method, the test person is classified to the class receiving the largest number of votes.

5. Experimental results

All FMRI scans contain 200 repetitions of a high resolution EPI scan. In this paper, after preprocessing using PCA method, the number of repetitions is reduced to 10. The PCA method not only reduces the number of repetitions but also maps data onto a new space. After that, in order to further data analysis, an ICA method is used for extraction of independent components (ICs) of the PCA results. Although the dimension of data has

been reduced significantly, but data still have a very high dimension. It may causes over-fitting in classification step. Therefore, to obtain a set of features with large discrimination power, LBP operator is used for all ICs, which transform each IC into a spatial histogram of LBP values. In this paper, LBP operator is used in 3×3 mode (see Figure 3), which transforms each brain slice onto a histogram with 256 bins. Then, GA is employed for finding a subset of histogram bins with acceptable discrimination power. In this paper, Fisher criteria are used as a GA fitness function. The GA tries to find a subset of histogram bins from train data, which have most or near most amount of Fisher value. Figure 4 shows examples of increasing fitness value for different generations of GA. Best chromosome in GA is represented by a binary vector with the length of 256. If a bit is equal to 0, it means that the corresponding feature is not selected, and if the bit is equal to 1, it means the feature is selected. After finding an optimal subset of bins, LDA maps these data onto a one-dimensional space. It should be noted, all brain slices in all independent components of the test subjects are classified completely separately. For example, for first slice in first IC of the test subject, the training set includes the only first slices of first ICs of the remaining 17 subjects. When a subject is given for classification, it is preprocessed using mentioned methods, and LBP operator is used for histogram extraction. Then, for each brain slices in each IC, an optimal subset of bins is selected using related best chromosome of GA, and LDA maps these features onto the new space. Finally, comparing the mean distances between a slice of test subject and related slices of each training group will label this slice of test person labeled as a member of nearest group. This process is repeated for all brain slices in all ICs. Then, using a majority vote method, the test subject is assigned to the group, which has maximum votes. As mentioned, GA is a random search, and for performance evaluation of the proposed method, we apply the classification process in three different runs. Table 1 shows classification results in all runs of the proposed method. As can be seen in table 1, all normal subjects in all runs are classified correctly, which causes the sensitivity of 100% in all cases. In the SZ group, “case01018” is classified incorrectly in all runs. In run #1, only 4 SZ subjects were classified correctly. It causes about 78% (14/18) accuracy and 71% specificity. In run #2, in addition to subject “case01018”, the subject with number “case01015” is classified incorrectly. Thus, obtained accuracy and

specificity are about 89% and 83%, respectively. In run #3, 3 SZ subjects are classified incorrectly, and accuracy 83% and specificity 77% were achieved.

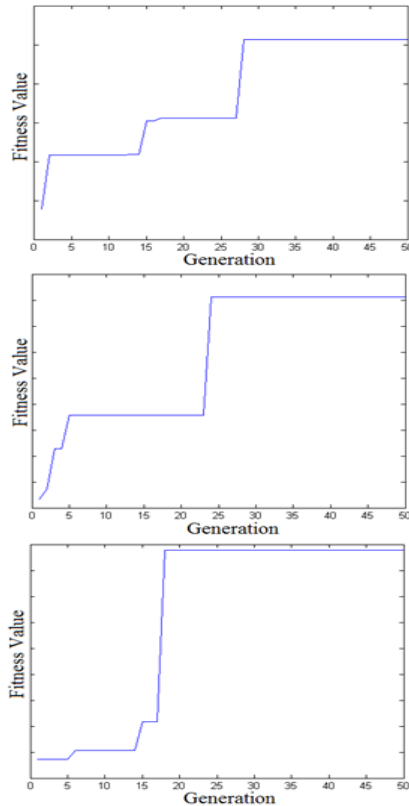


Figure 4. An example of fitness value increasing in order to different generations of GA.

Table 1. Classification results using proposed method for different runs of GA method (SZ = 1 & NC = -1).

Case number	Diagnosis	Run #1	Run #2	Run #3
case01019	NC	-1	-1	-1
case01020	NC	-1	-1	-1
case01025	NC	-1	-1	-1
case01026	NC	-1	-1	-1
case01029	NC	-1	-1	-1
case01033	NC	-1	-1	-1
case01034	NC	-1	-1	-1
case01035	NC	-1	-1	-1
case01041	NC	-1	-1	-1
case01104	NC	-1	-1	-1
case01011	SZ	-1	1	-1
case01015	SZ	-1	-1	1
case01018	SZ	-1	-1	-1
case01028	SZ	1	1	1
case01039	SZ	1	1	1
case01042	SZ	1	1	1
case01044	SZ	-1	1	-1
case01045	SZ	1	1	1
case01017	There is not fMRI scans for these subjects			
case01073	There is not fMRI scans for these subjects			

Table 2 shows the classification performance in different runs of the proposed method. Also, table 2 shows the importance of each step in the

proposed method. When some parts of our method are eliminated, obtained accuracy is lower than complete form of the proposed method. For comparing of the proposed method with state-of-the-art methods, the overall accuracy should be calculated. It is done using an averaging procedure, and the results are shown in table 2. The results prove that our method is comparable with other methods in this area. In order to prove the effectiveness and compatibility of the proposed method, we have compared the proposed method with several state-of-the-art methods including, Ford et al. [26], Pokrajac et al. [27], and Georgopoulos et al. [28] methods, and the results are shown in table 3.

Table 2. Evaluation of classification performance.

GA runs	Sensitivity	Specificity	Accuracy
Run #1	1.00	0.71	0.78
Run #2	1.00	0.83	0.89
Run #3	1.00	0.77	0.83
Average	1.00	0.77	0.83
without PCA	0.67	0.60	0.61
without ICA	0.75	0.64	0.67
without GA	0.20	0.46	0.38
without LDA	0.57	0.64	0.61
without GA & LDA	1.00	0.59	0.60

Table 3. Comparison of obtained accuracy (%) using proposed method with a list of previous research.

Method	Accuracy
Ford	60-80
Pokrajac	68-80
Georgopoulos	77
Proposed method	83

6. Conclusion

This paper proposed a GA-based method for classification of schizophrenia using fMRI data. Preprocessing step includes several steps. First, the fMRI scans are realigned, normalized and smoothed using SPM8 software. Then, PCA is used for dimension reduction, and ICA is used for independent components estimation. In feature extraction step, LBP method is used for transforming ICs into spatial histograms of LBP values. For feature selection, GA and LDA are used for spatial histograms for finding the histogram bins with most discrimination power. Finally, a Euclidean-based classifier is used for classification of subjects into predefined groups (SZ or NC). Performance evaluation using the leave-one-out cross validation proved the superiority of the proposed method. The experimental results demonstrate that the proposed method is comparable to other state-of-the-art work.

References

- [1] Kandel, E. R. (2000). Disorders of thought and volition: schizophrenia. *Principles of neural science*, vol. 4, pp. 1188–1208.
- [2] Jalili, M. & Knyazeva, M. G. (2011). EEG-based functional networks in schizophrenia. *Computers in Biology and Medicine*, vol. 41, pp. 1178–1186.
- [3] Du, W., Calhoun, V. D., Li, H., Ma, S., Eichele, T., Kiehl, K. A., Pearlson, G. D. & Adali, T. (2012). High classification accuracy for schizophrenia with rest and task fMRI data. *Frontiers in human neuroscience* 6.
- [4] McCarley, R. W., Faux, S. F., Shenton, M. E., Nestor, P. G. & Adams, J. (1991). Event-related potentials in schizophrenia: their biological and clinical correlates and new model of schizophrenic pathophysiology. *Schizophrenia research*, vol. 4, pp. 209–231.
- [5] Ford, J. M. (1999). Schizophrenia: the broken P300 and beyond. *Psychophysiology*, vol. 36, pp. 667–682.
- [6] Levin, J. M., Ross, M. H. & Renshaw, P. F. (1994). Clinical applications of functional MRI in neuropsychiatry. *The Journal of neuropsychiatry and clinical neurosciences*, vol. 7, pp. 511–522.
- [7] Calhoun, V. D., Maciejewski, P. K., Pearlson, G. D. & Kiehl, K. A. (2008). Temporal lobe and “default” hemodynamic brain modes discriminate between schizophrenia and bipolar disorder. *Human brain mapping*, vol. 29, pp. 1265–1275.
- [8] Demirci, O., Clark, V. P., Magnotta, V. A., Andreasen, N. C., Lauriello, J., Kiehl, K. A., Pearlson, G. D. & Calhoun, V. D. (2008). A review of challenges in the use of fMRI for disease classification/characterization and a projection pursuit application from a multi-site fMRI schizophrenia study. *Brain imaging and behavior*, vol. 2, pp. 207–226.
- [9] Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, vol. 24(6), pp. 417–441.
- [10] Calhoun, V. D., Kiehl, K. A. & Pearlson, G. D. (2008). Modulation of temporally coherent brain networks estimated using ICA at rest and during cognitive tasks. *Human brain mapping*, vol. 29, pp. 828–838.
- [11] Arribas, J. I., Calhoun, V. D. & Adali, T. (2010). Automatic Bayesian classification of healthy controls, bipolar disorder, and schizophrenia using intrinsic connectivity maps from FMRI data. *Biomedical Engineering, IEEE Transactions*, vol. 57, pp. 2850–2860.
- [12] Shinkareva, S. V., Ombao, H. C., Sutton, B. P., Mohanty, A. & Miller, G. A. (2006). Classification of functional brain images with a spatio-temporal dissimilarity map. *NeuroImage*, vol. 33, pp. 63–71.
- [13] Ford, J., Shen, L., Makedon, F., Flashman, L. A. & Saykin, A. J. (2002). A combined structural-functional classification of schizophrenia using hippocampal volume plus fMRI activation. *Engineering in Medicine and Biology*, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference, 2002. Proceedings of the Second Joint (IEEE), pp. 48–49.
- [14] Ojala, T., Pietikäinen, M. & Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, vol. 29, pp. 51–59.
- [15] Holland, J. H. (1975). *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence.* (U Michigan Press).
- [16] SPM8 - Statistical Parametric Mapping Available: <http://www.fil.ion.ucl.ac.uk/spm/software/spm8/> [Accessed June 30, 2013].
- [17] MIDAS - Collection NAMIC: Brain Mutlimodality Available: <http://insight-journal.org/midas/collection/view/190> [Accessed July 2, 2013].
- [18] Sidhu, G. S., Asgarian, N., Greiner, R. & Brown, M. R. G. (2012). Kernel Principal Component Analysis for dimensionality reduction in fMRI-based diagnosis of ADHD. *Frontiers in systems neuroscience*, vol. 6:74.
- [19] Van der Maaten, L., Postma, E. & van den Herik, H. (2007). Matlab toolbox for dimensionality reduction. *Proceedings of the Belgian-Dutch Artificial Intelligence Conference*. Vol. 2007, pp 439–440.
- [20] Hyvärinen, A. & Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural computation*, vol. 9, pp. 1483–1492.
- [21] Hyvarinen, A. (1998). *The FastICA MATLAB toolbox.* Helsinki Univ of Technology.
- [22] Unay, D., Ekin, A., Cetin, M., Jasinschi, R. & Ercil, A. (2007). Robustness of local binary patterns in brain MR image analysis. *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE (IEEE)*, pp 2098–2101.
- [23] Tan, F., Fu, X., Zhang, Y. & Bourgeois, A. G. (2008). A genetic algorithm-based method for feature subset selection. *Soft Computing*, vol. 12, pp. 111–120.
- [24] Marinakis, Y., Dounias, G. & Jantzen, J. (2009). Pap smear diagnosis using a hybrid intelligent scheme focusing on genetic algorithm based feature selection and nearest neighbor classification. *Computers in Biology and Medicine*, vol. 39, pp. 69–78.
- [25] Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics*, vol. 7, pp. 179–188.

[26] Ford, J., Farid, H., Makedon, F., Flashman, L. A., McAllister, T. W., et al. (2003). Patient classification of fMRI activation maps. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2003* (Springer), vol. 2879, pp 58–65.

[27] Pokrajac, D., Megalooikonomou, V., Lazarevic, A., Kontos, D. & Obradovic, Z. (2005). Applying spatial distribution analysis techniques to classification of 3D medical images. *Artificial Intelligence in Medicine*, vol. 33, pp. 261–280.

[28] Georgopoulos, A. P., Karageorgiou, E., Leuthold, A. C. Lewis, S. M., et al. (2007). Synchronous neural interactions assessed by magnetoencephalography: a functional biomarker for brain disorders. *Journal of neural engineering*, vol. 4:349.

انتخاب ویژگی توسط الگوریتم ژنتیک برای دسته‌بندی افراد مبتلا به بیماری اسکیزوفرنی با استفاده از داده‌های اف ام آر ای

حسین شهامت* و علی اکبر پویان

دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه شاهرود، شاهرود، ایران.

ارسال ۲۰۱۳/۱۲/۱۷؛ پذیرش ۲۰۱۴/۰۴/۳۰

چکیده:

در این مقاله روش جدیدی برای دسته‌بندی افراد به دو گروه سالم و بیمار مبتلا به اسکیزوفرنی با استفاده از داده‌های تصویربرداری تشدید مغناطیسی عملکردی ارائه شده است. در مرحله پیش پردازش ابعاد بالای داده‌ها با استفاده از روش تحلیل مولفه‌های اصلی کاهش داده شده است. در ادامه برای تحلیل بیشتر داده‌ها و مستقل سازی آنها از الگوریتم تحلیل مولفه‌های مستقل استفاده شده است. برای استخراج ویژگی از مولفه‌های مستقل تولید شده در مرحله پیش‌پردازش، از روش الگوهای دودویی محلی استفاده شده است. این روش هر مولفه مستقل را به یک نمودار فراوانی از کدهای دودویی تبدیل می‌کند. برای یافتن زیر مجموعه‌ای از میله‌های نمودار فراوانی که بیشترین جدا شونده‌گی بین دو گروه را نشان دهد، از الگوریتم ژنتیک استفاده شده است. در ادامه برای استخراج ویژگی‌های مفیدتر، روش آنالیز افتراقی خطی بکار گرفته شده است. این روش براساس پراکنندگی درون کلاسی نسبت به پراکنندگی بین کلاسی عمل می‌کند. در نهایت داده‌های مغزی مربوط به یک شخص خاص توسط یک طبقه‌بند بر پایه فاصله اقلیدسی و روش اکتساب بیشترین آرا، به یکی از دو گروه سالم یا بیمار اختصاص داده می‌شود. در این مقاله برای ارزیابی کارایی روش معرفی شده، از روش اعتبارسنجی جدا کردن یک نمونه استفاده شده است. نتایج بیان شده نشان می‌دهد که روش ارائه شده دارای دقت قابل قبولی می‌باشد.

کلمات کلیدی: اسکیزوفرنی، تحلیل مولفه‌های مستقل، استخراج ویژگی، الگوهای دودویی محلی، آنالیز افتراقی خطی.