

Bi-objective approach for computer-aided diagnosis of schizophrenia patients using fMRI data

Indranath Chatterjee¹  · Manoj Agarwal² ·
Bharti Rana² · Navin Lakhyani³ · Naveen Kumar¹

Received: 15 May 2017 / Revised: 25 January 2018 / Accepted: 15 March 2018 /
Published online: 4 April 2018
© Springer Science+Business Media, LLC, part of Springer Nature 2018

Abstract Computer-aided diagnosis (CAD) of schizophrenia based on the analysis of brain images, captured using functional Magnetic Resonance Imaging (fMRI) technique, is an active area of research. The main problem lies in the identification of brain regions that contribute to differentiating between a healthy subject and a schizophrenia affected subject. The problem becomes complex due to the high dimensionality of the fMRI data on the one hand and the availability of data for only a small number of subjects on the other hand. In this paper, we propose a three-stage evolutionary based framework for feature selection. It comprises application of general linear model, followed by statistical hypothesis testing, and finally application of Non-dominated Sorting Genetic Algorithm (NSGA-II) to arrive at a small set of about fifty features. Experiments show that the feature set generated by the proposed approach yields accuracy as high as 99.5% in classifying fMRI dataset of healthy and schizophrenia subjects, and can identify the relevant brain regions that are affected in schizophrenia.

✉ Indranath Chatterjee
indranath.cs.du@gmail.com

Manoj Agarwal
agar.manoj@gmail.com

Bharti Rana
bhartirana.it@gmail.com

Navin Lakhyani
navinlakhyani@gmail.com

Naveen Kumar
nk.cs.du@gmail.com

¹ Department of Computer Science, University of Delhi, Delhi, 110007, India

² Department of Computer Science, Hans Raj College, University of Delhi, Delhi, 110007, India

³ Department of MRI, Saral Diagnostics, New Delhi, 110034, India

Keywords Functional magnetic resonance imaging (fMRI) · Schizophrenia · Computer-aided diagnosis (CAD) · Feature selection · NSGA-II · Bi-objective evolutionary algorithm

1 Introduction

Schizophrenia is a chronic brain disorder that disrupts the process of normal thinking, speech, and behavioral characteristics of a person. Functional magnetic resonance imaging (fMRI) plays a pivotal role in the design of automated tools for diagnosis of schizophrenia. It is a neuro-imaging technique that captures brain activity in small units of the brain volume called voxels, by measuring the change in blood-oxygen-level dependent (BOLD) [41] signals over time. The brain activities are closely linked to the supply of oxygen to various regions of the brain. As the blood oxygenation level of a brain region varies according to the neural activity, these differences play an important role. The difference in the magnetic properties causes small differences in the magnetic resonance (MR) signal of blood depending on the degree of oxygenation.

Functional magnetic resonance imaging is used to detect biomarkers within the brain for different types of task-related activations enabling detection of several brain disorders such as schizophrenia, Parkinson's disease, Alzheimer's disease, mild traumatic brain injury, addiction, and bipolar disorder. Several models based on machine learning techniques have been proposed [17, 18, 26, 39, 45] for investigating the fMRI data to identify different ailments. High dimensionality poses a major challenge in applying machine learning techniques to fMRI data. The fMRI data are typically 4-dimensional consisting of 3-D images across time. A 3-D fMRI image may be thought of as a sequence of 2-D images (slices) across the whole brain. Further, each slice comprises small units of brain volumes, called voxels. Thus, a voxel represents a particular position in the brain. Another issue that confronts the researchers is the non-availability of sufficient number of subjects/ data samples. The curse of dimensionality [5] and the small sample size render most models very sensitive to changes in data. To deal with high-dimensional fMRI data, most models employ some feature reduction/ selection techniques for the problem under investigation.

In this paper, we propose a three-stage feature selection model to classify schizophrenics and healthy subjects using fMRI Data. The study is based on fMRI data, acquired during auditory oddball (AUD) task. The first stage deals with the application of General Linear Model (GLM) followed by paired student's t-test in the second stage, and finally, we employ the Non-dominated Sorting Genetic Algorithm (NSGA-II) [14] to generate a feature set (set of voxels) that has low cardinality and yields high classification accuracy. The proposed model achieves classification accuracies in the range 92.6% - 99.5% for FBIRN phase-II dataset [30] having healthy and schizophrenia subjects. Using the proposed model, we are able to identify relevant regions in the brain affected by schizophrenia. To the best of our knowledge, evolutionary approach has not been used in the bi-objective framework for fMRI data to build a computer-aided diagnosis model for schizophrenia subjects.

The rest of the paper is organized as follows: in Section 2, we summarize the related work; in Section 3, we describe the data sets and the details of the proposed methodology; in Section 4, we describe the experimental settings and the results, and finally in Section 5, we summarize the conclusions and outline the scope of future work.

2 Related work

Acquisition of fMRI data is a complex process that generates huge volumes of data. Knowledge extraction from this data involves several steps including preprocessing, feature reduction, and modelling, often using machine learning techniques. Several machine learning algorithms like Principal Component Analysis (PCA) [8, 16, 18, 28, 49], Fisher Linear Discriminant (FLD) analysis [17, 45], Singular Value Decomposition (SVD) [4, 27], deep neural networks [31, 52], Convolution Neural Network (CNN) [42] are often used for feature extraction and feature selection.

Ford et al. [17] combined both structural and functional MRI scans for classifying the schizophrenia and healthy individuals. They extracted hippocampal formation by applying a mask and used Fisher linear discriminant analysis (FLDA) to reduce the feature set with the objective of maximizing the ratio of between-class and within-class variability. Using Leave-One-Out Cross-Validation (LOOCV), they obtained an accuracy of 83% - 87% on a group of 23 subjects (15 schizophrenic and 8 healthy). In another study, Ford et al. [18] proposed the application of Principal Component Analysis (PCA) to lower the dimensionality of the data, and applied FLD to distinguish between healthy subjects and schizophrenia patients, obtaining an accuracy of 60% - 80% for different principal components. They also demonstrated the effectiveness of the approach for differentiating the healthy subjects from Alzheimer's disease patients and patients with a mild traumatic brain injury. Shi et al. [45] used regional homogeneity [54] as a measurement of regional coherence of brain spontaneous activity. They used the anatomical template on ReHo map to organize it into 116 brain regions. Mean and standard deviation of ReHo values in each region were used as features for the classification model. Pseudo Fischer linear discriminant (PFLD) was applied in LOOCV manner to classify the healthy subjects and schizophrenia patients achieving correct prediction rate of 80%. Dermici et al. [15] proposed projected pursuit (PP) algorithm for feature selection and used Independent Component Analysis (ICA) for separating the data into maximally independent groups to identify the networks which are related to the schizophrenia. They applied three group ICA operations on the data from three different tasks and obtained 20 independent spatial components. The classification was performed using LOOCV. Arribas et al. [4] used a two-step method – one-sample t-test, followed by Singular Value Decomposition (SVD) to reduce the number of features of the fMRI scans with AUD task for classification of healthy subjects, patients with bipolar disorder, and patients with schizophrenia. They trained four classifiers using stochastic gradient learning rule and obtained average three-way correct classification rate (CCR) in the range 70% - 72%. Using the resting state and task-related fMRI data, Du et al. [16] classified the schizophrenia patients and healthy control. They used three-level feature selection approach. In the first step, they used hypothesis testing based on t-test. In the second step, they used the kernel principal component analysis (K-PCA) to compute a low-dimensional representation of significant voxels, and finally applied FLD to further extract features which maximize the ratio of the between-class variability to the within-class variability. Classification was done using LOOCV approach. Using majority voting, they achieved accuracy of 98% and 93% for the AUD task and the rest data, respectively.

In a study, Castro et al. [10] used a combination of Multiple Kernel Learning (MKL) machines and proposed a new MKL (v-MKL) algorithm for achieving a tunable sparse selection of feature sets which resulted in improvement in the classification accuracy while using functional brain imaging dataset. They obtained a classification accuracy of 85% and

90% using l_p -norm and L-norm, respectively. Juneja et al. [27] have used pattern recognition techniques for dimension reduction for fMRI data to classify schizophrenia and healthy subjects. They proposed a three-phase method for analysing the fMRI data. In the first phase, they generated 3-D spatial maps using GLM and ICA to generate independent components. In the second phase, they used clustering to retain local spatial contiguity followed by singular value decomposition (SVD) on each cluster, thus reducing the number of features substantially. In the third phase, a novel hybrid multivariate forward feature selection method was used to extract the features. Finally, schizophrenia and healthy control were classified using SVM with LOOCV policy, achieving 92.6% and 94% classification accuracy for the two fMRI datasets from FBIRN.

In another study, Juneja et al. [26], applied statistical paired t-test on the contrast map images created by SPM to develop a computer-aided diagnosis (CAD) tool to distinguish between the schizophrenic patients and the healthy controls. Having obtained the minimal set of features from statistical significance testing, they used the selected features for the classification task using Support Vector Machine (SVM). Using the LOOCV method, they obtained an accuracy of around 80% and 88% on the two fMRI datasets from FBIRN. In another study, Juneja et al. [28] proposed a three-phase dimension reduction technique comprising segmentation of 3-D spatial maps (ICA and β maps) into anatomical brain regions, followed by feature extraction carried out using fuzzy kernel PCA, and finally used the filter cum wrapper feature selection for finding reduced set of features. In their model, classification of schizophrenia and healthy subjects was done in LOOCV manner using SVM, resulting in accuracy of 95.6% and 96% on two fMRI data set from FBIRN.

Some works have also been reported for computer aided diagnosis of schizophrenia using resting state (rs) fMRI. Chyzyk et al. [13] used Pearson's correlation based features selection method, followed by application of genetic algorithm, to find an optimal set of features. Subsequently, they applied ensemble of extreme learning machine classifiers resulting in an accuracy of around 86%. Savio et al. [43], worked on rs-fMRI data of schizophrenia subjects and healthy controls. They computed different local activity measures, followed by application of three feature selection algorithms, namely, Pearson's correlation measure, Bhattacharyya distance [6] and Welch's t-test [50]. Finally, they used SVM to carry out the classification task and obtained maximum accuracy of around 80%.

Recently, multi-objective optimization approaches have been used for analysis of fMRI data. Aaberg et al. [1] proposed an evolutionary approach to select the features for multivariate pattern analysis. They used a single subject fMRI dataset having task conditions of brushing and resting state alternatively. Multiple Linear Regression (MLR) classifier was applied to the subjects individually using only five voxels to obtain an accuracy of 74.3%. Niiniskorpi et al. [39] used particle swarm optimization (PSO) in conjunction with simple MLR classifier and SVM with the linear kernel for the classification task for identifying the brain regions. They built two datasets, one having a single subject healthy control (a brushing task and resting state alternately), and another dataset comprising nine healthy controls (fingertapping task). They achieved a classification score of 83.5% on a group level 3D fMRI data from the fingertapping study. Ulker et al. [47] used a combination of an active method [38] and genetic algorithm for feature selection. Using a set of 300 voxels they obtained classification accuracy of around 90%. A genetic algorithm was also used by Shahamat et al. [44] for feature selection on fMRI images, followed by Linear Discriminant Analysis (LDA) to classify schizophrenia patients and healthy controls. The authors obtained an average classification accuracy of 83.0%, but they did not identify the regions in the brain that are responsible for the schizophrenia. Smart et al. [46] studied the application of Genetic Programming (GP) in feature selection using intracranial

Table 1 Demographic details of the dataset D1 (1.5 Tesla)

	No. of Subjects	Age Group(Mean \pm Std. Dev)	Male/Female
Healthy	30	40.4 \pm 12.29	20/10
Schizophrenia	30	42.3 \pm 10.81	24/6

electroencephalography (iEEG) and fMRI data of epilepsy patients. They observed the need for patient-specific feature selection for better classification results. Using nearest-neighbour classification and 30 GP generations, they achieved over 60% median sensitivity and over 60% median selectivity for fMRI data. Ma et al. [37] carried out Multi-Voxel Pattern Analysis (MVPA) as a Multi-Objective (MO) pattern classification problem. They integrated a hierarchical heterogeneous PSO (HHPSO) scheme with SVM to propose a feature interaction detection framework for voxel selection. In this framework, the first stage finds a subset of interacted features while the second stage further eliminates interaction (or connectivity) redundancy, improving the classification accuracy.

3 Materials and methods

3.1 Dataset

All the data used for this study were obtained from the Function BIRN Data Repository. FBIRN repository contains the multi-site fMRI dataset which includes schizophrenia and healthy subjects. The data was acquired using 1.5T and 3T scanners keeping all other parameters same for the subjects. In this study, we have used BOLD fMRI data of Auditory oddball (AUD) task, where all subjects had regular hearing levels, sufficient eyesight, and were able to perform cognitive task. Healthy subjects were excluded if they had a current or past history of head injury or major medical illness. Only those subjects with schizophrenia and schizoaffective disorder were allowed who met the criteria as per the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) [20].

3.1.1 Dataset details

In our study, we have used two datasets, namely, D1 and D2. The dataset D1 contains fMRI data of 30 schizophrenia patients and 30 healthy subjects (available at site 0009 and site 0010 of FBIRN repository), which were acquired with 1.5T scanner. Four runs of each subject's scan have been used for the experiments. Table 1 shows the demographic details of the dataset.

The dataset D2 comprises fMRI data of 25 schizophrenia patients and 25 healthy subjects (available at site 0005, site 0006 and site 0018 of FBIRN repository) acquired with 3T scanner. Four runs of each subject's scan have been used for the experiments. Table 2 shows the demographic details of the dataset.

Table 2 Demographic details of the dataset D2 (3 Tesla)

	No. of Subjects	Age Group(Mean \pm Std. Dev)	Male/Female
Healthy	25	40.08 \pm 11.35	15/10
Schizophrenia	25	44.8 \pm 10.77	19/6

3.1.2 Task details

Auditory oddball task is a common task [3, 25, 29, 36, 40] used to detect alterations in brain activation patterns that help to differentiate between schizophrenic and healthy subjects. A subject is presented with a continuous stream of sound, and he/she must identify the sequence of discrete stimuli comprising standard tones and deviant (i.e. oddball) tones. Standard tones, i.e., 1000 Hz appear for 95% of trials. Deviant (i.e. oddball) tones (1200 Hz) that are distinct from standard tones, appear occasionally (5% of trials). The FBIRN conducted the Auditory oddball task consisting of four experimental runs, each having duration of 280 seconds. During the experiment, in each run, the subjects were asked to see a gray screen with a black fixation cross in the middle. They were asked to press button '1' each time they heard a deviant tone while focusing on the cross and listening to the tones. The task began with a fixation block of the silence of 15 seconds. Then a sequence of standard tones (duration = 100 ms) were presented. The deviant tone (duration = 100 ms) was presented every 6 to 15 seconds. A period of silence (duration = 15 seconds) ended each task run. In each experimental run, 140 brain scans were acquired with repetition time (TR) of 2 seconds.

3.1.3 Imaging parameters

According to FBIRN repository, the functional scans were T2*-weighted gradient EPI (Echo Planar Imaging) sequences. Pulse sequence parameters were closely matched based on pilot studies carried out by FBIRN research group: Orientation: anterior commissure-posterior commissure line; the number of slices: 27; slice thickness: 4 mm; TR: 2 seconds; time to echo: 40 ms for 1.5 T scanners; matrix: 64 × 64; field of view: 22 cm; and flip angle: 90°.

3.2 Theoretical background

Genetic algorithms (GA), often used for solving optimization problems, are evolutionary algorithms based on natural or biological evolution processes. They follow Darwin's "survival of the fittest" concept and evolve to find the optimal solution from a set of candidate solutions. A genetic algorithm starts with an initial population of candidate solutions represented by vectors of strings or alphabets, mainly binary alphabets (0,1). These vectors, also called chromosomes, are randomly initialized. Once the chromosomes are generated, the genetic algorithm finds the fitness values of each of them for the optimization problem at hand. The next generation of solutions (also called child chromosomes), is created using selection, crossover, and mutation operations. The selection step imitates the *survival of the fittest* by giving preference to the better individuals. The selected chromosomes are placed in a common mating pool. In the crossover step, a crossover point is randomly selected and the crossover is done by recombining the portions of the two individuals to create two new offspring. The mutation step involves flipping one or more bits of the individuals. The purpose of the mutation is to maintain diversity amongst the chromosomes with the objective of avoiding premature convergence. The steps are repeated until no significant improvement is observed in successive generations, or the time-out condition is reached.

A bi-objective optimization problem is modelled using two conflicting objective functions f^1 and f^2 as:

- f^1 : To be maximized or minimized
- f^2 : To be maximized or minimized

The optimal solutions to the above problem can be modelled as a vector valued objective function f as:

$$f : X \rightarrow \mathbb{R}^2, f(x) = (f^1(x), f^2(x))^T$$

where a point $x^* \in X$ denotes a feasible solution, and $Y \in \mathbb{R}^2$ (solution space) denotes the image of X (decision space). Since the objectives are conflicting in nature, no single solution can optimize both the objective functions simultaneously. A solution $x^* \in X$ is said to be Pareto optimal [14] if and only if there is no other solution $x \in X$ that is equally good or better than x^* on both the objectives.

The fMRI dataset dimensions are too large for a classification model to distinguish between healthy and schizophrenic patients. Therefore, one needs to select an appropriate feature set for the efficacy of a decision model. To the best of our knowledge, evolutionary approaches have not been effectively applied to select relevant features that help to differentiate between schizophrenic and healthy subjects. Moreover, there is a conflicting relationship between classification accuracy and feature set size. This paper is the first attempt towards bi-objective modelling of the fMRI data analysis in schizophrenia to address the above mentioned conflicting issues. In this paper, we make use of Non-Dominated Sorting Genetic Algorithm (NSGA-II) [2, 14] to arrive at the Pareto optimal front. It is an evolutionary algorithm to solve the bi-objective optimization problem that aims

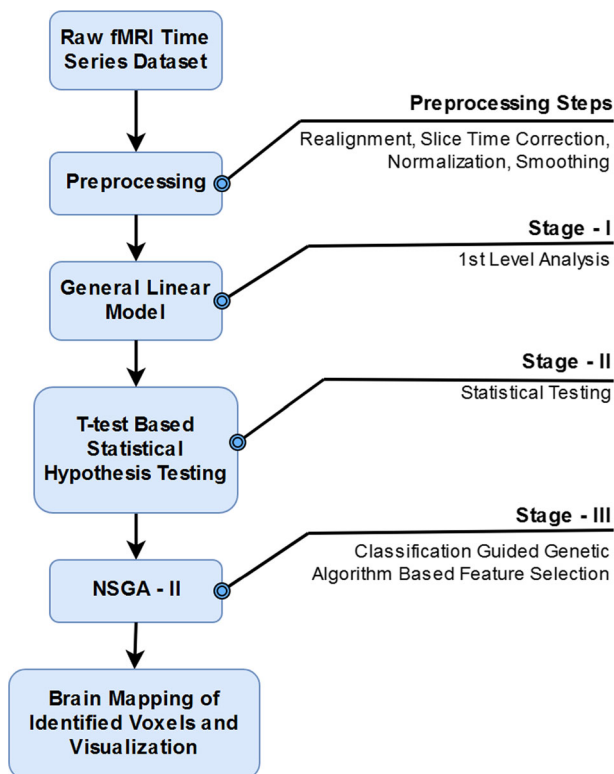


Fig. 1 Stages in the proposed approach

at improving the fitness and adaptability of the population of candidate solutions towards the Pareto front.

The runtime complexity of NSGA-II mainly lies in the non-dominated sorting – the most expensive part of the algorithm, and the little time spent in computing the objective functions is insignificant. Thus, the runtime complexity of the algorithm is of the order $O(mN^2)$, where m is the number of objective functions and N is the population size [14]. As NSGA-II dominates the computation time of the proposed approach, the run-time complexity of the overall approach is also $O(mN^2)$. The space complexity of the NSGA-II is of the order $O(mN + N^2)$ [14, 19]. In this study, we are proposing a three-step feature selection algorithm. In the first step, we use standard general linear model (GLM) [21] approach. The second step involves the application of the paired Students' t-test. Finally, we apply the NSGA-II to select the features useful for the classification task.

In this study, we have identified the following two conflicting objective functions:

- f^1 : Maximization of classification accuracy
- f^2 : Minimization of number of features

In the next section, we will discuss each step of our feature selection methodology in detail.

3.3 Our approach

For classification of schizophrenic patients and healthy subjects, we adopted a three-stage approach as shown in Fig. 1. Each stage of the proposed approach is described in the following subsections.

Algorithm 1 Non-dominated Sorting Genetic Algorithm (NSGA-II)

INPUT:

1. S = Number of chromosomes (population size)
2. $MaxGen$ = Maximum number of generations allowed
3. f^1 = Objective function to be maximized
4. f^2 = Objective function to be minimized

OUTPUT: Pareto-optimal set of chromosomes

- 1: Randomly generate initial binary population, P_0 , of size S
 - 2: Compute fitness values for $f^1(P_0)$ and $f^2(P_0)$ // For each chromosome belonging to P_0
 - 3: **for** $i = 0$ to $MaxGen$ **do**
 - 4: $S_i \leftarrow$ BinaryTournamentSelection (P_i) // S_i represents the population after tournament selection
 - 5: $C_i \leftarrow$ Crossover (S_i) // C_i represents the population after one-point crossover
 - 6: $M_i \leftarrow$ Mutation (C_i) // M_i represents the mutated population
 - 7: Compute fitness values for $f^1(M_i)$ and $f^2(M_i)$ // For each mutated chromosome belonging to M_i
 - 8: $T_i \leftarrow P_i \cup M_i$ // T_i denotes the pooled population
 - 9: $P_{i+1} \leftarrow$ NonDominatedSort (T_i)
 - 10: **end for**
 - 11: Generate set of pareto-optimal solutions x^*
-

3.3.1 Data pre-processing

The raw datasets taken from FBIRN repository have been preprocessed using Statistical Parametric Mapping (SPM) toolbox version 8 (SPM8, Wellcome Trust Centre for Neuroimaging, University College London, UK).¹ Raw scans were collected at voxel size of $3.4 \times 3.4 \times 4 \text{ mm}^3$. These are realigned with the first scan as a reference. The slice timing correction is done to correct the possible errors by temporal variations during the acquisition of fMRI datasets. Subsequently, the fMRI scans are spatially normalized into standard Montreal Neurological Institute (MNI) space using an EPI template available in SPM8. This transforms the initial voxel's dimension to $3 \times 3 \times 3 \text{ mm}^3$ and yields each volume of $53 \times 63 \times 46$ voxels. Finally, spatial smoothing is done with a $9 \times 9 \times 9 \text{ mm}^3$ full width at half maximum (FWHM) Gaussian kernel to get the smoothed volumes.

3.3.2 Stage-1: 1st level analysis

The 4-D fMRI scans of each subject obtained from the preprocessing steps are analyzed by employing general linear model (GLM) using SPM8 toolbox in MATLAB. GLM analysis is carried out by specifying the condition pair of deviant tone response versus standard tone response.

GLM analysis generates a 3-D contrast map, also called activation map. In a contrast map, the value at a particular voxel estimates the difference between the activation of that voxel amongst the conditions. Zero value at a voxel indicates that the particular voxel is not activated during the task condition [26]. GLM analysis is carried out for each of the four runs corresponding to each subject. Thereafter, for each subject, an average 3-D contrast map having only the activated voxels, is generated by averaging the contrast maps obtained for each of the four runs. Though, this stage reduces the feature vector size considerably, the dimensionality is still too large to efficiently distinguish the two classes.

3.3.3 Stage-2: the statistical testing

We linearised each subject's data into a one-dimensional vector. In the second stage of analysis, we have used the popular two sample t-test for selecting the relevant features. T-test is applied feature-wise to find the statistical significance of a feature between the two groups of data. The null hypothesis (H_0), mean value of a feature between the two groups being the same, is tested at $\alpha = 0.01$. Let d be the total number of features selected in the first stage, then the t-test value will be computed for each i^{th} feature as:

$$t_i = \frac{\mu_{s_i} - \mu_{h_i}}{\sqrt{\frac{\sigma_{s_i}^2}{n_s} + \frac{\sigma_{h_i}^2}{n_h}}} \quad (1)$$

$$\forall i = 1, 2, 3, \dots, d$$

where μ_{s_i} , $\sigma_{s_i}^2$ are the mean and variance for the schizophrenia patients and μ_{h_i} , $\sigma_{h_i}^2$ denote the mean and variance values for healthy subjects respectively, corresponding to i^{th} feature. n_s and n_h are the number of schizophrenia and healthy subjects respectively. Higher t-test value signifies higher relevance of a feature. The t-test values are considered for ranking

¹SPM Version 8: <http://www.fil.ion.ucl.ac.uk/spm/software/spm8>

the features and they have been sorted accordingly. Based on experimental exploration with selection of different numbers of features, finally top 300 features (rank wise) were passed to stage-3 of our approach.

3.3.4 Stage-3: application of GA

The third stage of the proposed approach involves the application of the non-dominated sorting genetic algorithm (NSGA-II) [14] which is outlined in Algorithm 1. Based on the features selected in stage-2, we have created a population of binary chromosomes. Each chromosome is 300 bits long. A one (zero) at a position in the chromosome indicates the presence (absence) of the corresponding feature. Initial chromosome is randomly generated with 20% of the bits being one. For our experiments, the population size (S) is fixed at 200. The fitness value of a chromosome for the first objective function (f^1 , maximization of classification accuracy), is independently evaluated by employing three different classifiers, namely, support vector machine (SVM) with linear kernel, SVM with sigmoid kernel and k -NN classifier (with $k=1$). The fitness value of the second objective function (f^2 , minimization of number of features) for a chromosome is computed by counting the number of ones in the chromosome. Offspring population (M_i) is generated using binary tournament selection, followed by one-point crossover and mutation. The mutation is applied at the rate of 0.01. The fitness value of the child population (M_i) generated after mutation step is computed, and a pooled population (T_i) of the initial (P_i) and child population (M_i) is formed. The pooled population (T_i) is then sorted to find the set of non-dominated solutions along the Pareto-Front. The chromosomes representing the trade-off solutions (P_{i+1}) are passed to the next generation. The maximum number of iterations ($MaxGen$) has been set to 100.

We have used the LOOCV scheme for feature selection. In LOOCV, one data sample is used for testing and rest are used for training purpose. This process is repeated N times (where N is the sample size) in such a way that each sample is chosen as a test sample exactly once. The feature selection process is carried out only on the training data to avoid the danger of double dipping [33]. We have repeated each experiment 10 times to capture the variability of the evolutionary approach. The feature selection process is shown in Algorithm 2.

Algorithm 2 Proposed feature selection approach

INPUT: GLM analysis generated N contrast images comprising healthy subjects (HS) and schizophrenia patients (SP).

OUTPUT: Set of voxels that distinguish between healthy and schizophrenic subjects.

```

1: for all  $i$  in  $NumOfRuns$  do
2:   for each fold  $j$  in LOOCV cross-validation do
3:     Select the  $j^{th}$  element as a test sample and the others as training samples
4:      $k = N - j$ 
5:      $F_j \leftarrow T\text{-Test}(HS_k, SP_k)$  //  $F_j$  is the feature set selected by applying Two-sample t-test at  $\alpha = 0.01$ 
6:     Rank  $F_j$  according to t-score and select top ranked features
7:      $\mathbf{x}_j^* \leftarrow \text{NSGA-II}(HS_k(F_j), SP_k(F_j))$  //  $\mathbf{x}_j^*$  is the optimal solution
8:     Store  $\mathbf{x}_j^*$ 
9:   end for
10: end for
11: Obtain the final set of relevant features or voxels by extracting unique voxels across all the solution sets

```

Table 3 Classification accuracy, mean and standard deviation values of feature sets for each of the 10 runs of the experiments with i) SVM using Sigmoid kernel, ii) SVM using Linear kernel and iii) 1-NN classifier respectively for dataset D1

	Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8	Run9	Run10	Mean
SVM (Sigmoid)											
Accuracy	98.33%	100%	98.33%	98.33%	93.33%	100%	98.33%	98.33%	100%	100%	99.0%
Feature Set Size Mean±Std Dev	46 ± 4.2	48 ± 5.5	45 ± 6.5	47 ± 4.5	47 ± 4.5	47 ± 4.2	46 ± 4.8	46 ± 2.9	46 ± 5.1	46 ± 4.1	46
SVM (Linear)											
Accuracy	100%	100%	98.33%	100%	100%	98.33%	100%	100%	100%	98.33%	99.5%
Feature Set Size Mean±Std Dev	48 ± 4.9	47 ± 5.3	47 ± 5.0	47 ± 4.4	48 ± 5.6	48 ± 6.0	48 ± 7.4	47 ± 4.0	49 ± 6.5	47 ± 4.3	48
k-NN (k=1)											
Accuracy	95%	95%	95%	93.33%	93.33%	95%	91.67%	96.67%	96.67%	98.33%	95.0%
Feature Set Size Mean±Std Dev	45 ± 5.0	46 ± 5.5	46 ± 5.1	46 ± 4.3	45 ± 5.0	45 ± 4.8	46 ± 5.8	46 ± 7.0	46 ± 5.3	46 ± 5.2	46

Table 4 Classification accuracy, mean and standard deviation value of feature sets for each of the 10 runs of the experiments with i) SVM using Sigmoid kernel, ii) SVM using Linear kernel and iii) 1-NN classifier respectively for dataset D2

	Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8	Run9	Run10	Mean
SVM (Sigmoid)											
Accuracy	96%	98%	98%	98%	96%	98%	98%	98%	96%	98%	97.4%
Feature Set Size Mean±Std Dev	47±5.24	48±5.0	48±4.47	48±5.132	47±5.09	47±5.9	47±4.56	48±4.95	47±5.25	48±4.41	47
SVM (Linear)											
Accuracy	94%	96%	94%	94%	94%	96%	96%	96%	96%	96%	95.20%
Feature Set Size Mean±Std Dev	46±4.41	47±4.30	48±4.15	47±3.31	49±4.16	45±5.48	47±3.91	49±5.99	46±6.93	47±4.16	47
<i>k</i> -NN (<i>k</i> =1)											
Accuracy	90%	90%	94%	94%	92%	94%	94%	92%	94%	92%	92.6%
Feature Set Size Mean±Std Dev	48±5.86	48±3.31	48±6.00	48±5.92	47±4.80	46±4.58	47±4.90	48±4.39	48±4.28	47±5.21	48

4 Experimental results and discussion

Experiments are carried out using MATLAB-R2014a (Mathworks Inc., Natick, MA, USA) in Ubuntu 14.04LTS environment on a machine having Intel[®] Xeon having 2.10GHz x17 processor with 32GB RAM. We have used SPM8 toolbox for preprocessing and general linear modeling; libsvm [11] package for the classification task; Talairach Daemon for mapping; and Multi-image Analysis GUI (Mango) [35] for visualizing the mapped brain regions.

We have used C-Support Vector Classification (C-SVC) [7], available in libsvm tool for Matlab, by fine tuning its parameters. The regularization parameter C was fine tuned at $C=100$ after evaluating the values of C from 0.01 to 1000 in steps of 10. C-SVC uses the loss function,

$$\min_{\omega, b, \xi} \quad \frac{1}{2} \omega^T \omega + C \sum_{i=1}^l \xi_i \quad (2)$$

subject to $y_i(\omega^T \phi(x_i) + b) \geq 1 - \xi_i$,

$$\xi_i \geq 0, \quad i = 1, \dots, l$$

where $\phi(x_i)$ maps x_i into a higher-dimensional space, ω is the vector variable, ξ_i are the slack variables, and $C > 0$ is the regularization parameter [11].

4.1 Experimental results

Each experiment in our three-stage evolutionary based approach was repeated ten times to capture the variability. For dataset D1 (1.5 Tesla), we obtained mean classification accuracies of 99.0%, 99.5%, and 95.0% using SVM with sigmoid kernel, SVM with linear kernel, and 1-NN classifier, respectively. Table 3 shows the results for each run of the experiments on D1. It shows the mean and standard deviation of the number of features obtained in each run along with classification accuracies of the models. Similar experiments were carried out on dataset D2 (3 Tesla). For dataset D2, we obtained accuracies of 97.4%, 95.2% and 92.6% using SVM with sigmoid kernel and linear kernel, and 1-NN classifier, respectively. Table 4 shows the results for dataset D2.

Figures 2 and 3 show the variability in the number of relevant features selected for classification using linear SVM across 10 runs for dataset D1 and D2, respectively.

To evaluate the relevance of the proposed methodology, we have conducted experiments without incorporating any feature selection method for datasets D1 and D2. The feature set obtained after the GLM analysis was used for classification using SVM with linear kernel and sigmoid kernel, and 1-NN classifiers in LOOCV manner. For datasets D1 and D2, we obtained the mean classification accuracies of 45.0% and 44.0% in case of linear SVM, accuracies of 53.33% and 40% in case of SVM with sigmoid kernel, and accuracies of 53.33% and 42% for the 1-NN classifier, respectively (see Table 5).

We have also conducted the experiments using principal component analysis (PCA) tool available in Matlab2014b for feature selection. The feature set obtained on carrying out GLM and t-test from the stage-1 and stage-2 analysis, was used as an input to the PCA. For the purpose of classification, we have used linear kernel SVM and L1-regularized L2-loss SVC in linear SVM in a LOOCV manner. We have experimented with the cost parameters by changing the value of C from 0.01 to 1000 in an interval of multiple of 10 to obtain the optimal accuracy. For dataset D1, using linear SVM and L1-regularized L2-loss SVC, we obtained highest mean classification accuracy of 65% and 60% respectively. For dataset

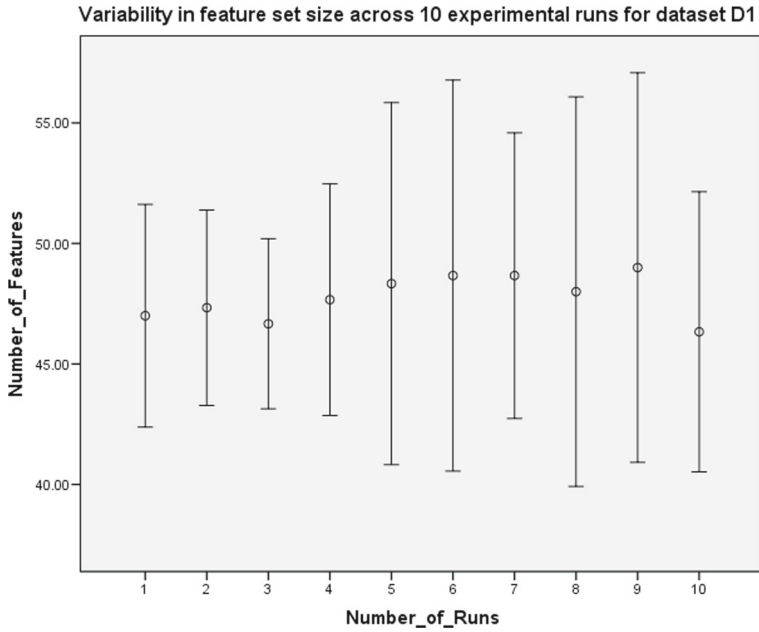


Fig. 2 Variability in the number of selected features by linear SVM across the ten runs for dataset D1

D2, we have obtained highest mean accuracy of 50% and 52% with linear SVM and L1-regularized L2-loss SVC respectively.

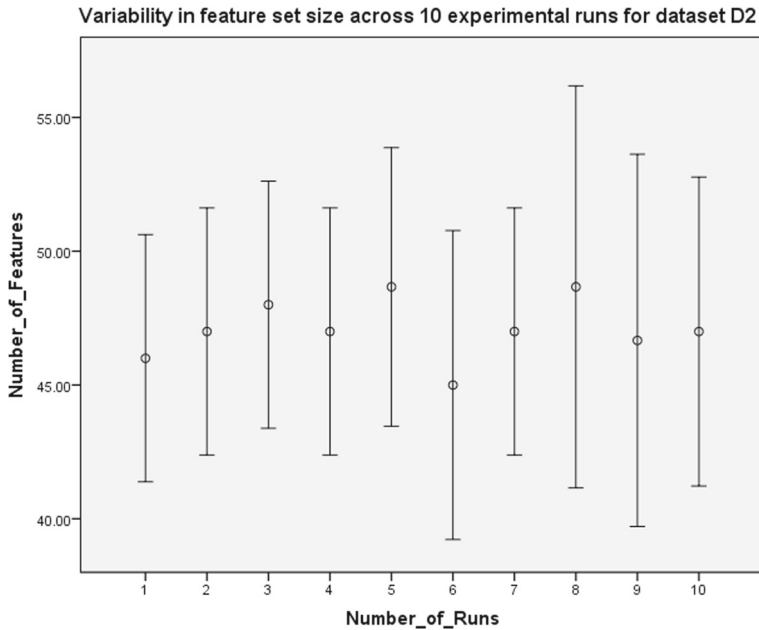


Fig. 3 Variability in the number of selected features by linear SVM across the ten runs for dataset D2

Table 5 Mean classification accuracy for dataset D1 and D2 without feature selection

Classifier	Dataset D1 (1.5T)	Dataset D2 (3T)
SVM with Linear Kernel	45.0%	44.0%
SVM with Sigmoid Kernel	53.33%	40.0%
k -NN where $k = 1$	53.33%	42.0%

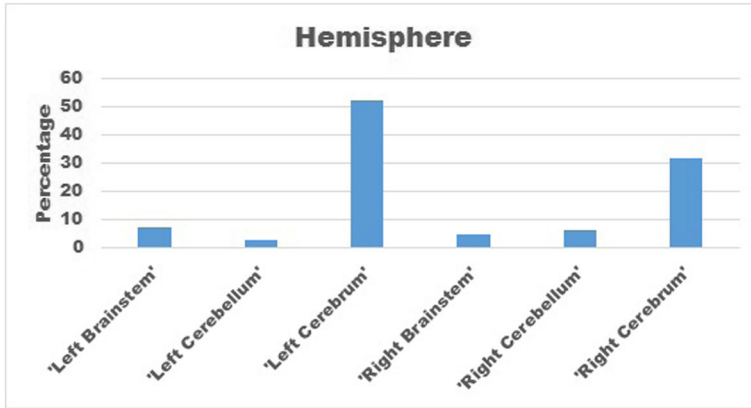


Fig. 4 Percentage wise distribution of affected voxels covering hemisphere regions for dataset D1

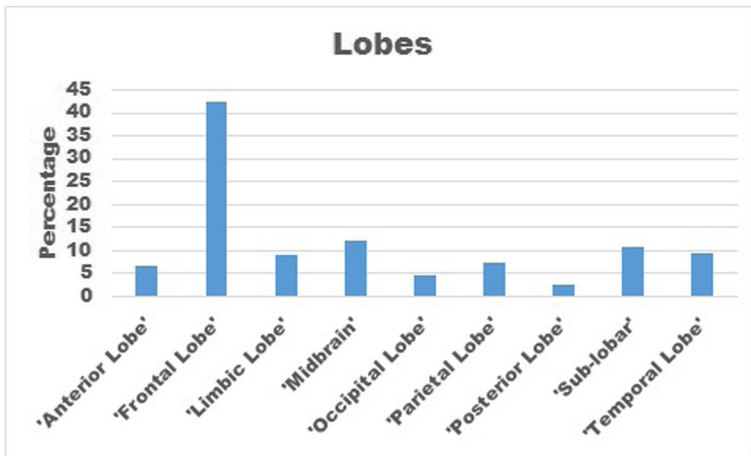


Fig. 5 Percentage wise distribution of affected voxels covering the lobes for dataset D1

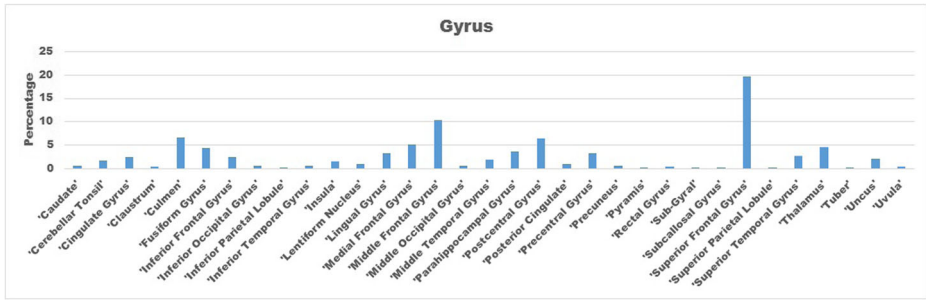


Fig. 6 Percentage wise distribution of affected voxels covering gyral regions for dataset D1

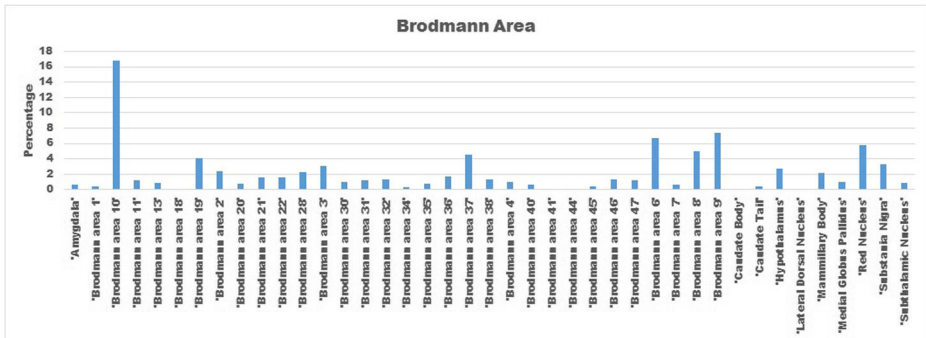


Fig. 7 Percentage wise distribution of affected voxels covering Brodmann's areas for dataset D1

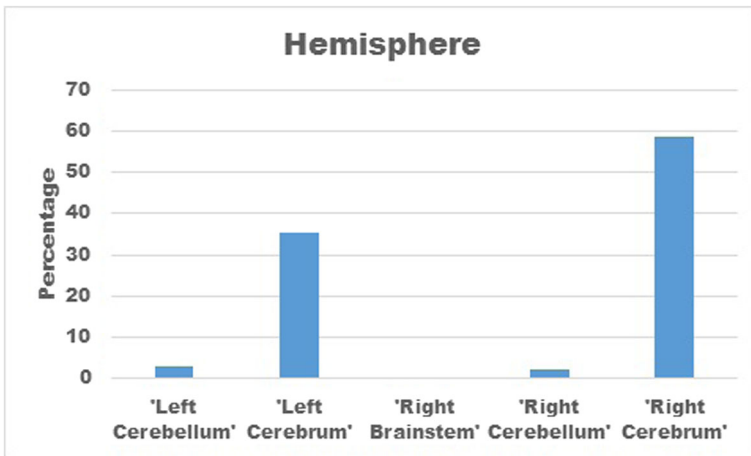


Fig. 8 Percentage wise distribution of affected voxels covering hemisphere regions for dataset D2

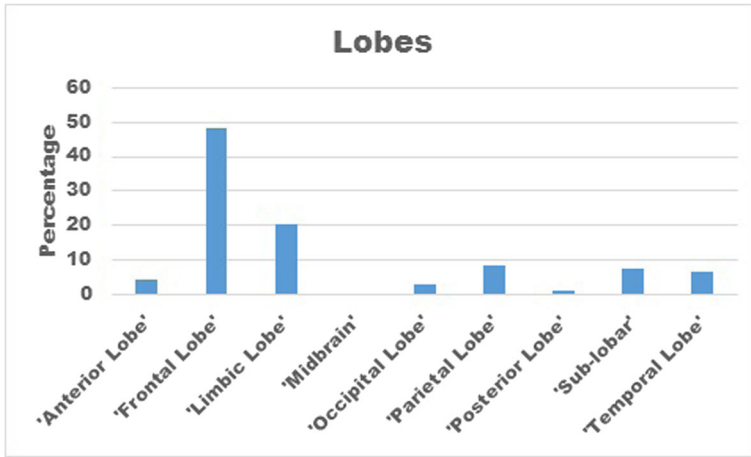


Fig. 9 Percentage wise distribution of affected voxels covering the lobes for dataset D2

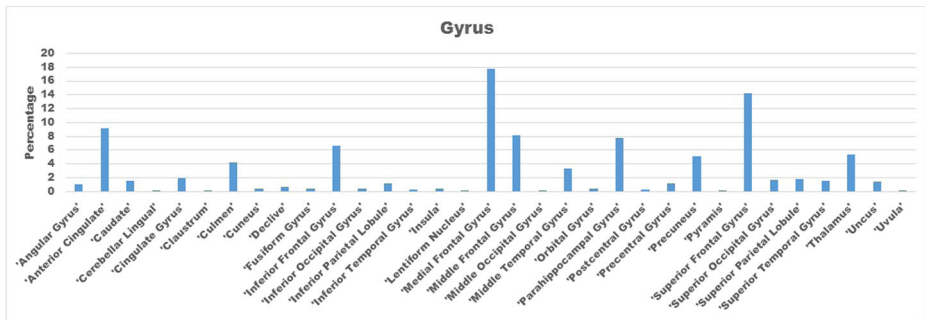


Fig. 10 Percentage wise distribution of affected voxels covering gyral regions for dataset D2

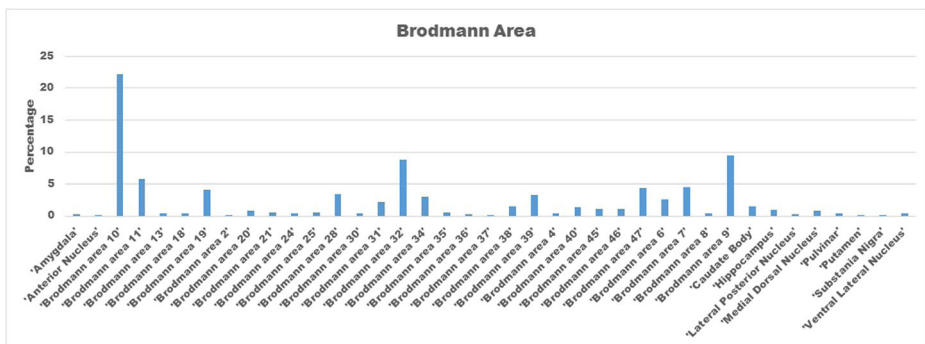


Fig. 11 Percentage wise distribution of affected voxels covering Brodmann's areas for dataset D2

4.2 Discussion

Different runs of the experiment on dataset D1 (see Table 3) resulted in about 40-50 features for each fold of the LOOCV technique. The experiments, repeated ten times, yielded about 800 distinct voxels. These identified voxels represent the regions that help in distinguishing

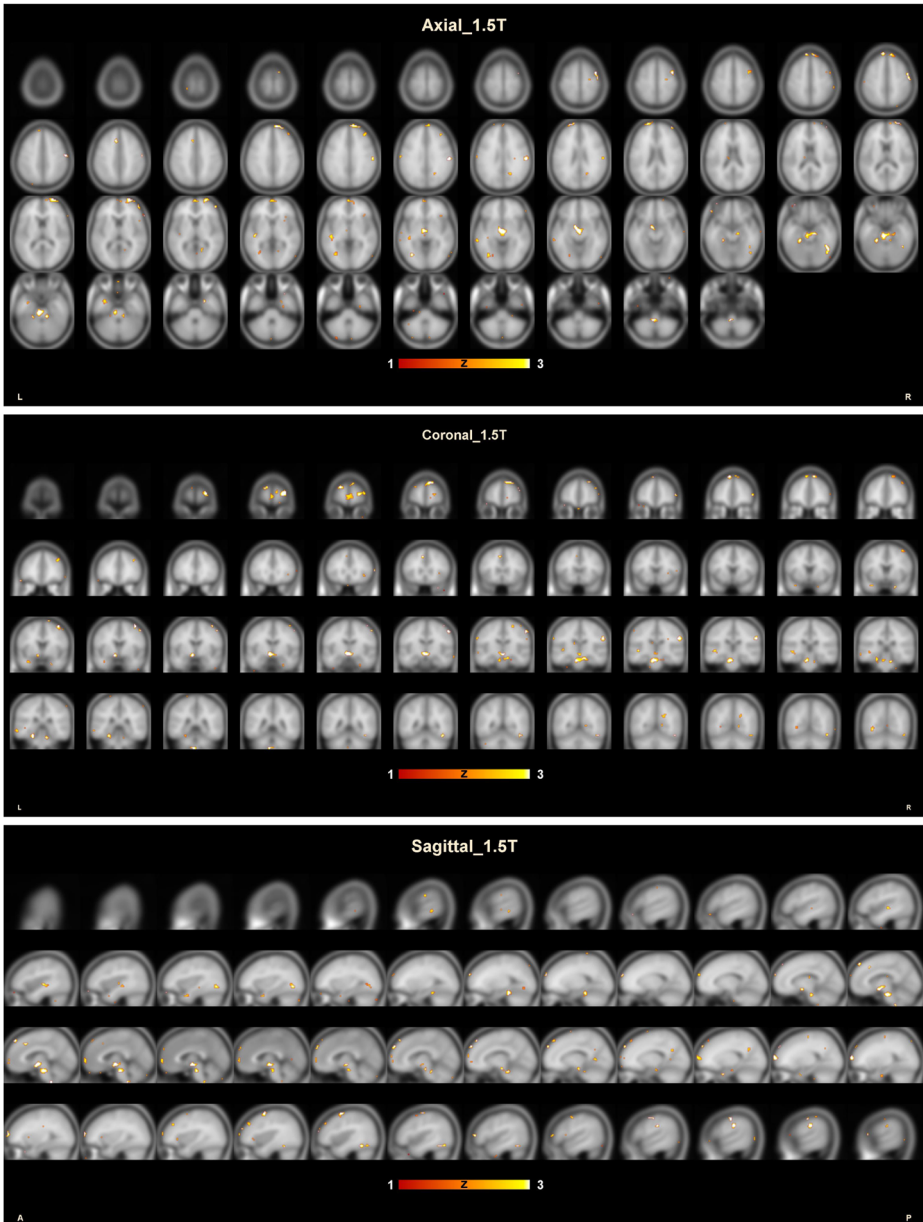


Fig. 12 Voxels identified by the proposed approach with variation in t-test values in different views of the brain for dataset D1 (1.5 Tesla)

between the schizophrenia patients and the healthy subjects. Brain regions, to which these voxels belong, are identified using the Talairach Daemon [34] for carrying out multi-level analysis – hemisphere level, lobe level, gyrus level and cell type (Brodmann Area) of the human brain in Talairach’s space. Figures 4, 5, 6 and 7 represent the selected regions for

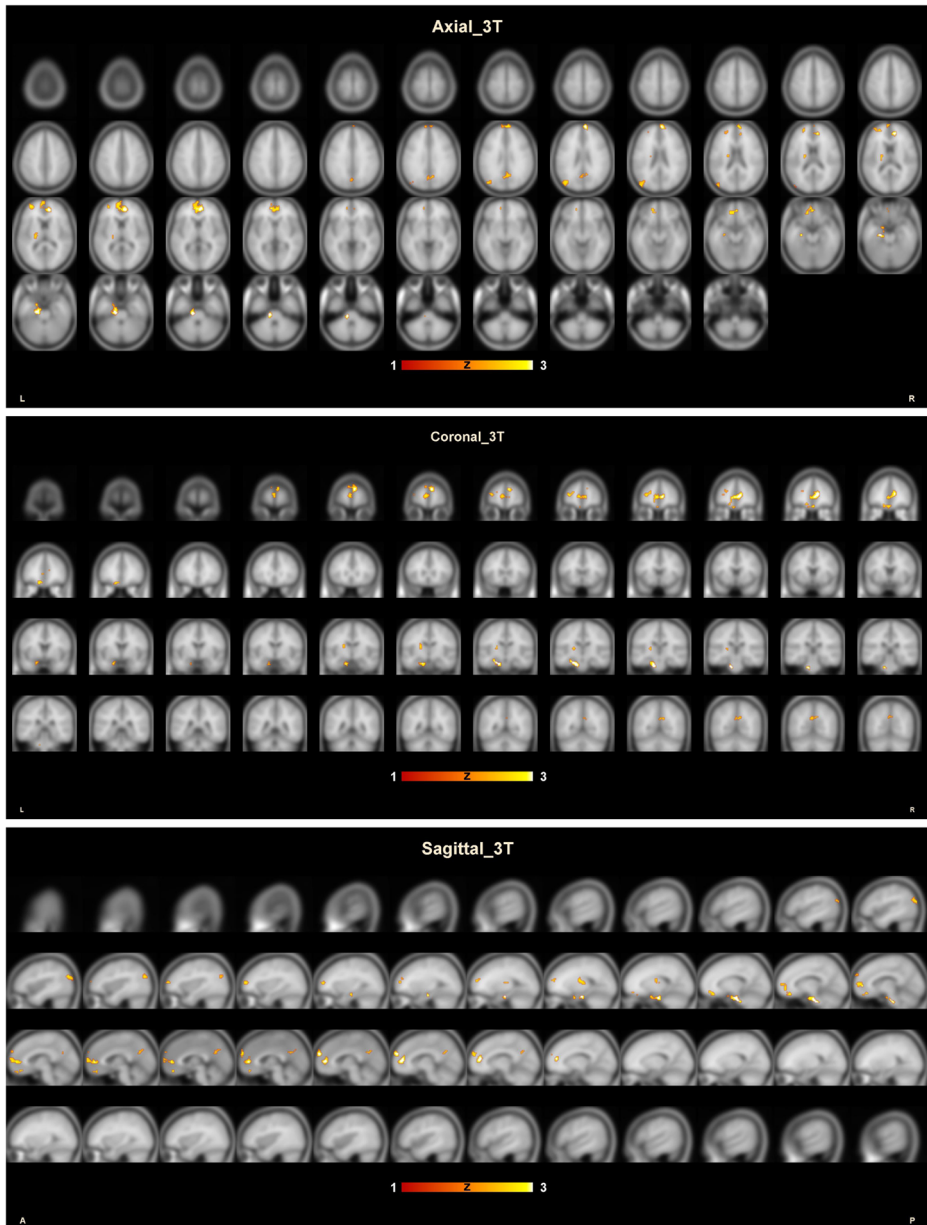


Fig. 13 Voxels identified by the proposed approach with variation in t-test values in different views of the brain for dataset D2 (3 Tesla)

dataset D1. As shown in Fig. 4, majority of the selected voxels either belong to the left cerebrum or right cerebrum region of the brain, or to the right brainstem. Figure 5 shows the percentage wise of distribution of voxels among the lobes. It can be observed that the majority of the voxels either lie in the frontal lobe, limbic lobe, mid brain or the temporal lobe. Figure 6 represents gyrus level analysis. It can be observed that the identified voxels either lie on the superior frontal gyrus, medial frontal gyrus, middle frontal gyrus, culmen, post-central gyrus and thalamus. Figure 7 shows the percentage wise distribution of the identified voxels among the Brodmann Areas (BA). We can observe that the majority of the identified voxels either lie in BA 10, 6, 37, 8, 9, 2, 3, 19, substantia nigra, red nucleus or hypothalamus regions.

Like the results on dataset D1, results on dataset D2 map to similar regions as shown in Figs. 8, 9, 10 and 11 except for red nucleus and hypothalamus regions. In addition, voxels from anterior cingulate, parahippocampal gyrus, inferior frontal gyrus and precuneus regions are also present in the results on dataset D2.

The regions identified by our proposed approach, are similar to previous studies on schizophrenia [10, 24, 30, 48, 51, 53]. Several comparative studies between schizophrenia patients and healthy subjects were made to localise the brain regions, responsible for the diseased state. Kim et al. [30] identified the regions like culmen, superior temporal gyrus, middle temporal gyrus, inferior frontal gyrus, postcentral gyrus, parahippocampal gyrus, precuneus, angular cingulate gyrus, and so on. In a similar study, Garrity et al. [22], showed that the patients with schizophrenia exhibited similar brain connectivity between the regions comprising posterior cingulate, precuneus and cingulate gyrus. The experiments also marked the precuneus and the middle frontal gyrus regions that are involved in selective attention, which is an important characteristic of schizophrenia [48]. Neuropathological alteration in substantia nigra region has been noticed in a previous study [51], and some task-evoked hyperactivity in this region has also been observed in schizophrenia patients [53]. In another study, Honea et al. [24] analysed the structural segment of MRI with the objective of distinguishing between healthy and schizophrenia patients. Their findings showed that the patients with schizophrenia had highly significant decreases in the frontal cortex, mainly the bilateral medial frontal cortex and inferior frontal gyri regions. They noted that the changes in the gray matter volume in the prefrontal and medial frontal cortices were more evident. In another study, decrease in hypothalamus volume was noted from structural point of view [32].

Figures 12 and 13 show the identified voxels in the brain for three different views of the brain for dataset D1 and D2, respectively. Our study also points to the differences in these regions between the schizophrenia patients and healthy subjects, as seen in Figs. 6, 7, 10, and 11. Moreover, several studies [9, 12, 23] have also identified regions similar to our study.

The main contribution of the paper lies in the application of a bi-objective optimization framework in classification of fMRI data. It uses NSGA-II to select a small set of features (voxels) that improves the classification accuracy. Our study also identifies the relevant regions of the brain which are potentially affected in schizophrenia. This study may throw light on the conventional line of treatment of the disorder.

5 Conclusion and future scope

In this paper, we have addressed the problem of feature selection in fMRI data to improve the classification accuracy in a bi-objective framework. We have proposed a three-stage

approach comprising of GLM analysis, statistical hypothesis testing, and NSGA-II to obtain a small set of relevant features that yields high classification accuracy. Thus, using a small set of 40 to 50 voxels, we achieved a mean classification accuracy of 99.5% over ten runs of the experiment. Using brain atlases in the Talairach space, we have successfully identified the regions of the brain that are mostly affected in schizophrenia patients. Specifically, we were able to identify the regions that are helpful to make a distinction between healthy subjects and schizophrenia patients. In future, one may explore the applicability of other evolutionary approaches like differential evolution, particle swarm optimization, and ant-colony optimization for identifying the brain regions affected by schizophrenia. Further, this study may be extended to incorporate the effect of different co-variables like age, gender, smoking habit, and anti-psychotic medication. It may also be interesting to explore the applicability of the proposed methodology to structural MRI analysis to find the volumetric changes in brain.

Acknowledgements We are thankful to Prof. R. K. Agrawal, School of Computer & Systems Sciences, Jawaharlal Nehru University, Delhi, India for his insightful comments. Indranath Chatterjee is thankful to the Council of Scientific & Industrial Research (CSIR), India for his research fellowship with grant number 09/045(1323)/2014-EMR-I. Naveen Kumar is thankful to University of Delhi for the research grant RC/2015/9677. Data used here for this study were downloaded from the Function BIRN Data Repository (<http://fbirnbr.birncommunity.org:8080/BDR/>), i.e., Biomedical Informatics Research Network under the following support: for function data, U24-RR021992, Function BIRN and U24 GM104203, Bio-Informatics Research Network Coordinating Center (BIRN-CC). These data were obtained from the Function BIRN Data Repository, Project Accession Number 2007-BDR-6UHZ1.

Compliance with Ethical Standards

Conflict of interests The authors declare that they have no conflict of interest.

References

1. Åberg MB, Löken L, Wessberg J (2008) An evolutionary approach to multivariate feature selection for fmri pattern analysis. In: BIOSIGNALS, vol 2, pp 302–307
2. Agarwal M, Kumar N, Vig L (2014) Non-additive multi-objective robot coalition formation. *Expert Syst Appl* 41(8):3736–3747
3. Aliakbaryhosseinabadi S, Kamavuko EN, Jiang N, Farina D, Mrachacz-Kersting N (2017) Influence of dual-tasking with different levels of attention diversion on characteristics of the movement-related cortical potential. *Brain Res* 1674:10–19
4. Arribas JJ, Calhoun VD, Adali T (2010) Automatic bayesian classification of healthy controls, bipolar disorder, and schizophrenia using intrinsic connectivity maps from fmri data. *IEEE Trans Biomed Eng* 57(12):2850–2860
5. Bellman RE (1961) *Adaptive Control Processes: A Guided Tour*. Princeton University Press, Princeton
6. Bhattacharyya A (1946) On a measure of divergence between two multinomial populations. *Sankhyā: Indian J Statist*, 401–406
7. Boser BE, Guyon IM, Vapnik VN (1992) A training algorithm for optimal margin classifiers. In: *Proceedings of the fifth annual workshop on Computational learning theory*. ACM, pp 144–152
8. Caprihan A, Pearlson GD, Calhoun VD (2008) Application of principal component analysis to distinguish patients with schizophrenia from healthy controls based on fractional anisotropy measurements. *Neuroimage* 42(2):675–682
9. Castro E, Martínez-Ramón M, Pearlson G, Sui J, Calhoun VD (2011) Characterization of groups using composite kernels and multi-source fmri analysis data: application to schizophrenia. *Neuroimage* 58(2):526–536
10. 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* 87:1–17

11. Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2:27:1–27:27. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>
12. Chen J, Xu Y, Zhang J, Liu Z, Xu C, Zhang K, Shen Y, Xu Q (2013) A combined study of genetic association and brain imaging on the *daoa* gene in schizophrenia. *Amer J Med Gen Part B: Neuropsych Gen* 162(2):191–200
13. Chyzyk D, Savio A, Graña M (2015) Computer aided diagnosis of schizophrenia on resting state fmri data by ensembles of elm. *Neural Netw* 68:23–33
14. Deb K, Pratap A, Agarwal S, Meyarivan T (2002) A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Trans Evolut Comput* 6(2):182–197
15. Demirci O, Clark VP, Magnotta VA, Andreasen NC, Lauriello J, Kiehl KA, Pearlson GD, Calhoun VD (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 Imag Behav* 2(3):207–226
16. Du W, Calhoun VD, Li H, Ma S, Eichele T, Kiehl KA, Pearlson GD, Adali T (2012) High classification accuracy for schizophrenia with rest and task fmri data. *Front Human Neurosci* 6:145
17. Ford J, Shen L, Makedon F, Flashman LA, Saykin AJ (2002) A combined structural-functional classification of schizophrenia using hippocampal volume plus fmri activation. In: *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, vol 1*. IEEE, pp 48–49
18. Ford J, Farid H, Makedon F, Flashman LA, McAllister TW, Megalooikonomou V, Saykin AJ (2003) Patient classification of fmri activation maps. In: *Medical image computing and computer-assisted intervention-MICCAI 2003*. Springer, pp 58–65
19. Fortin FA, Grenier S, Parizeau M (2013) Generalizing the improved run-time complexity algorithm for non-dominated sorting. In: *Proceedings of the 15th annual conference on Genetic and evolutionary computation*. ACM, pp 615–622
20. Frances A et al (1994) Diagnostic and statistical manual of mental disorders. DSM-IV. American Psychiatric Association
21. Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, Frackowiak RS (1994) Statistical parametric maps in functional imaging: a general linear approach. *Hum Brain Mapp* 2(4):189–210
22. Garrity AG, Pearlson GD, McKiernan K, Lloyd D, Kiehl KA, Calhoun VD (2007) Aberrant “default mode” functional connectivity in schizophrenia. *Amer J Psych* 164(3):450–457
23. Gur RE, Gur RC (2010) Functional magnetic resonance imaging in schizophrenia. *Dial Clin Neurosci* 12(3):333
24. Honea RA, Meyer-Lindenberg A, Hobbs KB, Pezawas L, Mattay VS, Egan MF, Verchinski B, Passingham RE, Weinberger DR, Callicott JH (2008) Is gray matter volume an intermediate phenotype for schizophrenia? A voxel-based morphometry study of patients with schizophrenia and their healthy siblings. *Biolog Psych* 63(5):465–474
25. Iragui VJ, Kutas M, Mitchiner MR, Hillyard SA (1993) Effects of aging on event-related brain potentials and reaction times in an auditory oddball task. *Psychophysiology* 30(1):10–22
26. Juneja A, Rana B, Agrawal RK (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, ICVGIP '14*, pp 37:1–37:6
27. Juneja A, Rana B, Agrawal R (2016) A combination of singular value decomposition and multivariate feature selection method for diagnosis of schizophrenia using fmri. *Biomed Signal Process Control* 27:122–133
28. Juneja A, Rana B, Agrawal R (2017) fmri based computer aided diagnosis of schizophrenia using fuzzy kernel feature extraction and hybrid feature selection. *Multimed Tools Appl*. 1–27
29. Kiehl KA, Liddle PF (2001) An event-related functional magnetic resonance imaging study of an auditory oddball task in schizophrenia. *Schizophren Res* 48(2):159–171
30. Kim DI, Mathalon D, Ford J, Mannell M, Turner J, Brown G, Belger A, Gollub R, Lauriello J, Wible C et al (2009) Auditory oddball deficits in schizophrenia: an independent component analysis of the fmri multisite function birn study. *Schizophren Bull* 35(1):67–81
31. Kim J, Calhoun VD, Shim E, Lee JH (2016) Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *Neuroimage* 124:127–146
32. Koolschijn PCM, van Haren NE, Pol HEH, Kahn RS (2008) Hypothalamus volume in twin pairs discordant for schizophrenia. *Eur Neuropsychopharmacol* 18(4):312–315
33. Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI (2009) Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci* 12(5):535–540
34. Lancaster JL, Woldorff MG, Parsons LM, Liotti M, Freitas CS, Rainey L, Kochunov PV, Nickerson D, Mikiten SA, Fox PT (2000) Automated talairach atlas labels for functional brain mapping. *Hum Brain Mapp* 10(3):120–131

35. Lancaster JL, Laird AR, Eickhoff SB, Martinez MJ, Fox PM, Fox PT (2012) Automated regional behavioral analysis for human brain images. *Front Neuroinform* 6:23
36. Linden DE, Prvulovic D, Formisano E, Völlinger M, Zanella FE, Goebel R, Dierks T (1999) The functional neuroanatomy of target detection: an fmri study of visual and auditory oddball tasks. *Cereb Cortex* 9(8):815–823
37. Ma X, Chou CA, Sayama H, Chaovalitwongse WA (2016) Brain response pattern identification of fmri data using a particle swarm optimization-based approach. *Brain Inform*, 1–12
38. Mitchell TM, Hutchinson R, Niculescu RS, Pereira F, Wang X, Just M, Newman S (2004) Learning to decode cognitive states from brain images. *Mach Learn* 57(1-2):145–175
39. Niiniskorpi T, Åberg MB, Wessberg J (2009) Particle swarm feature selection for fmri pattern classification. In: *BIOSIGNALS*, pp 279–284
40. O'Brien JL, Lister JJ, Fausto BA, Clifton GK, Edwards JD (2017) Cognitive training enhances auditory attention efficiency in older adults. *Front Aging Neurosci* 9:322
41. Ogawa S, Lee TM, Kay AR, Tank DW (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci* 87(24):9868–9872
42. Riaz A, Asad M, Al-Arif SMR, Alonso E, Dima D, Corr P, Slabaugh G (2017) Fcnet: a convolutional neural network for calculating functional connectivity from functional mri. In: *International workshop on connectomics in neuroimaging*. Springer, pp 70–78
43. Savio A, Graña M (2015) Local activity features for computer aided diagnosis of schizophrenia on resting-state fmri. *Neurocomputing* 164:154–161
44. Shahamat H, Pouyan AA (2015) Feature selection using genetic algorithm for classification of schizophrenia using fmri data. *J AI Data Min* 3(1):30–37
45. 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*. Springer, pp 136–143
46. Smart O, Burrell L (2015) Genetic programming and frequent itemset mining to identify feature selection patterns of ieeg and fmri epilepsy data. *Eng Appl Artif Intell* 39:198–214
47. Ülker CC, Aytekin T (2013) Improving the performance of active voxel selection in the analysis of fmri data using genetic algorithms. In: *Proceedings of the 6th Balkan conference in informatics*. ACM, pp 129–136
48. Ungar L, Nestor PG, Niznikiewicz MA, Wible CG, Kubicki M (2010) Color stroop and negative priming in schizophrenia: an fmri study. *Psychiatry Res Neuroimaging* 181(1):24–29
49. Viviani R, Grön G, Spitzer M (2005) Functional principal component analysis of fmri data. *Hum Brain Mapp* 24(2):109–129
50. Welch BL (1947) The generalization of student's' problem when several different population variances are involved. *Biometrika* 34(1/2):28–35
51. Williams M, Galvin K, O'Sullivan B, MacDonald C, Ching E, Turkheimer F, Howes O, Pearce R, Hirsch S, Maier M (2014) Neuropathological changes in the substantia nigra in schizophrenia but not depression. *Eur Arch Psychiatry Clin Neurosci* 264(4):285–296
52. Wu L, Shen C, van den Hengel A (2017) Deep linear discriminant analysis on fisher networks: a hybrid architecture for person re-identification. *Pattern Recogn* 65:238–250
53. Yoon JH, Minzenberg MJ, Raouf S, D'Esposito M, Carter CS (2013) Impaired prefrontal-basal ganglia functional connectivity and substantia nigra hyperactivity in schizophrenia. *Biol Psych* 74(2):122–129
54. Zang Y, Jiang T, Lu Y, He Y, Tian L (2004) Regional homogeneity approach to fmri data analysis. *Neuroimage* 22(1):394–400



Mr. Indranath Chatterjee is currently a Ph.D scholar at Department of Computer Science, University of Delhi, Delhi, India. He has received his B.Tech degree in Computer Science and Engineering from West Bengal University of Technology, Kolkata, India. His area of research includes statistical analysis of fMRI data, medical image processing, schizophrenia, multiobjective optimization.



Dr. Manoj Agarwal is working as an assistant professor at Hansraj College, University of Delhi, Delhi, India. He has obtained his Ph.D in Computer Science from University of Delhi. His current area of research is medical image processing and multi-objective optimization.



Dr. Bharti Rana is working as an assistant professor at Hansraj College, University of Delhi, Delhi, India. She obtained her Ph.D in Computer Science from SC&SS, JNU, New Delhi, India. Her current area of research is pattern recognition, image processing, and medical image analysis.



Dr. Navin Lakhyani is working as a senior consultant and head at Department of MRI, Saral Diagnostics, Delhi, India. He has obtained MD in Radio diagnosis, Rabindra Nath Tagore (RNT), Udaipur, India. His area of research is neuroradiology, and neurological problems.



Prof. Naveen Kumar is working as a Professor at Department of Computer Science, University of Delhi, Delhi, India. He has received his Ph.D in Computer Science from Indian Institute of Technology (IIT), Delhi, India. His current area of research is medical image processing, multiobjective optimization, process mining and security.