Asian Journal of Information Technology 15 (10): 1607-1620, 2016
ISSN: 1682-3915

Functional Mribold Interpretation of Brain Functions Using Independent Component Analysis

[1]P. Suresh and [2]K. Bommannaraja
Department of Electronics and Communication Engineering,
Eranad Knowledge City Technical Campus, Cherukulam, Manjeri, Malappuram, Kerala, India
[2]Department of Biomedical Engineering,
KPR Institute of Engineering and Technology, Coimbatore, Tamil Nadu, India

Abstract: Functional MRI (fMRI) is a widely used technique to study about the brain activation and its implications through Blood Oxygenation Level Dependent (BOLD) interpretations. Data-driven analysis methods, in particular Independent Component Analysis (ICA) have proven quite useful for the analysis of fMRI data. A promising approach to multi-subject analysis is Group Independent Component Analysis (GICA), which identifies group components and reconstructs activations at the individual level. In this research, a robust model-free technique is proposed for detecting the fMRI activations during the tasks activating the language areas of the Brain. We evaluated the performance of the proposed method on a moderate size real time fMRI data acquired under three different tasks. This results in a set of spatial maps and time courses which are common to the whole group, together with an individual response activation map for each of the subjects in the group. The results show that, ICA components are involved in the direct correlation between language based tasks and their spatial/time course maps.

Key words: Functional neuroimaging, Functional magnetic resonance imaging (fMRI), Independent Component Analysis (ICA), inter-task fMRI analysis, Principal Component Analysis (PCA).

INTRODUCTION

Functional neuroimaging aims at finding brain regions specifically involved in the performance of cognitive tasks. In particular, Functional Magnetic Resonance Imaging (fMRI) is based on the detection of task-related Blood Oxygen-Level Dependent (BOLD) effect in the brainwhich is based on the differing magnetic properties of oxygenated (diamagnetic) and deoxygenated (paramagnetic) blood (Thirion, *et al*., 2007). The fMRI is a technique that provides the opportunity to study brain function noninvasively and is a powerful tool that has been used in both research and clinical arenas since the early 1990s (Kwong *et al*., 1992; Bai *et al*., 2007). fMRI data analysis methods can be broadly divided into two groups: model-driven and data-driven. When an expected response can either be obtained or estimated, a model-driven method and in the other case data-driven analysis methods present a more appropriate alternative (Friston *et al*., 1995). The data-driven methods have the advantage of data exploration, since no assumptions need to be made about the spatial or temporal characteristics of the neuronal responses (Lee *et al*., 2011). Data-driven analysis methods, in particular Independent Component Analysis (ICA) have proven quite useful for the analysis of fMRI data (Vince *et al*., 2006). By using a simple generative model based on linear mixing, ICA can minimize the constraints imposed on the temporal or the spatial dimension of the fMRI data and hence provides valuable new insights, especially when studying paradigms for which reliable models of brain activity are not available (Li *et al*., 2010).

FMRI pre-processing: FMRI time-series contain a number of systematic sources of variability that are not due to the BOLD effect of interest. These sources of systematic variability may be removed as a preprocessing step. The sources of variability include factors due to the physics of MRI, subject motion, heart beat and breathing, other physiological processes, random thermally generated noise and intersubject anatomical variability. The main preprocessing steps involved in fMRI time series analysis are Slice-Timing Correction, Spatial Realignment, Co-registration, Normalisation and Spatial Smoothing

Corresponding Author: P. Suresh, Department of Electronics and Communication Engineering,
Eranad Knowledge City Technical Campus, Cherukulam, Manjeri, Malappuram, Kerala, India

(Turner and Twieg, 2005; Lee *et al*., 2007). In this study, several preprocessing steps will be discussed that are commonplace in the analysis of fMRI data. BOLD fMRI data typically are acquired as echoplanar images and as such are likely to be distorted in space to some extent as a result of static magnetic field inhomogeneities produced by the concentration of magnetic field lines. A single volume of BOLD fMRI data, collected during one repetition time (TR) is assembled from multiple planar acquisitions or slices. One slice is collected at a time, either sequentially or in an interleaved fashion with the result that each slice samples will have a slightly different point in time. Slice acquisition correction compensates for this staggered order of slice acquisition (Richards and Berninger, 2008). A variety of methods are used to minimize head motion during scanning. This process is named as motion correction. These include foam padding around the head; bite bars; custom-designed, thermo-plastic face masks; and so on. This generally is done by realigning the image of the brain obtained at each point in time back to the first image acquired at the start of the scanning session (Cox and Savoy, 2003).

Analysis of a defined area of the brain across subjects, is done by computationally warping the anatomical structure of the brain of one subject to match a template brain within a standard defined space which is known as spatial normalization. It is a common practice to digitally smooth BOLD fMRI data in space prior to statistical analysis (Spiers and Maguire, 2007). By spatially smoothing the data in space, the number of independent statistical tests that are being performed is reduced.

ICA of FMRI time series: From a multivariate perspective, an fMRI data set may be thought of in two ways. In the usual way, the data are represented by a $^{p\times v}$adata matrix [M] in which each row represents one 3D voxel data set taken at a fixed scan time. That way, [M] represents p measurements of a v-dimensional quantity, the spatial image. The other way is to represent the data by a$^{v\times p}$ matrix $[Y] = [M]^T$ in which each column represents one 3D voxel data set taken at a fixed scan time. Viewed this way we have v measurements of a p-dimensional quantity, the voxel intensity time course. An ICA of [M] produces spatially independent components (component images) and is termed spatial ICA (sICA) and the ICA of [Y] produces temporally independent components (component time courses) and is termed temporal ICA (tICA) (Calhoun and Adali, 2006). The generation of components and the procedure using ICA is shown in Fig. 1. More specifically, let [W] be the estimate of the mixing matrix and [C] be the matrix containing the components as rows. The sICA approach gives:

$$[M] = [\hat{W}]^{-1}\,[C] \tag{1}$$

and the columns of [W] give the time courses corresponding to the spatial components in the rows of [C]. In general the number of rows in [C] is chosen to be <the number of rows in [M].The tICA approach which generally is used only on an ROI basis to limit the number of computations required, gives:

$$[Y] = [\hat{W}]^{-1}\,[C] \tag{2}$$

and the columns of $[W]^{-1}$ give the spatial images corresponding to the time course components in rows of

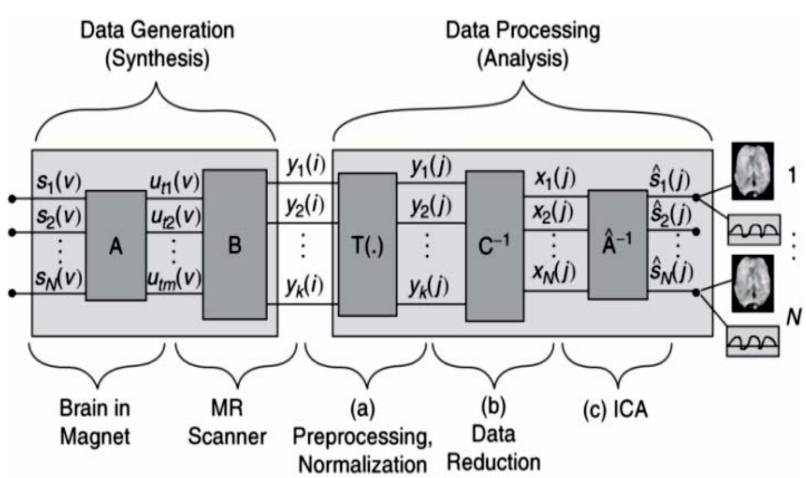

Fig. 1: A model for applying ICA to fMRI data (Calhoun *et al*., 2001a)

[C]. The principle behind ICA is that the data are considered to be mixtures of sources, represented in the rows of [S] according to:

$$[M] = [A][S] \tag{3}$$

where, [A] Is the mixing matrix. The goal is to find the unmixing matrix $[W] = [A]^{-1}$ so that:

$$[S] = [W][M] \tag{4}$$

ICA will produce estimates [C] and $[w]^{-1}$ of [S] and [W] respectively to give:

$$[C] = [\hat{W}][M] \tag{5}$$

The sources are assumed to be statistically independent. Such statistical independence is a stronger condition of independence than orthogonality (Calhoun *et al*., 2004; McKeown *et al*., 2002). The spatial ICA (sICA) refers to the statistical distribution of signals across the sampled hemodynamic locations, while temporal ICA (tICA) refers to the statistical distribution of source signals across the sampled time-points (Esposito *et al*., 2005). In ICA approach individual task related components are important as the group component maps and time courses. The components related to each task are back reconstructed from the group maps as follows. The data in a voxel sampled at two time points from the subjects while performing the three tasks (Task 1, Task 2 and Task 3) can be represented as:

$$d_1 = [x_1, x_2], d_2 = [y_1, y_2] d_3 = [z_1, z_2] \tag{6}$$

each of which is a normalized linear mixture of two hemodynamic sources. Concatenating these subjects into a single vector will result in $d = d\,[d_1\ d_2\ d_3]$. Assuming the correct number of sources is two, the data reduction and ICA analysis result in a mixing matrix:

$$W = [W1\ \ W2\ \ W3] = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \end{bmatrix} \tag{7}$$

and an estimate of the original sources s = Wd where W is partitioned to depict the submatrices corresponding to the original mixed sources. To back-reconstruct the individual subject maps, the partition of W corresponding to the desired subject's data is multiplied with the corresponding partition of d (e.g. $s_x = W_x\, d_x$). As the goal of ICA is to yield independent components, the rows of s will be approximately statistically independent. Additionally, data from each subject is expected to be independent of each other. Then the expression for s can be written as:

$$\hat{s} = \begin{bmatrix} s_{11} + s_{12} + s_{13} \\ s_{21} + s_{22} + s_{23} \end{bmatrix} \tag{8}$$

where: $s_{11} = a_1x_1+a_2x_2$, $s_{12} = a_3y_1+a_4y_2$,$s_{13} = a_5z_1+a_6z_2$,$s_{21} = b_1x_1+b_2x_2$,$s_{22} = b_3y_1+b_4y_2$, $s_{23} = b_5z_1+b6_6\ z$. As the ICA algorithm minimizes the dependence among the signals, the dependence between s_{11} and s_{21}, s_{12} and s_{22} and s_{13} and s_{23} will be minimized by more heavily relying on the data within that subject, forcing the parameters for each subject to be primarily determined by that subject's observations. Thus, the individual unmixing matrices will be approximately separable across subjects (partitions) and the back-reconstructed data will be a function of primarily the data within subjects rather than across subjects (Calhoun *et al*., 2001b; Ghasemi and Mahloojifar, 2010).

Experiment data sets: In this study, we used the real-time fMRI data sets provided by the Department of Imaging Sciences and Interventional Radiology, Sree Chithra Thirunal Institute for Medical Sciences and Technology, Trivandrum. In this fMRI recording experiment, the subjects were scanned for 4 min with a TR of 2000 ms and the condition switched every 30s between "rest" and "task," starting with "rest". The tasks paradigms employed in this experiment were Verb generation (Task 1), Story listening (Task 2) and Semantic processing (Task 3). The data set of each subject consists of 100 volumes with a resolution of 64 X 64 X 36 voxels. The dimension of each voxel volume is 3.75 X 3.75 X 5.50 mm. The real-time 3D fMRI volumes, each consisting of 36 axial planes of 64 X 64 voxels were imported to MATLAB 11.0 in ANALYSE format and the volume is spatially realigned to the first slice and co-registered using SPM99. The threshold value to identify the regions of interest was set to 10 to avoid the computational complexity due to the denoising of the background.

Proposed method and data processing: Data pre processing was performed using SPM5 (Friston *et al*., 1995; Genovese *et al*., 2002) including slice-timing,

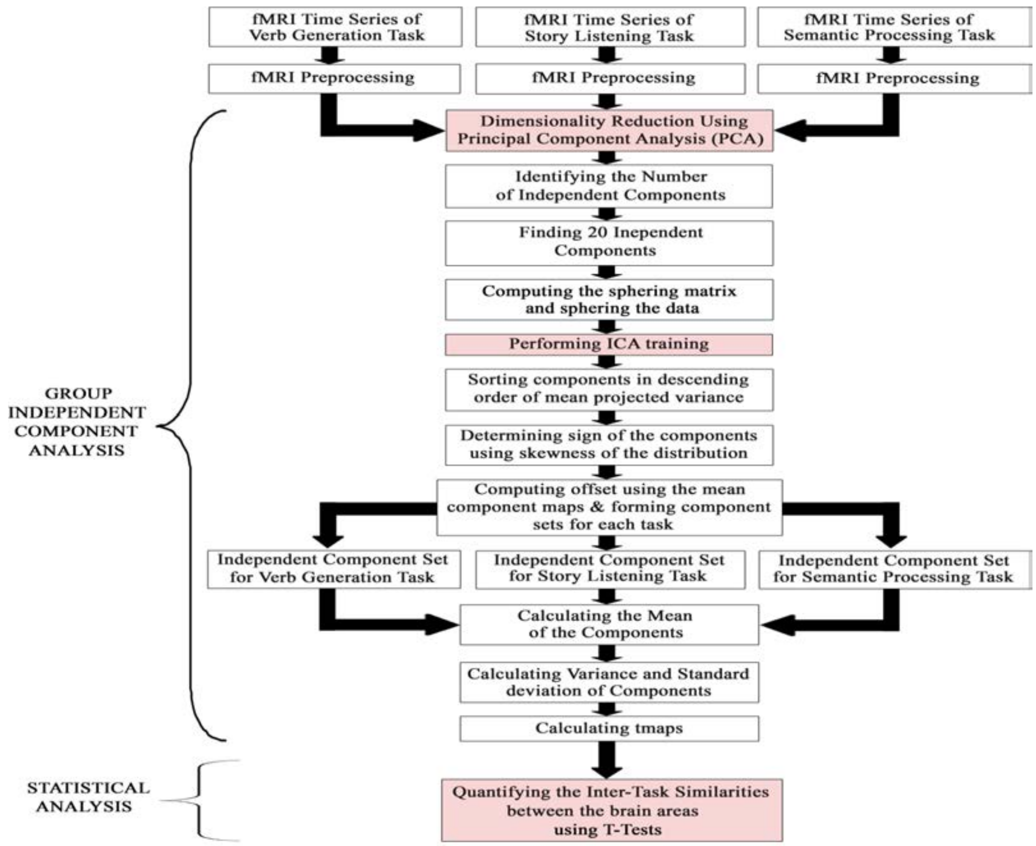

Fig. 2: Detailed procedure of the proposed method

motion correction, spatial normalization and spatial smoothing using a Gaussian kernel (FWHM = 9mm). Multiple Linear Regression [MLR] was applied to correct data sets for residual motion and global signal changes, before applying band pass filtering (0.08-0.9Hz) based on the recommendation of a recent study (Weissenbacher *et al*., 2009). The real time fMRI data acquired during three different tasks activating the language areas of the brain are considered as three groups. Group 1 is considered as the fMRI data acquired through Verb generation task in which the verbs were generated by the subjects after hearing the meaningful information related to the verbs. The fMRI series scanned while the subjects are subjected to listen a story is considered in Group 2 and Group 3 represents the functional brain volumes acquired while the subjects are directed to produce the sound corresponding to selected alphabets. The postprocessing was performed using group ICA implemented in GIFT (Dea *et al*., 2011) version 2.0.b, as a MATLAB toolbox. The Detailed procedure of the proposed Method is given in Fig. 2.The complete work flows through three main steps: 1) Dimensionality Reduction using PCA 2) Independent Component Analysis of fMRI data and 3) Quantifying the Inter task similarities between the brain areas using the T Tests

Dimensionality reduction using pca: Estimating the number of components to be processed and analysed is one of the important step to be considered before conducting any component based analysis, as this will vary with respect to the voxel based BOLD activations of the fMRI data. In this research, the components are estimated from the fMRI data using the Minimum Description Length (MDL) criteria (Rissanen, 1978) and fixed to 20 components. In ICA algorithms, the dimensionality reduction of the data is performed as the first step (Svensen *et al*., 2002). Ideally, the reduced representation should have a dimensionality that corresponds to the intrinsic dimensionality of the data (Lai and Fang, 1999; Zhang *et al*., 2011). Although, there exist various techniques to do so, Principal Components Analysis (PCA) is by far the most popular unsupervised linear technique (Maaten *et al*., 2009; Tsatsishvili *et al*., 2013) and hence employed in this research. PCA constructs a low-dimensional representation of the data that describes as much of the variance in the data as possible (Schopf *et al*., 2010). In this research, the dimensionality reduction of the three groups is carried out using the Standard type PCA algorithm in two steps. Firstly the groups were concatenated to form three new groups each consisting of 100 BOLD time points and then the size of these groups were then reduced to 30 principal components. Secondly, these three new concatenated groups, which contain 30 principal components each, were again concatenated to form a new group with a total of 90 stacked principal components. This group will then be reduced to 20 principal components. While performing PCA in each step the covariance matrix in time dimension was calculated from the observed data and eigen values were computed from the covariance matrix. The eigen values were arranged in decreasing order and only the eigen vectors that correspond to non-zero and first 20 components were taken into account. These components were orthogonal to each other and then passed to the whitening step. Whitening matrix was computed by solving the square root of diagonal matrix of eigen values and transpose of the eigen vectors. This matrix is then multiplied to the data to make the variances equal for all the components. The condition number of covariance matrix obtained in the first step for group 1, 2 and 3 were 66470.51, 68431.66 and 64328.66 respectively. 38.6058% of non-zero eigenvalues were retained as a total after completing all the steps. Finally the dimension of the data was reduced from 64 X 64 X 36 X 100 to 64 X 64 X 36 X 20, where 100 refers to the time points or scans.

Independent component analysis of fmri data: Several studies have already compared the most commonly used algorithms for ICA, Infomax and FastICA, using second order statistics (Esposito *et al*., 2002; Naik and Kumar, 2011; Correa *et al*., 2007). The majority of applications of ICA to fMRI use Infomax as the sources of interest and as already shown by Correa *et al*.(2007), Infomax based ICA approach give the best overall performance when being applied on fMRI data. Therefore, this study was limited to the use of Infomax algorithm. Infomax (Calhoun *et al*., 2005) maximizes the information transfer from the input to the output of a network using a non-linear function. The mean per time point was removed from the 20 components as a pre-processing step. GIFT, the group ICA toolbox for MATLAB, rely on temporal concatenation of single-task data and the mixing matrix is divided into subject-specific parts. After the estimation of the mixing matrix, single-task component maps was computed via projection of the single-task data onto the inverse of the individual task-specific part of the mixing matrix (Calhoun *et al*., 2001c) and the spatial maps are transformed to z-scores voxel-wise. In addition, a back-projection step was implemented, which returns subject-specific maps and time courses. The ICA maps from individual subjects are back reconstructed from the aggregate mixing matrix. This was based on the classical ICA framework assumption that a data set is characterized by the mixing-matrix and estimated sources. This back-reconstructed data will be a function of primarily the data within subjects rather than across subjects. The resulting group ICA maps were thresholded using a Z-threshold criterion (Kelly *et al*., 2010). Spatial maps were normalized using the maximum intensity value and the maximum intensity value is multiplied to the time courses. Spatial maps were scaled using the standard deviation of time courses and time courses are scaled using the maximum spatial intensity value (Sui *et al*., 2009). Component images are overlaid on the anatomical image of the concerned subject performing the three tasks. The components are classified using sorting crieteria both spatially and temporally. Temporal sorting is a way to compare the model's time course with the ICA time course whereas spatial sorting classifies the components by comparing the component's image with the template (Woods, 1996; Mckeown *et al*., 1998). Sorting criteria have been implemented using Multiple Linear Regression (MLR) algorithm which is a very useful method in separating the two task related components. As the ICA analysis produces estimates of hemodynamic sources, there is a physiologic meaning in calculating the mean and standard deviation of each component across subjects which is implemented in GIFT. In addition to the spatial maps, this analysis produces a representative time course. For each independent component, spatial maps and time courses are generated. The unreduced fMRI time courses were reconstructed and these time courses reflect the hemodynamics of the fMRI experiment and were inspected separately for each subject and are averaged across subjects to create a group time course.

Quantifying the inter task similarities using the t tests: Statistical group inferences are made through calculation of the probability for activation based on the mixture model for the voxel-wise z-score. Mean and variance of each component across subjects are calculated and the variance across subjects is used as an estimator of the population variance (Dronkers *et al*., 2007). Resulting single-task time courses are scaled using the original raw data to reflect percent signal change, followed by normalization to z-score values. Statistical group inferences are performed by testing all voxels of an ICA component set using a one-sample t-test. In this research, three groups were selected:

- Verb generation Vs Story listening
- Verb generation Vs Semantic processing and
- Story listening Vs Semantic processing

The weights of all voxels are treated as random variables and are tested against the null hypothesis of zero weight. The significance of the component time courses are also analyzed by doing statistics on the beta weights which are obtained after doing temporal sorting on all data-sets. After temporal sorting, the beta weights, ie. the slopes of regressors and regression values are computed. These beta weights were further used to determine the task-relatedness of the components. Thereafter, the obtained fit regression parameters are used in performing statistical tests to evaluate the task-relatedness of the components or to perform tests between groups. The fit regression parameters are:

- A number for each component
- Regressor
- Subject
- Session

The numbers represent the 'fit parameter' from the regression. In order to use these numbers, it is important to scale the ICA data. A separate test was carried out for each component and the components which shows significant difference were determined.

RESULTS AND DISCUSSION

ICA provides a method to ''blindly'' separate the data into spatially independent components, enabling exploratory analysis on fMRI data (Bloch *et al*., 2009). In this research, the results were inspected manually, to detect the task related components. Figure 3 shows the independent components from the single task ICA, associated with the time courses showing the highest correlation with the performance of the verb generation task, overlaid on the anatomical image. The corresponding time courses are shown in Fig. 12a and b, plotted together with the maximum time course representing the performance of the task. From the response pattern of the spatial ICA of verb generation task, most of the response is located to the components 18 and 13. In the images resulting from ICA of the individual tasks, the spatial responses are more similar across the group.

Components 18 and 15 showed pronounced activation out of the 20 independent components of the fMRI data acquired while undergoing story listening task which is shown in Fig. 4. The time courses of these components are given in Fig. 12c and d. Spatial activation map of top two components (12 and 8) responding to semantic processing task is given in Fig. 5 and the corresponding time courses are shown in Fig. 12e and f. Figure 3, 4, 5 and 12 clearly shows the similarity and differences in the activation of brain areas and the strength of time courses in these components. Group analysis of fMRI is important to study specific conditions within or between groups of subjects or tasks. Group analysis is needed to quantify the comparison statements between the components representing the language areas of the brain with respect to the tasks.

Figure 6 shows the composite Spatial and Temporal activation maps of the above said components with BOLD activity in the Language area (Just *et al*., 1996) with real coordinate: -32 52 22 for all the three tasks. The two components were distinguished using red and blue colours in Fig. 6a, b and c which correspond to the verb generation, story listening and semantic processing tasks, respectively. One slice corresponding to each category of task is presented in zoomed mode to view the activation more clearly in Fig. 6. The mean component mapped using the group ICA method is more significant as it explains the depth of activations in the interested areas more clearly as shown in Fig. 7.

All the selected 20 components of the mean activation map are presented in Fig. 7 through which the components commonly shares the similarity is more evident. The bold signal at the selected voxel coordinate for the tasks is given in Fig. 8. The main aim of this research is to compare the activations in brain areas related to language processing using ICA on fMRI data recorded while performing tasks 1, 2 and 3. Figure 9 represents the comparison of spatial map of the components with maximum activation between task

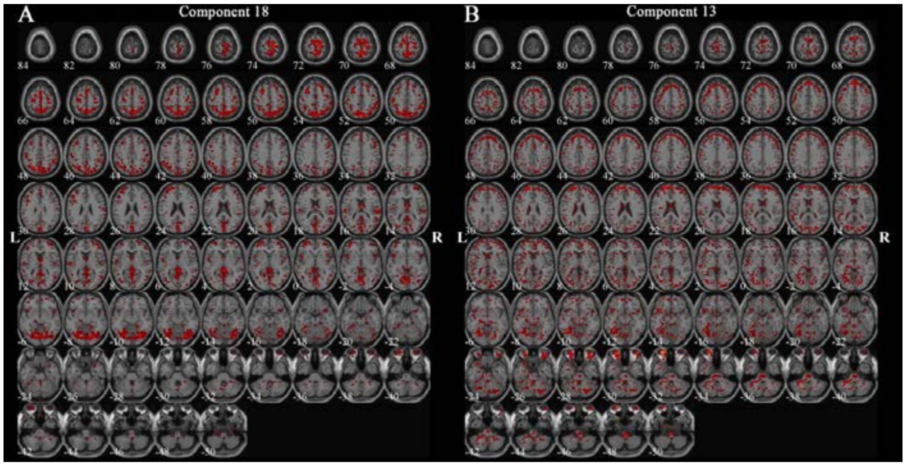

Fig. 3: Spatial activation map of top two components responding to verb generation task: a) Component 18; b) Component 13

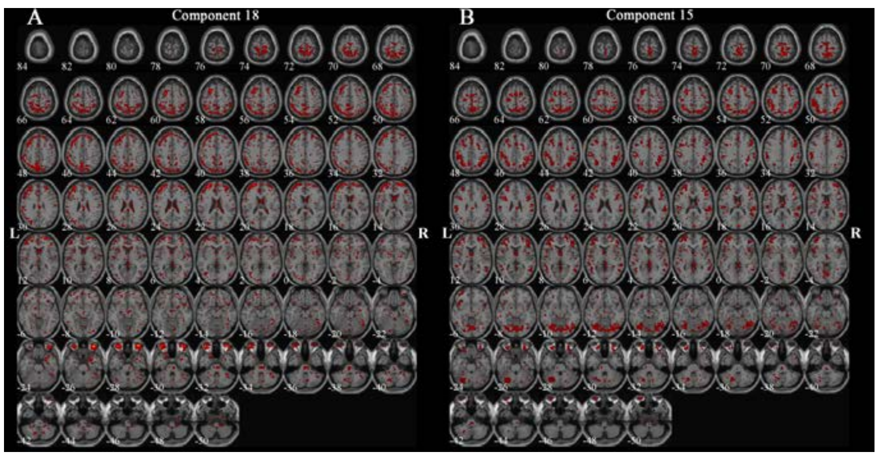

Fig. 4: Spatial activation map of top two components responding to story listening task: a) Component 18; b) Component 15

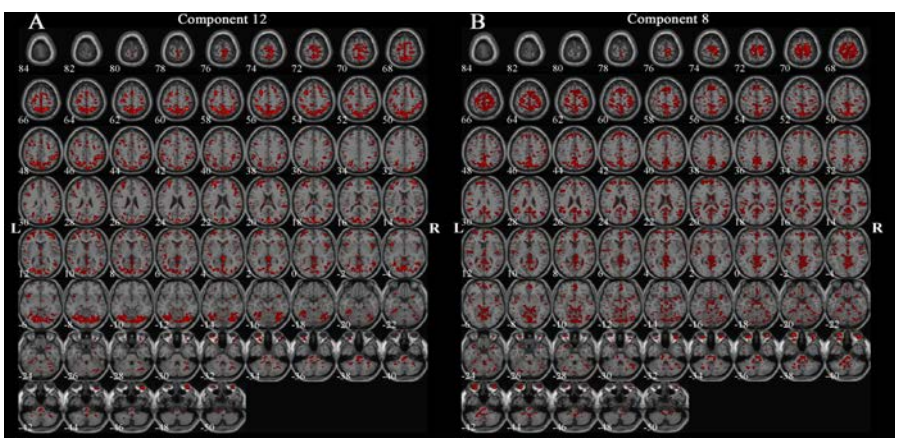

Fig. 5: Spatial activation map of top two components responding to semantic processing task: a) Component 12; b) Component 8

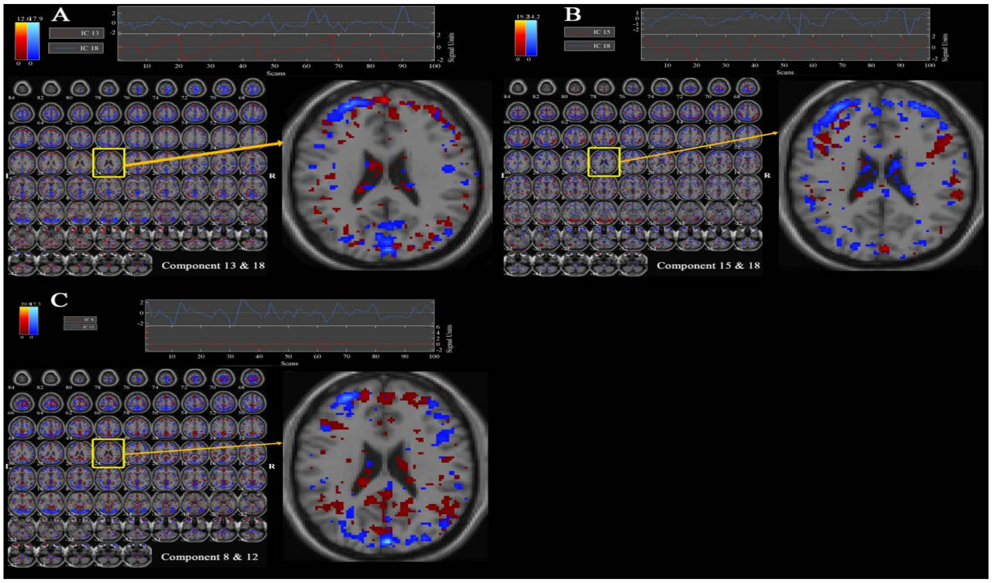

Fig. 6: Composite spatial and temporal activation maps of the two components with maximum BOLD activity in the language area (Real coordinate: -32 52 22) for all the three tasks: a) Components 13&18 on verb generation task; b) Components 15&18 of story listening task; c) Components 8&12 of semantic processing task

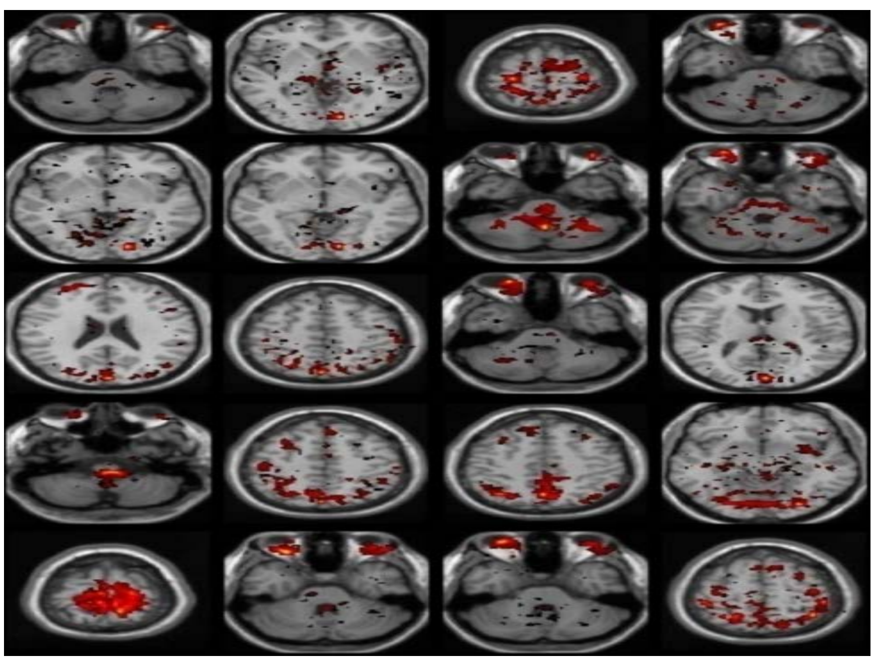

Fig. 7: Mean components with maximum voxel values of all the tasks

1 and task 2. BOLD activation in the Prefrontal Cortex is evident during verb generation task in components 18 and 13 as shown in Fig. 9a and b and the same during story listening task is given in Fig. 9c and d. The activation in the language area during task 2 and task 3 was compared visually from Fig. 10 and during task 1 and task 3 is given in Fig. 11. One sample t test was carried out to study about the similarities and dissimilarities among the components for the evaluated tasks. Table 1 shows the components

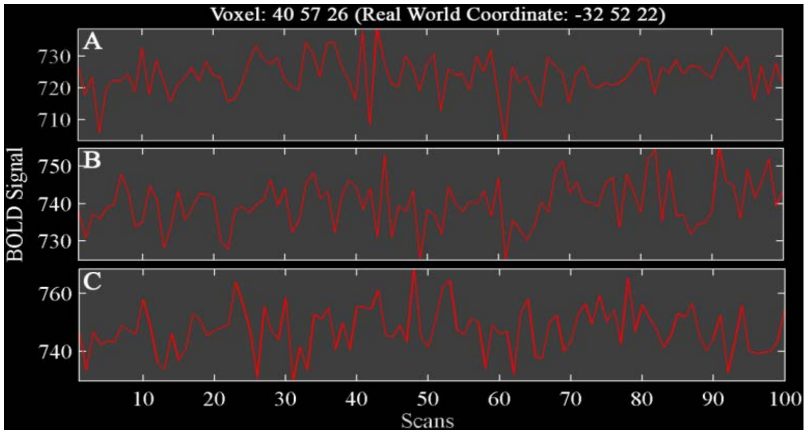

Fig. 8: BOLD signal time course at -32 52 22: a) Verb GENERATION Task; b) Story listening task; c) Semantic processing task

Table 1: Components showing similarity

Groups	ICs	p-value	t-value	Mean	Std
Verb gen.	9	0.31	0-1.88	0-1.38	1.04
Vs	15	0.24	0-2.47	0-0.77	0.44
Story List.	18	0.36	0-1.56	0-2.76	2.50
Verb gen.	11	0.16	0-3.88	0-6.18	2.25
Vs	13	0.08	0-7.60	0-1.95	0.36
Sem Proc.	17	0.10	0-6.61	0-5.67	1.21
Story List.	4	0.10	0-6.01	0-1.88	0.44
Vs	9	0.16	0-3.92	0-2.83	1.02
Sem Proc.	11	0.02	0-30.33	0-4.74	0.22

Table 2: Components showing difference

Groups	ICs	p-value	t-value	Mean	Std
Verb gen.	2	0.60	0.72	0.87	1.72
Vs	6	0.48	1.06	0.67	0.89
Story List.	20	0.17	3.56	1.74	0.69
Verb gen.	5	0.56	0.82	0.30	0.52
Vs	6	0.48	1.05	0.73	0.98
Sem Proc.	19	0.30	1.99	1.54	1.09
	20	0.40	1.36	1.28	1.34
Story List.	2	0.18	3.48	1.62	0.66
Vs	6	0.03	21.73	1.36	0.09
Sem Proc.	7	0.44	1.21	1.01	1.18
	12	0.35	1.63	0.21	0.18
	16	0.08	7.47	1.15	0.22
	20	0.33	1.74	0.79	0.65

with similar activations during the tasks and Table 2 shows the components with significant dissimilarities in the activation during the tasks.

We have demonstrated an approach for applying Independent Component based method Inter task analysis to fMRI data. In this research, the results were analyzed in two forms: Individual task based and inter task based. The subjects when under gone verb generation task paradigm showed pronounced activations in components 18 and 13 as shown in Fig. 3. Figure 3a the voxels with confine BOLD activation in component 18 is evident. This component presents the maximum response to the verb generation paradigm. Targeting one IC opens the possibility for tracking the cumulative history of the underlying activation phenomenon. This is done practically by producing increasingly accurate spatial maps over the entire time of the real-time fMRI session. Out of the remaining 19 components IC13 presented a sound activation to this task which is given in Fig. 3b. Figure 4 and 5 also shows the top two components which responded with good degree of activation to story listening and semantic processing task paradigms, respectively. Another way to visualize the task based characterization of the components is to view the composite frame where both the two components with maximum voxel values can be compared with respect to the strength of response to the tasks as shown in Fig. 6. The component with heaviest response in denoted by the blue color and the second highest by the red color. Finding the components with similar behavior for all the three tasks, i.e. which are closely related to language processing, is the main objective of this work. Comparative analysis mostly rely on a reference map which is a mean spatial map in which the mean activation towards the entire set of tasks are visible which is given in Fig. 7. All the 20 components are shown in Fig. 7 with the mean activation computed from the activations obtained from individual tasks.

In this study, the comparison is limited to the Brocas area which is one of the vital brain area involved in language processing as given in literatures. The BOLD signal which is considered to be one of the direct representations of the brain activation, at this real coordinate is visually studied and compared using Figure 8. The maximum BOLD value for Verb generation task at (-32 52 22) is obtained during the scan number 42 with an approximate value of 740 as shown in Fig. 8a. Correspondingly for story listening and semantic processing paradigms, scan numbers 82 and 91 and scan number 48 shows maximum BOLD signal at this coordinate respectively which is given in Fig. 8b and c. This clearly gives the difference in the BOLD signal generated during different tasks and the different scan times.While comparing the ICA results in this research,

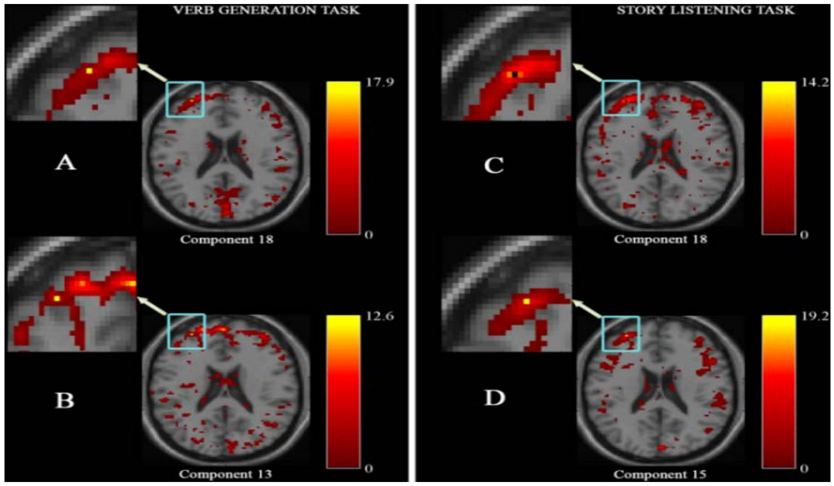

Fig. 9: Comparison of spatial activation map of top two components responding to verb generation task (A&B) and story listening task (C&D)

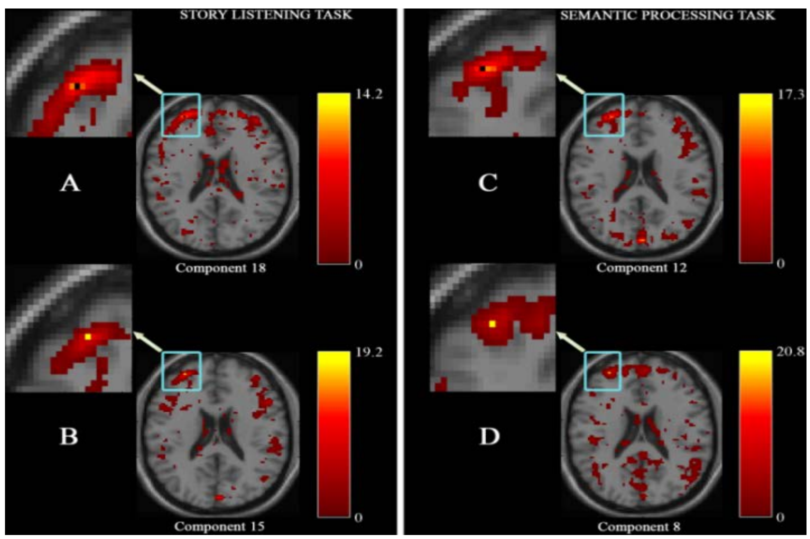

Fig. 10: Comparison of spatial activation map of top two components responding to story listening task (A&B) and semantic processing task (C&D)

the top two components responding to the tasks are taken into account. The activation pattern produced due to each task at Brocas area is visualized as shown in Fig. 9, 10 and 11. Even though, the patterns are not much similar, the activation strength defines a close relation of this brain area to the verb generation, story listening and semantic processing tasks which can be categorized under the group of language processing task. Figure 9, 10 and 11 activation pattern at the ROI is presented separately. In Fig. 9a and c the component IC18 scores the maximum correlation between verb generation and story listening tasks and IC13 shows a good degree of similarity with IC18 and IC15. The comparison of auditory and visual based language processing tasks is given in Fig. 10. From Fig. 10a-d components 18, 15 and components 12, 8 themselves explain the resemblances at the ROI. IC18 and IC13 computed for verb generation task presents sound similarity with IC12 and IC8 obtained from semantic processing task as shown in Fig. 11. The combined study of time courses at the ROI also gives the information about the criteria for selecting the top two components for each task which is shown in Fig. 12. The one sample t test results shown in Table 1 and Table 2 clearly says that components 9, 15 and 18 shows similar activations during the task 1 and task 2 while components 2, 6 and 20 shows significant dissimilarities. Major similarities in comparing task 1 and 3 are shown by the components 11, 13 and 17. Components 5 6 and

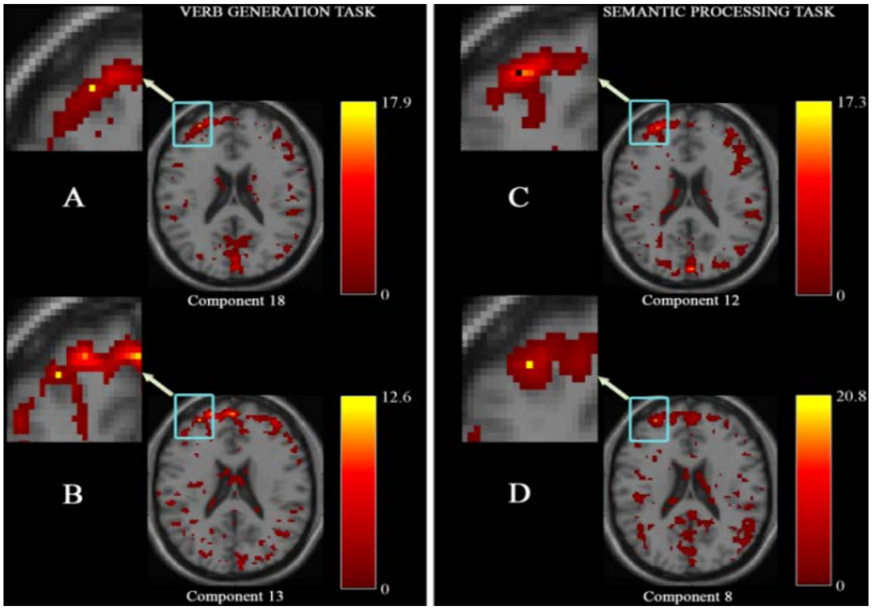

Fig. 11: Comparison of spatial activation map of top two components responding to verb generation task (A&B) and semantic processing task (C&D)

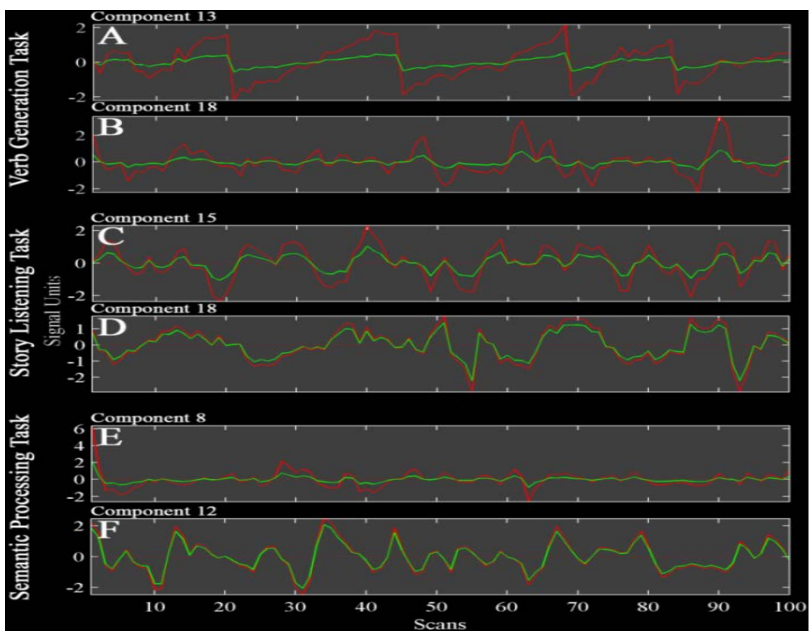

Fig. 12: Comparison of ICA time courses of the top two components with maximum activation in the language area (Real coordinate: -32 52 22) during the tasks with green representing the corresponding component time course and red representing the maximum time course

19 differs in their activation for task 1 and 3. While performing task 2 and 3, components 4, 9 and 11 shows similar activations in the voxel of interest.

CONCLUSION

Independent Component Analysis (ICA) has emerged as a powerful technique for the analysis of functional neuroimaging data. The development of Group ICA (GICA) frameworks has further increased the utility and popularity of this data-driven approach, permitting straightforward application of ICA to multi-subject datasets. In a group study, we are interested in which components are consistently contributing significantly. A hypothesis test can then be used to provide a "random effects" inference: the magnitudes or weights of the voxels within a set of ICA components are treated as random variables and a one-sample -test with the null hypothesis of zero magnitude is performed. As the ICA analysis produces estimates of hemodynamic sources, there is a physiologic meaning to the results and such an approach is justified.

We have implemented a method for making group inference maps through the application of independent component analysis to fMRI data. The application of a preprocessing stage in which the global average for each subjects' data is normalized to the same value as well as normalization of the back-reconstructed spatial maps by the time course variance is emphasized. We have extended independent component analysis of fMRI data to provide for group inferences during various tasks. Our method has general applicability is straightforward to apply and should be computationally reasonable for many fMRI group studies. This research may prove useful as an addition to the collection of ICA approaches for Inter task analysis of fMRI data.

REFERENCES

Bai, B., P. Kantor, A. Shoukoufandeh and D. Silver, 2007. FMRI brain image retrieval based on ICA components. Proceedings of the ENC 2007 Eighth Mexican International Conference on Current Trends in Computer Science, September 24-28, 2007, IEEE, Michoacan, Mexico, ISBN: 978-0-7695-2899-1, pp: 10-17.

Bloch, C., A. Kaiser, E. Kuenzli, D. Zappatore and S. Haller *et al*., 2009. The age of second language acquisition determines the variability in activation elicited by narration in three languages in Brocas and Wernickes area. Neuropsychologia, 47: 625-633.

Calhoun, V., G. Pearlson and T. Adali, 2004. Independent component analysis applied to FMRI data: A generative model for validating results. J. VLSI. Signal Process. Syst. Signal Image Video Technol., 37: 281-291.

Calhoun, V.D. and T. Adali, 2006. Unmixing fMRI with independent component analysis. IEEE. Eng. Med. Biol. Mag., 25: 79-90.

Calhoun, V.D., T. Adali, G.D. Pearlson and J.J. Pekar, 2001a. A method for making group inferences from functional MRI data using independent component analysis. Hum. Brain Mapp., 14: 140-151.

Calhoun, V.D., T. Adali, G.D. Pearlson and J.J. Pekar, 2001b. Spatial and temporal independent component analysis of functional MRI data containing a pair of task-related waveforms. Hum. Brain Mapp., 13: 43-53.

Calhoun, V.D., T. Adali, M.C. Stevens, K.A. Kiehl and J.J. Pekar, 2005. Semi-blind ICA of fMRI: A method for utilizing hypothesis-derived time courses in a spatial ICA analysis. Neuroimage, 25: 527-538.

Calhoun, V.D., T. Adali, V.B. McGinty, J.J. Pekar and T.D. Watson *et al*., 2001. FMRI activation in a visual-perception task: Network of areas detected using the general linear model and independent components analysis. NeuroImage, 14: 1080-1088.

Correa, N., T. Adalý and V.D. Calhoun, 2007. Performance of blind source separation algorithms for fMRI analysis using a group ICA method. Magn. Reson. Imaging, 25: 684-694.

Cox, D.D. and R.L. Savoy, 2003. Functional Magnetic Resonance Imaging (FMRI) brain reading: Detecting and classifying distributed patterns of FMRI activity in human visual cortex. Neuroimage, 19: 261-270.

Dea, J.T., M. Anderson, E. Allen, V.D. Calhoun and T. Adalý, 2011. IVA for multi-subject FMRI analysis: a comparative study using a new simulation toolbox. Proceedings of the 2011 IEEE International Workshop on Machine Learning for Signal Processing, September 18-21, 2011, IEEE, Santander, Spain, ISBN: 978-1-4577-1621-8, pp: 1-6.

Dronkers, N.F., O. Plaisant, I.M.T. Zizen and E.A. Cabanis, 2007. Paul Brocas historic cases: High resolution MR imaging of the brains of Leborgne and Lelong. Brain, 130: 1432-1441.

Esposito, F., E. Formisano, E. Seifritz, R. Goebel and R. Morrone *et al*., 2002. Spatial independent component analysis of functional MRI time-series: To what extent do results depend on the algorithm used?. Hum. Brain Mapp., 16: 146-157.

Esposito, F., T. Scarabino, A. Hyvarinen, J. Himberg and E. Formisano *et al*., 2005. Independent component analysis of fMRI group studies by self-organizing clustering. Neuroimage, 25: 193-205.

Friston, K.J., C.D. Frith, R. Turner and R.S. Frackowiak, 1995. Characterizing evoked hemodynamics with FMRI. Neuroimage, 2: 157-165.

Genovese, C.R., N.A. Lazar and T. Nichols, 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage, 15: 870-878.

Ghasemi, M. and A. Mahloojifar, 2010. FMRI data analysis by blind source separation algorithms: A comparison study for nongaussian properties. Proceedings of the 2010 18th Iranian Conference on Electrical Engineering, May 11-13, 2010, IEEE, Isfahan, Iran, ISBN: 978-1-4244-6760-0, pp: 13-17.

Just, M.A., P.A. Carpenter, T.A. Keller, W.F. Eddy and K.R. Thulborn, 1996. Brain activation modulated by sentence comprehension. Sci., 274: 114-116.

Kelly, R.E., G.S. Alexopoulos, Z. Wang, F.M. Gunning and C.F. Murphy *et al*., 2010. Visual inspection of independent components: Defining a procedure for artifact removal from fMRI data. J. Neurosci. Methods, 189: 233-245.

Kwong, K.K., J.W. Belliveau, D.A. Chesler, I.E. Goldberg and R.M. Weisskoff *et al*., 1992. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. Proc. National Acad. Sci., 89: 5675-5679.

Lai, S.H. and M. Fang, 1999. A novel local PCA-based method for detecting activation signals in fMRI. Magn. Reson. Imaging, 17: 827-836.

Lee, C.W., D.Y. Chen, C.W. Wu and J.H. Chen, 2007. Comparing the spatial and temporal reproducibility of brain activation using three fMRI techniques: BOLD, FAIR and VASO. Proceedings of the NFSI-ICFBI 2007 Joint Meeting of the 6th International Symposium on Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging, October 12-14, 2007, IEEE, Hangzhou , China, ISBN: 978-1-4244-0949-5, pp: 258-261.

Lee, S., F. Zelaya, Y. Samarasinghe, S.A. Amiel and M.J. Brammer, 2011. Data-driven fMRI group classification using connected components and Gaussian process classifiers. Proceedings of the 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), May 22-27, 2011, IEEE, Prague, Czech Republic, ISBN: 978-1-4577-0538-0, pp: 717-720.

Li, H., T. Adal, N. Correa, P.A. Rodriguez and V.D. Calhoun, 2010. Flexible complex ICA of FMRI data. Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, March 14-19, 2010, IEEE, Dallas, Texas, ISBN: 978-1-4244-4295-9, pp: 2050-2053.

Maaten, L.V.D., E. Postma and J.V.D. Herik, 2009. Dimensionality reduction: A comparative. J. Mach. Learn. Res., 10: 66-71.

McKeown, M.J. and T.J. Sejnowski, 1998a. Independent component analysis of fMRI data: Examining the assumptions. Hum. Brain Mapp., 6: 368-372.

McKeown, M.J., T.P. Jung, S. Makeig, G. Brown and S.S. Kindermann *et al*., 1998b. Spatially independent activity patterns in functional MRI data during the stroop color-naming task. Proc. National Acad. Sci., 95: 803-810.

McKeown, M.J., V. Varadarajan, S. Huettel and G. McCarthy, 2002. Deterministic and stochastic features of fMRI data: Implications for analysis of event-related experiments. J. Neurosci. Methods, 118: 103-113.

Naik, G.R. and D.K. Kumar, 2011. An overview of independent component analysis and its applications. Inf., 35: 63-81.

Richards, T.L. and V.W. Berninger, 2008. Abnormal fMRI connectivity in children with dyslexia during a phoneme task: Before but not after treatment. J. Neurolinguistics, 21: 294-304.

Rissanen, J., 1978. Modeling by shortest data description. Autom., 14: 465-471.

Schopf, V., C. Windischberger, C.H. Kasess, R. Lanzenberger and E. Moser, 2010. Group ICA of resting-state data: A comparison. Magn. Reson. Mate. Phys. Biol. Med., 23: 317-325.

Spiers, H.J. and E.A. Maguire, 2007. Decoding human brain activity during real-world experiences. Trends Cognit. Sci., 11: 356-365.

Sui, J., T. Adali, G.D. Pearlson and V.D. Calhoun, 2009. An ICA-based method for the identification of optimal FMRI features and components using combined group-discriminative techniques. Neuroimage, 46: 73-86.

Svensen, M., F. Kruggel and H. Benali, 2002. ICA of fMRI group study data. NeuroImage, 16: 551-563.

Thirion, B., P. Pinel, A. Tucholka, A. Roche and P. Ciuciu *et al*., 2007. Structural analysis of FMRI data revisited: Improving the sensitivity and reliability of FMRI group studies. IEEE. Trans. Med. Imaging, 26: 1256-1269.

Tsatsishvili, V., F. Cong, T. Puolivali, V. Alluri and P. Toiviainen *et al*., 2013. Dimension reduction for individual ICA to decompose FMRI during real-world experiences: Principal component analysis vs. canonical correlation analysis. Proceedings of the European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, April 24-26, 2013, Jyvaskyla University, Bruges, Belgium, ISBN: 978-2-87419-081-0, pp: 137-142.

Turner, G.H. and D.B. Twieg, 2005. Study of temporal stationarity and spatial consistency of FMRI noise using independent component analysis. IEEE. Trans. Med. Imaging, 24: 712-718.

Weissenbacher, A., C. Kasess, F. Gerstl, R. Lanzenberger and E. Moser *et al*., 2009. Correlations and anticorrelations in resting-state functional connectivity MRI: A quantitative comparison of preprocessing strategies. Neuroimage, 47: 1408-1416.

Woods, R.P., 1996. Modeling for intergroup comparisons of imaging data. Neuroimage, 4: S84-S94.

Zhang, J., X. Tuo, Z. Yuan, W. Liao and H. Chen, 2011. Analysis of fMRI data using an integrated principal component analysis and supervised affinity propagation clustering approach. IEEE. Trans. Biomed. Eng., 58: 3184-3196.