A NEW APPROACH FOR FEATURE EXTRACTION FROM FUNCTIONAL MR IMAGES

Güzin ÖZMEN[1,*], Seral ÖZŞEN[2, +]

[1] Selcuk University, Faculty of Technology, Department of Biomedical Engineering

Konya TURKEY

[2]Konya Technical University, Faculty of Engineering and Natural Sciences

Konya TURKEY

*gozmen@selcuk.edu.tr

Abstract

The functional MR images consist of very high dimensional data containing thousands of voxels, even for a single subject. Data reduction methods are inevitable for the classification of these three-dimensional images. In this study in the first step of the data reduction, the first level statistical analysis was applied to fMRI data and brain maps of each subject were obtained for the feature extraction. In the second step the feature selection was applied to brain maps. According to the feature selection method used in the classification studies of fMRI and which is called as the active method, the intensity values of all brain voxels are ranked from high to low and some of these features are presented to the classifier. However, the location information of the voxels is lost with this method. In this study, a new feature extraction method was presented for use in the classification of fMRI. According to this method, active voxels can be used as features by considering brain maps obtained in three dimensions as slice based. Since the functional MR images have big data sets, the selected features were once again reduced by Principal Component Analysis and the voxel intensity values were presented to the classifiers. As a result; 83.9% classification accuracy was obtained by using kNN classifier with purposed slice-based feature extraction method and it was seen that the slice-based feature extraction method increased the classification.

Keywords: Classification, Feature extraction, fMRI, SPM

+ This paper has been presented at the ICENTE'18 (International Conference on Engineering Technologies) held in Konya (Turkey), October 26-28, 2018.

1. Introduciton

fMRI, a special application of magnetic resonance imaging (MRI), is a noninvasive method. With this method, researchers record three-dimensional brain images while a subject performs a cognitive or sensory task in the MR device [1]. fMRI data analysis consists of image acquisition, pre-processing and statistical analysis. Due to its noisy structure, some preprocessing operations should be applied to fMR images. These processes are realignment, slice-timing, co-registration, segmentation and spatial normalization [2]. The most common filtering method used after standard pre-processing steps applied to fMRI images is known as spatial smoothing. With the spatial smoothing, the intensity value of each voxel is replaced with the mean values of neighboring voxels. This process corresponds to the use of a low-pass filter to suppress high-frequency signals in the frequency domain. Various filters are used to perform this operation. The most common spatial filtering method is the convolution of the image with the Gaussian filter [3].

Statistical analysis of pre-processed fMRI data can be used to detect activations in brain regions [4]. In the statistical analysis of fMRI data, the voxel intensity values can be determined in the brain regions by comparing the changes in the control state with the changes in the task state. The most commonly used method to evaluate the statistical significance of neural correlations in the brain is known as Statistical Parametric Mapping (SPM). SPM is a voxel-based and univariate approach based on the General Linear Model (GLM) [5]. Activation maps are the result of statistical analysis applied to functional MR images are evaluated as characteristics that characterize the brain functions of each subject. SPM [6] based analysis produces voxel values under the null hypothesis distributed according to a known probability density function. This is accomplished by obtaining the spm.T maps expressing the activation maps.

Functional MR data is composed of very high dimensional images containing thousands of voxels. These three-dimensional images need to be reduced for use in classification applications. The activation maps obtained as a result of analysis of fMRI data are one of the most important steps to reduce the size of fMRI. Because of the fact that fMRI consists of three-dimensional images, even the activation maps have quite a number of voxels for the classifier inputs. For this reason, the number of features should be reduced by applying feature selection processing to the reduced data. Many of the feature selection methods can be applied to spm.T maps obtained from fMRI data.

However, [7] emphasized that some special feature selection methods would yield better results. In this context, voxel-based and transformation-based feature selection methods are possible to use. In transformation-based feature extraction, data is transferred from the original high-dimensional coordinate system to a new low-dimensional coordinate system. Principal Component Analysis (PCA) is one of the transformation- based feature extraction methods. In PCA applications, vectors are found as eigenvectors of the estimated covariance matrix of the data. Thus, the original fMRI data matrix is converted into rows in a transformation matrix with these vectors.

Most often, the dimensionality of the data is reduced by eliminating the basic components with the lowest variance. PCA is used to create a new feature space from the high-dimensional data. Each brain voxel is expressed in the new feature space with a vector [8]. Voxel-based feature extraction methods keep the data in the original coordinate system. However, all the voxels are sorted and used only as high-density valuable voxels as a result of feature extraction.

"The most active voxel method" is the most commonly used method in the voxel-based feature extraction According to this method, voxels in the activation map are ranked from high to low according to their density values and the highest n voxels are selected as features. The disadvantage of this method is that it lost location informations of voxels. Another feature selection method which [7] suggested is "the most distinctive voxels" method. In this method, the accuracy of the selected voxels on the training data of the classifier is taken as a measure of the distinctive power of the voxel and the highest rated voxels according to this measurement are selected. Another feature selection method is "The most effective voxels according to their interest" method. This method is similar to the active method but allows a homogenous selection of voxels from all regions of interest (ROI) within the brain. This method applies the active method to each ROI and selects the most active voxels from each.

In this study, an alternative method against to active voxel method is proposed by using all brain voxels. With the proposed method which is called slice-based feature extraction, the density values of the active voxels in each slice of the three-dimensional image are evaluated according to sagital slice. Thus, all voxels are used as features without losing the position information.

Support Vector Machines (SVM) [9, 10], k nearest neighbor (kNN) [11], Naive Bayes (NB) [12] and Random Forest (RF) [13] classifiers were used with Orange Canvas 3.4 to classify fMRI data. Classifier performances were evaluated by classifying accuracy (CA), Precision, Recall, F score, and area under ROC curves (AUC). In order to measure the performance of the classification process, training and test data were prepared by k-fold cross-validation method. When the classifier results were compared to the active voxel method and the slice-based method, it was seen that the slice-based method improves the classification results.

2. Material and Method

In this study, fMR images, recorded from 20 healthy and 20 depression patients, were used. fMR images were obtained with positive and negative visual stimulus by block design recording procedure [14] . The processing steps applied to the recorded fMRI data are shown in Figure 1. In this study, re-alignment, slice-timing, co-registration, and spatial normalization operations were applied to the fMRI data with the SPM12. With re-alignment, the image's volumes are re-aligned to a single reference volume to correct errors due to head movements. Slice timing used to correct the timing difference between slices. The fMRI data does not provide a very clear image from the anatomical point of view. For this reason, the functional images were co-registered with the structural images. After co-registration, high-resolution T1 images were normalized to a T1-weighted standard brain template by spatial normalization. With the last step which is called smoothing, the distorting effect on the image was eliminated.

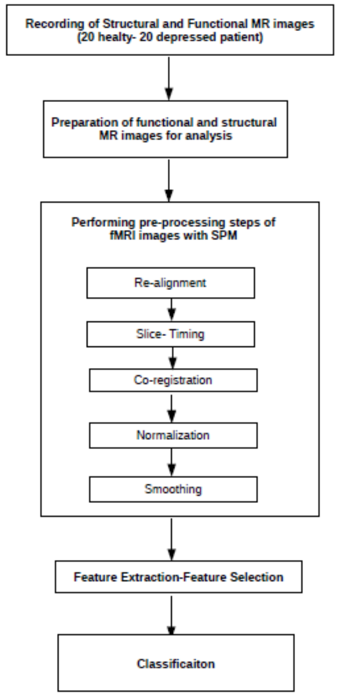

Figure 1. The workflow of the study

Examples of the preprocessing steps performed in this study are shown in Figure 2 for a subject selected from the control group.

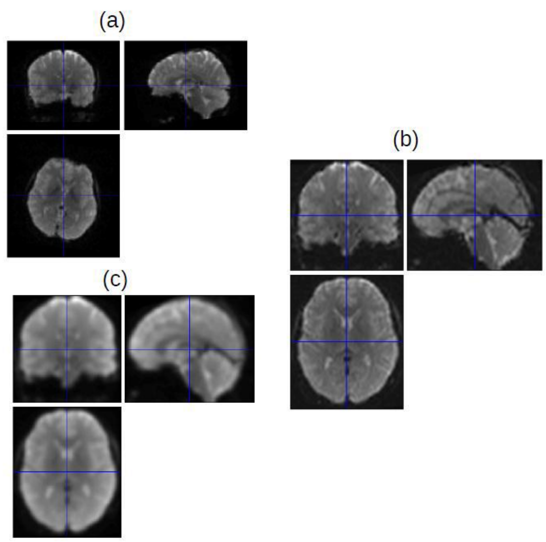

Figure 2. Preprocessing steps applied to a functional MR image of a subject selected from the control group a) Realignment b) Normalization c) Smoothing

In the next steps of this study are feature extraction, feature selection and classification of pre-processed fMRI data. The first operation of feature extraction is statistical analysis. In this study, fMRI data obtained from one subject contained 79x95x78 voxels. The total number of data consists of 120 fMRI volumes recorded for 4 seconds each. There are a total of 79x95x78x120 voxels in 120 volumes. In such conditions, it is impossible to use such high-size data that accommodate all brain voxels in classifier inputs. For this purpose, in order to reduce the size of fMRI data and also to obtain the density values of the voxels used as determinants of the classification problems, first level statistical analysis was applied to fMR images. As a result of this process, spm.T maps were created for all subjets. In this study, the first stage of feature extraction was completed for each subject by using spm.T maps. So that fMRI data with 79x95x78x120 dimension was reduced to 79x95x78 by using the spm.T map. This process was carried out for all 40 subjects. Figure 3 shows a spm.T maps of a healthy and a patient subject which are randomly selected.

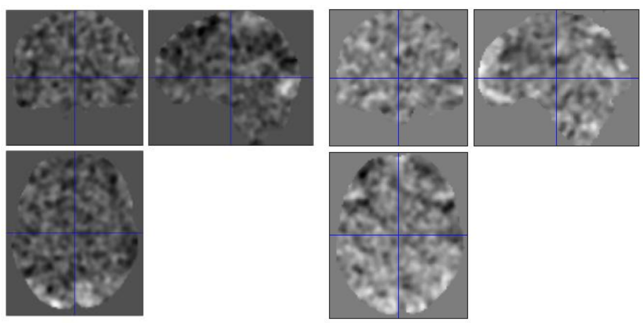

Figure 3 (a) Spm.T map of a subject from the control group (b) Spm.T map of a subject from the patient group

Figure 4 shows the colored map of the spm.T map of the control group given in Figure 3.

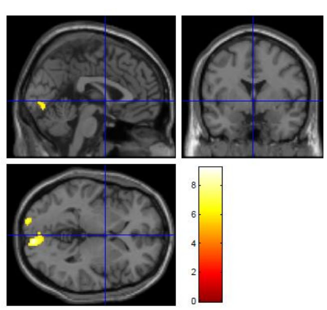

Figure 4. Color image of the activation map of a subject belonging to the control group

3. Findings and Discussion

3.1. Evaluation of classifier results by active voxel feature selection method

In this study, a total of 585390 voxels were present in the activation map (79x95x78) of a subject. Primarily, all of these voxels were presented to the classifiers as features but the result was not obtained. Secondly, according to the active voxel method, voxels were ranked from the highest to the lowest value and the first 1000 voxels features were chosen. Thus, a total of 1000 features for one subject and 1000x40 for 40 subjects were presented to the classifiers. The classification results are shown in Table 1.

Table 1. Classification results for the active voxel method

	AUC	**CA**	**F1**	**Precision**	**Recall**
SVM	57,50	57,50	66,70	60,80	57,50
NB	57,50	55,00	57,10	55,10	55,00
kNN	57,50	62,50	66,70	63,30	62,50
RF	52,50	47,50	48,80	47,50	47,50

AUC given in Table 1 is expressed the area under the ROC curve. CA is the classification accuracy. Precision is the positive prediction success and Recall shows how successful the positive samples are estimated. F1; indicates the harmonic mean of precision and recall values. In order to measure the performance of the classification process, training and test data were prepared by k-fold cross-validation method. The highest classification accuracy for Table 1 was obtained by kNN with 62.5% and the lowest classification accuracy was obtained by RF with 47.5%.

3.2. Evaluation of classifier results by slice-based feature selection method

In this study, a new feature selection method that can be used for functional MR images is suggested as an alternative to the active voxel method. With this method, the density values of the active voxels in each slice are evaluated according to the sagittal section. This method allows data transformation without losing voxels location information. According to the slice-based feature selection method, (79x95x78 → 7505x78) voxels in 78 different sagital slices were obtained for each subject. An important property of the slice-based method is that it increases the number of samples used for one person by using 78 slices for a subject. This method can be used to increase the number of samples in studies where the number of subjects is insufficient. In the method of slice-based feature selection, PCA was used to reduce the number of features after data transformation was completed. In this study, 20 principal component features were used. Thus, 20x78 features for one subject and 20x3120 for 40 subjects were presented to the classifiers. The classification results are shown in Table 2.

Table 2. Classification results for the slice-based method

	AUC	CA	F1	Precision	Recall
SVM	67,9	67,9	62,1	69,8	67,9
NB	75,7	67,8	67,0	67,9	67,8
kNN	92,7	83,9	83,9	83,9	83,9
RF	87,2	78,5	78,2	78,5	78,5

The highest classification accuracy for Table 2 was obtained by kNN with 83.9% and the lowest classification accuracy was obtained by RF with 78.5%. According to the feature selection method used in the classification studies of fMRI and which is called as the active method, the intensity values of all brain voxels are ranked from high to low and some of these features are presented to the classifier. However, the location information of the voxels are lost with this method. Especially in the classification of psychiatric diseases, it is known that activations occur in different brain regions of healthy persons and patients due to mood changes. In this study, a new feature selection method is presented for use in classification studies using all brain voxels. According to this method, active voxels can be used as features by considering brain maps obtained in three dimensions as slice based. Since the functional MR images have a very large size of data, the selected features were once again reduced by Principal Component Analysis and the obtained voxel intensity values were presented to the classifier. As a result, it was seen that the slice-based feature extraction method increased the classification accuracy against the active method.

References

[1] Billor N, Godwin J. Variable Selection for Functional Logistic Regression in fMRI Data Analysis.Turkiye Klinikleri Journal of Biostatistics 2015; vol. 7, no. 1

[2] Chen JE, Glover GH. Functional magnetic resonance imaging methods.Neuropsychology review 2015; vol. 25, no. 3, pp. 289-313.

[3] Friston K, Ashburner J, Frith CD, Poline JB, Heather JD, Frackowiak RS. Spatial registration and normalization of images. Human brain mapping 1995; vol. 3, no. 3, pp. 165-189.

[4] Francis S, Panchuelo RS. Physiological measurements using ultra-high field fMRI: a review. Physiological measurement 2014; vol. 35, no. 9, p. R167.

[5] Myers RH, Montgomery DC. A tutorial on generalized linear models. Journal of Quality Technology 1997; vol. 29, no. 3, p. 274.

[6] Friston KJ, Holmes AP,Worsley KJ, Poline JP, Frith CD, Frackowiak ,RS. Statistical parametric maps in functional imaging: a general linear approach. Human brain mapping 1994; vol. 2, no. 4, pp. 189-210.

[7] Mitchell TM. Learning to decode cognitive states from brain images. Machine learning 2004; vol. 57, no. 1-2, pp. 145-175.

[8] Suma H, Murali S. Principal Component Analysis for Analysis and Classification of fMRI Activation Maps. International journal of computer science and network security 2007; vol. 7, no. 11, pp. 235-242.

[9] Mourão-Miranda J. Patient classification as an outlier detection problem: an application of the one-class support vector machine. Neuroimage 2011; vol. 58, no. 3, pp. 793-804.

[10] Sacchet MD, Prasad G, Foland-Ross LC, Thompson PM, Gotlib IH. Support vector machine classification of major depressive disorder using diffusion-weighted neuroimaging and graph theory. Frontiers in psychiatry 2015, vol. 6, p. 21.

[11] Misaki M, Kim Y, Bandettini PA, Kriegeskorte N. Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. Neuroimage2010; vol. 53, no. 1, pp. 103-118.

[12] Kuncheva LI and Rodríguez JJ. Classifier ensembles for fMRI data analysis: an experiment. Magnetic resonance imaging 2010; vol. 28, no. 4, pp. 583-593.

[13] Tripoliti EE, Fotiadis DI, Argyropoulou M, Manis G. A six stage approach for the diagnosis of the Alzheimer's disease based on fMRI data. Journal of biomedical informatics 2010; vol. 43, no. 2, pp. 307-320.

[14] ÖZMEN G, Fonksiyonel Mr Görüntülerini Filtrelemede Yeni Bir Yaklaşim Ve Depresyon Hastalarinin Siniflandirilmasi Üzerine Etkileri, PhD. thesis,, Selcuk University, Graduate School of Natural and Applied Sciences 2018.