Diagnosis Of Autism Using Voxel Based Morphometry Features Of Brain Images

Praveen K, Rajeswari R

Abstract— The advancements in neuroimaging have resulted in huge amount of data and hence automatic tools are required to identify patterns, reduce noise and enhance the knowledge about the functioning of brain. This paper also proposes a new approach for detecting major differences in brain structures and activities between autistic and healthy individuals using Autism Spectrum Disorder (ASD) from structural Magnetic Resonance Imaging (sMRI) and functional MRI (fMRI) magnetic Resonance Imaging MRI. The preprocessing of sMRI and IMRI is carried out using Statistical Parametric Mapping (SPM) toolbox. The voxel based morphometry analysis results obtained from preprocessing are used as features which help in differentiating the autistic and healthy individuals. These features are given as input to Support Vector Machine or Random Forests (RF) to classify the autistic and healthy individuals. The sMRI and fMRI images used in this research work are obtained from the Autism Brain Imaging Data Exchange (ABIDE) database. The experimental results prove that the proposed computer aided diagnosis (CAD) system gives good performance in terms of classification accuracy.


1 INTRODUCTION

Autism spectrum disorder (ASD) or autism is a group of developmental brain disorders, which have an impact on the behaviour and communication capabilities of a person (Beheshti et al., 2005; Sungji et al., 2015). Some children are mildly affected by their symptoms, but others are severely disabled. Autism is a multiplex condition that impacts brain development and affects a person’s social interaction, communication issues, interests, and behaviour. Autism Speaks, one of the leading organizations the USA which supports autism and other autism related systems are working hard to educate parents and physicians so that children with autism are recognized as early as possible (Subburajuja et al., 2015; Tehila et al., 2016).

Most of the ASD children have trouble engaging with everyday social interactions. It is one of the most common symptoms in ASD according to Diagnostic and Mental Disorders (Ashburner et al.,2000). Some of the symptoms in children with ASD are (1) their eye contact is little; (2) they listen to people in their environment less and fail to respond; (3) they do not share their enjoyment or activities with other people; (4) some of them answer unusually and others show violence, suffering or affection; (5) they may have trouble understanding a different person’s point of view; and (6) they are playing games in alone and did not share anything with others.

The investigation and analysis of autism in neuroimages of infants will aid in diagnosis, prognosis and decision-making during treatment. Imaging features may also aid to predict the common brain neuropathology that the autistic individual has and how autistic individuals are responding to the treatment strategies. Quantitative feature extraction from the neuroimages may also provide as a metric for the biological efficacy of potential behavioural or pharmacologic interventions. This paper intends to use voxel based morphometry features from brain magnetic resonance images machine learning to efficiently diagnose autism and help in the differentiation of autistic individuals from healthy individuals. The rest of this paper is organized as follows. Section 2 gives an overview of existing work carried out to diagnose autism from brain images. Section 3 describes the proposed method to classify brain images of autistic and healthy individuals using voxel based morphometry features and machine learning techniques viz., support vector machine and random forests classifiers. Section 4 presents the experimental results obtained. Section 5 gives the conclusion of this paper.

2 RELATED WORK

The voxel based morphometry (VBM) study was conducted to investigate the differences in the volumes of grey matter (GM) and white matter (WM) between the autism and the control group (Calderoni S et al., 2012). The T1-weighted volumetric images were analysed using the VBM protocol with modulation. They had implemented the Diffeomorphic Anatomical Registration using the Exponentiated Lie algebra (DARTEL) algorithm, where a diffeomorphic warping is implemented to achieve an accurate inter subject registration with an improved realignment of small inner structures and to
generate a study-specific template. The analysis of images has been performed using the Statistical Parametric Mapping (SPM) software package. SPM is very helpful in preprocessing the brain images. The output of VBM analysis is the maximum intensity projection (MIP), which is a map that highlights the voxels that show a significant difference in the composition of the brain (Wenjing et al., 2013). These VBM detected brain regions are used as masks for feature extraction. VBM has been used to assess entire-brain structure with voxel-by-voxel comparisons, which has been developed to analyse issue concentrations or volumes between subject groups to distinguish degenerative diseases such as dementia (Beheshti et al., 2015).

A semi-automatic method to segment voxels/regions of interest (ROI) based on Euclidean distance of these voxels from seed pixels has been presented (Matsuda H et al., 2012). The dataset considered in this work included images from autism brain imaging data exchange (ABIDE) and other multiple sites. Based on data from the 7266 generated ROIs, a connectivity matrix giving the correlation between each ROI was calculated. They have used generalized linear model (GLM) to fit the connectivity matrix with the VBM result of autistic and healthy individuals. A number of methods are available in the literature which extract information from brain images in order to classify the images of autistic and healthy individuals or to predict the presence/progression of ASD in the patients.

3 PROPOSED METHOD:

The proposed system consists of three main stages: (i) preprocessing; (ii) feature extraction; and (iii) classification. Figure 1 illustrates the proposed CAD system for diagnosis of ASD from brain images. Preprocessing is the term used for all the steps taken to improve the data and prepare it for statistical analysis. Three levels of pre-processing have been performed in this work which are (1) segmentation (2) alignment using Dartel’s method (3) normalization and (4) smoothing. In this work, segmentation is done with statistical parametric mapping’s (SPM’s) segmentation (Ashburner and Friston, 2000). All the sMRI and fMRI images which are in the NIH (.nii) are segmented into different regions including gray matter, white matter and CSF regions. The idea behind Dartel is to increase the accuracy of inter-subject alignment by modeling each shape brain using so many parameters particularly three parameters for each voxel (Ashburner and Friston, 2000). Dartel works on grey matter among the images, while concurrently aligning white matter. This is achieved by generating an average template data, to which the data are iteratively aligned. This step generates “u_rcl” files, as well as a series of template images. After spatial alignment using Dartel six different templates are obtained. Normalization uses the “u_rcl” files (which encode the shapes), to generate smooth, spatially normalized and scaled grey matter images in Montreal Neurological Institute (MNI) space. Every person’s brain is slightly different from every other’s. Brains differ in size and shape. Before comparing the brain image of one person with the brain image of another person, both the brain images have to be converted to uniform shape and size and this process is called normalization. Structural and functional MRI images are smoothed by using filters. Smoothing increases noise ratio of the data by filtering the highest frequency. That is, removing the smallest scale changes among the voxels. Smoothing will help to make the larger-scale changes more apparent. The smoothing helps to reduce spatial differences between the subjects and aids in comparing the multiple subjects. The smoothing can cause regions that are functionally different to combine. Once preprocessing is complete, general linear model is used to fit through the processed data and inferences are made about where there are systematic differences among the processed data. In particular, the model which helps in identifying the differences among the processed data is used. The first level of VBM can be performed through SPM. Initially, the process considers the files generated by preprocessing step to generate a ‘.mat’ file which is then considered for VBM analysis. The results on the top left show the highlighted regions in brain images.

Figure 1: Block diagram of proposed system using Support Vector Machine/ Random Forest for diagnosis of ASD from brain images.
which differ among all the brain images considered for the analysis. The darker the region, the more the pixels in that region differs.

In the present work, either support vector machine/ random forests are used for classifying the obtained features from voxel based morphometry analysis. Support Vector Machines (SVMs) is one of the machine learning methods which is widely used for classification or regression (Beheshti et al., 2015). In general SVM tries to generate a hyperplane which discriminates the samples of different categories. In case the number of dimensions of the data is two-dimension, the SVM generates a line to separate the data. One of the main parameters responsible for generating the line/ hyperplane by SVM is kernel. Linear kernels help in generating the line whereas polynomial and exponential kernels help in generating the hyperplane. Random forests belong to the category of ensemble methods which make use of more than one method to classify data. Generally, random forests are based on more than one decision tree. Random forests are based on random sampling and subsets (David S 2005 ). Initially random forest generates a random sample from the data. The samples are generated are given as input to the available decision trees. Every decision tree classifies and gives a different result. Every result is given a weight and the result which gets the highest weight is selected as the classification result (Ecker C et al., 2010).

### 4 EXPERIMENTAL RESULTS:

Autism spectrum disorder images are taken from Autism Brain Imaging Data Exchange (ABIDE) dataset (Ashburner et al., 2000; Ahmadloua M et al., 2012). ABIDE dataset includes two large-scale collections, one is structural MRI and the other is functional MRI images. The ABIDE dataset is a collection from 24 international brain imaging laboratories. In the present work, 79 different structural magnetic resonance imaging (sMRI) and functional magnetic resonance imaging (fMRI) images of autistic and healthy individuals of children in the age group of one to nine years are considered (Jeffrey D et al., 2012). All the brain images are pre-processed using statistical parametric mapping 12 (SPM12) in the MATLAB R2018a environment. Diagnosis of autism from brain images are useful in the early detection and treatment planning in children suffering with this disease (Boddaert N et al., 2004).

<table>
<thead>
<tr>
<th>Images</th>
<th>Cross Validation Method</th>
<th>SVM Accuracy %</th>
<th>Random Forest Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>sMRI</td>
<td>LeaveMOut</td>
<td>89.00</td>
<td>88.20</td>
</tr>
<tr>
<td></td>
<td>HoldOut</td>
<td>90.00</td>
<td>86.70</td>
</tr>
<tr>
<td>fMRI</td>
<td>LeaveMOut</td>
<td>89.30</td>
<td>82.90</td>
</tr>
<tr>
<td></td>
<td>HoldOut</td>
<td>87.50</td>
<td>88.80</td>
</tr>
</tbody>
</table>

Table 1 Accuracy values for Support Vector Machine and Random forest based classification

Two different cross validation methods viz., LeaveMOut and HoldOut are used evaluate the classification of images of autistic and normal people using SVM and RF. In HoldOut method, 80% of total images i.e. 80% of 79 3D `.nii` images are used for training and the remaining images are used for testing. In the case of LeaveMOut, 60 3D `.nii` images are used for training and the remaining 19 images are used for testing. The results obtained for these cross validation methods using support vector machine and random forest for sMRI and fMRI images are presented in table 1.

![Classification](image)

Figure 1 Bar graph of accuracy values obtained using SVM/RF based classification

The visualization of accuracy values obtained for SVM/RF based classification of images of autistic and normal people is given as a bar graph in figure 1. The graph also presents the accuracy values obtained for LeaveMOut and HoldOut cross validation methods. It can be observed that the accuracy values obtained for sMRI and fMRI images using SVM and RF are almost same and above 85%. The accuracy values obtained for fMRI images are approximately 88% and are slightly higher compared to sMRI images.

### 5 CONCLUSION

Autism spectrum disorder is still in its infant stage and it requires a lot of theoretical and experimental studies to develop a completely automated system for the diagnosis of autism. This paper propose and implements a method to diagnose autism using differences in brain images of autistic and healthy individuals. In the first step preprocessing of brain images is performed using SPM. This stage results in voxel based morphological analysis of brain images of autistic and healthy individuals. The result is actually maximum intensity projection which helps in identifying the significant difference in gray matter between autistic and healthy individuals. The stage is the classification of this maximum
intensity projection using support vector machine and Random Forests/ Tree Bagger. The results of the proposed method are evaluated using sMRI and fMRI images obtained from ABIDE database. The experiments are conducted based on cross validation. In this research work, Hold Out and Leave-M-Out methods of cross validation are used. The experimental results of the proposed method for classifying autistic and healthy individuals are promising in terms of accuracy. In the present work in the classification techniques Support Vector Machine have based classification obtained as 88 percentage of accuracy Random forest have classified the above 85 percentage of accuracy. In future these methods can be replaced by even more effective or hybrid methods so that the classification accuracy can be improved.

ACKNOWLEDGMENT
The authors are thankful to bharathiar university for the support provided.

REFERENCES