Classification of Autism based on Feature Extraction from Segmented Brain MRI

B.J. Bipin Nair, Gopi Krishna Ashok, N.R. Sreekumar

Abstract--- The Autism Spectrum Disorder is a neurological irregularity with multiple behavioral symptoms. It includes Asperger syndrome and pervasive developmental disorders. It is called as "spectrum" disorder because an individual with ASD might have a wide range of symptoms. People with ASD will have communication trouble, low eye contact, limited attentiveness and tedious behaviors. Various researches on structural MRI have mostly concentrated on the detection of autism in people with ASD. This study's aim is to classify the type of Autism Spectrum Disorder for various body movements. Here we use supervised classification algorithms like ID3.For our study, we are considering the datasets which consists of 50 normal and 50 autistic brain MRI. Here, we are mainly focusing on effective classification of ASD using a classifier with a class label.

Keywords--- ASD (Autism Spectrum Disorder), SVM- (Support Vector Machine), ELM-(Extreme Learning Machine), H-ELM-(Hierarchical Extreme Learning Machine), GPC-(Gaussian Process Classification), GM-(Grey Matter).

I. INTRODUCTION

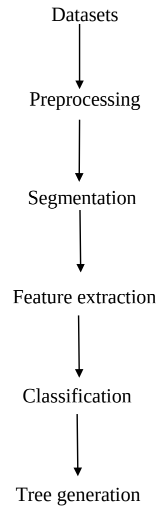
Digital Image Processing is defined as the manipulation of computer algorithms to conduct processing of digital image to get a magnified image.

Medical imaging is a method of generating an image depiction of internal fragment of the body for analytical introspection and medical intercession. Machine learning is a subdivision of Artificial Intelligence in the area of computer science that generally utilizes statistical method to provide computers the capability to "learn".

The data is trained without any specific use of coding in order to predict or make decisions according to the input that is given.

In the existing system, it is tough to detect the group of ASD Classification and hold many disadvantages. As some of the algorithms were not effective, computational forecasting was more compared to clinical prediction and solution of the ASD were minute as there was no standard result for the ASD victims. Hence, we have proposed a system with a productive classification approach in order to solve this issue.

Flow of Work



II. LITERATURE SURVEY

Anibal Sólon Heinsfeld et al[1] says that they applied categorization techniques of ASD with the help of ABIDE datasets and in the subsequent works they have implemented deep learning method which includes both supervised and unsupervised machine learning (ML) techniques. In the feature selection phase, functional association was used for segregation of studies like ASD and TC. Here they have trained two stacked de-noising auto-encoders for the unsupervised pre-training step in order to withdraw a lower-dimensional type from ABIDE data. They discovered 70% of precision for the recognition of ASD. Bipin Nair et al[2] says that they are predicting the presence of proteins inside the cell. Here they had used ANN and calculation of the threshold energy of RNA. The author has used back propagation technique in which they have reduced the error by altering the weights of input and achieved an overall accuracy of 95%. N Shobha Rani et al[3] says that they have used character recognition to achieve the classification of telugu handwritten text. They have used techniques such as zone-based feature extraction and classification using nearest neighbor, naïve Bayesian, SVM algorithm and achieved an over all accuracy of 84.4%.Guillaume Chanel et al[4] says that Multivariate Pattern Analysis (MVPA) was deployed and the grouping depends on emotion, anger, face recognition and body perception. Support Vector Machine (SVM) and Recursive Feature Elimination (RFE) are two algorithms that are made use over here.

Manuscript received February 01, 2019

B.J. Bipin Nair, Department of Computer Science, Amrita School of Arts and Sciences, Mysuru, Karnatka, Amrita Vishwa Vidyapeetham, India. (e-mail: bipin.bj.nair@gmail.com)

Gopi Krishna Ashok, Department of Computer Science, Amrita School of Arts and Sciences, Mysuru, Karnatka, Amrita Vishwa Vidyapeetham, India. (e-mail: gopikrishna71000@gmail.com)

N.R. Sreekumar, Department of Computer Science, Amrita School of Arts and Sciences, Mysuru, Karnatka, Amrita Vishwa Vidyapeetham, India. (e-mail: sreekumarnedumpallymana@gmail.com)

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

The Region of Interest (ROI) is a method that's used to choose a non-identical affected area of the anatomy. Discriminative scheme of arrangement was detected in several areas of the social brain. The method has grouped the autism affected patients from controls with an efficiency ranging from 69% to 92.3%.Muhammad Naveen Iqbal Qureshi et al[5] says that they have classified the data with the help of Hierarchical Extreme Learning Machine (HELM) algorithm and later compared the algorithm with SVM to check which is more efficient for the classification of ADHD. Here, they have taken the datasets of 159 MRI images by the help of scanner 3.0-Tesla and preprocessed the image with the help of Free Surfer 5.3.0 and finally they subjected it with the AFNI program for smoothing the images. They acquired a result in which HELM was more efficient than SVM for classification of ADHD. In the future for the classification of ADHD, H-ELM algorithm can be mostly used rather than SVM. Lena Lim et al[6] says that they have taken advantage of applying Gaussian process classification (GPC) to the grey matter (GM), to ingress whether the ADHD affected individuals could be separated from healthy controls. The data file consist of 29 ADHD boys and 10 ASD boys. GPC and SVMwas applied for the classification in order to predict based on structural pattern of brain images. As a result, overall precision of 79.3% has been detected for the classification of the ADHD patients. The usage of SVM algorithm has made the classification of ASD and the characterization of pattern of brain images much easier. With the help of MRI measures this can be used for the diagnosis of classification of ADHD in the future. Here, the limitation of the existing system was that the diagnosis and detection of ADHD was very difficult. Xiaolong Peng et al[7] says that they compared the productiveness across ELM and SVM algorithm. Here the drawback is the utilization of structural MRI data rather than usage of large-scale brain sub-network dysfunction. For finding classification of ADHD large scale brain sub network is used. They have calculated multiple brain parameters (cortical thickness, etc.) With Free Surfer and gained the result of ADHD prediction with effectiveness of 90.18% in ELM technique and efficiency for SVM algorithm of 84.73%. They got a result which says that ELM function is more accurate than SVM function and also ELM can be applied to detect brain disease.Cheng W et al[8] says that they have utilized a systematic approach to segregate between affected people and controls. Here, they direct the issue of classifying the brain state precisely going by individually for an immense data set. They have applied SVM for the grouping purpose and selection feature from each Pearson and partial correlation of the functional association to accomplish excellence in taxonomy. They obtained correctness of 76.15% for the categorization of ADHD with the help of advanced pattern recognition technique.Sara Calderon et al[9] says that the neuroanatomical phenotype of female children with ASDpoint out an ignored area of exploration. Every whole brain volumes were held and differentiated using Voxel Based Morphometry (VBM). The SVM-RFE technique was used to find the most separating voxels in GM parts. Here, they have taken the data file of 38 victims to compare the non-verbal IQ and controls. As a result VBM conceals arise of gray matter in the left modest frontal gyrus in ASD. Rong Chen et al[10] says that the Magnetic-resonance (MR) inspection produces a strong instrument for the detection of cerebrum systematic variations in adolescent with ASD. Here they have applied procedures like ROI based morphometry, voxel-based morphometry (VBM) etc. for the investigation and grouping of the ASD. The ROI-based volumetry technique tells that the young kids with ASD having variations improved total brain volume. In most of the SVM, works of ASD described the improved cortical opacity in the cerebral cortex and the DTI studies of ASD have systematically reported the corpus-callosum irregularity across a vast area. Christine Eckerd et al[11] says that the ASD is led by segregated imbalance in brain structure that are tough to find with conventional mass-univariate methods so it's not able to explore the kind of ASD. Here they are guessing the ASD patients with the help of SVM algorithm and by using whole-brain classification approach to know the structural pattern of the whole brain MRI. The Comparison was done with SVM and VBM to find efficiency in prediction and grouping of ASD. They have taken the datasets of 44 male adults as datasets in which 22 were analyzed with ASD and 22 were healthy controls. They got the result that SVM is more efficient for the prediction and classification of ASD rather than VBM. Christine Eckerd et al[12] says that multi-parameter distribution method is used here to differentiate the complex structure of grey matter anatomy. Different multiple parameters have been acquired in order to differentiate the ASD patients. They have used SVM algorithm for the classification. The outcome points that autism is truly multifaceted, also it affects multiple areas. The patterns which are found with the help of SVM may assist for the future survey of the inherited and neuropath logical under-pinning of ASD.

III. METHODOLOGY

First we collected 500 datasets of autistic patients then the brain MRI is preprocessed using Histogram Equalization technique where the noises are removed from MRI images. Later the preprocessed images are segmented using Fuzzy C Means technique in order to find autism affected regions. Subsequently, the datasets features are extracted using multiple-parameters to classify ASD using ID3.

Step1: Preprocessing using histogram equalization technique.

Step 2: Segmentation using Fuzzy C Means method.

Step 3: Extracting various features from thesegmented image.

Step 4: Classificationusing ID3 with the help of class label.

Step 5: Tree generation

IV. DATASETS

In our proposed system we have made use of 500 brain MRI of autistic patients which we have collected from ABIDE.

V. MATHEMATICAL MODEL

Preprocessing

Preprocessing is the technique of removing noises from the MRI images.

- **Histogram equalization**

Histogram equalization is a preprocessing technique used to increase the intensity of images.

K is the range

gray scale image {x}

conversion of form y=T(x) to generate a new image{y} CDF is Cumulative Distribution Function.

$$cdf_{y(\acute{y})} = cdf_y(T(k)) = cdf_x(k)$$

$$\acute{y} = y.(max\{x\} - min\{x\}) + min\{x\}$$

Segmentation

Segmentation is the extracting of region of interest from autistic MRI.

- **Fuzzy-c means**

Fuzzy-c means clustering technique which is used to segment a specific region from autistic MRI.

Where $X=\{x_1,\ldots,x_n\}$

Cluster centers $d=\{d_1,\ldots,d_d\}$

$F=F_{ij}\in [0,1]$, $i=1,\ldots,n$, $j=1,\ldots d$

$$\underset{d}{\arg\min} \sum_{i=1}^{n} \sum_{j=1}^{d} F_{ij}^{p} \|x_i - d_j\|^2,$$

$$F_{ij} = \frac{1}{\sum_{k=1}^{d} \left(\frac{\|x_i-d_j\|}{\|x_i-d_k\|}\right)^{\frac{2}{m-1}}}$$

Feature extraction

Extracting 11 feature from the segmented MRI images.

E = entropy(I);
M=mean2(I);
Contrast = mean([stats.Contrast]);
Homogeneity = mean([stats.Homogeneity]);
Correlation = mean([stats.Correlation]);
Energy = mean([stats.Energy]);
Area=bwarea(I);
Standard deviation=std2(I);
Variance = var(double(I(:)));
Skewness=skewness(I2(:));
kurtosis =kurtosis(I2(:));

Classification

Classification is a technique used to classify the trained dataset using a class label.

- **Decision tree induction algorithm**

Decision tree induction algorithm is used to classify based on information gained and entropy value.

Step 1: Train the dataset

Step 2: Calculate Entropy of the dataset.

Step 3:For every feature

1. Determine entropy for all parameters.
2. Choose average information entropy of current parameter.

Step 4: Find the gain of current parameter.

Step 5: Repeat the procedure until a desired trees is generated.

ID3 Tree Generation

$$G(K,B) = I(K) - \sum_{t\epsilon T} p(t)I(t) = I(K) - I(K|B)$$

$I(K)$=Entropy of K
I(t)= Entropy of t
p(t)=quantity of number
T=The subsets created from splitting set K by attribute B

VI. EXPERIMENTAL RESULT

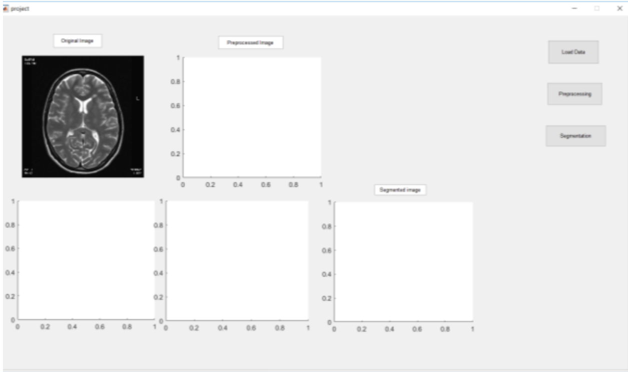


Fig. 1: Loading of image

The above fig:1 result shows the loading of autistic MRI as well the above figure window demonstrate with various buttons perform different operations.

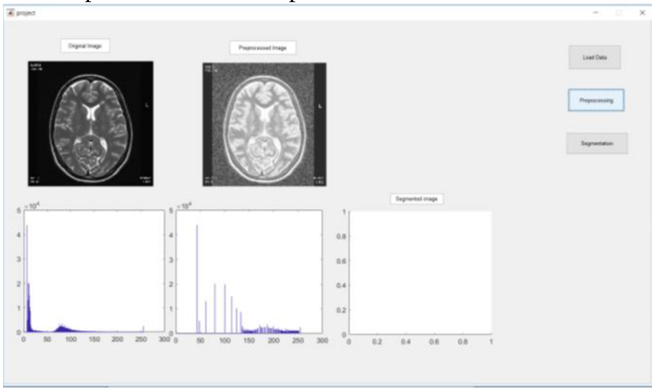


Fig. 2: Preprocessing

During preprocessing, the noises from the image were removed using histogram equalization technique. The histogram equalization method is used to intensify the contrast of image.

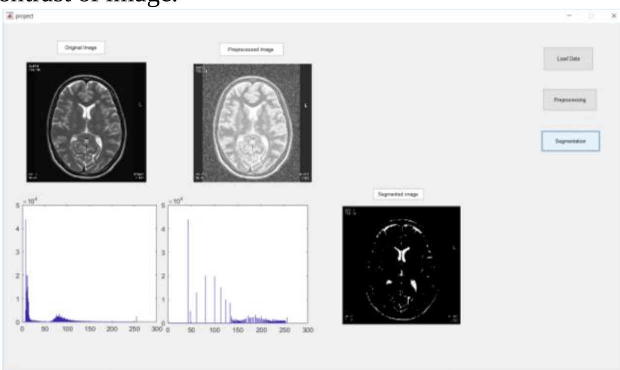


Fig. 3: Segmentation

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

In segmentation, fuzzy c means is the method used to find the autism effected regions. It is an iterative technique used for clustering to extract the autism affected region.

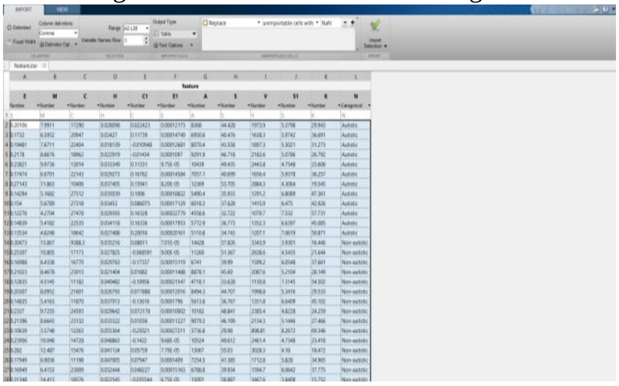


Fig. 4: Feature Extraction

The above result demonstrate with 11 features like entropy, energy, mean, kurtosis, variance, standard deviation, skewness, contrast, homogeneity, correlation and area from 28 brain MRI data set and converted in CSV file format.

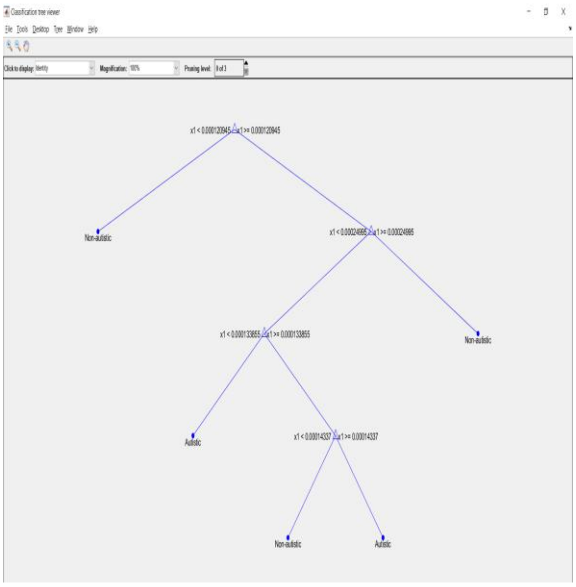


Fig. 5: Decision tree

From theabove classification result we used ID3 algorithm over a trained dataset which consist of parameters that were extracted from MRI images to create a decision tree and achieved an overall accuracy of 77%. Later the ID3 decision tree was used for classifying the dataset into autistic and non-autistic way.

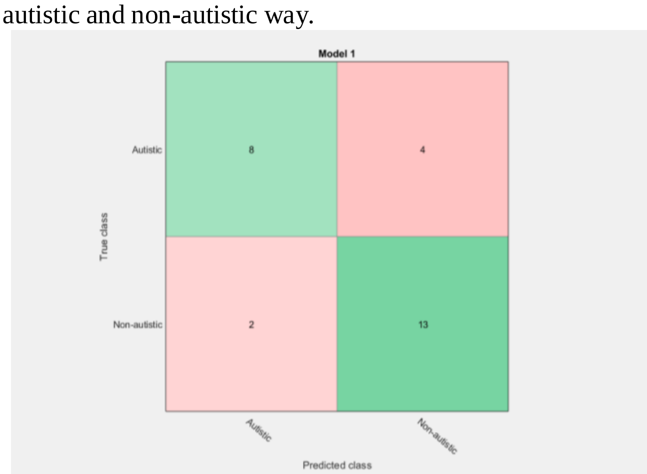


Fig. 6: Confusion Matrix

The above confusion matrix will give the accuracy of classifier as well as how many will be in autistic and non-autistic form. In our experiment we got the accuracy of 77 percentages.

VII. CONCLUSION

In our proposed system,we have collected around 500 MRI datasets and applied preprocessing technique using histogram equalization and then proceeded with segmentation using Fuzzy C Means method. After feature selection, we are classifying the data with the help of class label and later it is represented as a tree structure. From our work we can classify limited data set with normal classification .but in future we can take bulky data set to get better accuracy we are using various deep learning technique

REFERENCES

1. Heinsfeld, A. S., Franco, A. R., Craddock, R. C., Buchweitz, A., &Meneguzzi, F. (2018). Identification of autism spectrum disorder using deep learning and the ABIDE dataset. *NeuroImage: Clinical*, *17*, 16-23.
2. Nair, B. B., &,Rahul Raghunath (2018). Coding and functional defect region prediction of placental protein in an embryo cell of first trimester using ANN approach. In *Recent Findings in Intelligent Computing Techniques* (pp. 167-170). Springer, Singapore.
3. Rani, N. S., &Vasudev, T. (2016). A Comparative Study on Efficiency of Classification Techniques with Zone Level Gabor Features towards Handwritten Telugu Character Recognition. *International Journal of Computer Applications*, *148*(1).
4. Chanel, G., Pichon, S., Conty, L., Berthoz, S., Chevallier, C., &Grèzes, J. (2016). Classification of autistic individuals and controls using cross-task characterization of fMRI activity. *NeuroImage: Clinical*, *10*, 78-88.
5. Qureshi, M. N. I., Min, B., Jo, H. J., & Lee, B. (2016). Multiclass classification for the differential diagnosis on the ADHD subtypes using recursive feature elimination and hierarchical extreme learning machine: structural MRI study. *PloS one*, *11*(8), e0160697.
6. Lim, L., Marquand, A., Cubillo, A. A., Smith, A. B., Chantiluke, K., Simmons, A., ... &Rubia, K. (2013). Disorder-specific predictive classification of adolescents with attention deficit hyperactivity disorder (ADHD) relative to autism using structural magnetic resonance imaging. *PloS one*, *8*(5), e63660.
7. Peng, X., Lin, P., Zhang, T., & Wang, J. (2013). Extreme learning machine-based classification of ADHD using brain structural MRI data. *PloS one*, *8*(11), e79476.
8. Cheng, W., Ji, X., Zhang, J., &Feng, J. (2012). Individual classification of ADHD patients by integrating multiscale neuroimaging markers and advanced pattern recognition techniques. *Frontiers in systems neuroscience*, *6*, 58.
9. Calderoni, S., Retico, A., Biagi, L., Tancredi, R., Muratori, F., &Tosetti, M. (2012). Female children with autism spectrum disorder: an insight from mass-univariate and pattern classification analyses. *Neuroimage*, *59*(2), 1013-1022.
10. Chen, R., Jiao, Y., & Herskovits, E. H. (2011). Structural MRI in autism spectrum disorder. *Pediatr Res*, *69*(5 Pt 2), 63R-8R.

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

12. Ecker, C., Rocha-Rego, V., Johnston, P., Mourao-Miranda, J., Marquand, A., Daly, E. M., ... & MRC AIMS Consortium. (2010). Investigating the predictive value of whole-brain structural MR scans in autism: a pattern classification approach. *Neuroimage*, *49*(1), 44-56.
13. Ecker, C., Marquand, A., Mourão-Miranda, J., Johnston, P., Daly, E. M., Brammer, M. J., ... & Murphy, D. G. (2010). Describing the brain in autism in five dimensions—magnetic resonance imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. *Journal of Neuroscience*, *30*(32), 10612-10623.