IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 09, Issue 6 (June. 2019), ||S (III) || PP 56-65

Multimodal Neuroimaging for the Detection of Alzheimer's Disease

Gayathri V S [1], Deepa P L[2]
Dept. of ECE, MBCET, Thiruvananthapuram, India
Corresponding author: Gayathri V S

Abstract: Alzheimer's disease (AD) is a most common form of dementia which causes severe memory loss and decline in other cognitive functions. The early diagnosis of Alzheimer's disease is essential in order to avoid severity in the disease. The early stage of AD called as Mild cognitive impairment should be diagnosed early for better timely treatment. The combination of neuroimaging modalities like Magnetic resonance imaging (MRI) and positron emission tomography (PET) shows effective way for the diagnosis of Alzheimer's disease. The technique using polynomial neural network can be effectively used for the better diagnosis of Alzheimer's as early as possible. This paper explains an efficient method to detect Alzheimer's disease using a polynomial network by incorporating GMDH and multimodal neuroimaging techniques.

Keywords: Alzheimer's disease, Mild cognitive impairment, Deep learning, multimodal neuroimaging, GMDH, polynomial neural network.

Date of Submission: 11-06-2019 Date of acceptance: 28-06-2019

I. INTRODUCTION

Alzheimer's disease, which is an usual form of dementia, is a degenerative brain disorder which ends up with gradual memory loss. In US, it is the seventh leading reason behind the death and also it affects 5.3 million Americans. The familial AD affects the individuals with age less than 65, if we account within the United States over 500000 AD cases. Other classification is termed as isolated AD which occurs in adults having age greater than 65 and older. AD will make progress in step by step which may last for long years. There are three main stages of the disease which can differentiate mainly in terms of the severity. The symptoms and severity is different for each and every type. The doctors will examine the patients with the current symptoms and predict the stage in which they exists and also predicts the symptoms which can be shown in the future and its severity and also the possible course of treatment available for the stage. The classification of AD will vary ac-cording to many other reasons including age, genetics, education level and co-morbidities. For the effective diagnosis of Alzheimer's disease, the most definite way is to do an autopsy. There is no complete cure for Alzheimer's disease. But hopefully promising researches, analysis, and development for early detection is underway. Over some past decades MRI, Functional MRI, PET, CT has been used with the advancement in neuroimaging (Liu SD et al., 2015). Biomarkers have been used for diagnosing of Alzheimer's disease since it has capabilities of visualization, functional information of the brain (Jack CR et al., 2008). In some recent years, Computer Aided Diagnosis (CAD) has considerable achievements in neuroimaging system (Liu SQ et al., 2015). Also the different brain imaging modalities and multimodal analysis improves the diagnostic performance to a better step. The main aim in all diagnosis framework are to detect, analyze and classify the AD (Brookmeyer R et al., 2013) . In the case of CAD, since it needs feature representation, feature extraction will be a crucial step. In recent terms Deep Learning provides many applications by achieving great success. It was first introduced in 2006 by Hinton (Jack CR et al., 2008). By considering the conventional learning architectures, Deep learning provides better data representation of high level features with layered and hierarchical architecture (Shen DG et al., 2015). In recent years deep learning forms a good domain in medical field with great applications in detection, segmentation and classification (Payan A, Montana G et al., 2015).Also it can be used for the diagnosis of other brain disorders. Suk et al. have proposed a multimodal deep Restricted Boltman machine for learning features from neuro images having 3D huge patches (Szegedy C et al., 2015). Brosch et al. have used a Deep Belief Network for the effective diagnosis of AD. Gupta et al. have used a Stacked Auto Encoder (SAE) followed by the application of a Convolution network (Arica Nafiz et al., 2018). Liu et al. proposes a stacked auto encoder network with the help of multimodal neuroimaging for Region of Interest (ROI) based feature learning. ROI based approach is using widely since it covers the whole brain. But the features extracted using ROI cannot be reflected with small changes involved in the disease (Shen DG et al.,

2014). Payan et al. involved with AD prediction using local patches. However here creates a new space for feature learning with the help of learning algorithms (Tofighi G et al., 2016).

II. POLYNOMIAL NEURAL NETWORK

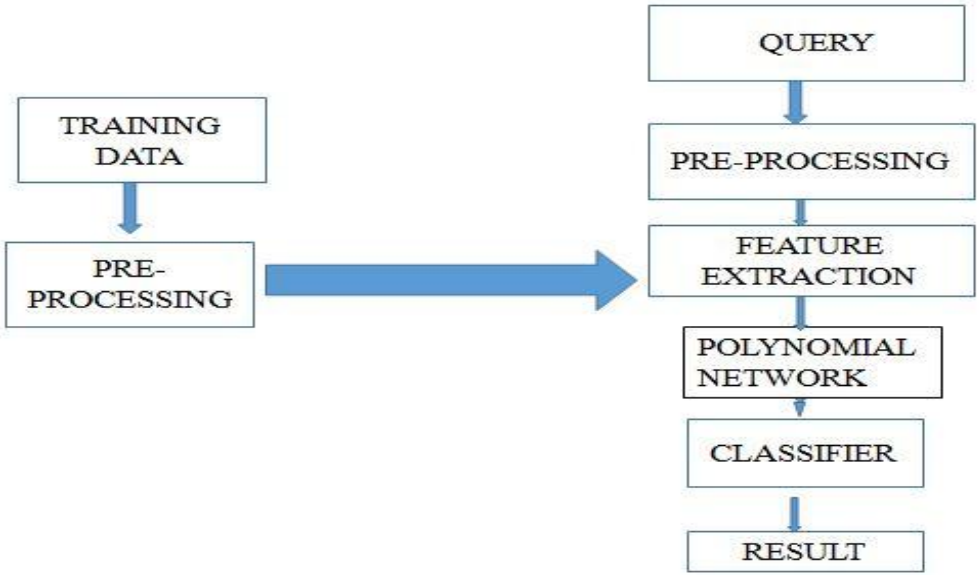


Figure 1: Block schematic of "Alzheimer's Disease" diagnosis supported on polynomial network

In the case of early diagnosis system of AD, here first preprocess the multimodality information given. Which are the MRI and PET images. The preprocessing is finished for training data in correspondence. Once preprocessed then the aim is to extract feature by employing a deep learning network and classify the given images to understand whether the patient get affected by Alzheimer's or not. The foremost common preprocessing steps for MRI and PET images includes demising/filtering ,image enhancement, normalization, morphological operations ,anterior commissure and posterior commissure correction, bias field correction corner detection, intensity inhomogeneity by N3 algorithm. Then MR images are divided into three completely different tissues, called white matter, grey matter and cerebrospinal fluid using FAST algorithm. Each and every neuro-image is partitioned into 93 ROIs. For each ROI, volume of grey matter tissue are calculated as a feature. For PET images, the average intensities of same ROIs are calculated as the feature.

I. Feature Extraction

After preprocessing features are extracted from the available dataset. Features are extracted for limiting the available data accurately. Feature extraction is used for analyzing the visual contents of the given images for mainly retrieving and indexing. The low level features of images are such as color shape texture. Here for better accuracy here extracting edge and texture features. Edge features are extracting using the edge direction histogram and texture feature are extracted using grey level co-occurrence matrix.

a. Edge Direction Histogram.

Basically edge direction histogram is used to represent the distributions of the edge points in each and every direction. Commonly it is calculated by taking the number of pixels in each direction which is user defined. In terms of edge detection, the main aim is to find out the points in a digital image where the image brightness changes sharply or contain any discontinuities. The point where the image brightness changes abruptly or sharply is called as edges by organizing the points into a distribution of curved line segments. Among this the step detection is finding the discontinuities in the one – dimensional signals and change detection in which finding discontinuities over time. Also it can be used for the feature extraction and feature detection processes

b. Texture features extraction using GLCM...

The Grey level Co-occurrence matrix can be make use to extract the second order statistical features of textures. Mainly four parameters exist for using a GLCM which can be discussed on following sections. For the analysis of a process having large dataset, requires a more computational power, more memory. Thus the feature extraction came into exists by combining the variables to solve the above problems. The texture extraction is for better robustness by representing the texture features in a unique way. There exists many

algorithms and architecture for feature extraction since it plays a vital role in pattern recognition, image processing and analysis. Here using GLCM second order features like entropy, homogeneity, correlation and contrast. These four parameters helps for having high discrimination accuracy. Usually statistical texture analysis will compute the texture features by computing the statistical distribution of the intensities at unique positions relative to each other. Based on the number of pixels the statistics will classified into first, second and higher order statistics.

A GLCM is basically a matrix in which the number of columns, rows will be equal to the number of grey levels. It is represented as G. Here P(I , j /Δx ,Δy) is the matrix element, the relative frequency in which two pixels separated by a distance (Δx ,Δy) with intensity I for one and j for another pixels. P (i , j / d , θ) is a matrix element containing the probability values of intensity i and j with a particular distance d and particular angle θ. For each combination of (Δx,Δy) or (d, θ) a large dataset of intensities are temporarily stored in a G*G matrix. One thing about the number of grey level is reduced since GLCM is sensitive to the size of the texture features to be estimated. For example, in the figure below GLCM matrix formation is given having four different grey level. Here we are using one pixel offset. That is contains with a reference pixel and one neighboring pixels. The large offset is possible only if the window is large. The cell on the topmost corner have to fill with how many number of times does the combination of 0,0 .It will refers how many times a pixel having grey level 0 (neighbor pixel) falls to another pixel having grey level 0(reference pixel). Grey level co-occurrence matrix become popular for extracting the texture features on statistical ground. Eventually fourteen texture features can be extracted from the probability matrix. Here four important features are taken includes correlation, homogeneity, entropy and contrast.

(a) Homogeneity

Homogeneity is also known as Angular Second Moment (ASM) or Uniformity. It mainly measures the homogeneity of the image. It is calculated by taking the sum of squares of entries in the GLCM. High homogeneity occurs when having similar pixels.

It is represented by the following equation.

$$\text{Homogeneity /ASM} = \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} pij^2 \quad (1)$$

where p(i, j) is the function and i, j are the spatial co-ordinates of that function.

(b) Inverse Difference Moment

Inverse Difference Moment (IDM) represents homogeneity in local. It is even high when local grey level points are uniform or when inverse GLCM becomes high. It is represented by the following equation.

$$\text{Contrast/ IDM} = \frac{\sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} pij}{1+(i-j)\text{^}2} \quad (2)$$

The contrast weight is the reciprocal of IDM weight.

(c) Entropy

For image compression, the amount of information which is needed is given by the entropy. It will measures the information contained in an image and also the loss of message/ information.

It is represented by the following equation.

$$\text{Entropy} = \sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} -Pij * \log Pij \quad (3)$$

d). Correlation

The linear dependency of neighboring pixels based on grey levels is measured by correlation. Correlation also helps in tracking and registration process of digital images for three dimensional and two dimensional measurements. The main application is for the motion estimation of optical mouse. Also used for measuring strain, displacement, optical flow etc. It is represented by the following equation.

$$\text{Correlation} = \frac{\sum_{i=0}^{Ng-1} \sum_{j=0}^{Ng-1} (i,j)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (4)$$

If the size of the image for which texture to be extracted increases, the values of all the features of the image will also increases. Thus the optimal size for better accuracy is 128*128. This will also reduce the loss of information from the image.

I. GMDH

Group method of data handling (GMDH) is a family of computer based algorithms that can be used for mathematical modelling and identification. Its applications can spread over the fields of data mining, pattern recognition, forecasting, image processing etc. an object which is analyzed by GMDH method can be

represented by multiple inputs and at least one output. Also it can be modeled by accommodating the components of subset of the base functions. It is represented by the given equation,

$$Y(x_{1,}x_{2\ldots\ldots}x_n = a_0 + \sum_{i=1}^{k}(a_i f_i) \quad (5)$$

where Y is the output and x are outputs. f is a function depending on the given set of inputs also k denotes the number of base components. The main advantage of GMDH model is, it is noise resistant. Also there exists a parameter called as external criterion which will determines the accuracy of the model. The input output relationship of the self-organizing network having one output and multiple inputs is represented in terms of Volterra Kolmogorov Gabor polynomial as,

$$y_n = a_0 + \sum_{i=1}^{M} a_i x_i + \sum_{i=1}^{M}\sum_{i=1}^{M} a_{ij} x_i x_j + \sum_{i=1}^{M}\sum_{j=1}^{M}\sum_{k=1}^{M} a_{ijk} x_i x_j x_k \quad (6)$$

where $X=(x_{1,}\ x2_{\ldots\ldots}x_M)$ is the vector form of input variables and $A= (a_{0,}a_{1,}\ a_{ii},a_{iik\ldots\ldots})$ is the vector form of weights or coefficients. In the case of multilayer GMDH method, can create a bivariate activation function polynomial in the nodes and in leaves, the variables. Thus we can allocate higher order polynomial by composing the activation function which are placed in the hidden nodes of the network. In effect with the neural network, the GMDH can be a multilayered and feed forward neural network.[21] The weights which are arriving in a particular hidden node is estimated by the help of ordinary lease square fitting algorithm. A multilayered procedure is equivalent to the Artificial Neural Network having the activation function. Thus the GMDH in such an approach can be referred to as GMDH neural network or Polynomial Neural Network. GMDH can be used to find out the input output relationship. Its structure is determined during the training process. It will express the nonlinearity function in the model using polynomial without the problem of instability. Here training happens on each layer. The network takes the decision that which input variable is more relevant to the current process. The network will keep forming the layers in training process. There will be having two input variables in a neuron and the output will be a quadratic function of these two inputs. The quadratic function is obtained by the analysis of linear regression. The layer is trained before entering to the next layer. The best performing neuron is selected based on some criterions. The adding of layers will stops until the stopping criterion is achieved.

GMDH can be used as a modelling technique mainly for the identifying the higher order nonlinear system. GMDH can automatically learn the status of relation which dominates in the system variables during the training process. The neuron structure will be selected automatically so that it minimizes the prediction error and other criterion for better accuracy and the unknown neurons are eliminated on the stages. It fit the complexity of n on linear characteristics.

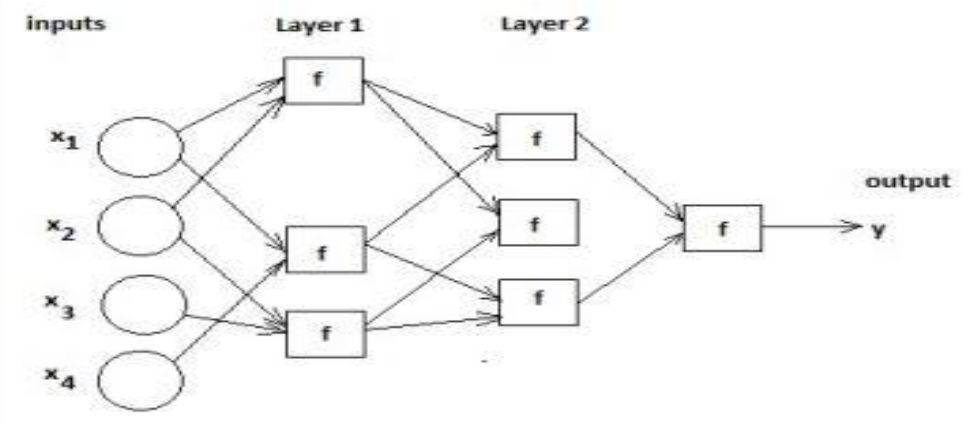


Figure 2: Network architecture of GMDH.

The basic GMDH algorithm can be explained on the following steps.

a. Divides the given data into two samples. Let it be A and B.
b. Structure for model is generated.
c. The coefficients are obtained using the least square method and data sample A.
d. The external criterion value is calculated for the data sample B.
e. Choose the model possessing minimum value of criterion.

The procedure for the implementation of GMDH is as follows.

Step 1: The separation of the available dataset into validation and training sets.

Step 2: generation of possible combinations of two inputs from the input variables

Step 3: the output of each polynomial is evaluated and tested using the data available on the validation set. After calculating the output, for each neuron, regularity criterion is calculated in the first layer. As the

criterion rule, the layer having the regularity value less than the predefined one allowed to be proceed to the next layer. Here the output of the current neuron will becomes the new values for inputs. This will happened to a number of selected neurons and other neurons will get eliminated from the network as it cannot reach the threshold values for regularity criterion.

Step 4: The whole process from the step 2 is repeating until the condition for an effective GMDH method gets satisfied.

The main advantage of GMDH network is that it provides an adaptive network for synthesis. It has threshold function as an objective function rather than transfer function. The layers are inductive. That is, by using the minimum regularity criterion, the number of nodes and layers are estimating. The data can be stored in the way of easy accessibility and can use repeatedly. The convergence is guaranteed globally. The optimization is done fast since it eliminates the unwanted neurons. We can evaluate features which are linear and nonlinear, dynamic and static. With the help of the fitting technique, by calculating the ordinary least square fitting, we will be able to eliminate the bad and provides good weights for better accuracy.

The limitation of GMDH method is that, it is not suitable for long range applications. Sometimes the over fitting problem can be caused. The partition of data will affect the optimum model since it aided for the utilization of regularity criterion.

Here for creating and training a GMDH network, The maximum number of neurons in a layer is taken as 40, the maximum number of layer is taken as 20, the selection feature in each layer is taken as 1.2 denoted by alpha, the validation ratio is taken as 0.85

The flow chart of the GMDH network is as shown below.

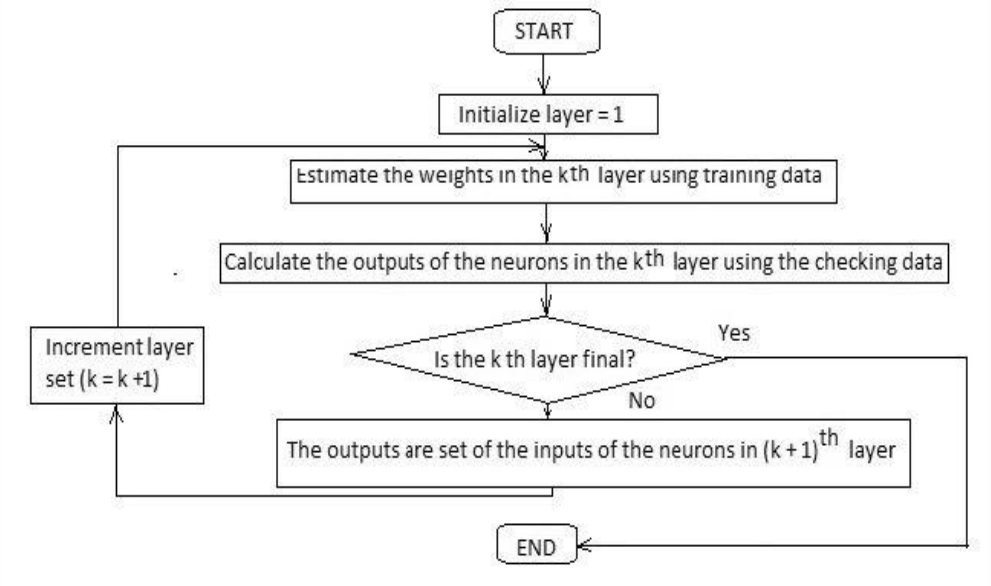


Figure 3: Flow chart of GMDH network.

II. Feature Selection

Reduction of dimension of the feature is more important in the case of statistical learning. In some bioinformatics data, the features will be in a large amount and the features for observation will be less in number. This will cause for over fitting problem due to noise. Among the number of features, the features which are useful for creating a learning process will be less. So it is necessary for the feature selection which will reflect the features which are needed For the general idea of separation of two groups by each feature, we can apply the t-test algorithm for each feature and compare the p- values (the absolute value) of t- statistical for each and every features to measure how effectively they can be separated. For that we plot the Cumulative Distribution Function (CDF) of the p- values.

a. Feature Selection Using T Test

The main aim is to find out some of the important features from the large feature set so that it will provide better classification performance. In the feature selection techniques, the t test has a very common form called as student t test. Basically it is used to find whether the mean of two classes are different on the basis of statistics. It is calculated by taking the difference between the two class mean and the two class variability, then taking the ratio between these two parameters. In the case of multiclass problem, the t

statistics is calculated by taking the difference between mean of one class to the mean of all other class. But the difference is standardizing by within the class standard deviation.
The t-statistics value for i-th in the c-th class is given by,

$$t_{ic} = \frac{x_{ic} - x_i}{M_c.(S_i + S_0)} \quad (7)$$

Where x_{ic} denotes the mean of i-th feature in the c-th class. And x_i represents the mean of i-th feature in all classes. And S_0 represents the median value of all within class standard deviation for all features. The within class standard deviation is given by the following equation,

$$S_I^2 = \frac{1}{N-C}\sum_{c=1}^{C}(x_{ij} - x_{ic})^2 \quad (8)$$

where x_{ij} represents i-th feature present in j-th sample.
The total mean of the class is given by,

$$M_c = \sqrt{(1/n_c + 1/N)} \quad (9)$$

where N is the total number of samples present and n_c is the number of samples present in class c. Here extends the t-test score of all the features and finding out the t- score of i- th feature is greater among the t-score values of all classes for the feature represented by i. It is given by,

$$t_i = \max\left\{\left|\frac{x_{ic} - x_i}{M_C S_i}\right|, c = 1,2,3 \ldots . C\right\}. \quad (10)$$

Thus we can generalize the feature with t-score as follows.

1. Consider the set of features, F= $\{f_1, f_2, \ldots\ldots\ldots f_i \ldots\ldots f_g\}$ and feature i have nominal values which is represented by, $f_i = \left\{x_i^{(1)}, x_i^{(2)} \ldots \ldots \ldots x_i^{(mi)}\right\}$.
2. Now we have to convert this nominal features into a vector form
3. Then replace the numerical value of features with the vectors obtained.
4. The ranking rule is such a way that the feature will be more relevant if the t-score is greater.

III. Sparse Classifier

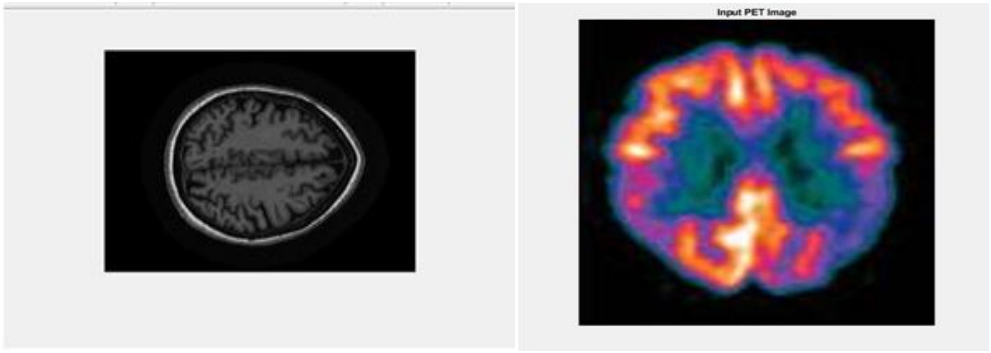
For better classification, rather than using a nearest neighborhood (NN) classifier, a sparse classifier can be defined in such a way that its cost function involves with continuous parameters (eg: support vector machine (SVM), neural network, logistic). Then we can create a sparse on the vector form of parameters. The sparse classifier provides its ability and tries to set zero value to most of the estimation parameters. This is done with the help of cost function. The cost function of a sparse classifier is given by the below equation.

$$J(\theta) = |y - X\theta|_2^2 + \lambda|\theta|_2^2 \quad (11)$$

The sparsity mentions that only a small portion of the input variable is responsible for the entire classification. So the duty of sparse classifier is to find out that portion of influence.

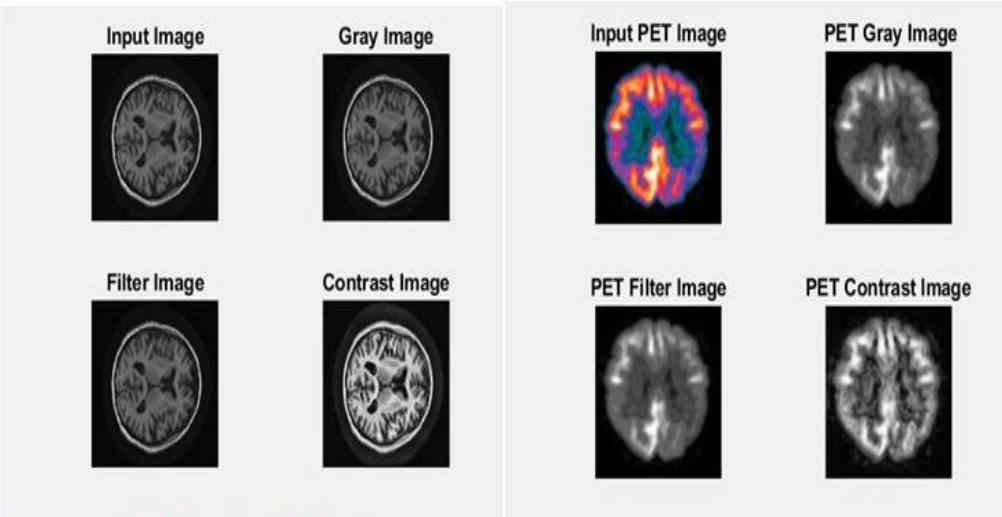
III. PERFORMANCE EVALUATION

The multimodal neuroimaging data is obtained from the 'Alzheimer's Disease Neuroimaging Initiative' (ADNI) database. The ADNI was first introduced on 2003 as a combined private and public partnership. The main goal of ADNI is to measure the progression in the stages of Alzheimer's disease mainly during the mild cognitive impairment and the early stages by incorporating the MRI, PET, clinical and bio markers. Here the simulations are done by using MATLAB 2018a.



(a) (b)

Figure 4: (a) Input MRI image (b) Input PET image



(a) (b)

Figure 5: (a) Preprocessed MRI (b) Preprocessed PET image

The stimulated results of preprocessing stages including filtering, normalization, image enhancement is given above.

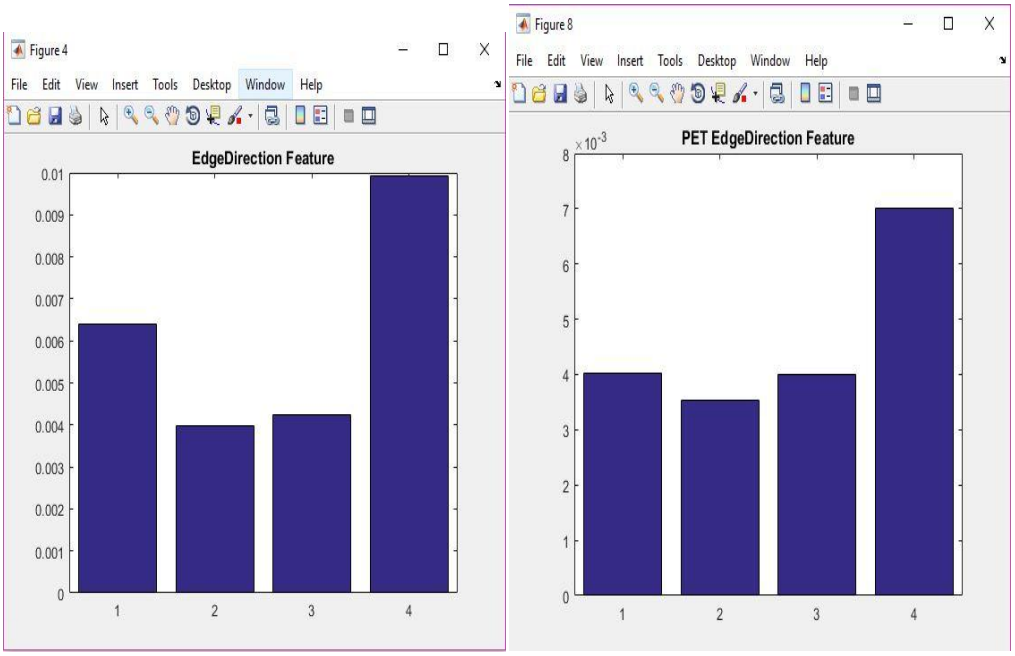


Figure 6: (a) MRI edge direction feature (b)PET edge direction feature

The simulated results for feature extraction for PET and MRI images are given above. As discussed texture and edge features are extracted from the available dataset.

The output obtained from the feature selection using t- test algorithm is given below. Here from the graph plotted againt the CDF value in y axis and p values in x axis, about 75% of the total features are observed with absolute values close to zero. It means that if the features in the data is is about 5000, then there exists 3750 features among the original 5000 features have strong discrimination power.

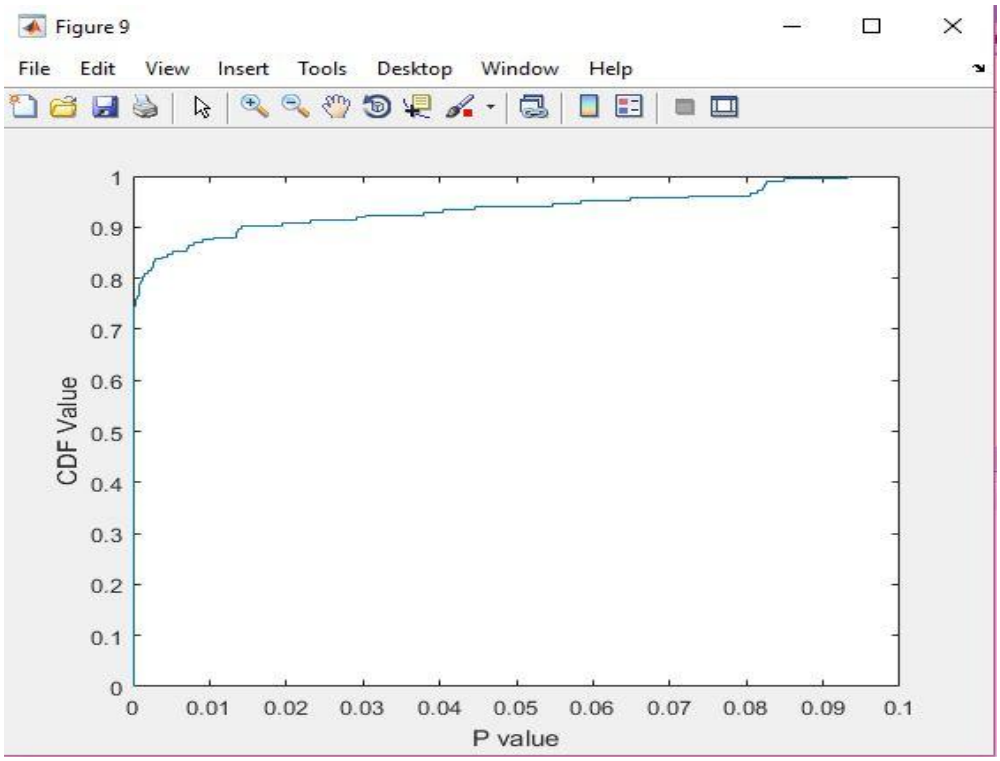


Figure 7: The graph plot between Cumulative Distribution Function and the absolute value (p)

The output obtained after using sparse classifier for the effective classification on neuroimaging techniques into two classes Alzheimer's and Non Alzheimer's disease.

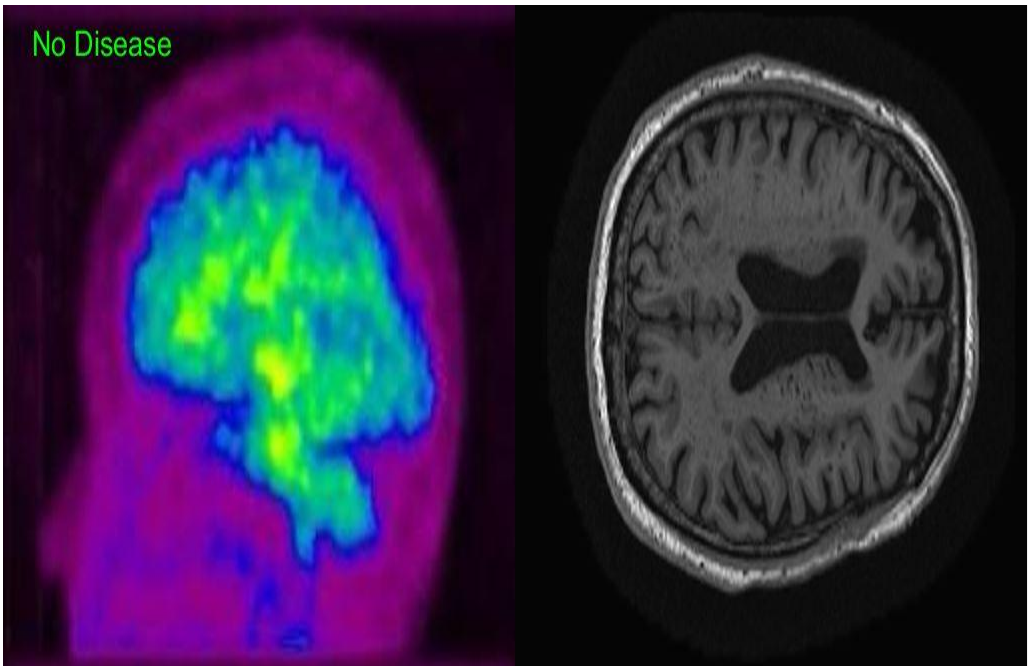


Figure 8: PET and MRI image having non alzheimer's disease is correctly classified under the class of no disease.

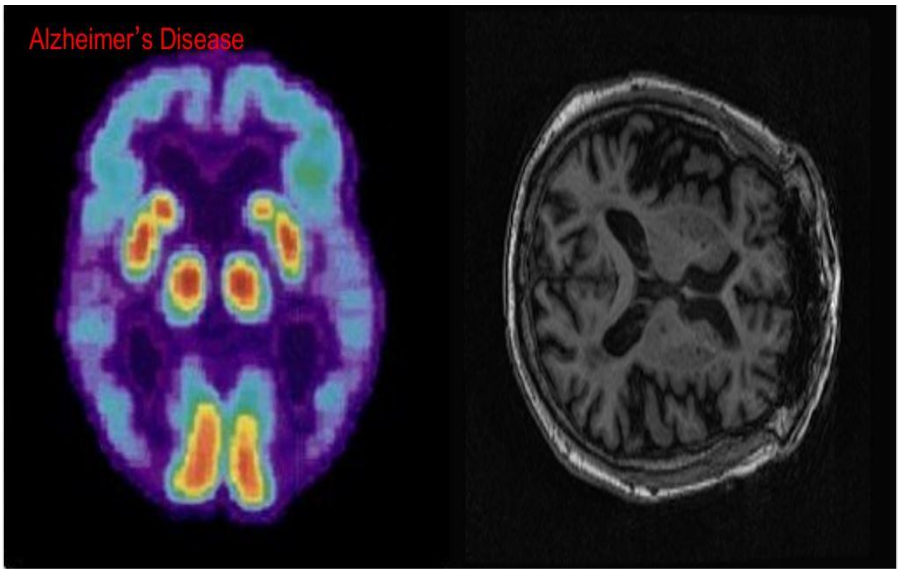


Figure 9: MRI and PET image having alzheimer's disease is correctly classified under the class of alzheimer's disease

Here the time taken for completing the whole process is about 14-20 seconds.

TABLE 1: Performance analysis for the detection of alzheimers disease using nn classifier and sparse classifier.

SL NO	ACTUAL STATUS OF THE MRI AND PET IMAGES	DETECTION OF NEUROIMAGES WITH ALZHEIMER'S BY USING NN CLASSIFIER IN PNN	DETECTION OF NEUROIMAGES WITH ALZHEIMER'S DISEASE USING SPARSE CLASSIFIER IN PNN
1	ALZHEIMER'S	UNKOWN/NO DISEASE	ALZHEIMER'S
2	ALZHEIMER'S	ALZHEIMER'S	ALZHEIMER'S
3	ALZHEIMER'S	UNKNOWN/NO DISEASE	ALZHEIMER'S
4	ALZHEIMER'S	ALZHEIMER'S	ALZHEIMER'S
5	NO ALZHEIMER'S	NO DISEASE	NO ALZHEIMER'S
6	NO ALZHEIMER'S	UNKNOWN/NO DISEASE	ALZHEIMER'S
7	NO ALZHEIMER'S	ALZHEIMER'S	NO ALZHEIMER'S
8	NO ALZHEIMER'S	UNKNOWN/NO DISEASE	ALZHEIMER'S

The results in terms of accuracy can be taken from the above table. Using NN classifier in PNN network , three among the eight images are misclassified. While using polynomial network with sparse classifier gives more efficient classification than the previous. Regarding the time consumption for the entire methodology, sparse classifier incorporating with PNN provides better time accuracy.

IV. CONCLUSION

This paper introduces a completely unique methodology for the effective early diagnosis of Alzheimer's disease. The diagnosis is using a Polynomial Neural Network incorporating GMDH. The early stage of Alzheimer's disease that is mild cognitive impairment is also detected using this method. If we are comparing this methodology to conventional methods of Alzheimer's disease classification based on machine learning techniques like Support vector machine, this will enable to classify not only in binary classification but also the multi class classification. Here it needs minimum knowledge about the previous sessions with respect to the improvement in the model used. There exists spatial dimensionality reduction in this method which is a disadvantage of traditional methods. Feature selection using the t test algorithm provides better classification rate. Also the data fusion is done simultaneously with dimensionality reduction to take advantages over data modalities. In performance wise, the method provides efficiency in both binary and four class classification. This method will pave a way for effective diagnosis of Alzheimer's disease in the biomedical field under the computer aided diagnosis.

REFERENCES

[1]. Brookmeyer R, Johnson E, Ziegler-Graham K, Michael Arrighi H (2007) Forecasting the global burden of Alzheimer's disease. Alzheimer's Dement. ,IEEE Transaction on medical imaging, 3:186–191

[2]. Brosch T, Tam R (2013) Manifold learning of brain MRIs by deep learning. In: Proc. Int. Conf. Med. Image Comput. Comput. Assist. Interv., pp. 633-640,.

[3]. Gupta A, Ayhan M, Maida A (2013) Natural image bases to represent neuroimaging data . In: Proc. Int. Conf. Mach. Learn., pp. 987-994.

[4]. Gürkaynak Cahit Deniz, Arica Nafiz, (2018) A case study on transfer learning in convolu-tional neural networks. In: IEEE 2018: 26th Signal Processing and Communications Ap-plications Conference (SIU)10.1109/SIU.

[5]. He K (2014) Spatial pyramid pooling in deep convolutional networks for visual recognition. In: Computer Vision ECCV 2014. New York: Springer International Publishing, pp 346-361.

[6]. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science 313:504-507

[7]. Jack CR, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowski B, Britson PJ, Whitwell JL, Ward C, Dale AM, Felmlee JP, Gunter JL, Hill DLG, Killiany R, Schuff N, Fox-Bosetti S, Lin C, Studholme C, DeCarli CS, Krueger G, Ward HA, Metzger GJ, Scott KT, Mallozzi R, Blezek D, Levy J, Debbins JP, Fleisher AS, Albert M, Green R, Bartzokis G, Glover G, Mugler J, Weiner MW (2008) The Alzheimer's disease neuroimag-ing initiative (ADNI): MRI methods. In:J. Magn. Reson. Imaging, 27:685-691

[8]. Lin M, Chen Q, Yan S (2013). Network in network [EB / OL]. Http: // arxiv.Org / abs / 1312. 4400

[9]. Liu SD, Cai WD, Liu SQ, Zhang F, Fulham M, Feng DG, Pujol S, Kikinis R (2015) Mul-timodal neuroimaging computing: a review of the applications in neuropsychiatric disor-ders. Brain Inform. 2:167-180

[10]. Liu SQ, Liu SD, Cai WD, Che HY, Pujol S, Kikinis R, Feng D, Fulham MJ (2015) Mul-timodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease. IEEE Trans. Biomed. Eng. 62:1132-1140

[11]. Payan A, Montana G (2015) Predicting alzheimer's disease: a neuroimaging study with 3d convolutional neural networks. In: arXiv preprint arXiv:1502.02506

[12]. Risacher SL , Saykin AJ (2013) Neuroimaging and other biomarkers for Alzheimer's dis-ease: the changing landscape of early detection. Annu. Rev. Clin (eds). Psychol. 9:621-648

[13]. Sarraf S, Anderson J, Tofighi G (2016) Deep ad: Alzheimer'sdisease classification via deep polynomial neural networks using mri and fmri. bioRxiv, page 070441

[14]. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw, 61:85-117

[15]. Shi J, Zheng X, Li Y, Zhang Q, Ying S (2018) Multimodal Neuroimaging Feature Learn-ing with Multimodal Stacked Deep Polynomial Network for Diagnosis of Alzheimer's Disease. IEEE Journal of Biomedical and Health Informatics 22

[16]. Shin HC, Orton MR, Collins DJ, Doran SJ, Leach MO (2013) Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D pa-tient data. IEEE Trans. Pattern Anal. Mach. Intell. 35:1930-1943

[17]. Suk HI, Lee SW, Shen DG (2014) Hierarchical feature representation and multimodal fu-sion with deep learning for AD MCI diagnosis. NeuroImage 101:569-582

[18]. Suk H, Shen D (2013) Deep learning-based feature representation for AD/MCI classifica-tion. In: Proc. Int. Conf. Med. Image Comput. Assist. Interv., pp. 583-590.

[19]. Szegedy C (2015) Going deeper with convolutions. Computer Vision and Pattern Recogni-tion (CVPR). In:IEEE Conference on. Boston, USA, pp 1-9.

[20]. Turner JA, Potkin SG, Brown GG, Keator DB, McCarthy G, Glover GH (2007) Neuroim-aging for the diagnosis and study of psychiatric disorders. IEEE Signal Process. Mag 24:112-117

[21]. Godfrey C. Onwubolu. (2014): Sprand understanding GMDH implementation and algorithm. In: srpringer International Publishing, pp 01-24.

[22]. Zhang WL, Li RJ, Deng HT, Wang L, Lin WL, Ji SW, Shen DG (2015) Deep convolu-tional neural networks for multi-modality isointense infant brain image segmentation," Neuroimage 108:214-224

Gayathri V S" Multimodal Neuroimaging For The Detection Of Alzheimer's Disease" IOSR Journal of Engineering (IOSRJEN), vol. 09, no. 06, 2019, pp. 56-65.