A Machine Learning Based Approach for Early Detection of Alzheimer's Disease by Extracting Texture and Shape Features of the Hippocampus Region from MRI Scans

Arpita Raut[1], Vipul Dalal[2]

Student of Master of Engineering, Department of Computer Engineering, Vidyalankar Institute of Technology,
University of Mumbai, India[1]

Associate Professor, Department of Computer Engineering, Vidyalankar Institute of Technology, University of
Mumbai, India[2]

Abstract: Alzheimer's Disease (AD) is a neurological disease and the most common cause of dementia in the age group 65 years and above. The accurate and timely diagnosis of the AD is crucial in order to prevent the progression of this irreversible disease. This paper concentrates on a method to detect Alzheimer's from MRI scans using machine learning approach. The proposed approach focuses on the hippocampal region of the brain. The texture features such as entropy, homogeneity, energy, contrast, correlation and variance are extracted from the hippocampus region using the Gray Level Co-Occurrence Matrix (GLCM). The area and shape feature are extracted using the Moment Invariants. The Error-back Propagation (EBP) in Artificial Neural Network (ANN) is used as the classifier for detection of various stages of AD. The proposed system gives an average accuracy of 86.8%.

Keywords: Alzheimer's Disease (AD), Gray Level Co-Occurrence Matrix (GLCM), Artificial Neural Network (ANN).

I. INTRODUCTION

Alzheimer's disease (AD) is a degenerative disorder of the brain that leads to memory loss. It is the most common form of dementia, caused by the build-up of beta amyloid plaques in the brain. The plaques and tangles are some of the main features of the disease. As the number of Plaques and tangles increases, the healthy neurons begin to function less effectively and gradually lose their ability to communicate and finally die which results in overall shrinkage of the brain tissues. The death of neurons particularly in the hippocampus restricts the ability to form new memories. The hippocampus is the first region in the brain which gets affected. It is the region in the brain that is responsible in forming memories and serves as a relay structure between the brain and the body.

As Alzheimer's is a progressive and irreversible disease it progresses gradually following a distinct pattern of brain damage and can last for decades. The disease progresses slowly into three main stages namely mild, moderate and severe, where each stage has its own symptoms and challenges. The mild stage generally lasts for about an average of 2 to 4 years in which a person may function independently but may have memory lapses. In the moderate stage a person may require assistance in carrying out day to day activities and may experience increased difficulty with memory. This stage is probably the longest stage of the disease and may last for a duration of about 2 to 10 years.

The final stage of the disease may last between a period of 1 to 3 years. In this stage, the memory and cognitive skills continue to worsen and the patient may require round the clock assistance as he/she loses the ability to respond to their environment. It is very important to diagnose the disease at an early stage before irreversible neural damage is done. The current non- automated methods such as Cognitive Impairment Testing, Mini-Mental State Examination (MMSE) and Clinical Dementia Rating (CDR) as well as Imaging techniques such as Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) and Single-Photon Emission Computed Tomography (SPECT) are used to track abnormal changes in the brain and diagnose AD.

In the proposed work, the detection of AD is done from MRI scans. The texture, area and shape features are extracted using Gray-Level Co-Occurrence Matrix (GLCM) and Moment Invariants from the hippocampus which is selected as the Region of Interest (ROI). GLCM extracts the second order statistical texture features and Moment Invariants define

the set of properties used for identification of shape. AD is then classified into various stages based on the features extracted from the ROI. using Artificial Neural Network (ANN) which trained using Error Back Propagation (EBP) algorithm.

The paper is organized as follows. Section 2 describes work related to detection of AD using MRI scans. The proposed approach is explained in section 3. Section 4 consists of results. Finally, section 5 concludes the proposed work.

II. RELATED WORK

Yang Han and Xing-Ming Zhao [6] proposed a feature selection based approach known as Hybrid Forward Sequential Selection (HFS) for detection of Alzheimer's Disease. This proposed approach combines the filter and wrapper approaches to detect informative features from the MRI data which was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. In this approach the features were ranked and top-k features are selected. The Support Vector Machine(SVM) was used as the classifier. The authors claim that proposed approach outperforms other feature selection methods, improves accuracy of diagnosis and reduces computational cost.

Devvi Srawinda and Alhadi Bustamam[8] proposed an Advanced Local Binary Pattern (ALBP) method. The ALBP method was introduced as 2D and 3D feature extraction descriptors. As ALBP produces large number of features, the feature selection was done using principal Component Analysis (PCA) and factor analysis. The Support Vector Machine was used for multiclass classification. The authors claim that their proposed work gives better performance and accuracy compared to the previous Local Binary Pattern(LBP) method. The average accuracy achieved for whole brain and hippocampus data was between 80% and 100%. It is also claimed that the uniform rotation invariant ALBP sign magnitude outperforms other approaches with an average accuracy of 96.28% for multiclass classification of whole brain image. It is concluded that the extracted feature vector has high dimensionality, requires high computation for processing and can be improved using parallel computing for feature extraction from large brain datasets of MRI.

M.Evanchalin Sweety and G.Wiselin Jiji[15] proposed a Particle Swarm Optimization (PSO) and Decision Tree Classifier based method for Alzheimer's disease detection. In the proposed approach, the processed images are normalized and the Markov random filter is used for noise reduction. Features are extracted from the normalized images using moments and Principal Component Analysis. The particle Swarm Optimization for reduction of extracted features and the classification is done using the Decision Tree Classifier. The authors claim that for similar work their proposed approach gives an accuracy of 92.07% on SPECT images and 86.71% on PET images.

In the work proposed by Lauge Sorensen et.al[13], the hypothesis that the hippocampal texture is associated with early cognitive loss was tested. They used three independent datasets from Australian Imaging Biomarkers and Lifestyle flagship study of ageing(AIBL), ADNI and metropolit 1953 for training the classifier. In this study it was found that hippocampal texture is a better bio-marker as compared to reduction in hippocampal volume for predicting MCI-to-AD conversion for ADNI dataset. The hippocampal texture was found to have superior differentiation capability between stable MCI and MCI-to-AD conversion than volumetric changes in the hippocampus region. The findings in their research supported the hypothesis that textural information of hippocampus region is more sensitive as compared to volume and can be used to detect AD in early stage.

Chetan Patil et.al[12] proposed an approach to estimate the possibility of early detection of Alzheimer's by evaluating the utility of image processing on MRI images. The authors demonstrated the applications of image processing techniques such as k-means clustering, wavelet transform, watershed algorithm and a customized algorithm implemented on open source platforms, OpenCV and Qt. In the proposed approach, the T1-weighted MRIs were used for image processing to evaluate the Hippocampal atrophy. A boundary detection algorithm was used to extract the region of interest(ROI) and k-means clustering was used for segmentation. The brain and hippocampal volume is implemented. The authors claim that the overall brain volume is less and the difference between the grey and white matter is higher in case of Alzheimer's. It is also claimed that their proposed work is useful to the technical as well as medical community.

III. PROPOSED APPROACH

The system is implemented using MATLAB R2014a. To perform experiments the OASIS dataset is used. From this dataset 235 MRI scans with different AD stages were downloaded. These scans are used to train the network.

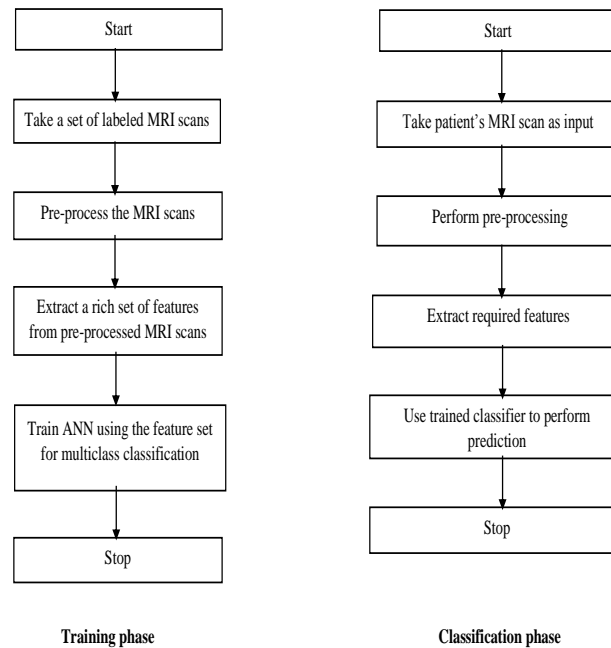The proposed approach is divided into two phases:

Figure1. Flowchart of Proposed Approach

A. Preprocessing

The pre-processing on the MRI scans is done using Contrast enhancement.

B. ROI Extraction

The hippocampus is extracted as the region of interest as it is the first region in the brain which gets affected in Alzheimer's disease.

C. Feature Extraction

The Texture, Shape and Area features are extracted from the Hippocampus region of the Brain for detection of AD. The Gray Level Co-occurrence Matrix is used to extract the texture features and the Shape, Area features are extracted using seven moment invariants. The features available with the dataset such as age, gender, education, socio-economic status, Mini-Mental Examination Score are also extracted. These extracted features are then used to generate the feature vector.

1) Gray Level Co-occurrence Matrix (GLCM):

GLCM is a method used for extracting the second order statistical texture features. It is a matrix in which the number of rows and columns are equal to the number of gray levels in the image. GLCM of an image is computed using a displacement vector $d(\delta, \theta)$. Six important texture features are extracted using GLCM:

I. Entropy: It measures the complexity in the image.

$$\mathbf{Entropy} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} -P_{ij} * \log P_{ij}$$

II. Energy: It is a measure of global uniformity in the image.

$$\mathbf{Energy} = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_{ij}^2$$

III. Local Homogeneity: It is a measure of local uniformity in the image.

$$\mathbf{Local\ Homogeneity} = \frac{\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_{ij}}{1 + (i-j)^2}$$

IV. Contrast: It measures intensity variation between a pixel and its neighborhood.

IJARCCE

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified
Vol. 6, Issue 6, June 2017

$$\mathbf{Contrast} = \sum_{i,j=0}^{N_g-1} P_{ij}(i-j)^2$$

V. Correlation: It is a measure of linear dependency between gray levels.

$$\mathbf{Correlation} = \sum_{i,j=0}^{N-1} P_{ij}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{(\sigma_i^2)\,(\sigma_j^2)}}\right]$$

VI. Variance: It is a measure of heterogeneity.

$$\sigma_i^2 = \sum_{i,j=0}^{N-1}(i-\mu_i)^2\,P_{ij}$$

$$\sigma_j^2 = \sum_{i,j=0}^{N-1}\left(j-\mu_j\right)^2 P_{ij}$$

2) Moment Invariants:
The moment invariants were first introduced by Hu in 1962. Six absolute orthogonal invariants and a skew orthogonal invariant based on algebraic invariants were derived. These derived invariants were independent of position, size, orientation as well as parallel projection. These invariants define the calculated set of properties of region that can be used for class identification as well as identification of shape.
Central Moments are given by:

$$\mu_{p,q} = \sum_{x,y}(x-x_c)^p\,(y-y_c)^q$$

Where, (x_c, y_c) is the center of the object.

Central moments can be standardized to make them scale-independent,

$$\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^\gamma}$$

Where,

$$\gamma = \frac{p+q+2}{2}$$

Hu brought forward seven moments independent of translation, rotation and scaling based on these moments.

$$\phi_1 = \mu_{2,0} + \mu_{0,2}$$
$$\phi_2 = \left(\mu_{2,0} - \mu_{0,2}\right)^2 + 4\mu_{1,1}^1$$
$$\phi_3 = \left(\mu_{3,0} - 3\mu_{1,2}\right)^2 + \left(\mu_{3,0} - 3\mu_{2,1}\right)^2$$
$$\phi_4 = \left(\mu_{3,0} + \mu_{1,2}\right)^2 + \left(\mu_{0,3} + \mu_{2,1}\right)^2$$
$$\phi_5 = \left(\mu_{3,0} - \mu_{1,2}\right)\left(\mu_{3,0} + \mu_{1,2}\right)\left[\left(\mu_{3,0} + \mu_{1,2}\right)^2 - 3(\mu_{\_2,1} + \mu_{\_0,3})^2\right]$$
$$+ \left(3\mu_{2,1} - \mu_{0,3}\right)\left(\mu_{2,1} + \mu_{0,3}\right)\cdot\left[3\left(\mu_{3,0} + \mu_{1,2}\right)^2 - \left(\mu_{2,1} + \mu_{0,3}\right)^2\right]$$
$$\phi_6 = \left(\mu_{2,0} - \mu_{0,2}\right)\left[\left(\mu_{3,0} + \mu_{1,2}\right)^2 - \left(\mu_{2,1} + \mu_{0,3}\right)^2\right] + 4\mu_{1,1}(\mu)\left(\mu_{2,1} + \mu_{0,3}\right)$$
$$\phi_7 = \left(3\mu_{2,1} - \mu_{0,3}\right)\left(\mu_{3,0} + \mu_{1,2}\right)\left[\left(\mu_{3,0} + \mu_{1,2}\right)^2 - 3\left(\mu_{2,1} + \mu_{0,3}\right)^2\right]$$
$$- \left(\mu_{3,0} - \mu_{1,2}\right)\left(\mu_{2,1} + \mu_{0,3}\right)\left[3\left(\mu_{3,0} + \mu_{1,2}\right)^2 - \left(\mu_{2,1} + \mu_{0,3}\right)^2\right]$$

All the extracted features are merged to form a feature vector. The training set consists of feature vectors extracted from the MRI scans. Each feature vector has 18 features.

D. Classification

Artificial neural network using error-back propagation is used for multi-class classification. The Scaled Conjugate Gradient (SCG) algorithm is used to train the network. The neural network is trained on 235 MRI scans. There are 135 CDR0 scans, 69 CDR0.5 scans, 29 CDR1 scans and 2 CDR2 scans. Once the network is trained it is ready for classification of AD's stages. The input layer consists of 21 neurons as it is the length of the feature vector. 4 hidden layers are considered and the output layer consists of 4 neurons which are used to classify the subjects into normal, mild, moderate and severe category.

IV. RESULTS

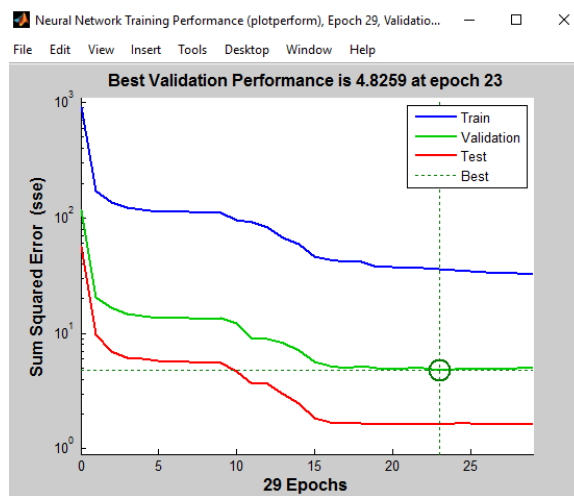The Neural Network training performance is measured using Sum Squared Error (sse).

Figure 2. Neural Network Training Performance

The figure 2 depicts the neural network training performance. The performance is shown for each of the training, validation and test sets. The x-axis shows the no.of epochs/iterations required for training and the y-axis shows the performance measured in terms of sum square error (sse) which is shown in log scale. The ANN fits the best value after 15 epochs. sse of the ANN is decreased with the number of epochs. The ANN was well trained as it presents a very low sse at the end of the training phase. The best validation performance found was 4.8259 at epoch 23.

The error histogram was used to obtain additional verification of network performance.
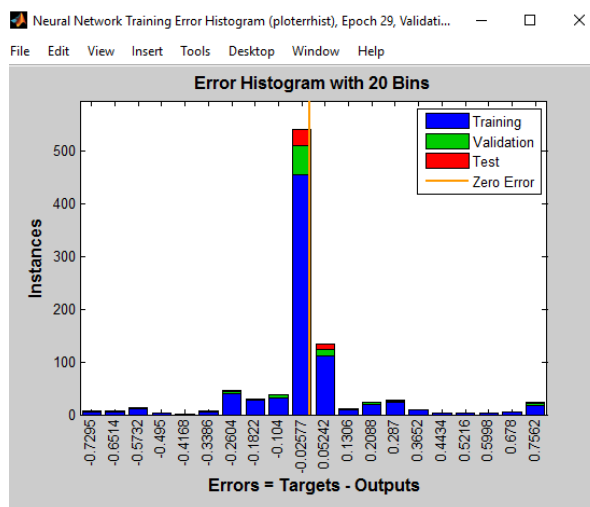

Figure 3. Error Histogram

The figure 3 depicts the error histogram. The x-axis represents various error values occurred during training, validation and testing phases. The y-axis represents number of times a particular error occurs. Most of the large values of errors have occurred during training phase.

324

The confusion matrix shows the various types of errors that occurred for the final trained network

Figure 4. Confusion Matrix

The confusion matrix is shown in figure 4. The overall accuracy found was 86.8% and overall error rate was 13.2%.

V. CONCLUSION

In this work, a machine learning based approach for detection of Alzheimer's disease is proposed. The OASIS dataset was used for experiments. The texture, area and shape features from hippocampus region of MRI scan are extracted. The textual features from OASIS were also extracted. These features were used to train the neural network with error-back propagation for classification. The proposed system has an average accuracy of 86.8 %.

REFERENCES

[1] Ali H. Al-nuaimi et.al "Changes in the EEG Amplitude as a Biomarker for Early Detection of Alzheimer's Disease",2016 30th Annual International Conference of the IEEE Engineering in Medicine and Biology society (EMBC).

[2] Jiehui Jiang et.al "A Computed Aided Diagnosis tool for Alzheimer's Disease based on 11C-PiB PET imaging technique", IEEE International Conference on Information and Automation Lijiang, China, August 2015.

[3] R. Chaves et.al "FDG and PIB Biomarker PET Analysis for the Alzheimer's Disease Detection Using Association Rules", 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC).

[4] Imene Garali et.al "Region-Based Brain Selection and Classification on PET Images for Alzheimer's Disease Computer Aided Diagnosis",2015 IEEE International Conference on Image Processing (ICIP).

[5] Ali H. Al-nuaimi et.al "Tsallis Entropy as a Biomarker for Detection of Alzheimer's Disease",2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC).

[6] Yang Han, Xing-Ming Zhao "A hybrid sequential feature selection approach for the diagnosis of Alzheimer's Disease", 2016 International Joint Conference on Neural Networks (IJCNN).

[7] M.RANGINI, Dr.G,WISELIN JIJI "Detection of Alzheimer's Disease Through Automated Hippocampal Segmentation", 2013 International Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMAC4s).

[8] Devvi Sarwinda , Alhadi Bustamam "Detection of Alzheimer's Disease Using Advanced Local Binary Pattern from Hippocampus and Whole Brain of MR Images", 2016 International Joint Conference on Neural Networks (IJCNN).

[9] S. Saraswathi et.al "Detection of onset of Alzheimer's Disease from MRI images Using a GA-ELM-PSO Classifier", 2013 Fourth International Workshop on Computational Intelligence in Medical Imaging (CIMI).

[10] Devvi Sarwinda, Aniati M. Arymurthy "Feature Selection Using Kernel PCA for Alzheimer's Disease Detection with 3D MR Images of Brain", 2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS).

[11] Mayank Agarwal, Javed Mostafa "Content-based Image Retrieval for Alzheimer's Disease Detection", 2011 9th International Workshop on ContentBased Multimedia Indexing (CBMI).

[12] Chetan Patil et.al "Early Detection of Alzheimer's Disease", 2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES).

[13] Lauge Sørensen et.al "Early detection of Alzheimer's disease using MRI hippocampal texture", US National Library of Medicine National Institute of Health, DOI:10.1002/hbm.23091,Epub 2015 Dec 21.

[14] RigelMahmood, BishadGhimire "Automatic Detection and Classification of Alzheimer's Disease from MRI Scans Using Principal Component Analysis and Artificial Neural Networks", 2013 20th International Conference on Systems, Signal and Image Processing (IWSSIP).

[15] M.Evanchalin,Sweety,G.Wiselin Jiji "Detection of Alzheimer Disease in Brain Images Using PSO and Decision Tree Approach", 2014 IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT).

[16] Arpita Raut, Vipul Dalal "A Survey of Recent Approaches for Detection of Alzheimer's Disease", International Journal of Engineering Research in Computer Science and Engineering (IJERCSE) Vol 4, Issue 2, February 2017.