Multistage Classification of Alzheimer's Disease

Neha Surendran[1], Ahammed Muneer K V[2]

M.Tech Scholar[1], Assistant Professor[2]
Department of Applied Electronics Engg., Govt. Engg. College, Kozhikode, Kerala, India.

Abstract: - **Alzheimer's disease is a type of dementia that destroys memory and other mental functions. During the progression of the disease certain proteins called plaques and tangles get deposited in hippocampus which is located in the temporal lobe of brain. The disease is not a normal part of aging and gets worsen over time. Medical imaging techniques like Magnetic Resonance Imaging (MRI), Computed Tomography (CT) and Positron Emission Tomography (PET) play significant role in the disease diagnosis. In this paper, we propose a method for classifying MRI into Normal Control (NC), Mild Cognitive Impairment (MCI) and Alzheimer's Disease(AD). An overall outline of the methodology includes textural feature extraction, feature reduction process and classification of the images into various stages. Classification has been performed with three classifiers namely Support Vector Machine (SVM), Artificial Neural Network (ANN) and k-Nearest Neighbours (k-NN).**

Keywords: **Alzheimer's Disease, Magnetic Resonance Imaging, Support Vector Machine, Artificial Neural Network, k-Nearest Neighbours.**

I. INTRODUCTION

Alzheimer's disease is a chronic neuro degenerative disease that usually starts slowly and worsens over time. It is the main cause of 60% - 70% of the cases of dementia. There are different stages of the disease like mild stage, moderate stage and crucial stage. Advanced medical imaging techniques like MRI, CT, PET etc. shows significant role in the diagnosis of the disease.AD typically destroys neurons in the brain areas involved in memory, including the entorhinal cortex and the hippocampus. Early symptom is difficulty in remembering recent events. Amyloid plaques, neurofibrillary tangles, synaptic loss and cell death are the striking features of Alzheimer's brain. More than 90% of the disease occur in people above age 60. Some people with memory problem may have MCI, a condition that may lead to AD. Major tools of the disease diagnosis includes analyzing medical history of the patient, a physical exam, and tests which measure memory, language skills and other abilities related to brain functioning. Neuro psychological tests such as Mini Mental State Examination (MMSE) are used for diagnosis as the screening test. Low MMSE score needs further evaluation such as brain imaging techniques.

Current diagnosis of AD is made by clinical, mental and neuro-physiological tests. Therefore, developing new approaches for early and specific recognition of Alzheimer's disease is of crucial importance. [1] used textural and morphological features for AD classification. The disease detection from the brain MRI can be carried out by extracting some relevant features of the diseased image. Feature values show variation for different stages of the disease. Machine learning is employed for the classification of given brain MRI into normal, MCI and AD stages.

A method for AD detection has been performed in [1] in which Haralick features, Gist features and morphological features are extracted for classification. They made use of voxel-based morphometry (VBM) for morphological feature extraction. Feature reduction is performed using svm rfe and PCA. They used ADNI database and classified the brain MRI using SVM.In [3] an automatic method for detecting AD patients using brain MRI is used. They combined VBM and SVM and detected AD mainly for clinical applications. Using VBM method they ex-tracted 20 features from the brain MRI of normal and diseased patients and reduced the dimension of features using PCA. Then classification was performed using SVM classifier. Results with PCA was slightly better than without using PCA. The accuracy of classifier was found proportional to the number of training samples. In [2], a new approach is developed based on mathematicaland image processing techniques. In order to categorizethe reduced features into various classes they employeda multiclass neural network classifier. The neural networkwas trained with 230 MRIs obtained from OASIS database.Results yielded an accuracy of 90% for AD detection.

A method of AD detection using brain SPECT image is performed in [4]. This Computer Aided Diagnosis system uses Empirical Mode Decomposition and Gaussian filters, intensity normalization, PCA feature extraction and an SVM Classifier. The method could improve the baseline Voxel-As-Feature (VAF) approach yielding up to 85.87% accuracy in separating AD and NC. For feature extraction process, they made use of adaptive image decomposition method.In [5] an application to detect AD from MRI is proposed which includes three sections for AD detection at different planes: frontal plane to extract the Hippocampus (H), Sagittal plane to analysis the Corpus Callosum (CC) and axial plane to work with the variation features of the Cortex (C). Their method of classification was based on SVM. Their system yielded an accuracy of 90.66% in the early diagnosis of the AD. In [6], presented a technique to do AD classification between healthy control subjects, amnesic mild cognitive impairment (a-MCI)subjects or AD subjects. Subject classification have been performed based on the functional connectivity scores of

resting state fMRI (rs-fMRI)brain scans. They used Gaussian process logistic regression mode for classification.

Many of the papers described above used VBM method for feature extraction and SVM for classification. There are papers which used ANN classifier for classification. This paper uses textural feature extraction and PCA feature reduction methods. Classification is performed using three classifiers, viz. SVM, ANN and k-NN. The detailed methodology is well explained in the following sections.

The rest of the report is organized as follows. Section II discusses the proposed method. Section III gives details of the implementation of the project. Section IV provides the results of the proposed method. Section V summarizes and concludes the method including future works.

II. PROPOSED SYSTEM

The general methodology involved in the image classification is depicted in the fig 1.

In general, image to be classified is preprocessed firstly and then features are extracted from the preprocessed image.

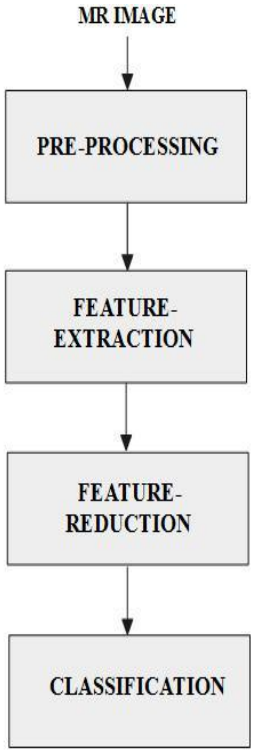

Figure 1: General block diagram of image classification

If the dimension of features is too high, most relevant features can be selected by any of the feature reduction mechanisms. And finally, image is classified into different classes based on the selected features. Fig 2 shows the detailed schematic diagram of the proposed method.

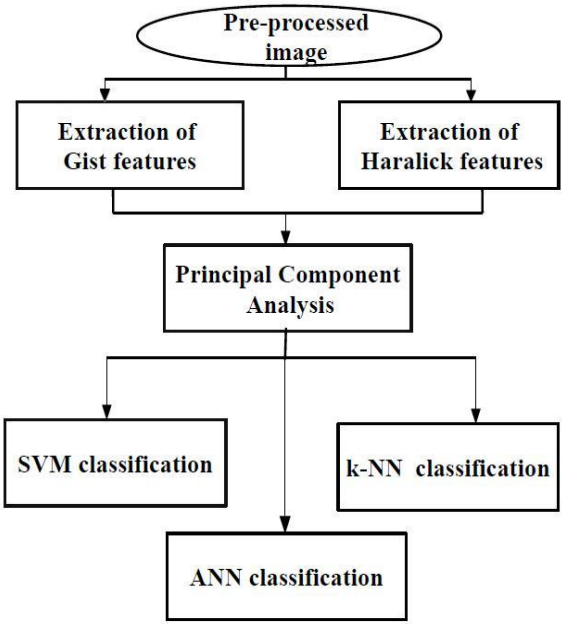

Figure 2: Block diagram of the proposed method

2.1 Pre-processing

The aim of pre-processing is to improve the characteristics of the image by suppressing distortions and enhance the properties as per the requirement.In our method, skull masking is performed in the pre-processing step.The removal of non-brain tissue from MRI brain images is called skull masking or skull stripping or brain extraction. There is need to remove skull portions from the brain MR images before feature extraction in order to reduce the computational complexity and numerical burden. Erosion is the main process behind skull masking. Erosion can be performed on original gray scale image or binary image.Brain extraction tools are available for skull stripping. But, we performed erosion of binary MRI and masked the eroded image with original image.

2.2 Feature extraction

Feature extraction is an inevitable step in the area of image processing. In this process, the most discriminating features are extracted from the raw data.A good feature set contains discriminating information, which can distinguish one image from others. It must be as robust as possible such as to generate comparable feature vectors for all the images belonging to same class and discriminating feature vectors for images in different classes.

2.3 Extraction of Haralick features

GLCM represents the distance and angular spatial Relationship of pixels of an image. Texture can be analyzed using Haralick features extracted by GLCM analysis. GLCM determines how often a pixel of a gray scale value i occurs adjacent to a pixel of the value j. Four angles can be considered for observing the pixel adjacency i.e., Θ= 0, 45,90

and 135 are used. Another parameter for creating GLCM is an offset value D, which defines pixel adjacency by certain distance. Figure 3 illustrates how to create GLCM from an image. After creating the GLCMs, it is possible to derive several statistics using different formula.

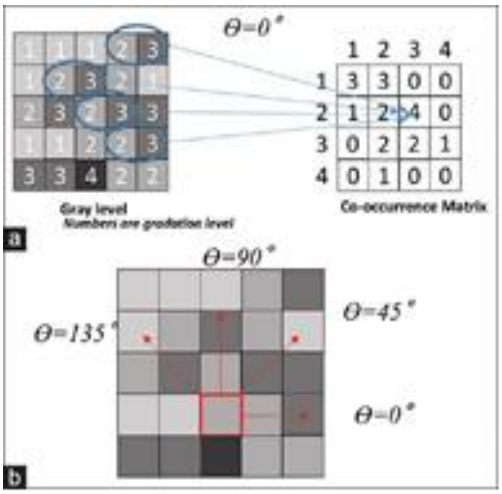

Figure 3: GLCM creation from the image

2.4 Extraction of Gist features

Gist features are global features which are extracted by the convolution of Gabor filter with the image. 2D Gabor filter is a Gaussian kernel function multiplied by a sinusoidal wave. Impulse response of Gabor filter is the product of sinusoidal function and Gaussian function. Gist features are global features which represents low dimensional representation of the image. Gist generates the gradient information of different parts of the image which provides a rough description of the image. A group of Gabor filters at different scales and orientation creates a Gabor filter bank. In computing Gist features, a Gabor filter bank with 32 Gabor filters at 4 scales and 8 orientations is created. Each image is convolved with 32 Gabor filters of the filter bank to produce 32 feature maps of the same size of the input image. Each feature map is then divided into 16 regions and averaged the feature values within each region. Finally 16 averaged values of all 32 feature maps are concatenated to compute 512 Gist descriptors. Thus each image has 512 Gist features.

2.5Feature selection

PCA is a well known feature reduction technique. Number of principal components can be selected less than original dimension of features by selecting the relevant features and omitting irrelevant features.

PCA Algorithm:

1. Input data matrix
2. Calculate mean
3. Calculate deviation from mean
4. Compute the co-variance matrix
5. Compute eigen values and eigen vectors of co-variance matrix
6. Rearrange the eigen values in the descending order
7. Arrange the eigen vectors in in the order of sorted eigen values.
8. Select L largest eigen values and corresponding eigen vectors.
9. Eigen vectors with highest eigen values are projected into a space.
10. Projection results in a vector represented by a fewer dimension (L < M) containing the essential coefficients.

2.6 Classifiers used

SVM classifier:

SVM is a supervised learning model which analyses the given data for classification or regression. In the case of classification, SVM finds an optimal decision plane which separates data into different classes. Basically SVM is a binary classifier which classifies given data into two classes. Marginal hyper plane is the plane through which support vectors pass through. By supervised learning, SVM tries to maximize the margin of separation between the marginal planes.

ANN classifier:

ANN is a kind of classifier based on supervised learning strategy, which is inspired by biological neural network. It is based on a collection of units called artificial neurons. These neurons are arranged in layers. The input, hidden, and output layers are different layers of the network that perform certain transformations on the input. The number of nodes or the number of neurons in the input layer is equal to the input dimension. The number of nodes in the output layer depends on the number of output classes. The number of hidden layers may be one or more, and the number of neurons in the hidden layer is usually chosen to be higher than the number of nodes in the input layer. Since ANN is a supervised learning model, it has some learning rules that modify connection weights based on the input patterns provided. More simply, when a neural network is initially presented with a pattern, it would be a random guess on it. Then it will see how far the actual output is, and make the appropriate adjustments to its connection weights.

k-NN classifier:

k-NN is a classifier where each pixel is classified in the same class as the training data with the closest intensity. Here the Euclidean distance, the difference d between the M descriptions of a sample, s and the description of a known texture, k is calculated. For M measurements of N known samples of textures and for O samples of each, will get an M-dimensional feature space that contains the N x O points. If we select the point in the feature space that is closest to the

current sample, then we can select the sample's nearest neighbor.

k-NN algorithm:

- Input the data
- Calculate the distance between test samples and all training samples and sort the distance vector in ascending order
- Choose a suitable value for k
- Select k samples which are closest to the test samples
- Test sample is allotted to the group or class which contains more number of nearest training samples

III. IMPLEMENTATION

The programming is performed in MATLAB (R2015a, 64-bit) from Mathworks, Inc.(Natick, MA; United States), with included Image Processing Toolbox. The operating system is windows 10 enterprise with Intel Core i3 processor.Database used in this project have been selected from the ADNI database provided by IDA(Image and Data Archive). IDA provide resources for searching, visualizing ad sharing diverse range of neuro science data. Demographic data of subjects in database is given in Table 1.

Table 1: Demographic data of subjects in database

Diagnosis	Number	Age	Gender	MMSE
NC	900	65-90	M,F	28-30
MCI	900	65-90	M,F	24-27
AD	900	65-90	M,F	20-23

IV. RESULTS AND ANALYSIS

T1 weighted axial MR images were skull stripped in order to remove extra meningial tissues of brain. Gist global features and Haralick textural features were extracted from the skull stripped image. To reduce the computational complexity and to select the relevant features, feature reduction is performed using PCA. Classifiers were trained using the reduced features and performance analysis of different classifiers were carried out. Figure 4 below shows the results of pre-processing. Original gray scale image is first converted to binary image. Then binary image is eroded. And finally, eroded image is masked with the original image to form the skull stripped image. We trained an SVM classifier to distinguish NC, MCI and AD. Classifier is trained with four different kernels: Linear kernel, Radial Basis Function(RBF)kernel, Polynomial and sigmoid kernel. Their comparison is given in Table 2.

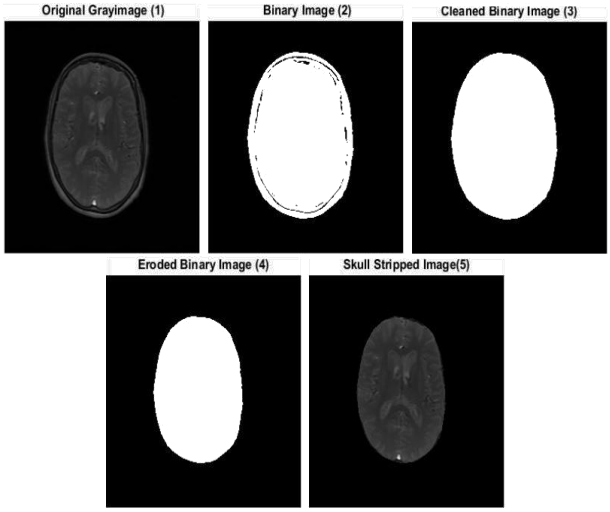

Figure 4: Results of skull stripping process

Table 2: Comparison of classification accuracy of SVM classifier with different kernels

Kernel	Classification accuracy(%)
RBF	88.518
Linear	86.667
Polynomial	81.4815
Sigmoid	50.1852

Table 3 shows the performance evaluation of SVM classifier with various values of C. While varying the value of C, highest accuracy was obtained at C=50.

Table 4 shows the variation in test accuracy of ANN by varying the number of hidden nodes. Performance analysis of k-NN classifier was analysed with various values of k and is given in Table 5. 'k' is a user defined function. In terms of accuracy, performance of classifier is better for k=40. In k-NN, distance from test samples to all other training samples was computed using three distance metrics. Difference in accuracy with the distance metrics is depicted in Table 6. It is clear from the table that performance is almost comparable, but when city block distance is used as thedistance metric k-NN performed well. Performance of classifier is analyzed with originally extracted features and with reduced features.There are a total of 608 features which combines Haralick features and Gist features. Classifierswere trained with 608 features and with 75%, 50% and 25% of the original features. For finalclassification 152 features were selected and classifier performance was analyzed. Table 7 showsthe final classification performance of the three classifiers with 152 features. It is clear from theTable 7 that, k-NN with city block distance metric outperformed other two classifiers.

Table 3: Performance of SVM with various values of C

Value of C	Accuracy (%)
1	77.777
2	79.259
3	79.444
4	80.582
5	81.851
10	85.452
15	86.114
20	86.148
25	85.612
30	87.626
40	88.153
50	88.518
60	88.4712

Table 4: Test accuracy of ANN with different number of hidden nodes

No.of hidden nodes	Test accuracy(%)
10	87
20	81.8
30	84.0
40	71.4
50	79.7
60	83.3
70	92.8
80	85.7
90	83.3
100	73.5
110	85.6
120	74.7

Table 5: Performance analysis of k-NN classifier with various values of k

Value of k	Accuracy(%)
5	91.1108
10	91.2560
15	91.6365
20	92.8940
25	93.5612
30	95.8546
35	95.3101
40	96.2962
45	96.1285
50	95.5463
55	96.1160
60	95.9824

Table 6: Classification accuracy of k-NN with different distance metrics

Distance metric	Classification accuracy(%)
Cityblock	96.2962
Correlation	95.2618
Euclidean	93.0627

Graph shown in figure 5 illustrates performance analysis of classifiers with and without feature reduction using PCA.

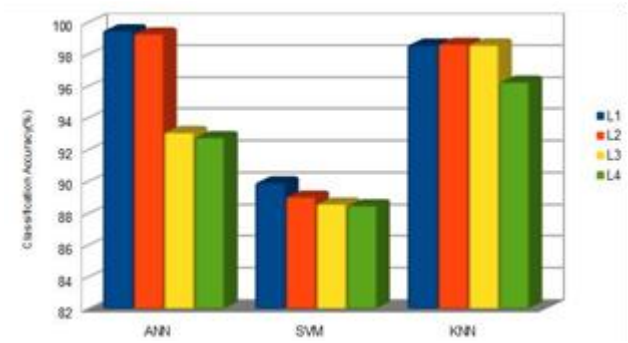

Figure 5: Performance comparison of classifiers with and without PCA

L1=608 features

L2=75% of total features

L3=50% of total features

L4=25% of total features

Table 7: Performance comparison of classifiers with reduced features

Classifier	Classification accuracy(%)
k-NN (Cityblock distance)	96.29
ANN	92.8
SVM (RBF kernel)	88.51

V. CONCLUSION

We have developed a method to compare the performance of SVM, ANN and k-NN classifiers for detecting AD. Brain MRI images are classified into three stages as NC,MCI and AD. SVM with RBF kernel yielded more accuracy than with other kernels. k-NN with cityblock distance metric provided more accuracy than with Euclidean distance. k-NN classifier with cityblock distance outperforms SVM and ANN with a

test accuracy of 96.29% for reduced features (152 features). ANN yielded second most accurate with reduced features. To reduce the numerical burden, we performed feature reduction technique on the extracted features. PCA reduced features lowered the classification accuracy to a slight extent, but computational complexity was reduced. Results show that Gist features perform well when compared to Haralick features.

For further analysis, segmentation can be performed before the extraction of features and thus improve the accuracy. Also MR images of different stages like Early MCI (EMCI), Late MCI (LMCI) etc. can be included in the dataset. Feature reduction can also be performed with Fisher Discriminant Analysis (FDA) and hence its performance can be compared with PCA.

BIBLIOGRAPHY

[1]. Yi Ding,Cong Zhang,Tian Lan,ZhiguangQin (2015). Classification of Alzheimer's Disease Based on the Combination of Morphometric Feature and Texture Feature,IEEE International Conference on Bioinformatics and Biomedicine.

[2]. RIgel Mahmood,Bishad Ghimire (2013).Automatic Detection and Classification of Alzheimer's Disease from MRI Scans Using Principal Component Analysis and Artificial Neural Network.

[3]. Jin Zhang, Bin Yan, Xin Huang, Pengfei Yang, Chengzhong Huang(2013),The diagnosis of Alzheimer's disease based on voxel-based morphometry and support vector machine,Fourth International Conference on Natural Computation.

[4]. A. Rojas, 1. M. Gorriz, 1. Ramirez, A. Gallix, I. A.Illan(1964), Empirical Mode Decomposition as a feature extraction method for Alzheimer's Disease Diagnosis.

[5]. Amira Ben Rabeh, Faouzi Benzarti, Hamid Amiri(2016),Diagnosis of Alzheimer's Diseases in Early Step Using SVM (Support Vector Machine),13th International Conference Computer Graphics, Imaging and Visualization

[6]. Edward Challis, Peter Hurley, Laura Serra, Marco Bozzali , Seb Oliver, Mara Cercig-nani(2015),Gaussian process classification of Alzheimer's disease and mild cognitive impairment from resting-state fMRI,NeuroImage 112 (2015) 232243

[7]. Jack CR, Petersen RC, Obrien PC, Tangalos EG. MR-based hippocampal volumetry in the diagnosis of Alzheimers disease. Neurology, 1992;42(1):183-8.

[8]. LONI. (2011). Retrieved November 5, 2011, from ADNI (Alzheimer's Disease Neuroimaging Initiative): adni.loni.ucla.edu

[9]. G. F. Busatto, B. S. Diniz, and M. V. Zanetti, "Voxel-based morphometry in alzheimer's disease." Expert Review of Neurotherapeutics, vol. 8, no. ll, pp. 169 1-l702, 2008.

[10]. El-Dahshan, E. A., Salem, A. M., Younis, T. H. (2009). A hybrid technique for automatic MRI brain images classification. Studia Univ. Babes-Bolyai, Informatica, 54 (1).

[11]. Xiaojing Long, Chris Wyatt, "An automatic unsupervised classification of MR images in Alzheimer's disease," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2010.

[12]. J. Y. Tou, Y. H. Tay, and P. Y. Lau, "Gabor filters and grey-level co-occurrence matrices in texture classification," in MMU International Symposium on InjiJrmation and Communications Technologies, 2007, pp. 197-202.

[13]. W.-Y. Wang, J.-T. Yu, Y. Liu, R.-H. Yin, H.-F. Wang, J. Wang, L. Tan, J. Radua, and L. Tan, "Voxel-based meta-analysis of grey matter changes in alzheimers disease," Translational neuro-degeneration, vol. 4, no. I, p. 6, 2015.

[14]. G. W. Jiji, G. E. Suji, and M. Rangini, "An intelligent technique for detecting alzheimer's disease based on brain structural changes and hippocampal shape," Computer Methods in Biomechanics and Biomedical Engineering: Imaging and Visualization, vol. 2, no. 2, pp. 12 1- 128,2014.

[15]. Chaplot, S., Patnaik, L., Jagannathan, N. R. (2006). Classification of magnetic resonance brain images using wavelets as input to support vector machine and neural network. Biomedical Signal Processing and Control, 1, 86-92.

[16]. Lee, J., Su, S., Huang, C., Wang, J. J., Xu, W., Wei, Y., et al. (2009). Combination of multiple features in support vector machine with principal component analsis in application for Alzheimer's disease diagnosis. Lecture Notes in Computer Science, 5863, 512-519.

[17]. Chan, C.,Lin, C. (2011). LIBSVM: a library for support vector machines. ACM Transactions on IntelligentSystems and Technology, 2 (3), 2:27:1-27:27.

[18]. Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehricy, S., Habert, M.O., Chupin,M., Benali, H., Colliot, O., et al., 2011. Automatic classification of patients with Alzheimers dis-ease from structural MRI: a comparison of ten methods using the ADNI database. Neuroimage 56, 766781.

[19]. D. Zhang, Y.Wang, L. Zhou, H. Yuan, and D. Shen, Multimodal classification of Alzheimers disease and mild cognitive impairment, NeuroImage, vol. 55, no. 3, pp. 856867, Apr. 2011.