Journal of Critical Reviews
ISSN- 2394-5125 Vol 6, Issue 6, 2019

# A HYBRID SEGMENTATION BASED MAJORITY VOTING AND CLASSIFICATION FRAMEWORK FOR ALZHEIMER DISEASE PREDICTION

**[1]I. Murali Krishna,[2] Dr. Challa Narsimham, [3]Dr. A. S. N. Chakravarthy.**

**[1]Research Scholar, JNTU, Kakinada, Andhra Pradesh, India**
imuralijntuk@gmail.com
**[2]Professor and Dean , Vignan's Institute of Information Technology, Vishakhapatnam, Andhra Pradesh, India.**
Narasimham_c@yahoo.com
**[3]Professor and HOD, Department of Computer Science and Engineering, University College of Engineering, JNTU, Vizianagaram, Andhra Pradesh, India**
drasnchakravarthy@yahoo.com

***Received: 23.11.2019*** ***Revised: 21.12.2019*** ***Accepted: 22.12.2019***

**Abstract:** Machine learning models play a vital role in the Alzheimer's disease (AD) prediction in neurodegenerative brain disease. Several high-dimensional, reliable classification methods have been implemented in recent years to automatically classify the Alzheimer's disease patterns. Since most conventional machine learning models such as random forest, random tree, neural network, multi-class SVM, etc. are hard to find the feature extraction process and difficult to detect essential features for classifying disease. However, as the number of features space increases, it is difficult to find the essential disease patterns on the training image datasets. Also, most of the conventional image prediction models have high false positive rate for disease classification. To minimize these issues, a hybrid segmentation-based disease prediction model is implemented on the Alzheimer disease database. In this work, a gaussian image features are used to filter the homogeneous and heterogeneous features for image classification process. Simulation results show that the present feature extraction-based segmentation and classification algorithm has high true positive, F-measure, recall and precision than the conventional models on Alzheimer database.

**Keywords:** ADNI feature selection, Alzheimer, classification model.


DOI: http://dx.doi.org/10.31838/jcr.06.06.57

## Introduction

Alzheimer's disease is most common disease due to (short-term) loss of memory and other behavioral changes caused by brain cell degeneration. "Alzheimer's disease" is often used in conjunction with "dementia," the two words are not synonymous. The learning algorithm analyzes training data and generates a model for every valid input image that should predict the correct label prediction. This requires the learning algorithm to predict the unknown data from the training data. The most common type of dementia is Alzheimer's disease. AD is not always an old-age disorder; many people younger than 65 will develop the disease as well. It is irreversible and it gradually destroys memory and mind, leading to memory and mental function decline. The precise diagnosis of Alzheimer's is difficult, as other causes of dementia can have the same symptoms. The diagnosis of AD is confirmed by severe cognitive deficiency and autopsy confirmation of histopathological brain changes. Memory issue, normal activities of daily life, normal general cognitive function; impaired memory for age are primary clinical characteristics of MCI. MCI participants showed similar memory loss in other areas but less damage than moderate AD patients. MCI subjects have reported less declines over time than patients with mild However, many people with MCI experience cognitive deficits and physical difficulties in conjunction with AD[1]. The heterogeneity of MCI subjects properly identified those MCI subjects that are likely to convert to AD, since they are a target group with the best bio-marker for early therapeutic intervention in AD. Although unbalanced knowledge is a problem for binary classification, it is even more difficult to classify it in multiclass. Machine learning is the process of identifying and analyzing the unknown hidden patterns and its relationships on large uncertain databases. Decision making system on different medical datasets. As the size of the medical databases increases, traditional machine learning models such as decision tree, SVM, neural networks, naïve bayes, fuzzy ensemble learning[1] etc become difficult to process the patterns due to noise, high dimensionality and non-relational instances in the medical databases. Also, the major challenge of the existing models include disease pattern discovery and the quality services. Feature selection and classification are the essential requirements for most of the medical disease pattern discovery models.

Generally, the SVM classification scheme is based upon the characteristics of statistical learning mechanism[2]. The classification of SVM classifier supports structural risk minimization to carry out the whole process of classification smoothly and effectively. Additionally, the SVM technique is quite efficient in case of small training data. It has many other applications such as, disease assessment, detection of exudates in digital data, Alzheimer's disease prediction, glioma recognition, and so on. Initially, two different statistical evaluation processes are carried out in order to verify whether variances are similar across healthy and diseased persons[3].Various research works have given emphasis on training systems along with respiratory data to find relevant and essential features for the process of diagnosis and prediction. All the traditional machine learning approaches can be categorized under two basic categories, those are: - linear machine learning approaches and non-linear machine learning approaches. Linear and non-linear machine learning approaches are analyzed and compared with each other. As logistic regression can be included under the category of linear and random forests can be included under the category of non-linear machine learning technique. A multi-scale filter bank is used in order to present the characteristics of local image texture and structure. Different efficient and effective classification schemes are implemented to train the system. All supervised techniques need manual annotated ROIs. This process is very much complex and expensive. There is another method in order to use weakly labeled medical images. A global image label is used in order to train an image classifier. When labeled ROIs are not present, this image label is propagated to its ROIs. The training of ROI classifier is performed in the traditional way[4]. The above technique is known as Simple MIL approach. From all scan reports of patients

having Alzheimer, a small group of ROIs are influenced. But in actual case, symptoms of Alzheimer can be found in various subjects those are not yet diagnosed with Alzheimer. It enhances the noise label in case of ROI classifier. The approach is implemented in order to handle learning. This proposed technique involves weakly labeled data and it is known as multiple instance learning technique. The prime objective of this classification scheme is to construct an efficient classifier for a group (bag) of feature vectors. These feature vectors are also known as instances. A bag is considered as positive, only when it contains at least a single positive instance. Pattern Recognition and Classification techniques give emphasis on the identification of masses and the classification of those masses into either benign or malignant lung tissue. The major objective of this feature selection and extraction phase is to select an optimum subset of features. Hence, the error rate must be minimized and performance must be enhanced. Additionally, it shows improved classification accuracy and reduced computation time.

**Related works**

Giraud*, et.al,* implemented a multi-level classification model to predict the ADNI disease variations [5]. In this paper, they considered chronic obstructive pulmonary disease. Alzheimer can be included under the category of lung disease. CT scans play important role in the diagnosis process of the above said disease. There are numbers of different Alzheimer approaches which are implemented in the diagnosis process of lung diseases. This method emphasizes on simple drawing movements in Alzheimer's disease. It is capable of differentiating healthy state from diseased state by simply drawing straight lines. There are certain other approaches which involve writing words, drawing Archimedes spiral, circles. White matter hyper-intensities normally require expert raters to do manual segmentation/classification on Fluid-attenuated Magnetic Resonance ADNI. It is difficult to perform consistent and accurate segmentation of white matter hyper intensities for a couple of reasons. Their patterns and texture are heterogeneous, and the borders between the intensities are not clear. The main problem to determine the border between the non-WMH and WMH tissue make it better to use the intra-rater and inter-rater agreements.

Meng*, et.al,* proposed a integrated automatic segmentation model on the CT-scan images [6]. This technique includes quantitative analysis of Alzheimer. Alzheimer usually affects lung tissues and it hampers the normal functions of lungs.

Huang et.al, proposed a segmentation approach to find the outliers in chest radiographs images [7]. In this research paper, the researchers have given emphasis on the characterization of coarse-level textural regions in case of chest radiographs. Here, a new cellular automation-oriented unsupervised segmentation approach is presented which is also known as unsupervised grow-cut (UGC). A region growing and model-based technique are implemented together in order to carry out the process of lung segmentation.

Tejani *et.al,* implemented a unsupervised learning algorithm to predict the severity level in Alzheimer disease [8]. In this paper, certain clinical decision support systems are thoroughly studied and analysed in order to prevent and diagnose Alzheimer. The empirical analysis of a representative sample tries to detect all crucial factors those have significant role in the diagnosis of Alzheimer.

Pang *,et.al,* developed a high segmentation and classification model to test the Alzheimer disease severity in disease rehabilitation [9]. GOLD classification scheme can be further divided into four phase classification. Retrospective analysis is carried out on 253 patients. Analysis of variance technique is usually implemented to identify the differences among all GOLD phases. Hence, the classification scheme is not at all beneficial to choose candidates for pulmonary rehabilitation.

Inglese al.[10] developed a feature selection model and classification model for disease prediction . Here, local binary pattern method and random forest approaches are used to classify the disease rate. Local Binary Pattern is computed on small medical disease datasets. They trained the decision tree using a rotated feature space hence proposed the rotation forest algorithm. In this method, samples from the main datasets are obtained. These samples form a new subset which is fed into a new feature space. Decision trees are used to train this new subset. The accuracy of this method in prostate ADNI diagnosis is impressive. Unfortunately, the originality of the data is lost which results in overfitting of data. If the number of selected characteristics is high, the accuracy of this system also becomes high.

Fast correlation-based filter method was used by Amoroso [11] in their proposed classification algorithm. The feature subsets are generated based on how these features correlate. The base classifier is the support vector machine which enables the algorithm to learn from the feature subsets presented to it. Voting is then done to determine the results. Sets of features are first created by subdividing the redundant features. This is in a bid to achieve diversity. This gives the classifier an opportunity to train from several subsets instead of one selected subset. This method is more accurate than bagging and the other discussed methods. It has the capability of comfortably dealing with various types of features. Most predictive algorithms don't use the FCBF because it may lead to instability of the algorithm developed.

In a proposal by Hajiesmaeili [12], classification of ADNI was done using neural networks. In their research, they allowed for several subsets of significant genes. Their algorithm had three stages before producing output. First, it has begun by selecting the significant gene subset. Then the base classier used was the neural networks. After the results were obtained, they were put together using the Bayesian methodology. A gene vector is used during the selection of gene for every class. The measure of similarities is used to choose between classes. Through this, the significant subset of genes can be selected. Subsets of significant genes are constructed basing on highest similarities. The results of the base are calculated and through the application of probability, the final results are obtained. In this approach, the originality of the features is preserved although it has no significant impact on performance.

The imbalanced data problem is relaxed in unsupervised self-organizing learning with support vector ranking as mentioned in (Gamberger et al. [13]). In this method variables are selected by the model adopted by support vector machines to deal with this problem. ESOM also known as Advanced Self-Organising Map is used to cluster the ranker features for unsupervised cluster based classification model. A Kolomogrov-Smirnov statistic based on decision tree method(K-S tree) (Kohlmann et al. [14]) is the latest method in which complex problem is divided into several easier sub problems, in that case imbalanced distribution becomes less daunting. This method is also used for feature selection removing the redundant ones. After division, a two-way resampling is employed to determine optimal sampling criteria and rebalanced data is used to incorporate into logistic regression models. If all the posterior probabilities are considered, simpleMIL will work impressively well. Some methods such as miSVM and EM-DD ignore some concepts but still work well. This data points out that a high density of disease patches, as well as a low density of normal patches, can be found in a region of the feature space. When you compare the performance of the other method to the bag based methods, their performances are lower. miSVM in noisy conditions performs less accuracy than miSVM with the averaging rule. The Gaussian texture features algorithm shows better performance as compared to intensity features. An efficient weighting technique which depends upon classifiers can distinguish among scans from various domains. This technique can definitely enhance the results of the traditional classification approaches.

**Proposed Framework**

In this work, a segmentation-based classification model is designed and implemented on the ADNI training image dataset. As shown in the figure 1, initially ADNI image data is taken as input .ADNI slices are taken from the URL[https://drive.google.com/file/d/0B5zQ7kc5shlgRG9laXMxcnlyMk0/view]. ADNI images are pre-processed and its features data is given as training data. These image slices are pre-processed using the non-linear optimization function and feature extraction method. Image segmentation is applied on the pre-processed image in order to classify the different types of ADNI images based on gender and disease severity.

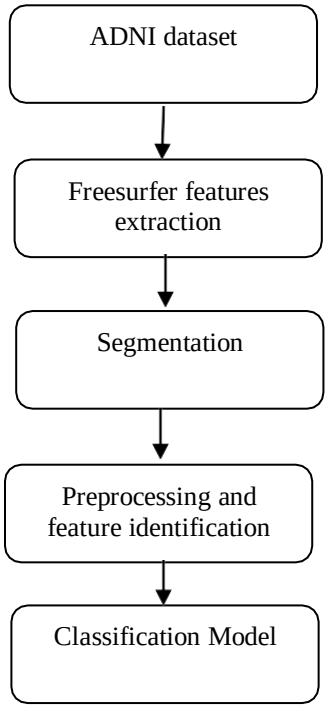



**Figure 1: Overview of proposed Framework**

These ensembles ranked features are used to predict the new Alzheimer test instances using the proposed ensemble multi-level classification model as shown in figure 2.

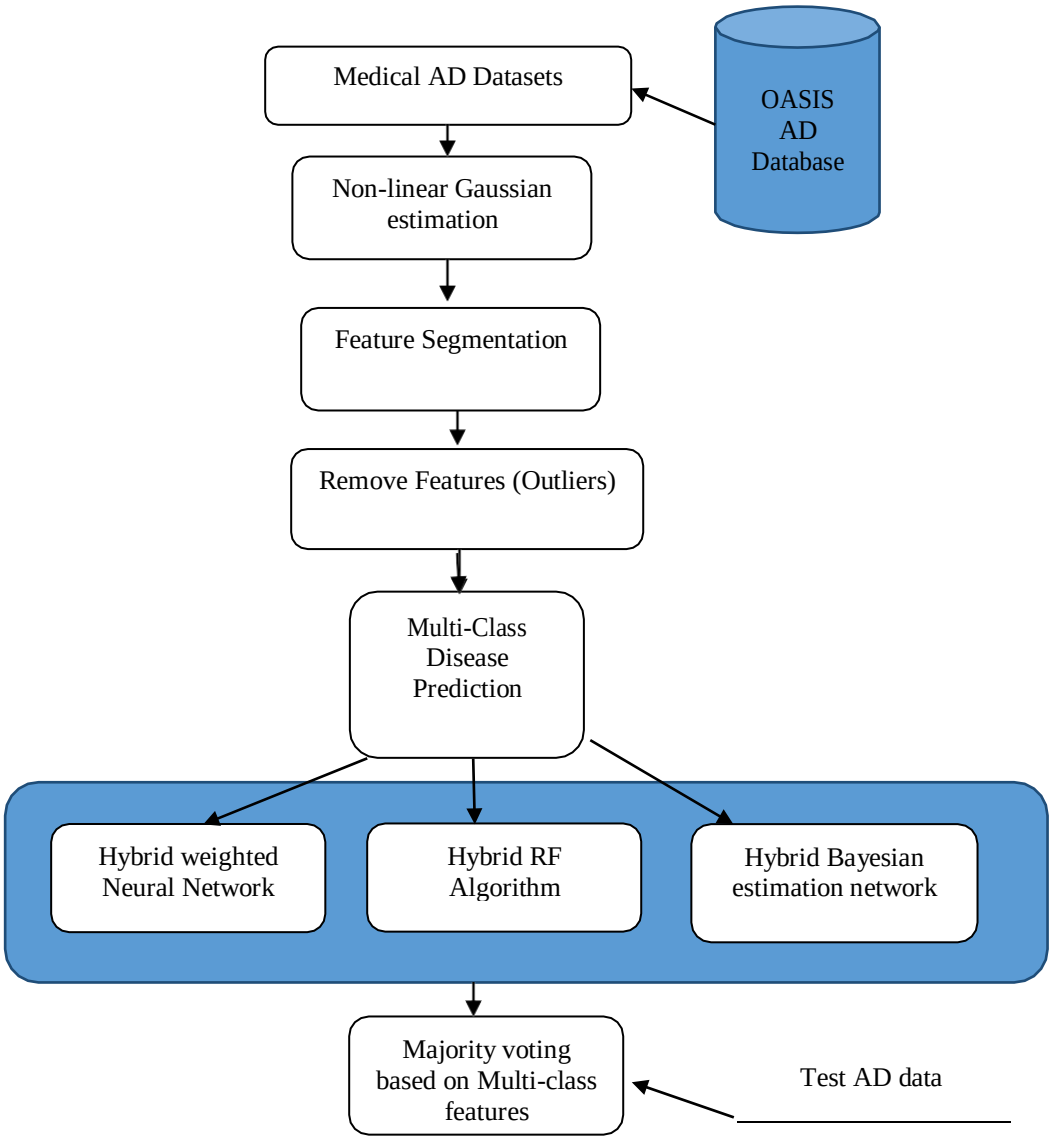



**Figure 2: Proposed Model**

**Algorithm1: ADNI data filtering with feature extraction**
Step 1: Input ADNI images.
Step 2: Filter the ADNI images using Subject ID.
Step 3: Non-linear gaussian estimation using histogram H is computed as:

$$NLG = \sum_{i=}^{N} \frac{1}{2\sqrt{\pi \log(H)}} . e^{-|I[i]-H|/2} \quad \text{----(1)}$$

Step 4: For each subject id in Image database
Do For each block in image id
Do Compute Block histogram intensities I.
Compute average block histogram value H using eq.(1).

Step 5: Apply Max correlation measure for efficient correlated feature extraction process as

$$MaxCorrelation = \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \max\left\{ \frac{\{i \times j\} \times P(i,j) - \{\mu_x \times \mu_y\}}{\sigma_x \times \sigma_y}, \sum_{i=0}^{G-1} \sum_{j=0}^{G-1} \{i-j\}^2 \times P(i,j) \right\}$$

Done
Step 5: Each 256 histogram intensity feature values of ROI and kernel density estimation are store in a csv file. Each subject id and class label of the ADNI either disease(1) or not disease (0) are stored in separate files as index and class along with given input features.

**Advanced Segmentation Model**
In this algorithm, labelled disorder data is filtered using the discretization and conditional probability measures. If the missing value of a feature in the feature-space is numerical then discretization operation is used to replace the missing value with the computed value.
Improved multi-level EM clustering algorithm**:**
Expectation maximization algorithm follows an iterative three step process:
Step 1: E-step: In this step, based on the model parameters, proposed model computes the probability of each data point as a segment.
In the M step, filling missing labels and find the model that improves similarity likelihood of the data.
Find the features using EM clusters as $S_{IF}$.
For each feature pair in F[]
Do

$$\eta_1 = \sum_{i=1}^{|S_r|} F_i[i].S(F_p, F_q)$$

$$\eta_2 = \sum_{i=1}^{|S_{IF}|} F_j[i].S(F_p, F_q)$$

$$MI = Mutual - Information(S_i, S_j)$$

$$Chebyshev = d(i, j) = \max_r | x(i, r) - x(j, r) |$$

$$S(F_p, F_q) = Max\{MI, Min\{Chebyshev(F_r(x, y)), Chebyshev(F(x, y))\}\}$$

Similar Segmented measure = $Max\{\eta_1, \eta_2\}$

Step 3: The probability estimator is used to improve the occurrence of disorder patterns in the given dataset.

**Optimization weights initialization to neural network**

It is the maximization of the correlation between the features, hybrid t-test and hybrid SNR ratio. This ranking measure is used to select the optimal binary class features in each cluster.

$$W[] = Max\{\frac{\mu_P - \mu_N}{\sqrt{\max\{\sigma_P^2/|P|, \sigma_N^2/|N|\}}}, \frac{|\mu_i - \mu_j|}{2(\sigma_P + \sigma_N)}\} \text{ ---- (3)}$$

Step 2: Defining the input, hidden and output layers to each iteration.

Step 3: To each hidden layer apply the logistic activation function for weights and error rate optimization.

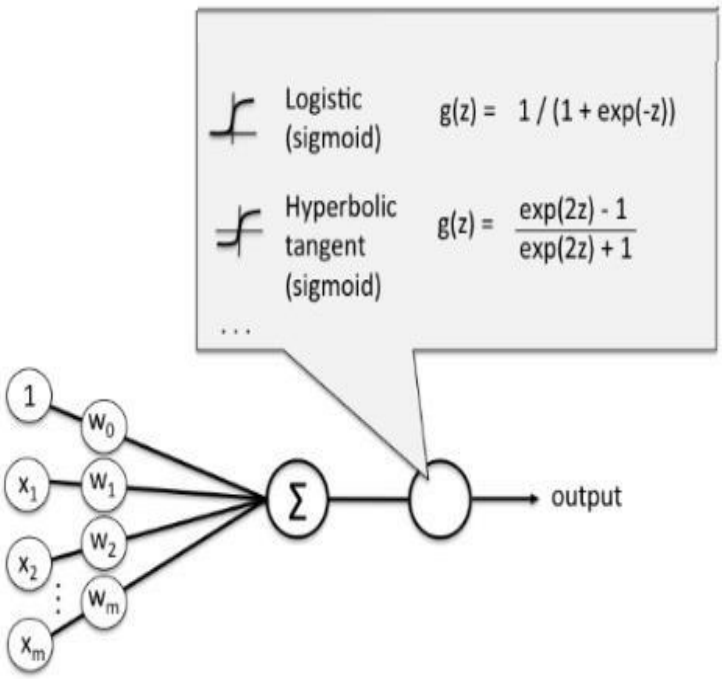



**Figure 3: Neural network structure**

**Random forest optimized entropy measure:**

Attribute ranking measure:

a) Enhanced entropy:

$$Pr = -Pr\ ob(D_i).\log(Prob(D_i))$$

$$Ent(D) = \sum_i Pr$$

Math.cbrt(entropy(data)*total*GHDSplitCriterion.computeHellinger (data))*Pr/ ( chiVal(data));

In the proposed boosting algorithm, a set of weak classifiers are used to improve the classification rate using the boosting mechanism. In this proposed approach, decision tree approach is used as weak classifier to train the samples in the Adaboost algorithm. A novel entropy and conditional entropy based decision trees are optimized using the modified attribute ranking measure for decision tree construction. In this algorithm, the classifier with low classification error rate is selected for instance prediction.

**Modified Bayesian score in Bayesian Network model:**

**Proposed Joint Bayesian Score**

$$\theta = \text{Conditional PriorProb}(s_i);$$

$$\phi = \text{Joint Pr ob}(D / s_i);$$

$$\text{PropBayesScore} = \log(\theta) + \log(\phi)$$

where

$$\text{Joint Pr ob}(D / s_i) = \prod_{i=0}^{n} \prod_{j=0}^{q_i} \frac{\Gamma(\sum_{k=1}^{r} \alpha_{ijk} \log(\alpha_{ijk}))}{\Gamma(\sum_{k=1}^{r} \alpha_{ijk} \log(\alpha_{ijk}) + \sum \log(N_{ij}))} \prod_{k=1}^{r} \frac{\Gamma(\sum_{k=1}^{r} \exp(\alpha_{ijk}) \log(\alpha_{ijk}) + \sum N_{ijk})}{\Gamma(\sum_{k=1}^{r} \alpha_{ijk} \log(\alpha_{ijk}))}$$

**Experimental results**

Experimental results for AD detection is performed on OASIS and ADNI databases which is an open access Imaging brain Magnetic Resonance images freely available to the scientific and research community [13]. The dataset comprises a collection of MR images along with computed data in CSV format from 416 patients between 18 to 96 years, where all the subjects are both men and women. 100 patients over 60 years have been clinically diagnosed with very mild to moderate Alzheimer 's disease (AD), rated using the Clinical Dementia Rating (CDR): very mild AD correspond to a CDR value of 0.5, while moderate AD has a CDR value of 2.0. OASIS data used in this study were retrieved from the Alzheimer's disease neuroimaging initiative (ADNI) database (http://adni.loni.usc.edu/)[14].

**Sample Alzheimer Data Slices**

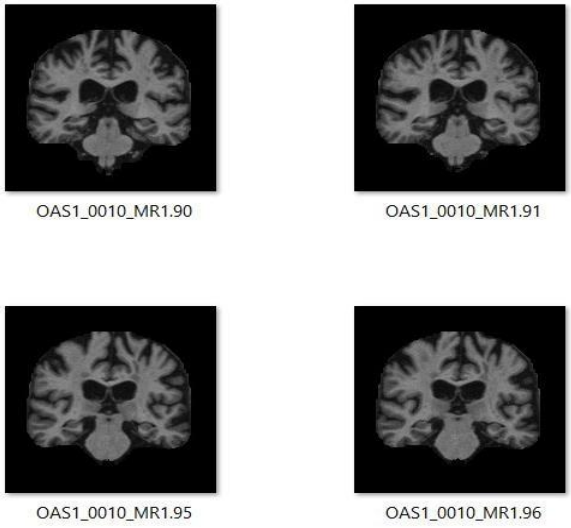



**Figure 4: Sample AD image slices of patient ID:10**

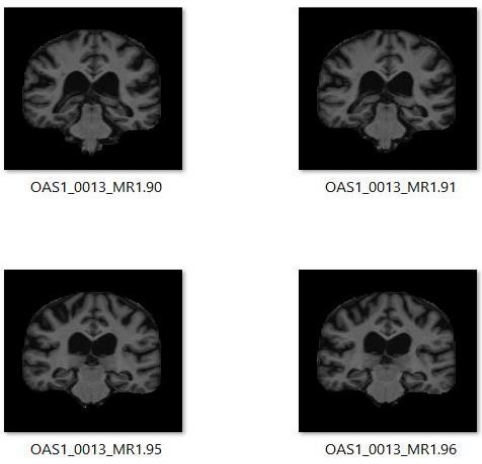
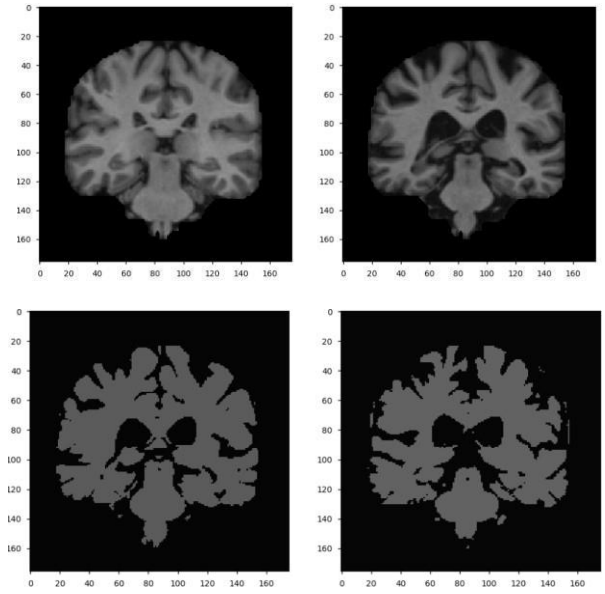



**Figure 5: Sample AD image slices of patient ID:13**

Figure 4 and 5 describes the sample image slices of patient ID 10 and 13. Here, different types of images are extracted using the freesurfer tool on the ADNI image database.

**Proposed Segmentation results on image slices**




```
<fs:region SegId="2" name="Left-Cerebral-White-Matter" hemisphere="left">
  <!--
    hidden_fields[fs_asegRegionAnalysis_region_id="433",regions_region_fs_asegRegionAna_id="OAS1_0011_MR1_ASEG"]
  -->
  <fs:NVoxels>205012.0</fs:NVoxels>
  <fs:Volume>205012.0</fs:Volume>
  <fs:normMean>108.2739</fs:normMean>
  <fs:normStdDev>9.6746</fs:normStdDev>
  <fs:normMin>42.0</fs:normMin>
  <fs:normMax>194.0</fs:normMax>
  <fs:normRange>152.0</fs:normRange>
</fs:region>
<fs:region SegId="3" name="Left-Cerebral-Cortex" hemisphere="left">
  <!--
    hidden_fields[fs_asegRegionAnalysis_region_id="434",regions_region_fs_asegRegionAna_id="OAS1_0011_MR1_ASEG"]
  -->
  <fs:NVoxels>195319.0</fs:NVoxels>
  <fs:Volume>195319.0</fs:Volume>
  <fs:normMean>69.2097</fs:normMean>
  <fs:normStdDev>10.8012</fs:normStdDev>
  <fs:normMin>12.0</fs:normMin>
  <fs:normMax>167.0</fs:normMax>
  <fs:normRange>155.0</fs:normRange>
</fs:region>
<fs:region SegId="4" name="Left-Lateral-Ventricle" hemisphere="left">
  <!--
    hidden_fields[fs_asegRegionAnalysis_region_id="435",regions_region_fs_asegRegionAna_id="OAS1_0011_MR1_ASEG"]
  -->
  <fs:NVoxels>9183.0</fs:NVoxels>
  <fs:Volume>9183.0</fs:Volume>
  <fs:normMean>25.7558</fs:normMean>
  <fs:normStdDev>13.4117</fs:normStdDev>
  <fs:normMin>9.0</fs:normMin>
  <fs:normMax>82.0</fs:normMax>
  <fs:normRange>73.0</fs:normRange>
</fs:region>
```


**Figure 6:Sample Segmented values for image OAS1_0011_MR1**

Figure 6, represents the segmented results on the image slices. Here, patient ID 11, is used to check the slices and segmentation result. In the above figure, different segmented regions and its properties are tabulated.


```
<fs:region SegId="2" name="Left-Cerebral-White-Matter" hemisphere="left">
  <!--
    hidden_fields[fs_asegRegionAnalysis_region_id="481",regions_region_fs_asegRegionAna_id="OAS1_0012_MR1_ASEG"]
  -->
  <fs:NVoxels>221514.0</fs:NVoxels>
  <fs:Volume>221514.0</fs:Volume>
  <fs:normMean>108.0963</fs:normMean>
  <fs:normStdDev>9.8403</fs:normStdDev>
  <fs:normMin>46.0</fs:normMin>
  <fs:normMax>137.0</fs:normMax>
  <fs:normRange>91.0</fs:normRange>
</fs:region>
<fs:region SegId="3" name="Left-Cerebral-Cortex" hemisphere="left">
  <!--
    hidden_fields[fs_asegRegionAnalysis_region_id="482",regions_region_fs_asegRegionAna_id="OAS1_0012_MR1_ASEG"]
  -->
  <fs:NVoxels>259052.0</fs:NVoxels>
  <fs:Volume>259052.0</fs:Volume>
  <fs:normMean>67.6321</fs:normMean>
  <fs:normStdDev>11.1187</fs:normStdDev>
  <fs:normMin>30.0</fs:normMin>
  <fs:normMax>133.0</fs:normMax>
  <fs:normRange>103.0</fs:normRange>
</fs:region>
<fs:region SegId="4" name="Left-Lateral-Ventricle" hemisphere="left">
  <!--
    hidden_fields[fs_asegRegionAnalysis_region_id="483",regions_region_fs_asegRegionAna_id="OAS1_0012_MR1_ASEG"]
  -->
  <fs:NVoxels>7812.0</fs:NVoxels>
  <fs:Volume>7812.0</fs:Volume>
  <fs:normMean>27.1687</fs:normMean>
  <fs:normStdDev>13.6426</fs:normStdDev>
  <fs:normMin>7.0</fs:normMin>
  <fs:normMax>80.0</fs:normMax>
  <fs:normRange>73.0</fs:normRange>
</fs:region>
```


**Figure 7: Segmented values for image OAS1_0012_MR1**

Figure 7, represents the segmented results on the image slices. Here, patient ID 12, is used to check the slices and segmentation result. In the above figure, different segmented regions and its properties are tabulated

.

**Decision Patterns:**

Initially, ensemble feature selection model is applied on the input Alzheimer dataset to find an essential weighted features from the feature space. These subset of weighted features in each iteration are used to generate decision patterns using the ensemble classification model. These generated patterns are shown below.

**[Group=Demented]: 146 ==> [Visit=1, Hand=R]: 64**
**[Hand=R, Group=Demented]: 146 ==> [Visit=1]: 64**
**[Group=Demented]: 146 ==> [Visit=1, MR Delay=0]: 64**
**[Group=Demented]: 146 ==> [MR Delay=0]: 64**
**[Group=Demented]: 146 ==> [Visit=1]: 64**
**[MMSE=29]: 91 ==> [Visit=1, MR Delay=0, Hand=R]: 40**
**[Hand=R, MMSE=29]: 91 ==> [Visit=1, MR Delay=0]: 40**
**[MMSE=29]: 91 ==> [MR Delay=0, Hand=R]: 40**
**[Hand=R, MMSE=29]: 91 ==> [MR Delay=0]: 40**
**[MMSE=29]: 91 ==> [Visit=1, Hand=R]: 40**
**[Hand=R, MMSE=29]: 91 ==> [Visit=1]: 40**
**[MMSE=29]: 91 ==> [Visit=1, MR Delay=0]: 40**
**[MMSE=29]: 91 ==> [MR Delay=0]: 40**
**[MMSE=29]: 91 ==> [Visit=1]: 40**
**[SES=1]: 88 ==> [M/F=F, Hand=R]: 39**
**[Hand=R, SES=1]: 88 ==> [M/F=F]: 39**
**[SES=1]: 88 ==> [M/F=F]: 39**
**[M/F=M]: 160 ==> [Hand=R, CDR=0.5]: 73**
**[M/F=M, Hand=R]: 160 ==> [CDR=0.5]: 73**
**[M/F=M]: 160 ==> [CDR=0.5]: 73**
**[EDUC=12]: 103 ==> [M/F=M, Hand=R]: 47**
**[Hand=R, EDUC=12]: 103 ==> [M/F=M]: 47**
**[EDUC=12]: 103 ==> [M/F=M]: 47**
**[Group=Nondemented]: 190 ==> [Hand=R, MMSE=30, CDR=0]: 88**
**[Hand=R, Group=Nondemented]: 190 ==> [MMSE=30, CDR=0]: 88**
**[Group=Nondemented]: 190 ==> [MMSE=30, CDR=0]: 88**
**[Group=Nondemented]: 190 ==> [Hand=R, MMSE=30]: 88**
**[Hand=R, Group=Nondemented]: 190 ==> [MMSE=30]: 88**
**[Group=Nondemented]: 190 ==> [MMSE=30]: 88**
**[SES=1]: 88 ==> [Hand=R, CDR=0, Group=Nondemented]: 41**
**[Hand=R, SES=1]: 88 ==> [CDR=0, Group=Nondemented]: 41**
**[SES=1]: 88 ==> [CDR=0, Group=Nondemented]: 41**
**[SES=1]: 88 ==> [Hand=R, Group=Nondemented]: 41**
**[Hand=R, SES=1]: 88 ==> [Group=Nondemented]: 41**
**[SES=1]: 88 ==> [Group=Nondemented]: 41**
**[EDUC=12]: 103 ==> [Hand=R, CDR=0.5]: 48**
**[Hand=R, EDUC=12]: 103 ==> [CDR=0.5]: 48**
**[EDUC=12]: 103 ==> [CDR=0.5]: 48**
**[CDR=0, Group=Nondemented]: 188 ==> [Hand=R, MMSE=30]: 88**
**[Hand=R, CDR=0, Group=Nondemented]: 188 ==> [MMSE=30]: 88**
**[CDR=0, Group=Nondemented]: 188 ==> [MMSE=30]: 88**
**[CDR=0]: 206 ==> [Hand=R, MMSE=30]: 97**
**[Hand=R, CDR=0]: 206 ==> [MMSE=30]: 97**
**[CDR=0]: 206 ==> [MMSE=30]: 97**
**[M/F=F, CDR=0]: 142 ==> [Hand=R, MMSE=30]: 67**
**[M/F=F, Hand=R, CDR=0]: 142 ==> [MMSE=30]: 67**
**[M/F=F, CDR=0]: 142 ==> [MMSE=30]: 67**
**[M/F=F, Group=Nondemented]: 129 ==> [Hand=R, MMSE=30, CDR=0]: 61**
**[M/F=F, Hand=R, Group=Nondemented]: 129 ==> [MMSE=30, CDR=0]: 61**
**[M/F=F, CDR=0, Group=Nondemented]: 129 ==> [Hand=R, MMSE=30]: 61**
**[M/F=F, Hand=R, CDR=0, Group=Nondemented]: 129 ==> [MMSE=30]: 61**
**[M/F=F, Group=Nondemented]: 129 ==> [MMSE=30, CDR=0]: 61**
**[M/F=F, CDR=0, Group=Nondemented]: 129 ==> [MMSE=30]: 61**
**[M/F=F, Group=Nondemented]: 129 ==> [Hand=R, MMSE=30]: 61**
**[M/F=F, Hand=R, Group=Nondemented]: 129 ==> [MMSE=30]: 61**
**[M/F=F, Group=Nondemented]: 129 ==> [MMSE=30]: 61**
**[Visit=2]: 144 ==> [Hand=R, CDR=0, Group=Nondemented]: 69**
**[Visit=2, Hand=R]: 144 ==> [CDR=0, Group=Nondemented]: 69**
**[Visit=2]: 144 ==> [CDR=0, Group=Nondemented]: 69**
**[Visit=1]: 150 ==> [MR Delay=0, Hand=R, CDR=0, Group=Nondemented]: 72**

```
[MR Delay=0]: 150 ==> [Visit=1, Hand=R, CDR=0, Group=Nondemented]: 72
[Visit=1, MR Delay=0]: 150 ==> [Hand=R, CDR=0, Group=Nondemented]: 72
[Visit=1, Hand=R]: 150 ==> [MR Delay=0, CDR=0, Group=Nondemented]: 72
[MR Delay=0, Hand=R]: 150 ==> [Visit=1, CDR=0, Group=Nondemented]: 72
[Visit=1, MR Delay=0, Hand=R]: 150 ==> [CDR=0, Group=Nondemented]: 72
[MR Delay=0]: 150 ==> [Hand=R, CDR=0, Group=Nondemented]: 72
[MR Delay=0, Hand=R]: 150 ==> [CDR=0, Group=Nondemented]: 72
[Visit=1]: 150 ==> [Hand=R, CDR=0, Group=Nondemented]: 72
[Visit=1, Hand=R]: 150 ==> [CDR=0, Group=Nondemented]: 72
[Visit=1]: 150 ==> [MR Delay=0, CDR=0, Group=Nondemented]: 72
[MR Delay=0]: 150 ==> [Visit=1, CDR=0, Group=Nondemented]: 72
[Visit=1, MR Delay=0]: 150 ==> [CDR=0, Group=Nondemented]: 72
[Visit=1]: 150 ==> [MR Delay=0, Hand=R, Group=Nondemented]: 72
```

Here, statistical classification measures are evaluated in the experimental results to find the true positive rate and kappa error. In the proposed model, 99.16% of accuracy is achieved on the high dimensional Alzheimer dataset using the proposed model.

```
Correctly_Classified Instances   4958        99.16  %
Incorrectly_Classified Instances   42         0.84  %
Kappa_statistic                          0.9829
Mean absolute error                      0.0139
  a b <-- classified as
 2792 8 | a = Normal
   34 2166 | b = Alzheimer
```

**Table 1: Comparison of ADNI image sets and its accuracy values of various classification models.**

| #Image Slices | NN | SVM | DT | Proposed Classifier |
|---|---|---|---|---|
| #50 | 0.934 | 0.9243 | 0.946 | 0.9894 |
| #100 | 0.938 | 0.953 | 0.967 | 0.978 |
| #150 | 0.924 | 0.956 | 0.963 | 0.984 |
| #200 | 0.954 | 0.975 | 0.9723 | 0.9934 |
| #250 | 0.954 | 0.968 | 0.981 | 0.996 |

Table 1, describes the performance comparison of traditional ADNI disease prediction models to the proposed segmented based classification model on various image slices. From the table, it is observed that the proposed classification accuracy is better than the conventional models.

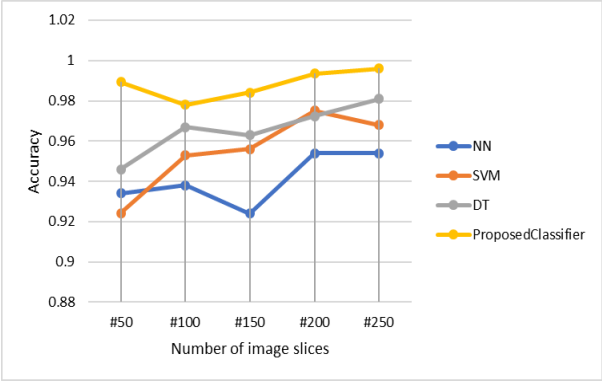



**Figure 8: Comparison of ADNI image sets and its accuracy values of various classification models.**

Figure 8, describes the performance comparison of traditional ADNI disease prediction models to the proposed segmented based classification model on various image slices. From the figure, it is observed that the proposed classification accuracy is better than the conventional models.

**Conclusion and future work**

In this work, a hybrid segmentation-based classification model is designed and implemented on the ADNI Alzheimer disease prediction. In this work, a hybrid feature extraction measure , segmentation and classification algorithms are implemented to each ADNI image slices. Experimental results proved that the proposed model has high computational efficiency than the traditional segmentation based classification models. In the future

work, a deep learning based biomarker ADNI classification model is designed on the large ADNI image slices.

**References**


1. L. I. Kuncheva, J. J. Rodriguez, Y. I. Syed, C. O. Phillips and K. E. Lewis, "Classifier Ensemble Methods for Diagnosing Alzheimer from Volatile Organic Compounds in Exhaled Air".
2. A. Corlateanu, N. Siafakas and V. Botnaru, "Defining Alzheimer: from simplistic approach to multilateral assessment of Alzheimer", "Curr Respir Care Rep (2012) 1:pp.177–182, 2012.
3. E. M. van Rikxoorta, P.A. de Jongb, O. M. Metsb, B. van Ginnekena, "Automatic classification of pulmonary function in Alzheimer patients using trachea analysis in chest CT scans, Proc. of SPIE Vol. 8315, 83150P · © 2012 SPIE 2012.
4. X. Artaechevarria A. Munoz-Barrutia C. Ortiz-de-Solorzano "Combination strategies in multi-atlas image segmentation: Application to brain MR data" IEEE Trans. Med. Imag. vol. 28 no. 8 pp. 1266-1277 Aug. 2009.
5. R. Giraud et al. "An optimized PatchMatch for multi-scale and multi-feature label fusion" NeuroImage vol. 124 pp. 770-782 Jan. 2016.
6. Y. Meng et al. "An improved label fusion approach with sparse patch-based representation for MRI brain image segmentation" Int. J. Imag. Syst. Technol. vol. 27 no. 1 pp. 23-32 2017.
7. M. Huang et al. "Brain extraction based on locally linear representation-based classification" NeuroImage vol. 92 pp. 322-339 May 2014.
8. A. Tejani R. Kouskouridas A. Doumanoglou D. Tang T.-K. Kim "Latent-class Hough forests for 6 DoF object pose estimation" IEEE Trans. Pattern Anal. Mach. Intell. vol. 40 no. 1 pp. 119-132 Jan. 2018.
9. S. Pang et al. "Hippocampus segmentation based on local linear mapping" Sci. Rep. vol. 7 Apr. 2017.
10. P. Inglese et al. "Multiple RF classifier for the hippocampus segmentation: Method and validation on EADC-ADNI harmonized hippocampal protocol" Physica Medica vol. 31 no. 8 pp. 1085-1091 2015.
11. N. Amoroso et al. "Hippocampal unified multi-atlas network (HUMAN): Protocol and scale validation of a novel segmentation tool" Phys. Med. Biol. vol. 60 no. 22 pp. 8851 2015.
12. Hajiesmaeili M. Amirfakhrian "A new approach to locate the hippocampus nest in brain MR images" Proc. 3rd Int. Conf. Pattern Anal. Image Anal. (IPRIA) pp. 140-145 Apr. 2017.
13. D. Gamberger B. Ženko A. Mitelpunkt N. Shachar N. Lavrač "Clusters of male and female Alzheimer's disease patients in the Alzheimer's disease neuroimaging initiative (ADNI) database" Brain Inform. vol. 3 no. 3 pp. 169-179 2016
14. P. Kohlmann, J. Strehlow, B. Jobst, S. Krass, J. Kuhnigk A. Anjorin , O. Sedlaczek, S. Ley, H. Kauczor and M. O. Wielpütz, "Automatic lung segmentation method for MRI-based lung perfusion studies of patients with chronic obstructive pulmonary disease".
15. http://adni.loni.usc.edu/