# Exploring of Classification Methods for Early Detection of Alzheimer's Disease

**G Stalin Babu, S N Tirumala Rao, R Rajeswara Rao**

*Abstract: Alzheimer's disease (AD) is a degenerative brain disease, a common health problem in elderly pesople which causes decline in memory and affected on nerve cells. AD has different stages like mild congestive impairment (MIC) (early stage), moderate (middle stage), severe (late stage) it is essential to detect AD early in MIC, so that pre-emptive measures can be taken. Significant research was carried out over the past century to diagnose and detect this disease early. The objective of the article is provide a review evaluation and critical analysis of the recent research work done to early diagnosis of AD using Machine Learning Strategies.*

*Keywords: Alzheimer disease, Classification, Machine Learning*

## I. INTRODUCTION

Alzheimer's disease (AD) is a brain degenerative disease. It is a effect on human thoughts and work. It is also caused by the other diseases and conditions. Like a decline in memory, effected on nerve cells. Neuronal damage in Alzheimer's disease consequently impacts parts of the brain that allow an person to perform general body activities such as walking, swallowing eating etc [22]. Dementia is a term describing a variety of disorders, and there are many various types of dementia, such as Alzheimer's disease, dementia with Lowy bodies and vascular dementia, and others. Dementia of the type (AD) of Alzheimer's disease is by far the most common cause of dementia. It is widely acknowledged that early detection of dementia can lead to more effective intervention and morbidity constraint Their research concludes that people who encounter MCI criteria can be differentiated from healthy control subjects and those with very mild Alzheimer's disease. It appears that this group of subjects is a clinical entity that can described for therapy measures.

To date, most mental disorder types have been diagnosed based on clinical observation. In specific, those include recognition of disorders that tend to cluster together, the timing of disease appearance and their tendency to cure, persist or be chronic. Currently there are no effective and efficient diagnostic tools not available to detect Alzheimer's.

**G Stalin Babu**, Dept of CSE, Aditya Institute of Technology and Management, Tekkali, A.P., India.
**S N Tirumala Rao**, Dept of CSE, Narasaraopeta Engg College. Narasaraopeta , A.P., India.
**R Rajeswara Rao**, Dept of CSE,JNTUK University College of Engineering, Vizianagaram, A.P., India

Boise et al. acknowledge prior research of low levels of clinical evaluation and diagnosis and conclude a possible explanation for this in the originality of the symptoms of dementia associated with the constraints encountered by physicians in their clinical practice. Through modern neuropsychological and cognitive testing, Alzheimer's disease is clinically diagnosed and other evidence of intellectual deficiency is verified [25]. According to Dubois et al., in addition to the clinical measures, the recommendations for the diagnosis of Alzheimer's disease accentuate the role that distinct biomarkers can perform. Like Computerized Tomography (CT), Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET), Single Photon Emission Computed tomography (SPECT) etc., Genetic risk profiles are analysed although they are costly and complex to measure to large amounts of evaluations. There is no specific test that can prove whether a individual has Alzheimer's disease or not [23]. While physicians always can evaluate whether an person has dementia, the actual cause can be difficult to determine. Alzheimer's disease immunology and risk parameters make up 50% to 60% of all dementia situations in the modern world today. Some of the other causes of dementia are vascular, lewd body and frontal lobe.The incidence of dementia rises with age, from 1% in 60-64 to 24-33% in those 85 years of age or elder. Around 24.3 million persons suffered from dementia in 2001, which is predicted to double every two decades as life expectancy also rises to 81.1 million in 2040. Western Europe, the USA and China are the nations and regions of the globe that are most impacted.By 2040, 55.7% of the total population impacted will be ho me to these nations and regions

## II. ALZHEIMER'S DIESES

*A. Alzheimer's early signs and symptoms [20]*
 ➢ Memory loss that disrupts everyday life
 ➢ Obstacles in planning or solving problems
 ➢ Difficulty in understanding visual pictures and partial relationships
 ➢ New issues with phrases in talking or writing
 ➢ Reduced or poorly judged
 ➢ losing the ability to retrace steps and Misplacing Things
 ➢ Withdrawal from job or social activities
 ➢ Difficulty in finishing familiar duties at home And at job or at leisure
 ➢ Changes in state of mind and character

## III. DATABASES

### A. ADNI Database:

ANDI [21] database having from known patient's data for Alzheimer's disease (AD) patients and the no disease patient's data and mild congestive implement (MIC) patients all information is there in the database. it mainly consider by the values are observed by the so many patients details of taken by the age, gender, education etc...In this database used to known the patient having Alzheimer's disease (AD) or not.

### B. OASIS Database:

The Open Access Series of Imaging Studies (OASIS) [19] is originates from Dr. R. Buckner which enables free availability of the brain MRI data sets available to the research community. The OASIS dataset contains two categories of data a) Cross-sectional MRI data b) Longitudinal MRI data of demented and non-demented elderly adults [25]. In this paper we use OASIS Dataset for studying the Alzheimer's dieses. This set is a cross-sectional compilation of 416 subjects between the ages of 18 and 96. Three or four person T1-weighted MRI scans are included for each subject acquired in single scan sessions. The subjects are all right-handed, including males and females alike. Clinically, 100 of the subjects over the era of 60 were diagnosed with Alzheimer's disease (AD) very mild to moderate.

## IV. MACHINE LEARNING STRATEGIES

"Machine learning is a data analysis technique that Automates the construction of analytical models. It is a division of artificial intelligence based on the concept that, with minimal human interference, systems can learn from data, identify patterns and make decisions."
ML is categorised into Supervised and Unsupervised these techniques are used to study of AD.
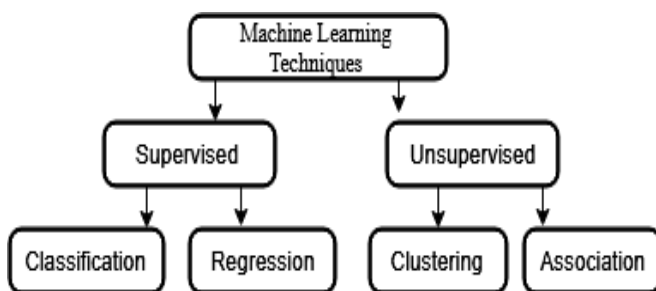


**Figure 1: Machine Learning Techniques**

### A. Supervised Learning:

The supervised learning concept is constructed on a supervisor, teaches the learning Model to Use the trainingdata set to reference unlabeled data to a class label [17], In this model provide the both input and desired output and mapped between them Mapped is based on the data features acquired.

### Support vector Machine:

"Support Vector Machine (SVM)" is a supervised algorithm for machine learning, used to solve the problems of classification and regression, the objective of SVM is to identify the ideal hyper plane for two classes between the training data. Kernel characteristics can be Use dto map the training set to a highlevel function space from input space.

### Logistic Regression:

Logistic Regression [17] is a supervised Machine Learning algorithm, it is used to solve the classification problems; LR Regression estimates the relationship between the dependent and two or more independent variables by estimating the probabilities using the Logistic function or also called Sigmoid function defined as

$$P(Y|X) = 1/e^{-(b_0 + b_1 X_1 + \cdots + b_n X_n)}$$

Where Xi is the independent variable
bi is The coefficients to be estimated in logistical regression by the fitting model.

In many AD research, LRC was used to justify discrimination against AD from other dementia types, early diagnosis, and AD progression.

### B. Unsupervised Learning:

In Unsupervised Learning separate objects in various groups, no supervisor is provided i.e. no training data is available to use. Unsupervised Learning validates a mapping function to discover masked patterns and comparisons from unlabeled input datasets

### K-Mean:

K-means clustering is unsupervised Machine learning strategy for grouping objects in a predefined set of K clusters. The process of K-Mean start with random center in data which are used as the beginning points for every cluster, and then tries to attach the nearest points to these canters. In several AD studies, K-means are used to cluster brain tissue, hippocampal and tumor segmentations.

## V. LITERATURE SURVEY ON CLASSIFICATION TECHNIQUES FOR ALZHEIMER'S DISEASE DETECTION

In Table I, Table II and Table III includes summary of various research articles of 2019,2018,2017 years for Classification techniques for early Alzheimer's disease Detection ,it cohere which type of classifier , data source ,type of features are extracted and outcome of the research.

**Table I: Summary of Classification approaches of Alzheimer's disease Diagnosis of 2017**

| Name of Author | Classifier | Modality | No of Images | Source | Features | Results |
|---|---|---|---|---|---|---|
| Gulhare, K. K 2017 [12] | DNN | MRI | AD+MCI+NC =150 | OASIS | Textural Features | Accuracy =96.6% |
| Alam, S 2017 [13] | multi-kernel learning support vector machine (SVM) | s-MRI | AD=89 MCI=102 | ANDI | volumetric feature | The detection rate from HCs and AD using volumetric features reaches 86.54 % and 75.12% respectively. |
| Raut, A 2017 [14] | ANN | MRI | AD+MCI+NC =416 | OASIS | Texture, Shape and Area features | CDR is 0 for no AD, 0.5 for mild AD, 1 for moderate AD and 2 for severe AD. |

**Table II: Summary of Classification approaches of Alzheimer's disease Diagnosis of 2018**

| Name of Author | Classifier | Modality | No of Images | Source | Features | Results |
|---|---|---|---|---|---|---|
| Ding, Y., 2018 [8] | Deep learning , CNN | 8F-FDG PET | AD=484 MCI=861 Non-AD/MCI=764 | ANDI | Structural features | 82% specificity |
| Islam, J.,2018 [9] | CNN | MRI | AD+MCI+NC =416 | OASIS | Texture FEATURES | Accuracy =0.93 positive predictive value (PPV) or precision=0.94 sensitivity or recall=0.93 harmonic mean of precision and sensitivity (f1-score)=0.92 |
| Mathew, N. A 2018 [10] | PNN ,SVM ,KNN | f MRI,PET and FMRI | AD+NC=600 | ADNI | Homogeneity, Correlation, Covariance, Contrast and edge direction | Classifier AUC SE SP<br>SVM 70% 35% 93%<br>KNN 77% 67% 86%<br>PNN 85% 83% 86% |
| Wang, S. H., 2018 [11] | convolution neural network (CNN) | MRI PET | AD Local Hospitals =70 AD =28 HC =98 | Local Hospitals and OASIS | NA | Accuracy =97.65 Sensitivity=97.96 Specificity=97.35 |

**Table III: Summary of Classification approaches of Alzheimer's disease Diagnosis of 2019**

| Name of Author | Classifier | Modality | No of Images | Source | Features | Results |
|---|---|---|---|---|---|---|
| Razavi, F 2019 [1] | Deep learning, multi-class classification 1.sparse filtering 2. SoftMax regression | MRI, FDG-PET, CSF | AD= 51 MCI=99 HC=52 | ADNI | GM, WM and CSF | Images Ad vs HC MCI and HC<br>MRI 0.901 0.872<br>PET 0.872 0.819<br>CSF 0.819 0.857<br>CONCAT 0.953 0.912<br>MK-SVM 0.983 0.912 |
| Varatharajah, Y 2019 [2] | Linear SVC (+ RBF kernel SVC, mixed kernel SVC, GLM) | CSF (all), MRI, FDG-PET, Amyloid PET, ApoE | MCI-P =35 MCI-NP=96 | ADNI | Information based feature | Classifier AUC SE SP<br>MKL-Linear 0.90 0.86 0.78<br>SVM-Linear 0.93 0.9 0.77<br>GLM-Elastic Net 0.92 0.91 0.76 |
| Lee, G 2019 [3] | Deep learning, multimodal recurrent neural network | Cognitive score, MRI, CSF biomarker, demographic data | MCI-C =307 MCI-NC =558 CN =415 AD =338 | ADNI | Single modal feature extraction and concatenation-based integration | Accuracy =0.81 Sensitivity=0.84 Specificity=0.80 |
| Mofrad, R. B 2019 [4] | Classification and Regression Tree (CART) | CSF (all), ApoE genotyping | AD=1004 MCI=363 SCD=442 | JPND | Regression tree allowed for learning of cut-off values that established the relative importance of the markers | Accuracy = 86% Sensitivity=86% Specificity=87% |
| Xu, L., 2019 [5] | Random Forest (RF) | MRI | positive samples $S^+$ =279 negative samples $S^-$ =1,743 | UniProt | Protein information | Accuracy =85.5% |
| Ge, C., 2019 [6] | 3D multi-scale convolutional networks | MRI PET | AD=600 NC=598 | ADNI | GM, WM and CSF | Accuracy=94.47% |
| Neffati, S., 2019 [7] | Multiclass SVM and DKPCA | MRI | AD+MCI+NC=416 | OASIS | NA | Accuracy =92.5% Specificity=95% Sensitivity= 88% |

## VI. PROCEDURE FOR CLASSIFICATION

The detail procedure of Classification of Brain MRI Images using machine Learning Techniques given in the below figure 2.

In first stage is Pre processing, it is an improvement of image data that suppresses the redundant and Unobtrusive distortions from image or enhance image features, different image pre processing techniques like linearization, Resizing, Thresholding, Normalization used in Study of AD.

In stage2 perform the image segmentation is a process of dividing image into various parts, the objective is change representation of image and faster to analyze, Segmentation of images used in numerous AD research, Sukanta Ghosh et al [15] Proposed a Fuzzy logic-based segmentation algorithm to solve the object Border detection of brain images, its topology-based algorithm is implemented with a novel Fuzzy rule-base and FIS targeted at detecting strong and weak edges of brain MRI images. The algorithm is helpful to early detection of AD by accurately identifying the of objects boundaries and tracing of abnormal tissue losses. The described fuzzy algorithm can be associated with a deep learning technique.
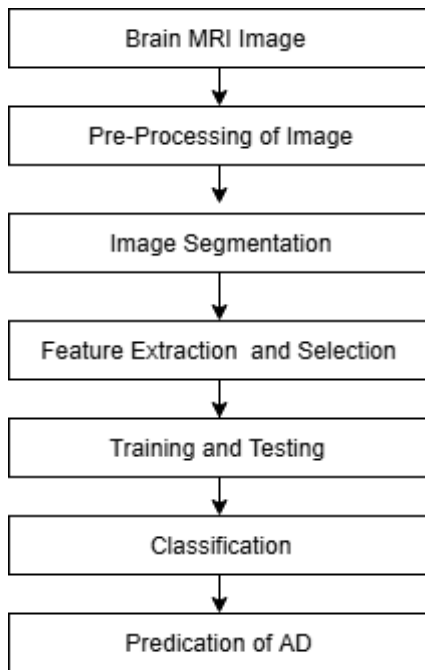


**Figure2: Classification process**

Yeliz Karaca et.al [16] implemented "Automatic RGB Color Segment Score Algorithm" It is an decision-making method that can be helpful for the diagnosis of brain MRI by decision rules. The technique implemented in 3 stages in stage 1 is Pre processing, to extract the brain from the MRI on a gray scale of AD, perform dilation and erosion morphological Operations. Stage2 is RGB segmentation stage, RGB values for color segmentation of threshold value (μi > 15 pixels) has been defined using Otsu's technique through multi-level thresholding. Stage3 Post-processing stage: By implementing the CART algorithm, the decision rules are acquired dilation and erosion was chosen from the morphological image processing techniques. Stage 3 is One of the significant step in the AD diagnosis is feature extraction and Selection. It extracts specific features from pre processed image of different abnormal categories [ 18 ] , It's technique of decrease the original data by evaluating certain important characteristics of image. Different types of features used to diagnosis of AD like Structural features [8], wavelet based, Texture Features [9, 12], volumetric feature [13], histogram based, color based, shape based, GLCM based from Brain MRI. In the classification stage it try to classify image based on their properties and assign a class label for the feature set. The classifiers into two categorized one is Binary classifiers ,other is Multi-Class classifiers ,In Binary classifier Classification with only 2 distinct classes or with 2 pre defined possible outcomes either presence or absence of abnormalities of brain image, Multi-Class classifiers is Classification with more than two distinct classes. In the Study of AD various classification methods are use like Support Vector Machine [7, 10, 13], convolution neural network [11], DNN [12], ANN [14], Decision tree, Naive Bayes K-nearest neighbor, Fuzzy methods are so useful to early detection of Alzheimer's detection.

The data set is divided into parts that are train and test data sets. Training data is used to calculate the patterns inherent the same. Testing is performed, and then observed if the patterns are in spot. For training and testing, similar data are required, generally extracted from same data set.

In the end, we have a anticipated disease based on the distinct algorithms with a final result of a disease predicted depending with certain symptoms of input.

### A. Evaluation Metrics:

Evaluation metrics plays a significant role to discriminate among various machine learning models. We use Accuracy, Sensitivity and Specificity metrics to evaluate the performance of the Classifier.

#### Accuracy:

Accuracy is one of the best metric to evaluating the performance of any classification method.

"*Accuracy is the ratio of number of correct predictions to the total number of input samples.*"

By using the below formula we will calculate the Accuracy of any classifier.

$$Accuracy = \frac{TruePositives + Flase\ Negatives}{Total\ No.of\ Samples}$$

#### Sensitivity:

Sensitivity (also called the True Positive Rate) is defined as the proportion of actual positives, which got predicted as the Positive, by considering all positive data points.

$$Sensitivity = \frac{True\ Positive}{True\ Positive + false\ Negative}$$

**Specificity**:

Specificity is defined is the proportion of true negative data points that are correctly considered as negative, by considering all negative data points.

$$Specificity = \frac{True\ Negative}{True\ Negative\ +\ False\ Positive}$$

**Results:**

In the table horizontally all the classifiers are mentioned and vertically all the performance Metrics mentioned on the basis of these parameters we can

Conclude the best approach to identify the Alzheimer's disease. All the above mentioned parameters in the table are derived by using WEKA tool. In the below graph we took classification methods in the X-Axis, it Represents Accuracy, Sensitivity and Specificity of each classifier, the vales of Accuracy, Sensitivity and Specificity can be measured based the above said formula. Based on the observations the accuracy of SVM is 97.6 % .When Compare to the remaining classifiers SVM gives best performance

**Table IV: performance metrics of various classifiers**

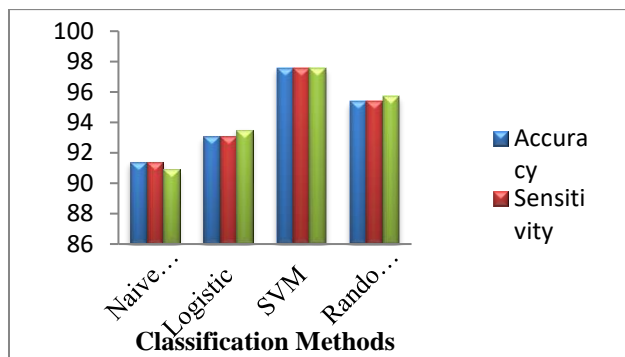| CLASSIFIER | TP RATE | FP RATE | PRECISION | RECALL | F-MEASURE | MCC | ROC AREA | PRC AREA |
|---|---|---|---|---|---|---|---|---|
| NAIVE BAYES | 0.914 | 0.053 | 0.909 | 0.914 | 0.911 | 0.868 | 0.978 | 0.944 |
| LOGISTIC | 0.93 | 0.06 | 0.935 | 0.93 | 0.916 | 0.89 | 0.986 | 0.973 |
| SVM | 0.976 | 0.021 | 0.976 | 0.976 | 0.975 | 0.961 | 0.979 | 0.958 |
| RANDOM FOREST | 0.954 | 0.04 | 0.957 | 0.954 | 0.948 | 0.93 | 0.999 | 0.995 |



**Figure 3: Comparison of various Classifiers**

**CONCLUSION:**

This article provided an extensive overview of Predication of Alzheimer's disease by using different machine learning based methods, also provided the process of classification of Brain images and a summary of the results acquired by various researchers to predict Alzheimer's disease. As recognized from the literature research, much research has been done in early detection of Alzheimer's disease, but the crucial need remains to identify Appropriate features that could detect Alzheimer are at an early stage .This article also provide various classifier and its performance evaluation on different metrics and finally concluded that SVM gives better results than other classifier.

Future research encompasses extracting reasonable set of features for early identification of Alzheimer's disease and also to decrease the insignificant and redundant conventional feature sets to enhance Alzheimer's methods effectively.
.

**REFERENCES:**

1. Razavi, F., Tarokh, M. J., & Alborzi, M. (2019). An intelligent Alzheimer's disease diagnosis method using unsupervised feature learning. *Journal of Big Data*, *6*(1), 32.
2. Varatharajah, Y., Ramanan, V. K., Iyer, R., & Vemuri, P. (2019). Predicting Short-term MCI-to-AD Progression Using Imaging, CSF, Genetic Factors, Cognitive Resilience, and Demographics. *Scientific reports*, *9*(1), 2235.
3. Lee, G., Nho, K., Kang, B., Sohn, K. A., & Kim, D. (2019). Predicting Alzheimer's disease progression using multi-modal deep learning approach. *Scientific reports*, *9*(1), 1952.
4. Mofrad, R. B., Schoonenboom, N. S., Tijms, B. M., Scheltens, P., Visser, P. J., van der Flier, W. M., & Teunissen, C. E. (2019). Decision tree supports the interpretation of CSF biomarkers in Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *11*, 1-9.
5. Xu, L., Liang, G., Liao, C., Chen, G. D., & Chang, C. C. (2019). k-Skip-n-Gram-RF: A Random Forest Based Method for Alzheimer's Disease Protein Identification. *Frontiers in genetics*, *10*.
6. Ge, C., Qu, Q., Gu, I. Y. H., & Jakola, A. S. (2019). Multi-stream multi-scale deep convolutional networks for Alzheimer's disease detection using MR images. *Neurocomputing*, *350*, 60-69.

*Retrieval Number F8215088619/2019©BEIESP*
*DOI: 10.35940/ijeat.F8215.088619*
*Journal Website: www.ijeat.org*

5210

*Published By:*
*Blue Eyes Intelligence Engineering*
*& Sciences Publication*

7. Neffati, S., Ben Abdellafou, K., Jaffel, I., Taouali, O., & Bouzrara, K. (2019). An improved machine learning technique based on downsized KPCA for Alzheimer's disease classification. *International Journal of Imaging Systems and Technology*, 29(2), 121-131.

8. Ding, Y., Sohn, J. H., Kawczynski, M. G., Trivedi, H., Harnish, R., Jenkins, N. W., ... & Behr, S. C. (2018). A Deep learning model to predict a diagnosis of alzheimer disease by using 18F-FDG PET of the brain. *Radiology*, 290(2), 456-464.

9. Islam, J., & Zhang, Y. (2018). Brain MRI analysis for Alzheimer's disease diagnosis using an ensemble system of deep convolutional neural networks. *Brain informatics*, 5(2), 2.

10. Mathew, N. A., Vivek, R. S., & Anurenjan, P. R. (2018, July). Early Diagnosis of Alzheimer's Disease from MRI Images Using PNN. In *2018 International CET Conference on Control, Communication, and Computing (IC4)* (pp. 161-164). IEEE.

11. Wang, S. H., Phillips, P., Sui, Y., Liu, B., Yang, M., & Cheng, H. (2018). Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. *Journal of medical systems*, 42(5), 85.

12. Gulhare, K. K., Shukla, S. P., & Sharma, L. K. Deep Neural Network Classification method to Alzheimer" s Disease Detection. International Journals of Advanced Research in Computer Science and Software Engineering ISSN: 2277-128X (Volume-7, Issue-6)

13. Alam, S., Kwon, G. R., & Alzheimer's Disease Neuroimaging Initiative. (2017). Alzheimer disease classification using KPCA, LDA, and multi-kernel learning SVM. *International Journal of Imaging Systems and Technology*, 27(2), 133-143.

14. Raut, A., & Dalal, V. (2017, July). A machine learning based approach for detection of alzheimer's disease using analysis of hippocampus region from MRI scan. In *2017 International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 236-242). IEEE.

15. Sukanta Ghosh , Abhijit ChandraRajani, K. Mudi," A novel fuzzy pixel intensity correlation based segmentation algorithm for early detection of Alzheimer's disease", Multimedia Tools and Applications,2018

16. Yeliz Karaca , Majaz Moonis, Abul Hasan Siddiqi, and Ba¸sar Turan,"Gini Based Learning for the Classification of Alzheimer's Disease and Features Identification with Automatic RGB Segmentation Algorithm.": ICCSA 2018, LNCS 10961, pp. 92–106, 2018

17. Mirzaei, G., Adeli, A., & Adeli, H. (2016). Imaging and machine learning techniques for diagnosis of Alzheimer's disease. Reviews in the Neurosciences, 27(8), 857-870.

18. Mamata Lohar., Rashmi Patange A Survey on Classification Methods of Brain MRI for Alzheimer's Disease. International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Vol. 7 Issue 05, May-2018

19. https://www.oasis-brains.org/

20. https://www.alz.org/alzheimersdementia/10_signs

21. http://adni.loni.usc.edu/

22. Alzheimer's, A. (2015). 2015 Alzheimer's disease facts and figures. Alzheimer's & dementia: the journal of the Alzheimer's Association, 11(3), 332.

23. Alzheimer's Association. (2017). 2017 Alzheimer's disease facts and figures. Alzheimer's & Dementia, 13(4), 325-373.

24. Antila, K., Lötjönen, J., Thurfjell, L., Laine, J., Massimini, M., Rueckert, D., ... & Hviid Simonsen, A. (2013). The PredictAD project: development of novel biomarkers and analysis software for early diagnosis of the Alzheimer's disease. Interface Focus, 3(2), 20120072.

25. Bansal, D., Chhikara, R., Khanna, K., & Gupta, P. (2018). Comparative analysis of various machine learning algorithms for detecting dementia. Procedia computer science, 132, 1497-1502.

## AUTHORS PROFILE

**G Stalin Babu ,** Research Scholar at JNTUK Kakinada in Department of CSE ,working as Assistant Professor in Aditya institute of technology and management , His area of research are image processing.

**Dr. S.N. Tirumala Rao,** M.Tech ., Ph.D . He is working as Professor and Head of the Department of CSE, NarasaraoPeta Engineering College, Narasaraopet, Guntur(AP)-India. His Areas of research are Data Mining, Multi-Core and Parallel Programming and Computer Networks. He Published more than 20 National and International Journals and organized several National and International Conferences. Books Published: Advanced Unix Programming HI-Tech Publisher.

**Dr.R.Rajeswara Rao,** Professor, Dept of CSE, JNTUK University college of Engg. Vizianagram(AP), India. His Areas of Interest: Speech Processing, Pattern Recognition, and Cloud Computing. He Published papers in several National and International Journals and also organized several National and International Conferences.