Alzheimer's Disease Detection Using Machine Learning

[1]Krishna Mundlik, Student, ICEM, Pune, India, krishnamundlik77@gmail.com
[2]Shubham Kolhe, Student, ICEM, Pune, India, shubhamkolhe962@gmail.com
[3]Prathmesh Vaste , Student, ICEM, Pune, India, prathmeshvaste9@gmail.com
[4]Akash Jagadale, Student, ICEM, Pune, India, jagadaleakash23@gmail.com

Abstract: Alzheimer disease is one of the most common and fastest growing neuro degenerative diseases in the western countries. Development of different biomarkers tools are key issues for diagnosis of Alzheimer disease and its progression. Prediction of cognitive performance of subjects from EEG and identification of relevant biomarkers are some of the research problems. EEG signal analysis can be well suited for automated diagnosis of Alzheimer's disease. Although, EEG based techniques are helpful in screening of Alzheimer and dementia; still there is a scope of improvement in terms of diagnostic accuracy, sensitivity and specificity. Thus, many issues are still left out in field of Alzheimer diagnosis using EEG signals related to the choice of features which can help in distinguishing the two or more subjects. This paper focuses on new features for diagnosis of Alzheimer's disease using EEG signals with effective increase in diagnostic accuracy. The use of new complexity based features is proposed in this paper which increases the diagnostic accuracy and helps in early Alzheimer's diagnosis

Keywords — Machine Learning, CNN Algorithm, MRI Images

I INTRODUCTION

Machine learning models are now widely used in medical diagnosis [16–19]. This paper compares different machine learning performances to diagnose Alzheimer's syndrome. Alzheimer's syndrome is an inherited, irreversible brain condition that steadily affects the ability to perform the necessary things, memory, and reasoning skills. A massive proportion of neurons stop working in Alzheimer's disease, losing synaptic connections. Alzheimer's diseases are infrequent in people aged between their 30s and mid-60s. Symptoms can include a shift in sleep habits, depression, anxiety, and difficulties doing basic tasks such as reading or writing and aggressive actions, and poor decision-making also happened in Alzheimer's disease. Alzheimer's disease and initial changes in the brain begin 10–20 years before the onset of symptoms [24].It progressively leads to memory damage and decreases thinking abilities [25].The leading cause of this disease is Dementia.

II LITERATURE SURVEY

a) Use of Non-linear and Complexity features for EEG Based Demen- tia Alzheimer disease Diagnosis

Authors: - Nilesh. N. Kulkarni1 , Saurabh. V. Parhad2 , Yasmin. P. Shaikh3

Alzheimer disease is one of the most common and fastest growing neu- rodegenerative diseases in the western countries. Development of different biomark- ers tools are key issues for diagnosis of Alzheimer disease and its progression. Pre- diction of cognitive performance of subjects from EEG and identification of relevant biomarkers are some of the research problems. EEG signal analysis can be well suited for automated diagnosis of Alzheimer's disease. Although, EEG based tech- niques are helpful in screening of Alzheimer and dementia; still there is a scope of improvement in terms of diagnostic accuracy, sensitivity and specificity. Thus, many issues are still left out in field of Alzheimer diagnosis using EEG signals related to the choice of features which can help in distinguishing the two or more subjects. This paper focuses on new features for diagnosis of Alzheimer's disease using EEG signals with effective increase in diagnostic accuracy. The use of new complexity based features is proposed in this paper which increases the diagnostic accuracy and helps in early Alzheimer's diagnosis.

b) Multivariate Analysis of Structural MRI and PET (FDG and 18FAV-45) for Alzheimer's Disease and Its Prodromal Stages

 DOI : 10.35291/2454-9150.2022.0176

Authors: - Qi Zhou, Mohammed Goryawala, Mercedes Cabrerizo, Warren Barker, David Loewenstein, Ranjan Duara and Malek Adjouadi*

A multivariate analysis method, orthogonal partial least squares to latent structures (OPLS), was used to discriminate Alzheimer's disease (AD), early and late mild cognitive impairment (EMCI and LMCI) from cognitively normal control (CN) using MRI and PET measures. FreeSurfer 5.1 generated 271 MRI features including 49 subcortical volumes, 68 cortical volumes, 68 cortical thicknesses, 70 surface areas and 16 hippocampus subfields. Subjects with all aforementioned MRI measures passing quality control and valid Fludeoxyglucose (18F) (FDG) and Flor- betapir (18F) PET scans were selected from ADNI database, resulting in a total of 524 participants (137 CN, 214 EMCI, 103 LMCI and 70 AD) for the study. Alto- gether 286 features including 15 significant PET uptake features (7 for FDG and 8 for AV-45) were utilized for OPLS analysis. Predictive power was evaluated by ( ), a quantifier of the statistical significance for class separation. The results show that MRI features ( ) =0.645), and PET features (( ) = 0.636) has comparable predictive power in separating AD from CN, and MRI features are better predictor of LMCI (( ) = 0.282) than PET (( ) = 0.294). Combination of PET and MRI has the most predictive power for LMCI and AD with ( ) of 0.294 and 0.721, respectively. While for EMCI, cortical thickness was found to be the best predictor with a ( ) of 0.108, suggesting cortical thickness may be the first structural change ahead of others and should be prioritized in prediction of very mild cognitive impairment.

c) A novel gene selection method using GA/SVM and Fisher criteria in Alzheimer's disease

Authors: - Seyede Zahra Paylakhi1 , Sadjaad Ozgoli1 , Seyed Hassan Paylakhi2

Identification of those genes which cause diseases can develop the pro- cess of diagnosis and the treatment of diseases. In this paper, a gene selection method based on genetic algorithm (GA) and support vector machines (SVM) is presented. At first, Fisher criteria is utilized in order to do filtration for those genes which are noisy and redundant in high dimensional microarray data. Then, GA/SVM model is used for selection of various subsets of maximally informative genes with the use of different training sets. The frequency of appearance of each gene in various sub- sets of genes is analyzed. Therefore, the last subset contains those genes which are highly informative. In fact, Fisher and GA/SVM methods have been merged in order to take benefit from a filtering method as well as an embedded method. The proposed method is tested on DNA microarray gene expression data of Alzheimer's disease. The results show that the proposed method has a good selection and classification performance, which can yield 100biological point of view, at least 8 (53Alzheimer associated genes. Thus, these genes not only can serve as predictors of the disease, but also can use as a means to find new candidate genes

d) Novel Gaussian Discriminant Analysis-based Computer Aided Diagnosis System for Screening Different Stages of Alzheimer's Disease

Authors: - Chen Fang1 , Chunfei Li1,Mercedes Cabrerizo1 Armando Barreto1 ,Jean Andrian1 , David Loewenstein2,3, Ranjan

This study introduces a novel Gaussian discriminant analysis (GDA)- based computer aided diagnosis (CAD) system using structural magnetic resonance imaging (MRI) data uniquely as input for screening different stages of Alzheimer's disease (AD) involving its prodromal stage of mild cognitive impairment (MCI) in relation to the cognitive normal control group (CN). Taking advantage of multiple modalities of biomarkers, over the past few years, several machine learningbased CAD approaches have been proposed to address this high-dimensional classifica- tion problem. This study presents a novel GDA-based CAD system on the basis of a tenfold cross validation and a held-out test data set. Subjects considered in this study included 187 CN, 301 MCI, and 131 AD subjects from the Alzheimer's Dis- ease Neuroimaging Initiative (ADNI) database. In the tenfold cross validation, the proposed system achieved an average F1 score of 97.20sensitivity of 99.14discrimi- nating together the MCI and AD groups from the CN group; and an average F1 score of 79.8287.43discriminating AD from MCI. By testing on the held-out test data, for discriminating MCI and AD from CN, an accuracy of 93.28were obtained. These results also show that by separating left and right hemispheres of the brain into two decisional spaces, and then combining their outputs, the GDA-based CAD system demonstrates a high potential for clinical application

e) Comparison Analysis of Machine Learning Algorithms to Rank Alzheimer's Disease Risk Factors by Importance

Authors: - Mohamed Mahyoub, Dr Martin Randles, Dr Thar

People have always feared aging, and the increasing rate of dementia dis- ease caused this fear to twofold. Dementia is irreversible, unstoppable and has no known cure. According to Alzheimer's Disease International 2015 and World

 DOI : 10.35291/2454-9150.2022.0176

Alzheimer Report 2015, the estimated financial cost for healthcare services of Alzheimer's Disease is Trillion in 2018. This paper discusses the importance of investigating Alzheimer's Disease using machine learning, the need to use both behavioural and biological markers data, and a computational method to rank Alzheimer's Disease risk factors by importance using different machine learning models on Alzheimer's Disease clinical assessment data from ADNI. The dataset contains Alzheimer's Dis- ease risk factors data related to medical history, family dementia history, demograph- ical, and some lifestyle data for 1635 subjects. There are 387 normal control, 87 significant memory concerns, 289 early mild cognitive impairment, 539 late mild cognitive impairment and 333 Alzheimer's Disease subjects. We deployed different machine learning models on the dataset to rank the importance of the variables (risk factors). The results show that some risk factors in subjects genetically, demogra- phy and lifestyle are more important than some medical history risk factors. Having APOE4, education level, age, weight, family dementia history, and type of work rank as more influential among Alzheimer's Disease subjects.

III EXISTING SYSTEM

Alzheimer's disease was only diagnosed with complete certainty after death, when examining the brain with a microscope revealed the characteristic plaques and tangles. Clinicians and researchers are now able to diagnose Alzheimer's disease during life with more certainty.

IV PROPOSED SYSTEM

A given 3D MR image is taken and is visualized in three orthogonal directions i.e., Axial, Coronal and Sagittal directions. The grey matter and white matter of the brain are separated from the 3D brain image and single slice extraction is performed. Skull stripping [15] is performed on these 2D

slices as a pre-processing step to remove non-cerebral tissues like skull, scalp, and dura from brain images. As part of the feature extraction, first-order statistical features are extracted

from the 2D slices, for both white matter and grey matter slices separately. The correlation matrix heatmap of all of the features is prepared to represent the interdependence between them. The principal component analysis is applied to these features as part of the feature reduction step to select the most prominent features. Pre-processing of the data is then performed. Four different classifiers are chosen here to classify the presence of Alzheimer's disease based on the prominent features selected, namely, Logistic Regression, SVM, Naïve Bayes and Adaboost classifier on both grey matter and white matter data individually in axial, coronal and sagittal directions.

Comparisons between the efficiencies of these classifiers are studied and analyzed in the last section. The schematic view of proposed method is shown in figure.

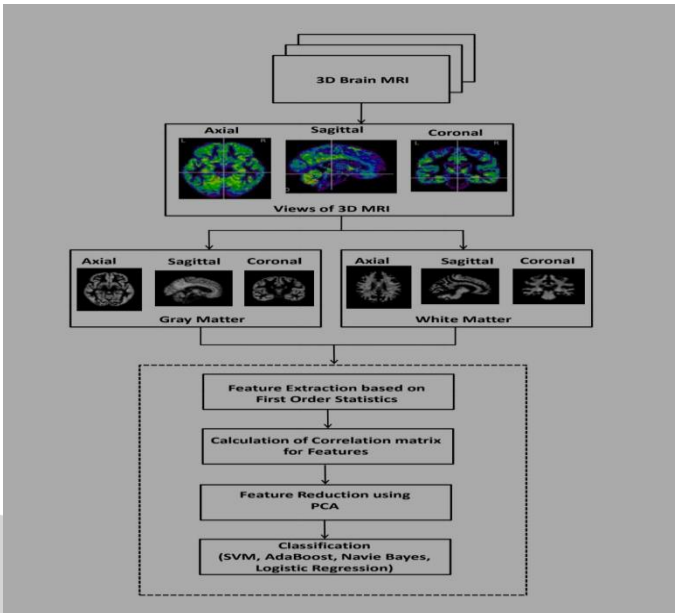

Fig. 1. Schematic view of Proposed Method

V METHODOLOGIES

This study proposes a new method for the detection of Alzheimer's Disease (AD) using first-order statistical features

in 3D brain Magnetic Resonance(MR) images. Alzheimer's disease is a neurodegenerative disorder that affects elderly people. This is a progressive disease and early detection and classification of AD can majorly help in controlling the disease. Recent studies use voxel-based brain MR image feature extraction techniques along with machine learning algorithms for this purpose. Grey and white matter of the brain gets affected and damaged due to AD and so studying these both prove to be more effective in predicting the disease. The proposed work uses 3D structural brain MR images to separate the white and grey matter MR images, extract 2D slices in the coronal, sagittal and axial directions and select the key slices from them for performing feature extraction on them. Feature extraction is applied on top of these slices to calculate the first-order statistical features and the prominent feature vectors generated by PCA are selected for further study. In the classification phase, different classifiers take the selected features as its input to predict the classes AD (Alzheimer's Disease) or HC (Healthy Control) based on the observations in the validation set. Experimental results show that the accuracy of 90.9 % compared to other techniques.

 DOI : 10.35291/2454-9150.2022.0176

VI ALGORITHM

Convolutional Neural Network(CNN):

- In deep learning a convolutional neural network (CNN/ConvNet) is a class of deep neural networks, most commonly applied to analyze visual imagery. Now when we think of a neural network we think about matrix multiplications but that is not the case with ConvNet. It uses a special technique called Convolution.
- It has four steps:
  - Convo 2D
  - Max Pooling
  - Flatten
  - Fully connected Network

VII. CONCLUSION AND FUTURE WORK

In this work, an effort has been made to study the 3D brain MR image slices for AD diagnosis. All the three different views of slices (Axial, Sagittal, and Coronal) of gray matter and the white matter has been used for this study. Based on several observations slice number 51 has been chosen and used for further analysis. The first-order statistical feature has been extracted from each slice. The correlation heat map is calculated from the features. Feature reduction is performed using PCA. Four classification methods logistic regression, Naive Bayes, SVM, and Adaboost has been used to study the performance of the proposed algorithm. The publicly available brain MR image data set OASIS is used for these experiments. The experimental results show that maximum accuracy has been achieved using white matter slices of coronal view. The proposed method shows 90:9% accuracy has been achieved in detection AD.

REFERENCES

[1] Dubois B, Padovani A, Scheltens P, Rossi A, DellAgnello G. Timely diagnosis for Alzheimers disease: a literature review on benefits and challenges. Journal of Alzheimer's disease. 2016 Jan 1;49(3):617-31

[2] World Health Organization. World health statistics 2010. World Health Organization,2010

[3] Jin K, Simpkins JW, Ji X, Leis M, Stambler I. The critical need to promote research of aging and aging-related diseases to improve health and longevity of the elderly population. Aging and disease. 2015 Feb;6(1):1

[4] Padilla P, Lpez M, Grriz JM, Ramirez J, Salas-Gonzalez D, Alvarez I. NMF-SVM based CAD tool applied to functional brain images for the diagnosis of Alzheimer's disease. IEEE Transactions on medical imaging. 2011 Sep 12;31(2):207-16.

[5] Song S, Lu H, Pan Z. Automated diagnosis of Alzheimer's diseaseusing Gaussian mixture model based on cortical thickness. In2012 IEEE Fifth International Conference on Advanced Computational Intelligence.

 DOI : 10.35291/2454-9150.2022.0176