ANALYSIS OF ALZHEIMER DISEASE PREDICTION USING MACHINE LEARNING TECHNIQUES

[1]B. Hemalatha, [2]Dr. M. Renukadevi.

[1]Research Scholar, Department of Computer Science, Sri Krishna College of Arts and Science, Coimbatore.
bharathihema17@gmail.com.

[2]Associate Professor, Department of Application, Sri Krishna College of Arts and Science, Coimbatore.
renuga.srk@gmail.com.

Abstract - Alzheimer's Disease (AD) is referred to as one of the highest non-unusual neurodegenerative disorders that inflict eternal harm to the memory-associated brain cells and wonder skills. There is a 99.6 percent failure rate in clinical trials of Alzheimer's disease pills, perhaps due to the fact that AD sufferers cannot be without early-stage complications. This observation analyzed machine learning knowledge of strategies to use empirical statistics to forecast the progression of AD in the years of fate. Diagnosis of AD is often difficult, particularly at an early stage in the disease system, due to the degree of mild cognitive impairment (MCI). However, it is at this point where treatment is much more likely to be successful, so there will be great benefits in enhancing the diagnosis process. Research in this area aims to identify the most complex mechanisms directly related to changes in AD. Various imaging methods are used to diagnose AD, and image modes play a key role in the diagnosis of AD. This paper uses a Positron Emission Tomography (PET) image to detect AD early. The PET image is often used to know how organs and tissues function in the human body. This research study analyses prediction approaches using various kinds of machine learning algorithms to solve AD diagnostic problems. Artificial Neural Networks are one of the many algorithms. Modern research has shown that deep learning is a proficient technique for solving numerous problems of image recognition, but most of these published approaches owe their performance to training on a very large number of data samples.
Keywords: - Alzheimer's disease, PET image, Image Modalities, Mild Cognitive Impairment (MCI) and NN.

1. INTRODUCTION

Alzheimer Disease (AD) is one of the most prevalent type of dementia, is a important health concern in the twenty-first century. A reported 5.5 million people aged 65 and older are living with AD, and AD is the sixth leading cause of death in the United States. The global expense of handling AD, including medical, social security, and wage reductions for patients' families, was $277 billion in 2018 in the United States, having a major impact on the national economy and pressuring the US health care system. [1]. AD is an irreversible, progressive brain condition characterized by a decrease in cognitive activity without a validated treatment-modifying disease [2]. Much effort has therefore been made to establish early detection strategies, especially at pre-symptomatic stages, to delay or prevent disease progression. [3, 4]. Advanced Neuroimaging techniques, such as magnetic resonance imaging (MRI) and positron emission tomography (PET), have been built and used to classify AD-related structural and molecular biomarkers [5]. Rapid development in Neuroimaging techniques has made it difficult to

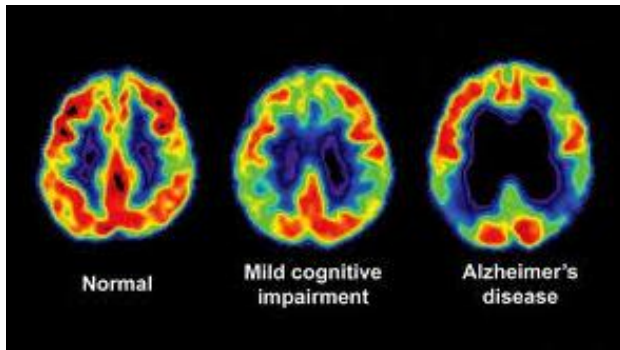incorporate large-scale, high-dimensional multimodal Neuroimaging data.

Figure 1: -Stages of AD

Alzheimer's disease is known to be a loss of neurons and synapses in the cerebral cortex and in some subcortical areas. This requires of neuron and disease synapses contributes to distinctly noticeable variations in brain tissues. Hippocampal atrophy, ventricle enlargement and cortex shrinkage are sensitive characteristics of Alzheimer's disease. Physicians therefore conduct brain scans, such as Computed Tomography (CT), Magnetic Resonance Imaging (MRI) or Positron Emission Tomography (PET), to rule out other potential causes of the symptom. [11].

Since this is time consuming and less precise, it is often difficult to diagnose Alzheimer's disease, particularly in the early stages. Moreover, the precision of the diagnosis depends primarily on the experience of the radiologists. A variety of studies have been performed in recent years on the automated diagnosis of Alzheimer's disease using various approaches. Numerous pattern classifiers have attempted to discriminate against subjects using different Neuroimaging evidence. Various extraction methods and classification methods have been used in these recent studies.

The rest of the paper will be structured as follows. In section 2, discuss the different methodologies for predicting AD. The problem statement for early detection of AD is prescribed in section 3. In section 4, clarify the overall performance analysis of the current technique. The methodologies are finally outlined in section 5.

2. RELATED WORK

Medical imaging data comes in a number of forms, such as structural or functional MRI, Diffusion Tensor Imaging (DTI), Positron Emission Tomography (PET), Computed Tomography (CT scan). All these types of scans have been used in different ways with different machine learning algorithms in the past for early diagnosis or disease detection.

Alzheimer's disease has some gradual pattern of brain tissue injury. It shrinks the brain's hippocampus and cerebral cortex and enlarges the ventricles. Some remarkable research has been conducted on the diagnosis of automated Alzheimer's disease. There are many methods found in this literature review to address this research problem. But each solution also has its own features and limitations. This section summarises the literature review of the topic.

Long, X., et al. (2017) Proposed to identify early morphological changes in the brain and early analysis of critical for Alzheimer's disease (AD). High resolution MRI can be used to help identify and predict the disease. In this work, a ML approach was proposed to differentiate between patients with AD or Moderate Cognitive Disability (MCI) and strong old patients and to predict AD conversion in MCI patients by computing and analysing area brain morphological variations among classes. The difference among every pair of subjects was quantified by symmetrical diffeomorphic registration, followed by an embedding algorithm and a classification knowledge approach. This deformation-based approach completed full use of the pair-wise

520

macroscopic shape distinction among groups and thus improved the capacity to discriminate. [6].

Dahiwade, D., et al. (2019) introduced about people are faced with different diseases appropriate to their environmental condition and their living habits. The prediction of a disease at an earlier stage is therefore an important task. But accurate predictions on the basis of symptoms are becoming too difficult for the doctor. Right disease prediction is the most difficult task. Data mining plays an important role in predicting the disease in order to solve this challenge. Medical science has a huge amount of data growth each year. Increased data development in the medical and healthcare industry means an accurate review of medical data that has gained from early patient care. Using disease data, data mining discovers secret pattern knowledge in a large amount of medical data. In general, the disease prediction is based on the patient's symptoms. For disease prediction, use the K-Nearest Neighbor (KNN) and Convolutional Neural Network (CNN) machine learning algorithm for accurate disease prediction. Dataset of disease symptoms necessary for disease prediction. Following general disease prediction, this method is capable of presenting a risk associated with general disease that is lower or higher risk of general disease. [7].

Lodha, P., et al. (2018) proposed about the Machine learning is commonly used in a number of medical fields. Advances in medical technology have provided access to better data to detect early-stage signs of various diseases. Alzheimer's disease is a chronic disorder that contributes to brain cell degeneration leading to memory enervation. Patients with cognitive mental disorders, such as uncertainty and forgetfulness, as well as other symptoms, including behavioural and psychological problems, are further recommended with CT, MRI, PET, EEG, and other Neuroimaging techniques. The goal of this research work is to use machine learning algorithms to process these data from Neuroimaging technologies for the detection of Alzheimer's at its early stage. [8].

Fisher, C. K., et al. (2019) proposed approaches for machine learning from electronic health data may only predict a single endpoint. The ability to model thousands of patient characteristics simultaneously is a critical step towards personalised Alzheimer's disease medicine. Use the Conditional Restricted Boltzmann Machine (CRBM) unsupervised machine learning model here to simulate detailed patient trajectories. Here, the use of 18-month trajectory data of 44 clinical variables from 1909 patients with moderate cognitive impairment or Alzheimer's disease to train a model for personalised disease progression prediction. Simulation of synthetic patient results, including the evolution of each sub-component of cognitive tests, laboratory tests and their correlation with basic clinical characteristics. Synthetic patient data produced by the CRBM accurately represent the means, standard deviations and correlations of each variable over time to the degree that synthetic data cannot be separated by logistic regression from real data. [9].

Khan, A., & Usman, M. (2015) proposed about Alzheimer's, an irreparable disease of the brain, impairs thought and memory, as the aggregate size of the mind shrinks, which at last induces death. Early diagnosis of AD is important for the advancement of more prevalent therapies. Machine learning (ML), a branch of artificial intelligence, employs a variety of probabilistic and optimising techniques that enable PCs to benefit from vast and complex datasets. As a result, researchers also concentrate on using machine learning to diagnose early stages of AD. This research work provides a study, analysis and critical assessment of current researches on early detection of AD using ML techniques. Several methods have achieved promising predictive accuracy; however,

521

they have been tested on various pathologically unproven data sets from different imaging modes, making it difficult to make a fair comparison between them. [10].

As a result, the research issue for the prediction of AD incorporates various preprocessing, feature extraction and classification algorithms for the processing of highly dimensional medical datasets.

3. PROBLEM FORMULATION

Alzheimer's disease (AD) is a progressive brain degeneration characterised by the accumulation of amyloid plaques in brain tissues and neurofibrillary entanglements. With more than 35 million people worldwide, this is the predominant type of dementia. The risk of disease doubles every five years after the age of 60 and is expected to have reached 135 million by 2050. Since 2000, Alzheimer's death rates have risen to 68%. There is an immense need to develop new approaches for early detection of AD. [1, 11].

For meaningful analysis and extraction of features, image processing techniques have been used for pre-processing of Positron Emission Tomography (PET) images. PET scans use radioactive tracers to highlight amyloid protein plaques in the brain that are a characteristic of Alzheimer's disease. They can encourage physicians to diagnose Alzheimer's earlier—even before any symptoms occur. Two of the most common biomarkers present in Alzheimer's are reduced glucose absorption and deposition of amyloid plaques in the brain. PET scans use various radioactive drugs, or radiotracers, to test these biomarkers in the brain tissues of patients with cognitive disability. The Alzheimer's disease prediction was achieved in the previous study using MRI images. One of the big disadvantages of using an MRI image is the early stages of Alzheimer's

disease, and an MRI scan of the brain may be common.

The selected features are then redirected to machine learning algorithms that classify PET images into one of the possible categories based on pattern recognition techniques. The importance of our approach lies in the fact that, in the classification process, only a few regions are selected which are at the greatest risk of atrophy and thus the computation time is reduced.

4. PROPOSED MODEL

First of all, it is believed that if there is an effective detection process, it will certainly yield successful pre-detection results as well. In addition, this analysis consists of two key sections. It is very difficult to compare recent research, because they have been extracted from different datasets and different methods [12]. Therefore, the latest literature review and the most effective detection methods assisted by Vector Machine have been checked (SVM). Pre-detection shall be carried out at any successful completion of the verification level, or any other method shall be suggested as a second step.
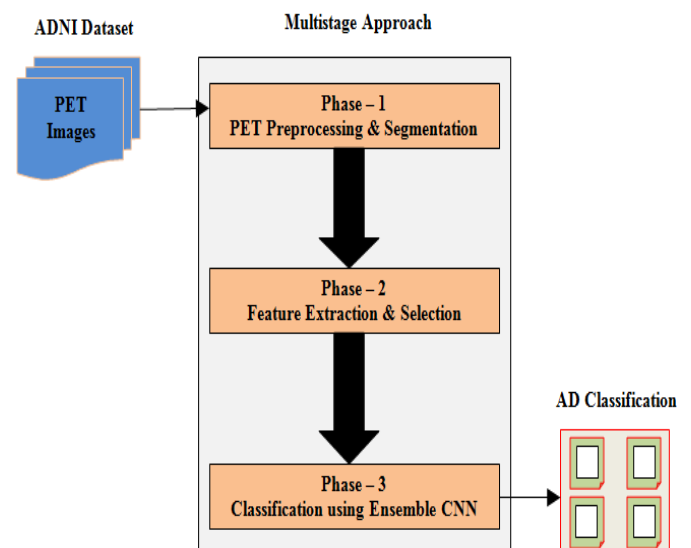

Figure 2: - Workflow of Proposed Model

522

Early detection of AD based on a prior knowledge of the image, such as the position and shape of the brain. They tested the prototype on PET scan images and concluded that even their software is not one hundred percent accurate and able to do better.

Phase – 1: - PET Image Preprocessing & Segmentation

Preprocessing is a sequence of data transformations designed to minimise noise in raw data. Pre-processing increases the sensitivity of the study (SNR) and certifies the validity of the statistical model. A median filter is chosen for the proposed work. In the first level, the noise pixels are identified. The edges are then identified in the second level. Noise pixels are restored around detected edges in the third step. The main objective of median filtering is simply to substitute each pixel value in an image with the median value of its neighbours, including themselves.

Image segmentation is among the most important tasks in the processing of medical images and is also the and perhaps the most essential thing in many clinical applications. Several segmentation methods used for the segmentation of PET image processes. The UNet Architecture method for the segmentation of the PET image is implemented in this proposed study.

Phase -2: - Feature Extraction & Selection

Feature extraction techniques are useful in several image processing applications, e.g. character recognition. As features define the behaviour of an image, they also display its position in terms of storage, efficiency in classification and, obviously, time consumption. A function is a piece of knowledge about the content of the image; generally about whether a certain area of the image has certain properties. Characteristics can be unique structures in the picture, such as points, edges or artefacts. The proposed work presents a novel method for extracting features such as Pyramid Histogram of Oriented Gradients (PHOG) and Gray Level Run Length Matrix (GLRLM) from the segmented region for PET images.

In this work, Feature Selection is a process that reduces the number of input variables when designing a predictive model. Various methods for selecting features use statistical tests to measure the association or dependency between input variables that can be filtered in order to select the most important features. The proposed work incorporates a filter and a wrapper-based feature selection facility called the Regularization Technique.

Phase – 3: - Classification using Ensemble of Deep CNN

Deep learning models have obtained impressive results in the field of computer vision, making them a modern trend in image recognition and classification. Our used Convolutional Neural Networks (CNN) algorithm appears in this context. CNN is divided into four steps: Convolution, Non-Linearity (ReLU), Pooling or Sub-Sampling and Classification (Fully Connected Layer). These activities are essential to the construction of the CNN system.

The proposed work use three dissimilar methods of classifiers (CNN, Support Vector Machine, and MLP) to evaluate the efficiency of our proposed strategy, as well as a selection of these classifiers working on majority voting. In most instances, the ensemble-based classifier performs better than individual classifiers. In this research work, a new approach to the automatic classification of Alzheimer's disease from MRI scans was explored. Comparing the proposed approach to existing methods reveals that the proposed approach has the highest specificity values relative to the current approach. Given the fact that the lesser function used has been set, the proposed approach has an overall accuracy value comparable to the current approaches. The suggested method is based on the extraction of two types of features

523

after the images have been pre-processed and segmented. Here, three classification models were used to evaluate our proposed approach on the basis of individual characteristics and a combination of all features.

5. PERFORMANCE ANALYSIS

This stage focuses primarily on the validation of a trained network and then compares the results with that or other pre-trained networks. A portion of the data is set back for validation and the classified performance of the qualified network is reviewed with the images in the validation set. Once adequate precision is not achieved, the network must adjust the data pre-processing of training options before the network performs well. The accuracy of the validity is achieved after the training has advanced. The ADNI clinical dataset contains clinical information on each subject, including recruitment, demographics, physical exams and cognitive assessment data. [13].

Any consequence of classification may have an error rate and may, on occasion, either fail to recognise an abnormality or identify an abnormality that is not present. It is normal to characterise this error rate in terms of true and false positive and true and false negative as follows: [14].

True Positive (TP): the classification value is positive in the case of clinical anomalies.

True Negative (TN): the impact of classification is negative in the absence of symptoms abnormality.

False Positive (FP): the result of classification is positive in the absence of clinical abnormality.

False Negative (FN): the effect of classification is negative in the presence of clinical abnormality.

The metrics applied to evaluate the model were: accuracy (ACC), sensitivity (SEN) and specificity (SPE) and the region under the ROC curve (AUC). Equations (1), (2) and (3) show the formulas describing these metrics. [15]:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \qquad \dots Equ(1)$$

$$SEN = \frac{TP}{TP + FN} \qquad \dots Equ(2)$$

$$SPE = \frac{TN}{TN + FP} \qquad \dots Equ(3)$$

6. CONCLUSION

Increased risk of AD can be associated with an increased age. Lower education and monthly family income, family history of dementia and physical inactivity may all contribute to the development of AD. In this research work, a new approach to the automatic classification of Alzheimer's disease from PET scans was examined.

Some findings can also be attested in the study of approaches that use deep learning. With regard to the attribute extraction criterion for this research issue, CNN provides better results compared to the auto-encoders observing all the metrics analyzed.

REFERENCES

[1] Alzheimer's Association. (2018). 2018 Alzheimer's disease facts and figures. Alzheimer's & Dementia, 14(3), 367-429.

[2] De Strooper, B., & Karran, E. (2016). The cellular phase of Alzheimer's disease. Cell, 164(4), 603-615.

[3] Schelke, M. W., Attia, P., Palenchar, D. J., Kaplan, B., Mureb, M., Ganzer, C. A., ... & Isaacson, R. S. (2018). Mechanisms of risk reduction in the clinical practice of Alzheimer's disease prevention. Frontiers in aging neuroscience, 10, 96.

[4]   Galvin, J. E. (2017). Prevention of Alzheimer's disease: lessons learned and applied. Journal of the American Geriatrics Society, 65(10), 2128-2133.

[5]   Veitch, D. P., Weiner, M. W., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., ... & Alzheimer's Disease Neuroimaging Initiative. (2019). Understanding disease progression and improving Alzheimer's disease clinical trials: Recent highlights from the Alzheimer's Disease Neuroimaging Initiative. Alzheimer's & Dementia, 15(1), 106-152.

[6]   Long, X., Chen, L., Jiang, C., Zhang, L., & Alzheimer's disease Neuroimaging Initiative. (2017). Prediction and classification of Alzheimer disease based on quantification of MRI deformation. PloS one, 12(3), e0173372.

[7]   Dahiwade, D., Patle, G., & Meshram, E. (2019). Designing disease prediction model using machine learning approach. In 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC) (pp. 1211-1215). IEEE.

[8]   Lodha, P., Talele, A., & Degaonkar, K. (2018). Diagnosis of alzheimer's disease using machine learning. In 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA) (pp. 1-4). IEEE.

[9]   Fisher, C. K., Smith, A. M., & Walsh, J. R. (2019). Machine learning for comprehensive forecasting of Alzheimer's disease progression. Scientific reports, 9(1), 1-14.

[10]   Khan, A., & Usman, M. (2015). Early diagnosis of Alzheimer's disease using machine learning techniques: A review paper. In 2015 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K) (Vol. 1, pp. 380-387). IEEE.

[11]   Zhang, D., Wang, Y., Zhou, L., Yuan, H., Shen, D., & Alzheimer's Disease Neuroimaging Initiative. (2011). Multimodal classification of Alzheimer's disease and mild cognitive impairment. Neuroimage, 55(3), 856-867.

[12]   Cuingnet, R., Gerardin, E., Tessieras, J., Auzias, G., Lehéricy, S., Habert, M. O., ... & Alzheimer's Disease Neuroimaging Initiative. (2011). Automatic classification of patients with Alzheimer's disease from structural MRI: a comparison of ten methods using the ADNI database. neuroimage, 56(2), 766-781.

[13]   Liu, M., Zhang, D., Shen, D., & Alzheimer's Disease Neuroimaging Initiative. (2012). Ensemble sparse classification of Alzheimer's disease. NeuroImage, 60(2), 1106-1116.

[14]   Tong, T., Gray, K., Gao, Q., Chen, L., Rueckert, D., & Alzheimer's Disease Neuroimaging Initiative. (2017). Multi-modal classification of Alzheimer's disease using nonlinear graph fusion. Pattern recognition, 63, 171-181.

[15]   Jain, R., Jain, N., Aggarwal, A., & Hemanth, D. J. (2019). Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. Cognitive Systems Research, 57, 147-159.