A MACHINE LEARNING PROPOSITION FOR CATEGORIZING ALZHEIMER'S DISEASE USING MAGNETIC RESONANCE IMAGES

Bandaru Shanmukha Priya
PG Student
Department of Information Technology
BVRIT HYDERABAD
College of Engineering for Women
Hyderabad, India

bandarushanmukhapriya@gmail.com

S.Rama Devi
Associate professor
Department of Information Technology
BVRIT HYDERABAD
College of Engineering for Women
Hyderabad, India

ramadevi.s@bvrithyderabad.edu.in

ABSTRACT

Alzheimer's disease is the sixth most typical autonomic disease in the world. Alzheimer's disease cannot be recognised at an early stage, it is identified at last stages of the disease. The effects of the disease are forgetfulness, dementia. There is no discernment of Alzheimer's disease in the early stages because, the symptoms of the disease are generally not seen and an extensive research is happening in the fields of computers and medicine. In existing work, algorithms like KNN-classification, SVM, Ada boost, XG Boost, convolutional neural networks etc.. are proposed for analysing the Alzheimer's disease. The limitations of the existing research don't isolate the Alzheimer's disease into different stages and they didn't perform all the parameters like accuracy, f1 score, recall, precision to measure the outcome. In this research, we propose for the detection of Alzheimer's disease using algorithms like Inception V3, Alexnet, Mobile Net V2, VGG16, Alzheimer Net. DenseNet201, ResNet50, Xception models which will classify the Alzheimer's disease into different stages like Subjective memory concern(SMC), Mild cognitive impairment(MCI), Early Mild cognitive impairment(EMCI), Late Mild cognitive impairment(LMCI) and Alzheimer disease(AD). The proposed model is manipulated with data augmentation techniques like flip, rotate and translate to overcome over-fitting of the data. The accuracy of our proposed model yields up to 99.5% when compared with already existing models.

Key words- Alzheimer's disease, Dementia, Cognitive Impairment, Magnetic Resonance Imaging(MRI), Convolutional neural networks, Xception, K- nearest neighbours(KNN), data augmentation.

1.INTRODUCTION

Alzheimer's disease is the neuro regressive disease which is probably diagnosed at the last stage of the disease [12]. This disease is caused by a protein called Amyloid β and its separation from the inter body of the neuron and forms lumps like structure called plaques [2]. A lot of research is going on in the medical field to identify the early detection of disease. The disease has no cure till date. Over 55 million people are diagnosed with Alzheimer's disease in 2020. There is a possible statistical diagnosis of over 139 million people will be affected by Alzheimer's by the year 2050 and over 79 million people will be affected by the year 2030 [4]. The effects of the disease [2] include hard to remember names, challenges in work place and social gatherings, having trouble finding valuable things, increasing difficulty in organising things. In this extensive research, the proposed works in this area of research are using algorithms like K-nearest neighbours (KNN), Support vector machine (SVM), Adaptive boosting(Ada boost), Decision tree, extreme gradient boosting (XG boost), Boot strap aggregation, Random forest, Dense Net-169, Res Net-50, etc. to test and train the data. The training and testing of different data sets are extracted from OASIS, ADNI etc. Most of the dataset is split into 70:30 in all the research work [4,5]. Each research work has different parameters taken into consideration to determine the output. All of the research work already proposed use various Machine learning and deep learning models to evaluate the dataset. The assessing metrics like accuracy, precision, F1-score and recall are evaluated. The image processing techniques like data augmentation, Contrast limited adaptive Histogram Equalization(CLAHE) are proposed for the better understanding of the pixel values of the datasets.

2.LITERATURE SURVEY

A profuse research is going on to determine the Alzheimer's disease in various age group people over the span of years. Numerous Machine learning algorithms like K-nearest neighbours(KNN), Decision tree, Support vector machine(SVM), Adaptive boosting (Ada boost), extreme gradient boosting (XG boost), Boot strap aggregation, Random forest, Dense Net-169, Res Net-50, etc. are some of the algorithms used in the research works to find detect the Alzheimer's disease. The algorithms have performed well to identify the diagnosis of the Alzheimer's disease.

Model proposed in this paper [3] put forward to diagnose whether the person has dementia or not. The classification techniques used in this paper

are Dense Net-169 and Res Net-50 CNN architectures to classify the Magnetic Resonance Images(MRI) as dementia or non-dementia. The Dense Net -169 out performs the Rest Net-50 in training and testing with more accuracy values. The accuracy values of Dense Net-169 are 96.77% whereas the accuracy value of the Res Net -50 is 88.9 The model can be used in real-time to classify the Magnetic Resonance Images(MRI) of a person in Health care.

The Journal [2] discusses about the basics of Alzheimer's disease and how it is formed in the human brain. The differences like enlarged ventricles, Atrophy of cerebral cortex, Atrophy of Hippocampus are shown in the diagram to understand about the Alzheimer's disease. Alzheimer disease is categorized as the 6th leading disease around the globe. [1,2] The various techniques in medical field to identify the Alzheimer's disease like Positron Emission Tomography(PET), MRI Biomarkers, Magnetic Resonance Imaging(MRI), etc. are discussed briefly in terms of health care. They discussed various algorithms and they are efficient towards the identification of Alzheimer's disease. The first one is Support Vector Machine(SVM) along with free surfer and SPM5. They performed quite well with 95% cross-validation and they showed 72.3% efficiency. They effectively proposed the paper [2] using the parameters like left hippocampus volume, Right anterior part of para hippocampal, Cortical thickness of left the per cuneus, Right hippocampus volume, left superior temporal sulcus. The[2] work reported 72.3% coherence on ADNI data.

Data sets[2] they used are 3d or 4d image data array, image metadata, affine array of data, neuropathology data set, minimum data set [2] . They reviewed on various other data sets from Harvard university of Medical sciences, Max plank institute Leipzig Mind-Brain-Body data set(LEMON), National health and ageing trends study(NHATS), Open access series of Imaging studies(OASIS)[1,2]. They concluded that this model will predict only the primary stage recognition of Alzheimer disease(AD)[2] .

The predictions in this paper [2] are done with the data record from Open access series of Imaging studies[OASIS]. Several machine learning techniques like Decision trees, Random forest(RF), Support vector machine, voting, Gradient Boosting classifiers have been deployed as algorithms to assess the dataset [2]. The variables like precision, recall, F1-score, accuracy are taken into consideration for the efficiency of the model. The validation is 83% for this model. The disease effects

more than 5.1 million [3] people every year all over the globe. They compared the results of different papers [2] and their accuracy. The different attributes in this data set are ID, Male or female, Hand, age, years of education, Socio economic status, Normalized brain volume, mental state examination, clinical dementia rating [2]. The training and testing data are in 80:20 ratios. In feature selection the filter methods, wrapper methods, embedded methods are used. Various data visualization techniques like confusion matrix, bar graphs, histograms [1,2] are used to determine and compare the attributes. All the algorithms are compared where as XG Boost and voting classifier determines the accuracy, F1 –score and precision.

The Journal [4] discusses about the dataset is extracted from the OASIS (Open Access Series of Imaging Studies). The input is classified into training and testing data in the ratio of 70:30. The training dataset are classified into morphometric features and graph theory features. In both the features the classifiers applied [4] are linear support vector machine(L-SVM), Decision Tree(DT), Extremely randomized Tree(ET), Random Forest(RF), Linear discriminant analysis(LDA), Cross Validation(CV), Logistic regression(LR), Matthews correlation coefficient(MCC), In both features the top training datasets are picked by the ranking metric upon applying the classifiers. The combined probability is taken from the testing dataset after processing through the classifiers is again classified as a feature set. This feature set is processed through the Linear SVM with 5-fold CV for C-parameter optimization. The performance values [4] in females are 0.749 and in males is 0.875.

In this paper [12] the dataset is extracted from OASIS(Open Access Series of Imaging Studies) of brain mapping which are accessible for research purposes. The customized network called Brain Net 2D is developed with 2DConvolutional layer and 2D Max pooling layer with an activation function called Soft max. To standardize the input values, the Batch Normalization technique is also used in the Brain Net 2D. The cyclical learning rate is used to find the optimal learning rate. The Brain Net 2D improvise the results up to 0.92% compared to Res Net 18 approach. The obtained table shows the results into CDR levels of 0,0.5,1,2 of several stages of Alzheimer's disease.

The article [13] discusses about early stage prediction of Alzheimer's disease. Deep learning models like Convolutional Neural Network(CNN), 18-Layered CNN, and 3D-CNN are applied. An accuracy of 80, 98 and 88% are

obtained and compared for the best result. The multi-layered CNN set off complex to identify pretentious areas of the brain. The result[13] illustrate that the 18-layered CNN has best exactness of 98% which makes the work for the radiologist easy. The models in this paper [13] doesn't classify dataset into NC, EMCI and LMCI.

The Journal [14] proposes the approach of applying Convolutional neural network by using Electroencephalogram (EEG) signals, extracts the features by using fast Fourier transform (FFT) and classifies the different stages of disease after feeding the extracted features to CNN. The layers in the neural network are convolutional, rectified linear Unit(ReLU), pooling layer, fully connected layer. The activation function used here[14] is rectified linear unit. This papers motive is to find the Alzheimer's disease from the cost effective test EEG rather than MRI.

The dissertations[15] proposes numerous algorithms like Support vector machine, Random forest, decision tree, logistic regression etc.. for the prediction of Alzheimer's disease. Dataset is extracted from OASIS, where the data is kept publicly for the purpose of research. The evaluation metrics of the results are accuracy, recall, precision, F1 score. The comparative analysis is shown to predict the best model. Among the proposed algorithms the support vector machine(SVM) gives an output precision of 92%, random forest with the accuracy of 81.3%. SVM is considered for testing the data of various people.

3.PROPOSED WORK

Figure 1 proposes the methodology for the categorizing of Alzheimer Disease into Subjective memory concern(SMC), Mild cognitive impairment(MCI), Early mild cognitive impairment(EMCI), Late mild cognitive impairment(LMCI) and Alzheimer disease(AD). For the prediction of disease, the records are gathered from the ADNI elements, then processed to enhance the pixel values, then given as input to the model algorithms and trained. The user then login to the site and upload the picture of the MRI. The prediction comes as one of the stage of Alzheimer's disease. The metrics used to review the result here are accuracy, recall, precision and F1-Score.

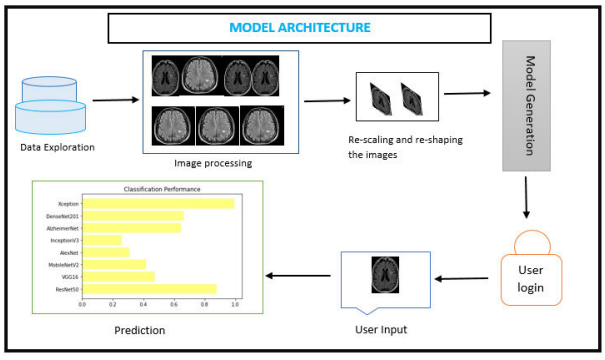

Figure 1: Model Architecture

The proposed work has 6 phases to predict the Alzheimer's disease. They are: (1)Data exploration, (2)Image processing, (3)Model generation, (4)User signup/Login, (5)User input and (6)Prediction. The phase detailing is as follows:

3.1 Data exploration

The data is explored from the Alzheimer's disease Neuroimaging Initiative(ADNI) data bases. [https://adni.loni.usc.edu/data-samples/access-data/] This database is a collection of Magnetic Resonance Images (MRI) of various people across the world. This dataset helps various researcher's in their research. In the ADNI databases the researchers utilize various data like MRI scans, PET images, genetics, cognitive tests, blood biomarkers etc. The data collected from the ADNI are used in training and testing of the proposed work. Total of 18,775 images are used, among them 16,200 are handed down for training the model and 2,575 images are put to test.

3.2 Image Processing

Image processing is necessarily used to improve the quality of image to train and test the data in the algorithm. The image processing techniques used here are:

(a)Re-scaling of the image: The images are rescaled to 1/255(between 0 and 1) to get more clarity on pixel values to get efficient training values while training the sample data.

(b)Zooming the Image: The images are enlarged or miniaturized to 0.2% to give the best outcome, to train the model irrespective of the size of the object and enhancing the optimization of the model generated.

(c)Horizontal Flip: Horizontal flipping of images means horizontal mirror of the images to add more diversity and features in the image processing during the model generation.

(d)Reshaping the image: The reshaping of the image involves in change in width and height of the images to [128,128] to present various dimensions in the training data samples to combat with the various types of images during testing the data.

3.3 Model generation

The models are essentially the different types of algorithms used in the proposed work during training and testing of the data instances. The training data is 70% and the experimental data is 30% of the overall dataset. Various libraries like tensor flow, keras, Matplotlib, pyplot, numpy are imported. The different algorithms mainly used in this work are Inception V3,Alex net, Mobile Net V2, VGG 16, Alzheimer Net, Xception, ResNet 50, DenseNet201. Out of these

Dense Net 201, Res Net 50, Xception algorithms performed well. Let us know the working of algorithm in detail:

3.3.1 Xception

The Xception algorithm performs the best in the training and testing the datasets. To the algorithm we feed the input with the size [128,128], then into the image net depth wise n*n convolution with size [128,128,3], then global average pooling, then to softmax activation function is applied and connected to a fully convolutional network to generated the output.

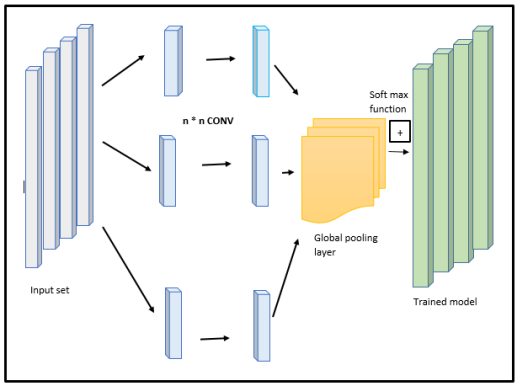

Figure 2: Working of Xception model

3.3.2 RES NET 50

The RESNET50 is one of the cutting edge convolutional neural network which, mainly used to solve the gradient descent problems. The layers in the network are convolutional layers(column wise) , identity block, convolution block, and global average pooling layers. After the images sent into global average pooling layer they are flattened and are then fed to soft max activation function to get the desired prediction.

3.4 User Signup/Login and User input

In the proposed work, by using flask application framework a light weight web based application is designed to login or signup Magnetic Resonance Imaging(MRI). The uploaded MRI is tested through various layers and the various stages of the Alzheimer disease is predicted.

3.5 Prediction

The prediction of the outcome is done by comparison between algorithms for classification of diseases into Subjective memory concern(SMC), Mild cognitive impairment(MCI), Early Mild cognitive impairment(EMCI), Late Mild cognitive impairment(LMCI) and Alzheimer disease(AD). The highest accuracy algorithm is taken and given as the output for testing data. Various metrics like precision, re-call, F1-Score and accuracy are measured to give the best outcome for the given input feed values.

4.RESULTS & DISCUSSION

The proposed work gives the comparison of various ML algorithms by training the data samples. The values like Precision, accuracy, F1 Score and recall are done to

give the absolute outcome. It also gives the prediction of different phases of Alzheimer's disease.

S. NO	ML Model	Accuracy	Precision	Recall	F1 Score
1	Inception V3	0.361	0.384	0.258	0.300
2	Alex Net	0.346	0.332	0.309	0.316
3	Mobile Net V2	0.468	0.497	0.417	0.444
4	VGG 16	0.474	0.474	0.474	0.474
5	Alzheimer Net	0.646	0.646	0.646	0.646
6	Dense Net 201	0.719	0.774	0.664	0.701
7	Res Net 50	0.888	0.919	0.878	0.858
8	Xception	0.995	0.995	0.995	0.995

Figure 2: Comparative results of all Models

4.1 Accuracy:

Accuracy is a metric of how often the model executes or performs correctly. It asks the model that how often the model is showing the correct value in its execution of the testing data set. The Accuracy metric is measured Number of predictions to the Total Number of predictions.

$$Accuracy = \frac{Number\ of\ Predictions}{Total\ Number\ of\ Predictions}$$

In comparison, the Xception model gives more accurate results than other models like ResNet 50 and Dense Net.

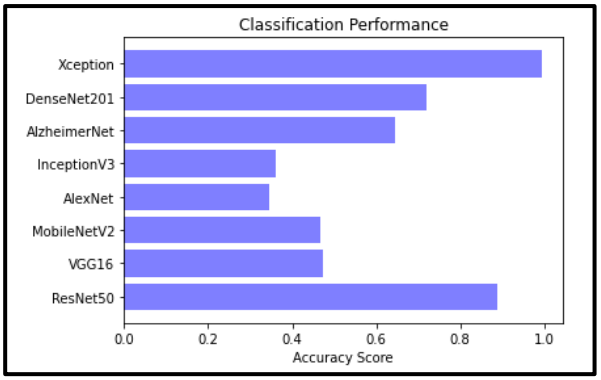

Figure 3: Accuracy Score of all Models

4.2 Precision:

Precision is a metric that is stipulated as the ratio of correctly classified positive samples to the total number of classified positive samples. In comparison to different the models, the Xception model gives more accurate results than other models like ResNet 50 and Dense Net.

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

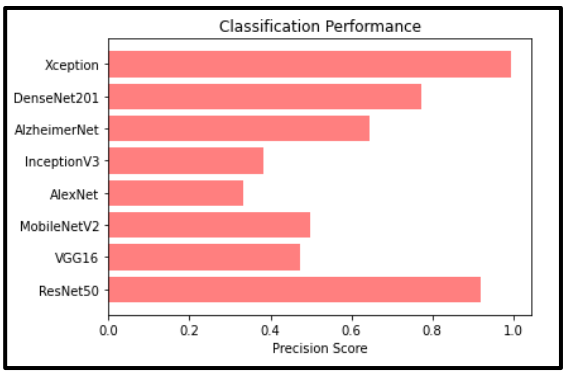

Figure 4: Precision in all ML models

4.3 Recall:

The recall is the metric which detects the ability to positive outcomes. It can be stipulated as the Number of positive samples to the total number of classified positive samples. Higher the recall the accurate the model. Xception model gives a recall value of 99.5% when compared with other models like DenseNet and ResNet 50.

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

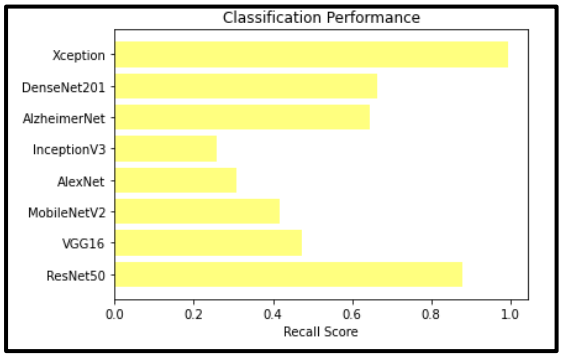

Figure 5: Recall score of all ML models

4.4 F1-Score:

F1 is the metric of the harmonic mean of precision and recall. Value of the F1-score is usually between o and 1. The F1-Score of Xception model gives 99.5% F1-score,when compared to ResNet50(85.8%) DenseNet(70.1%).

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

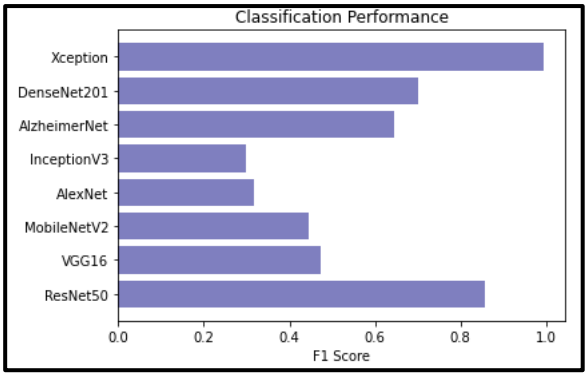

Figure 6: F1-Score of all ML models

The below are the results for the input given as the MRI. The uploaded image is predicted as different stages such as Subjective memory concern(SMC), Mild cognitive impairment(MCI), Early Mild cognitive impairment(EMCI), Late Mild cognitive impairment(LMCI) and Alzheimer disease(AD). The best result of the all the algorithms is produced as the output.

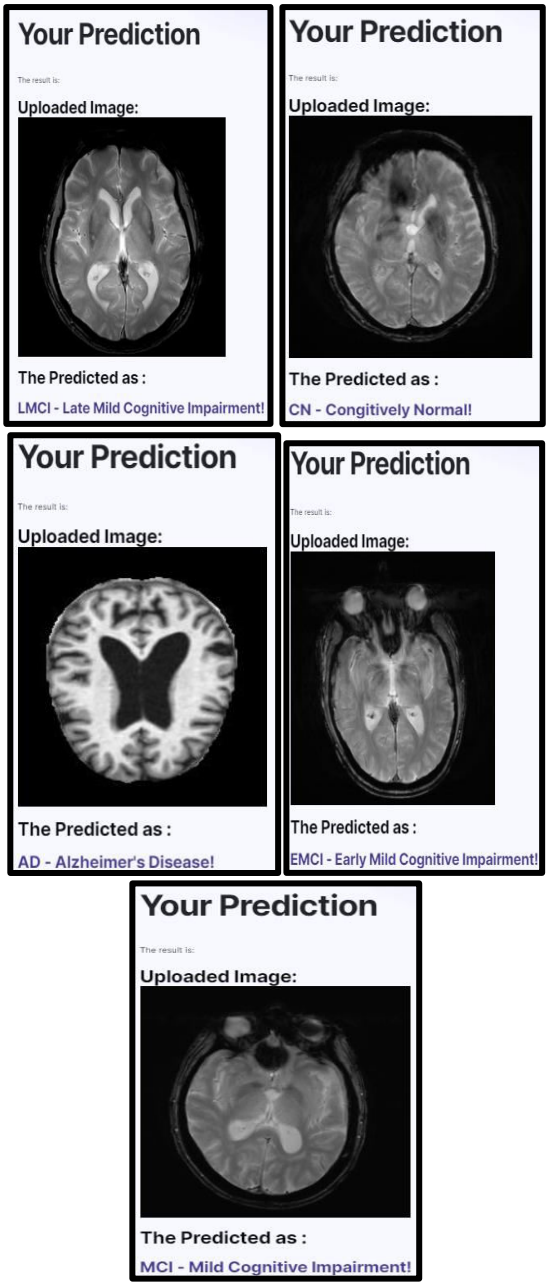

Figure 7: Various stages of Alzheimer Disease

5. CONCLUSION

Alzheimer's Disease is a non-curable disease, where we cannot retrieve the dead neurons in the brain. The disease can lead to death if we don't take proper care and attention. This paper discusses various algorithms like Rest net 50, Dense Net 201, Xception algorithms and its effectiveness to speculate the different levels of Alzheimer's Disease. The best performed model is Xception algorithm with an accuracy of 99.5 with Magnetic Resonance Images(MRI). The classification the MRI helps the radiologists to improve the interpretation of the report and makes the work easy. In future, we can implement the model in commercial health care for better evaluation of Alzheimer's disease. This also makes the work of radiologists and doctor's work easy and efficient.

Conflict-of-interest: The Machine learning algorithmic optimization makes life easier and gives accurate values. The interest here is to do stage classification of Alzheimer's disease to make the diagnosis easier in order to provide accurate treatment and medication.

6.REFERENCES

[1] **"Alzheimer's disease Diagnosis Using Machine Learning:A Survey"** – Omar Asgar Dara; Jose Manuel Lopez Guede; Hassan Issa Raheem; Javad Rahebi; Unnai Fernandez Gamiz; MDPI open access; Published on 18 July 2023.

[2] **"Alzheimer disease diagnosis and classification using deep learning techniques"** –Waaled Al shehn; National library of Medicine, National centre for

Biotechnology Information; published on 20th December 2020.

[3] **"Early stage Alzheimer's Prediction using Machine Learning Models"** – C.Kavitha, Vinodhini Mani, S.R. Srividhya, Carlos Andres, Tavera Romero; National Library of Medicine, National centre for Biotechnology Information; Published on 3rd March 2022.

[4] **"Early Diagnosis of Alzheimer's disease using Machine learning: a Multi diagnostic, generalizable approach"** – Vasco Sa Diogo, Hugo Alexandre feereira, Diana Prata; Biomedical Central, Part of Springer; published on 2022.

[5] **"Alzheimer's Disease Detection by Using Deep Learning Algorithms: A mini review"** –Suhad Al-Shoukry; Taha H Raseem; Nasrin M Makbol; Institute of Electricals and Electronic Engineering(IEEE); Published on 20th April 2020.

[6]**"Xception: Deep Learning With Depthwise Separable Convolutions"** – Francois Chollet; Institute of Electricals and Electronic Engineering(IEEE); Published on 9th November 2017.

[7]**"Predicting Alzheimer's Disease Progression using Deep Recurrent Neural networks"** –Minh Nguyen, Tonh He, Daniel T Alexander, B.T.Thomas Yeo, Jiashi Feng, Lijun An; Elsevier Science Direct Neuro Image Volume 222, Published on November 2020.

[8]**"Classification of Alzheimer's Disease using Convolutional Neural Networks"** –Lamis F. Samhan, Amjad H. Alfarra, Sammy S. Abu Nasser; International Jouranl for Academic Information Systems Research (IJAISR)6(3); 18-23; Published on 2022.

[9]"Evaluation of Neuro Images for the Diagnosis of Alzheimer's Disease using Deep Learning Neural Network" –Ahila A, Poongodi M, Mounir Hamdi, Sami Bourouis, Kulhanek Rastislav, Faizaan Mohamed; Front. Public Health, Volume 10-2022, Published on 2022.

[10]"Deep Convolutional Neural Network Based System for Early Diagnosis of Alzheimer's Disease" – R.R.Janghel, Y.K.Rathore; Volume 42, Issue 4; Published on August 2021.

[11]"**FDCNN-AS: Federated Deep Convolutional Neural Network Alzheimer detection Schemes for different Age Groups"** –Abdullah Lakhan, Mazin Abed Mohammed, Mohd Khanapi Abd Ghani, Karrar Hameed Abdul kareem, Hyder Abdulameer Marhoon, Jan Nedoma, Radek Martinek, Muhammet Devici, Volume 677,120833, Information Sciences, Elseiver, Published on August 2024

[12]"**MRI Deep Learning based Solution for Alzheimer's Disease Prediction"** – Cristina L. Sarataxaga, Iratxe Moya, Artzai Picon, Marina Acosta, Aitor Moreno-Fernandez-de-Lecetta, Estibaliz Garrote, Arantza Bericiartua Perez, Application of Artificial Intelligence in Personalized Medicine, Published on 9th September 2021.

[13]"**Early Prediction of Alzheimer's Disease using Convolutional Neural Network: a review"** –Vijeeta Patel, Manohar Magdi, Ajmeera Kiran, The Egyptian Journal of Neurology, Psychiatry and Neurosurgery, Volume 58 Springer Open, Published on 17th November 2022.

[14]"**An Intelligent Alzheimer Disease Prediction using Convolutional Neural Network"** –L.Darshana Deepthi, D.Shanthi, M.Buvana, International Journal of Advanced Research in Engineering and Technology (IJARET), Volume 11 Issue 4, Published on April 2020.

[15]"**A Comparative Analysis of Machine Learning Algorithms to Predict Alzheimer's Disease"** – Morshedul Bari Antor, A.H.M Shafayat Jamil, Maliha Mamtaz, Mohammad Monirujjaman Khan, Sultan Aljahdali, Manjit Kaur, Parminder Singh, Mehidi Masud, Journal of Health Care, Published on 3rd July 2021.

[16]"**Two stage deep learning Model for Alzheimer's disease Detection and Prediction of the Mild Cognitive Impairment time"** –Shaker El-Sappagh, Hager Saleh, Farman Ali, Eslam Amer, Tamer Abuhmed, Volume 34, Neural Computing and Applications, Springer Link, Published on 8th May 2022.

[17]"**Application of ML and DL Algorithms in the Prediction, Diagnosis and Prognosis of Alzheimer's Disease"** – Zeib Jahangir, Fiza Saeed, Ashish Shiwlani, Muhammad Umar, Volume 22 Issue 6, American Journal of Biomedical Science and Research, Published on 21st June 2024.