ISSN No:-2456-2165

Classifying Alzheimer's Disease and Predicting its Occurrence

Deepa B
Dept. of ECE
Jyothy Institute of Technology
Bangalore, India

Jayasri K R
Dept. of ECE
Jyothy Institute of Technology
Bangalore, India

Prathiksha S
Dept. of ECE
Jyothy Institute of Technology
Bangalore, India

Srushti M C
Dept. of ECE
Jyothy Institute of Technology
Bangalore, India

Dr. Chethana K
Associate Professor, Dept. of ECE
Jyothy Institute of Technology
Bangalore, India

Abstract:- Alzheimer's disease (AD) is the most common type of dementia which is an advancing brain disorder. AD destroys cells in the brain which causes memory related issues and day-to-day functioning. AD sometimes may result in fatal conditions. Since, there is no cure for AD presently, effective control of AD depends on early detection and/or prediction. Mild Cognitive Impairment (MCI) is said to be the precursor phase to AD. This project is divided into two sections. The first deals with detection and classification of AD with respect to MCI and the second section deals with prediction of AD in case if the person of interest is currently non-demented.

Keywords:- *Alzheimer's disease, CNN, Kaggle, MCI, OASIS, Logistic Regression.*

I. INTRODUCTION

One of the most prevalent conditions of dementia is known as the Alzheimer's disease (AD). This disease is a progressive neurodegenerative condition that causes the brain cells to deteriorate and amyloid plaques to form on the brain, which impairs memory, comprehension, orientation, and behavior. According to a recent study, 20% of the people over the age of 80 suffer from Alzheimer's disease. In the present circumstances, it is estimated that India is home to over 4 million people who suffer from dementia.

Considering the scenario worldwide, it is estimated that at the least 44 million people live with dementia around the globe, which makes it an international health crisis that must be focused on.

Initially, the disease affects that part of the brain which oversees functions such as language and memory. Memory loss, confusion, and issues with speaking, reading, or writing may be experienced by AD patients. AD may impair the region of the brain that regulates respiration and circulation, which ultimately may lead to death. Currently, there are no treatments available to slow down or stop the damage on the neurons in the brain caused by Alzheimer's disease. This brings the focus on Mild Cognitive Impairment (MCI) which is a presymptomatic phase of AD. Diagnosis of MCI can help improve the condition of the patient since early medical intervention is possible, which may reduce the number of cases by 30%.

In cases where dementia is not yet established in the brain, prediction algorithms may help in predicting the possibility of occurrence of AD. This helps in making early lifestyle changes so as to prevent dementia.

II. RELATED WORK

In [1], two methods have been applied to detect AD. They are CNN and Transfer learning. A comparison of these methods is done. The results show that CNN gives more efficient results than transfer learning. In [2], a novel 3D CNN was developed to perform the classification of images. The obtained was less when compared to other models that were developed in previous papers. In [3], uses CNN for the detection of AD. A high level of accuracy was achieved using this method. Datasets were downloaded from ADNI datasets. In [6], it was illustrated that only 17 coronal slices from the mid-brain are efficient for classification purposes. Variations between classifications can be shown by extracting learned local features. In [8], very precise methods of deep learning were used to detect Alzheimer's disease. In [10], convolutional encoders are used for the scrutinization of AD.

III. SOFTWARE REQUIREMENTS

The proposed system uses Anaconda software to run the python script on Jupyter Notebook. The libraries used to write the python script is NumPy, Pandas, MatplotLib, OpenCV, TensorFlow, Keras, Seaborn, and Flask.

IV. SOFTWARE IMPLEMENTATION

A. Detection and Classification

The block diagram of the detection and classification model is as shown in Fig 1. The system consists of five stages:

a) Dataset

The dataset used was obtained from the Kaggle website. The dataset consists of two folders namely, training and

testing, both containing a total of around 6400 images which are sorted based on the intensity of the disease. The data has four classes of images based on the dementia stage: Very Mild, Mild, Moderate and Non-demented.

b) Importing necessary libraries

Python language is used for importing and using necessary libraries. Necessary libraries such as Keras, sklearn, PIL, and other libraries such as pandas, NumPy, MatPlotLib, and TensorFlow are used.

c) Fetching the images

Images and their labels are retrieved. They are then resized to (224,224). This is performed to ensure that all the images are of the same size for recognition. The resized images are then converted into a NumPy array.

d) Splitting the dataset

The dataset is split into two folders namely, train and test. It is split in the ratio of 4:1 respectively.

e) Building the model

Sequential models from Keras library are made use of for building the model. Further, layers are added to the model to create a neural network, in this case a convolutional neural network. 32 filters and a kernel size of (5,5) are used to create the first two Conv2D layers. MaxPool2D layer is used for the selection of the maximum value in the area. In the MaxPool2D layer, a pool size of (2,2) is used. Hence the image dimensions are reduced by a factor of 2. The dropout rate in the dropout layer is set to 0.25 which results in 25% of the neurons present in the neural network to be removed randomly. This reduces the complexity and size of the model. The above mentioned layers are implemented additionally by changing the internal parameters. Lastly, a combination of dense and dropout layers are applied following a flatten layer which converts the 2-D data to a 1D vector.

f) Saving the trained model

The built model is then saved into a .h5 file.

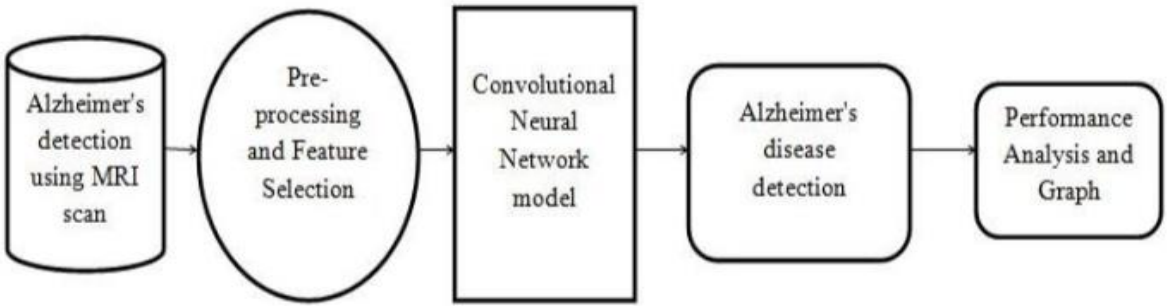

Fig 1: Block Diagram of classification system

B. Prediction

Fig 2 shows the block diagram of the prediction system. The system consists of five stages:

a) Acquiring dataset:

The dataset was obtained from OASIS. The OASIS_Longitudinal dataset is a .csv file which comprises 373 entries with 14 psychological features for each entry. The features of interest are gender, age, Socio-Economic Status (SES), Mini-Mental State Examination (MMSE), Clinical Dementia Rating (CDR), estimated Total Intracranial Volume (eTIV), normalised Whole Brain Volume (nWBV) and Atlas Scaling Factor (ASF).

b) Pre-processing the dataset

Data-mining methods were used to preprocess the dataset. In this stage, missing values were replaced with average values; features were extracted and transformed. 9 Missing values were found in the SES column. The missing values can be replaced by either dropping entire rows or by filling them.

c) Splitting the dataset

The dataset was split in the ratio 7:3, where 70% was allotted for training and 30% was allotted for testing the model.

d) Building the model

After the dataset was preprocessed and split, the model was built using Logistic Regression (LR). To build the LR model, the correlation coefficients between the features in the dataset and dementia were found before analyzing the dataset. Features with high correlation were extracted.

e) Giving inputs to the model

Features as provided as input to the trained model and output is observed.

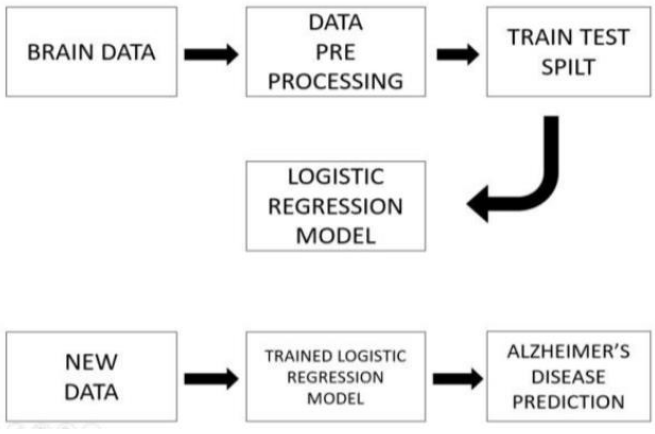

Fig 2: Block diagram of prediction system

ISSN No:-2456-2165

V. RESULTS

A. Detection and Classification Results

The model has testing accuracy and validation accuracy of 86.34% and 86.45% respectively. The user interface used to provide input to the model is shown in Fig 3. Fig 4 shows the results for one image from class when given as input.

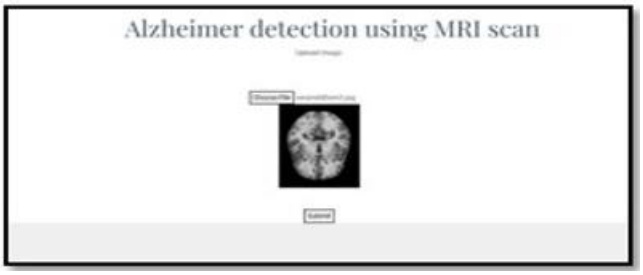

Fig 3: User input interface

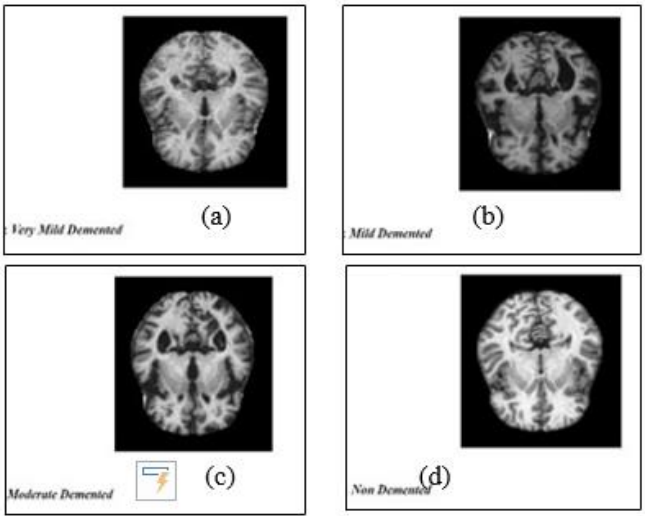

Fig 4: (a) Very Mild Demented (b) Moderate Demented (c) Mild Demented (d) Non-Demented

B. Prediction Results

The prediction model has training and testing accuracy of 93% and 96% respectively. Fig 5 shows the results obtained when input features are provided to the model.

```
0,68,12,2,27,0.5,1457,0.806,1.205
The Person is Demented
```

(a)

```
1,88,18,3,28,0,1215,0.710,1.444
The `Person is Non-Demented
```

(b)

Fig 5: (a) Demented (b) Non-Demented

The confusion matrix for the prediction results is plotted and is shown in Fig 6. It can be concluded from the matrix that there are 57 True-Negatives, 0 False-Positives, 4FalseNegatives and 51True-Positives.

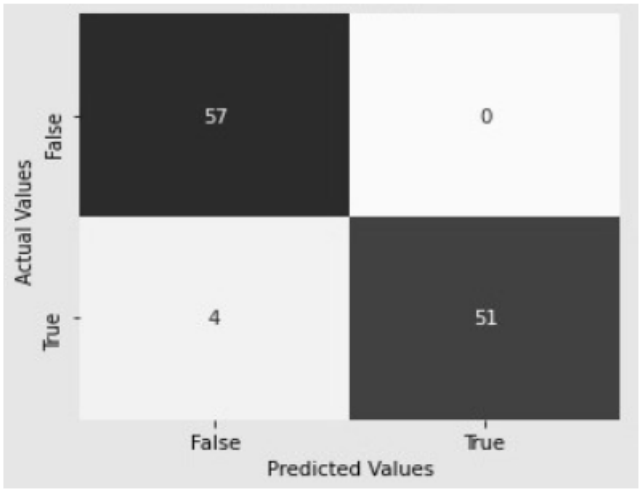

Fig 6: Confusion matrix for prediction results

VI. CONCLUSION

In the first section of this project, we have used basic Convolutional Neural Network (CNN) architecture model to classify AD from MRI scans. The CNN model was used to eliminate the complex procedure involved in training the model from the top. CNN is also used in order to achieve higher efficiency even with a finite dataset. The proposed model produced a training efficiency and validation accuracy of 86.34% and 86.45% respectively on the test data. Very minimal errors were found in the classification of non-demented and very mild demented cases. Future work can be performed on including data from other sources such as .csv file, PET scan results, and fMRI scans.

In the second section of this project, Logistic Regression was used to build a predictive model which predicts the occurrence of Alzheimer's disease beforehand. A total of 8 factors were used to predict AD which resulted in training and testing accuracy of 93% and 96% respectively. Future work includes considering large datasets and predicting AD before 3-5years with sufficient accuracy.

REFERENCES

[1]. M. Zaabi, N. Smaoui, H. Derbel and W. Hariri, "Alzheimer's disease detection using convolutional neural networks and transfer learning based methods," 2020 17th International Multi-Conference on Systems, Signals & Devices (SSD), 2020, pp. 939-943, doi: 10.1109/SSD49366.2020.9364155

[2]. Pan D, Zeng A, Jia L, Huang Y, Frizzell T, Song X. "Early Detection of Alzheimer's disease Using Magnetic Resonance Imaging: A Novel Approach Combining Convolutional Neural Networks and Ensemble Learning". Front Neurosci. 2020:14:259. Published 2020 May 13. doi:10.3389/fnins.2020.00259

[3]. L. Yue et al., "Auto-Detection of Alzheimer's Disease Using Deep Convolutional Neural Networks," 2018 14th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2018, pp. 228-234, doi: 10.1109/FSKD.2018.8687207

[4]. Feng W, Halm-Lutterodt NV, Tang H, Mecum A, Mesregah MK, Ma Y, Li H, Zhang F, Wu Z, Yao E, Guo X. "Automated MRI-Based Deep Learning Model for Detection of Alzheimer's Disease Process". Int J Neural Syst. 2020 Jun;30(6):2050032. doi: 10.1142/S012906572050032X. PMID: 32498641
[5]. K. A. N. N. P. Gunawardena, R. N. Rajapakse and N. D. Kodikara, "Applying convolutional neural networks for pre-detection of Alzheimer's disease from structural MRI data," 2017 24th International Conference on Mechatronics and Machine Vision in
[6]. Practice (M2VIP), 2017, pp. 1-7, doi: 10.1109/M2VIP.2017.8211486
[7]. C. D. Billones, O. J. L. D. Demetria, D. E. D. Hostallero and P. C. Naval, "DemNet: A Convolutional Neural Network for the detection of Alzheimer's Disease and Mild Cognitive Impairment," 2016 IEEE Region 10 Conference (TENCON), 2016, pp. 3724-3727, doi: 10.1109/TENCON.2016.7848755
[8]. Sheng Liu, Chhavi Yadav, Carlos Fernandez-Granda, Narges Razavian, "On the design of convolutional neural networks for automatic detection of Alzheimer's disease", Proceedings of Machine Learning Research 116:184–201, 2020 Machine Learning for Health (ML4H) at NeurIPS 2019
[9]. Emre Altinkaya, Kemal Polat, Burhan Barakli (2019). "Detection of Alzheimer's Disease and Dementia States Based on Deep Learning from MRI Images: A Comprehensive Review". Journal of the Institute of Electronics and Computer, 1, 39-53. https://doi.org/10.33969/JIEC.2019.11005
[10]. Helaly, H.A., Badawy, M. & Haikal, A.Y. Deep Learning Approach for Early Detection of Alzheimer's disease. Cogn Comput (2021). https://doi.org/10.1007/s12559-021-09946-2
[11]. Kavitha C, Mani V, Srividhya SR, Khalaf OI, Tavera Romero CA. Early-Stage Alzheimer's Disease Prediction Using Machine Learning Models. Front Public Health. 2022 Mar 3;10:853294. doi: 10.3389/fpubh.2022.853294. PMID: 35309200; PMCID: PMC8927715
[12]. J. Neelaveni and M. S. G. Devasana, "Alzheimer Disease Prediction using Machine Learning Algorithms," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 2020, pp. 101-104, doi: 10.1109/ICACCS48705.2020.9074248