International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
ISSN: 2457-0362

DEEP LEARNING FOR PRECISE ALZHEIMER'S DISEASE CLASSIFICATION USING NEUROIMAGING DATA

1. **Banoth Krishna Mohan Naik**, Department of Information Technology, Jawaharlal Nehru Technological University, krishnamohannaik2@gmail.com
2. **Dr. G Venkata Rami Reddy**, Professor, Department of Information Technology , Jawaharlal Nehru Technological University, gvr_reddi@jntuh.ac.in
3. **Mr. K. Balakrishna Maruthiram**, Assistant Professor, Department of Information Technology, Jawaharlal Nehru Technological University, kbkmram@jntuh.ac.in

Abstract: Alzheimer's disease is an unrepairable degenerative brain disease. Every four seconds, someone in the world is diagnosed with Alzheimer's disease. The result is fatal, as it leads to death. As a result, it's crucial to catch the disease early on. The leading cause of dementia is Alzheimer's disease. Dementia causes a reduction in reasoning abilities and interpersonal coping skills, which affects people's ability to function independently. The patient will forget recent events in the early stages. If the illness progresses, they will gradually forget whole events. It is essential to diagnose the disease as soon as possible. This paper proposes a model that takes brain MRI sample images as input and determines whether a person has mild, moderate, or no Alzheimer's disease as an output. We are using the ResNet50, Inception V3,Inception ResNet V2,DenseNet and MobileNet architectures for this classification, providing a comparative analysis of which architecture shows promising results.

Index Terms: *ResNet50, Inception V3,Inception ResNet V2,DenseNet and MobileNet*.

1. INTRODUCTION

The system architecture shows us a conceptual and behavioral view of the system. It is just a view that shows us how the database is used in taking the dataset and then how this data is used up in our project modules to train the different models. In the architecture diagram above, we can see the data is being taken from the training dataset and then provided to the models. Then it is validated against the test dataset to get the testing or validation accuracy. After the accuracy is compared, the diseased images are taken from the dataset, The classification done is of four types namely Mild Demented, Moderate Demented, Non Demented and Very Mild Demented. Also, the architecture diagram shows us the various modules working together in the project and how they are integrated to provide us the desired output, and how all the modules need to be interconnected to make the project work in unison.

This research compares two state-of-the-art deep learning models' detection accuracy in detecting Alzheimer's disease in an MRI image. The Keras module of tensor flow, an opensource library for implementing deep learning models, is used to implement VGG19. Using the Image Data Generator function, the data was augmented and loaded into the model. The training consisted of a batch size of 128, with 50 epochs with early stopping. Similarly, the Keras module is used to implement the Densenet Model, and the data is loaded into the densenet model via the Image Data Generator function. The densenet model is trained using a batch of 128 images each. Both models are trained on 3048 MRI images consisting of four classes, and both models are tested on a total of 2067 MRI images. All the work was done on Google colab.

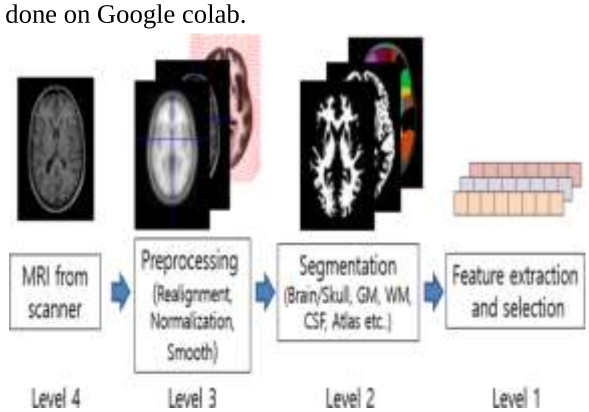

Fig 1 Example Figure

Machine Learning approach using biomarkers in their paper. They tested a personalized classifier for the disease using a method learning locally weighted and biomarkers. The methodology attempts to classify the subject first and then later decides which biomarker to order. They classified the patients who were MCI who advanced to AD inside a year against the individuals who didn't.

Three different neuroimaging techniques, EEG, MRI, and PET, allow us to explore and measure the insane amounts of activity going on in our brain; however, each comes with its own strengths and limitations, making the motivations behind using them very important. Commonly used brain imaging techniques are: functional magnetic resonance imaging (fMRI) computerized tomography (CT) positron emission tomography (PET). The technique that presently has the greatest spatial and temporal resolution is functional magnetic resonance imaging (fMRI), which relies on differences in the magnetic susceptibility of oxygenated and deoxygenated blood.

Neuroimaging falls into two broad categories:

- Structural imaging, which is used to quantify brain structure using e,g, voxel based morphometry.
- Functional imaging, which is used to study brain function, often using fMRI and other techniques such as PET and MEG

Neuroimaging includes the use of various techniques to either directly or indirectly image the structure, function/pharmacology of the brain. It is a relatively new discipline within medicine and neuroscience/psychology. The first neuroimaging technique ever is the so-called 'human circulation balance' invented by Angelo Mosso in the 1880s and able to non-invasively measure the redistribution of blood during emotional and intellectual activity. Then, in the early 1900s, a technique called pneumoencephalography was set. Several types of neuroimaging tools can help diagnose and treat mental illness. While these tools are powerful for detecting certain neurological conditions, they are limited in diagnosing mental illnesses like depression. Instead, they're mainly used to rule out other medical conditions.

2. LITERATURE REVIEW

Alzheimer Disease Detection Based on Deep Neural Network with Rectified Adam Optimization Technique using MRI Analysis.

Alzheimer is a memory depletion disease, which is widely recognized as dementia. The research on early detection of dementia has received huge interest among the researchers to help in reducing mortality rates of Alzheimer's patients. In recent years in the medical field, the deep learning techniques play an important role in computer aided diagnosis. In this research, the automatic recognition of Alzheimer Disease (AD) based on the Magnetic Resonance Imaging (MRI) is accomplished by implementing an unsupervised classification technique named Deep Neural Network (DNN) with the rectified Adam optimizer. At first, Histogram of Oriented Gradients (HOG) is utilized to extract the feature values from the brain images, which were acquired from National Institute of Mental Health and Neurosciences (NIMHANS) and Alzheimer disease Neuroimaging Initiative (ADNI) datasets. Next, the extracted features were given as the input to DNN with the rectified Adam optimizer to distinguish the healthy, AD and Mild Cognitive Impairment (MCI) patients. The experimental results have revealed that the HOG-DNN with the rectified Adam optimizer has achieved better performance in AD recognition and showed 16% enhancement in classification accuracy compared to other existing work; Landmark based features with support vector machine classifier.

An Improved Multi-Modal based Machine Learning Approach for the Prognosis of Alzheimer's disease

Alzheimer's disease (AD) is the most common type of neurological disorder that leads to the brain's cell death overtime. It is one of the major important causes of memory loss and cognitive decline in elderly subjects around the globe. Early detection and streamlining of diagnostic practices are the prime domains of the interest to the healthcare community. Machine learning (ML) algorithms and numerous multivariate data exploratory tools have been extensively used in the field of AD research. The primary purpose of this study is to present an automated classification system to retrieve information patterns. We proposed a five-stage ML pipeline, where each stage was further categorized in different sub-levels. The study relied on the Open Access Series of Imaging Studies (OASIS) database of MRI (Magnetic Resonance Imaging) brain images for the analysis. The dataset comprised of 343 MRI sessions involving 150 subjects. Three different scores namely, MMSE (Mini-Mental State

Examination), CDR (Clinical Dementia Rating), and ASF (Atlas Scaling Factor) were used in the analysis. The proposed ML pipeline constitutes a classifier system along with data transformation and feature selection techniques that have been embedded inside an experimental and data analysis design. Performance metrics for Random Forest (RF) classifier showed the highest output in the classification accuracy.

Usage Of Random Forest Ensemble Classifier Based Imputation And Its Potential In The Diagnosis Of Alzheimer's Disease.

To evaluate and compare the performance of Random Forest (RF) ensemble classifier in imputation and non-imputation method of missing data values, and its impact to diagnose Alzheimer's disease (AD) based on longitudinal MRI data. Method: We studied 373 MRI sessions involving 150 AD subjects aged 60 to 90 years [Mean age ± SD = 77.01 ± 7.64]. T1-weighted MRI of each subject on a 1.5-T Vision scanner were used for the image acquisition. The MRI dataset was taken from OASIS (Open Access Series of Imaging Studies) database. Based upon the MRI acquitted features in the dataset, we applied missing data imputation using RF ensemble to classify the subjects as demented or non-demented. We then compared them to determine which is more precise in the AD diagnosis Result: RF model-based imputation analysis outperforms with better accuracy than RF non-imputation method.

A multi- modal, multi- atlas- based approach for Alzheimer detection via machine learning

The use of biomarkers for early detection of Alzheimer's disease (AD) improves the accuracy of imaging-based prediction of AD and its prodromal stage that is mild cognitive impairment (MCI). Brain parcellation-based computer-aided methods for detecting AD and MCI segregate the brain in different anatomical regions and use their features to predict AD and MCI. Brain parcellation generally is carried out based on existing anatomical atlas templates, which vary in the boundaries and number of anatomical regions. This works considers dividing the brain based on different atlases and combining the features extracted from these anatomical parcellations for a more holistic and robust representation. We collected data from the ADNI database and divided brains based on two well-known atlases: LONI Probabilistic Brain Atlas (LPBA40) and Automated Anatomical Labeling (AAL). We used baselines images of structural magnetic resonance imaging (MRI) and ^{18}F-fluorodeoxyglucose positron emission tomography (FDG-PET) to calculate average gray-matter density and average relative cerebral metabolic rate for glucose in each region. Later, we classified AD, MCI and cognitively normal (CN) subjects using the individual features extracted from each atlas template and the combined features of both atlases. We reduced the dimensionality of individual and combined features using principal component analysis, and used support vector machines for classification. We also ranked features mostly involved in classification to determine the importance of brain regions for accurately classifying the subjects. Results demonstrated that features calculated from multiple atlases lead to improved performance compared to those extracted from one atlas only.

Alzheimer disease classification using KPCA, LDA, and multi- kernel learning SVM

Early diagnosis of Alzheimer disease (AD) and mild cognitive impairment (MCI) is always useful. Preventive measures might have an impact on reducing AD risk factors. Structural magnetic resonance (MR) imaging, one of the vital sensitive biomarkers for cerebral atrophy in the brain, is used to extract volumetric feature by FreeSurfer and the CIVET toolbox. All of the structural magnetic resonance imaging (s-MRI) data that we used were downloaded from the Alzheimer's disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) of imaging data. This novel approach is applied for the diagnosis of AD and MCI from healthy controls (HCs) combining extracted features with the MMSE (mini-mental state examination) scores, applying a two sample t-test to select a subset of features. The subset of features is fed to kernel principal component analysis (KPCA) module to project data onto the reduced principal component coefficients at higher dimensional kernel space to increase the linear separability. Then, the kernel PCA coefficients are projected into the more efficient linear discriminant space using linear discriminant analysis. A multi-kernel learning

International Journal For Advanced Research
In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

support vector machine (SVM) is used on newly projected data for stratification of AD and MCI from HCs. Using this approach, we obtain 93.85% classification accuracy when detecting AD from HCs for segmented volumetric features (using FreeSurfer) with high sensitivity and specificity. When distinguishing MCI from HCs and AD using volumetric features after subcortical segmentation, the detection rate reaches 86.54% and 75.12%, respectively.

3. METHODOLOGY

Escudero et al. proposed a ML approach using biomarkers in their paper. They tested a personalized classifier for the disease using a method learning locally weighted and biomarkers. The methodology attempts to classify the subject first and then later decides which biomarker to order. They classified the patients who were MCI who advanced to AD inside a year against the individuals who didn't.

Drawbacks:

- Such types can be vigorously standardized in human vision but need significant advances in figuring out how to dodge perils from overlooked and underrepresented measurable blunders.

Deep Learning is known to be learning a hierarchical set of representations such that it learns low mid and high-level features. Deep neural networks can adapt to more complex data sets. It's better in generalizing previously unseen data because of its multiple layers. Different algorithms use Deep Learning's fundamental expertise and use diverse datasets to train and test these algorithms. Like neurons in humans, deep learning has layers that help the model or algorithm learn and process the data. These layers process the data given to them as the input and learn by processing the input while traveling through the layers. When it passes out the last layer, an activation function is applied at last, and finally, we get the predicted output from the deep learning model. This gives us the training accuracy, and then when we take another similar type of dataset, we can predict or detect from the learned deep learning model whatsoever we want. Well, this is what deep learning does in simple working terms.

Benefits:

- Deep neural networks can adapt to more complex data sets. It's better in generalizing previously unseen data because of its multiple layers.
- we can predict or detect disease from the learned deep learning model.

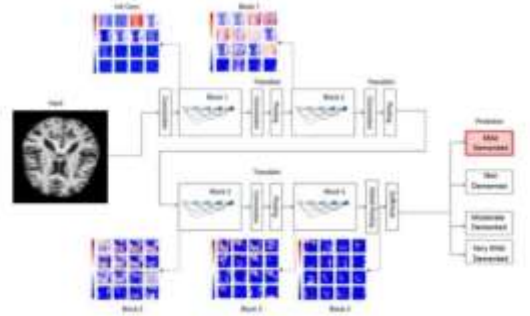

Fig 2 Proposed Architecture

Modules:

1.Data Collection

The composition of the dataset.understand the relationship among different features. A plot of the core features and the entire dataset.The dataset is further split into 2/3 for training and 1/3 for testing the algorithms. Furthermore, in order to obtain a representative sample, each class in the full dataset is represented in about the right proportion in both the training and testing datasets. The various proportions of the training and testing datasets used in the paper.

2.Data Preprocessing

The data which was collected might contain missing values that may lead to inconsistency. To gain better results data need to be preprocessed so as to improve the efficiency o the algorithm. The outliers have to be removed and also variable conversion need to be done. In order to overcoming these issues we use map function.

3.Model Selection

Machine learning is about predicting and recognizing patterns and generate suitable results after understanding them. ML algorithms study patterns in data and learn from them. An ML model will learn and improve on each attempt. To gauge the effectiveness of a model, it's vital to split the data into training and test sets first. So before training our models, we split the data into Training set which was 70% of the whole dataset and Test set which was the remaining 30%. Then it was important to implement a selection of performance metrics to the predictions made by our model.

4.Predict the results

The total number of features within the bank credit Defaulters dataset. However, not all have significant influence in determining the ability of a given customer in paying his/her loan or not. The designed system is tested with test set and the performance is assured. Evolution analysis refers to the description and model regularities or trends for objects whose behavior changes over time. Common metrics calculated from the confusion matrix are Precision; Accuracy. The mot important features since these features are to develop a predictive model using ordinary DenseNetmodel.

4. IMPLEMENTATION

Algorithms

DenseNet

DenseNet (Dense Convolutional Network) is an architecture that focuses on making the deep learning networks go even deeper, but at the same time making them more efficient to train, by using shorter connections between the layers. DenseNet is a convolutional neural network where each layer is connected to all other layers that are deeper in the network, that is, the first layer is connected to the 2nd, 3rd, 4th and so on, the second layer is connected to the 3rd, 4th, 5th and so on. This is done to enable maximum information flow between the layers of the network. To preserve the feed-forward nature, each layer obtains inputs from all the previous layers and passes on its own feature maps to all the layers which will come after it. Unlike Resnets it does not combine features through summation but combines the features by concatenating them. So the 'ith' layer has 'i' inputs and consists of feature maps of all its preceding convolutional blocks. Its own feature maps are passed on to all the next 'I-i' layers. This introduces '(I(I+1))/2' connections in the network, rather than just 'I' connections as in traditional deep learning architectures. It hence requires fewer parameters than traditional convolutional neural networks, as there is no need to learn unimportant feature maps. DenseNet consists of two important blocks other than the basic convolutional and pooling layers. they are the Dense Blocks and the Transition layers.

ResNet:

A residual neural network (ResNet) is an artificial neural network (ANN) of a kind that builds on constructs known from pyramidal cells in the cerebral cortex. Residual neural networks do this by utilizing skip connections, or shortcuts to jump over some layers. Typical ResNet models are implemented with double- or triple- layer skips that contain nonlinearities (ReLU) and batch normalization in between.[1] An additional weight matrix may be used to learn the skip weights; these models are known as HighwayNets.[2] Models with several parallel skips are referred to as DenseNets.[3] In the context of residual neural networks, a non-residual network may be described as a plain network. A reconstruction of a pyramidal cell. Soma and dendrites are labeled in red, axon arbor in blue. (1) Soma, (2) Basal dendrite, (3) Apical dendrite, (4) Axon, (5) Collateral axon. There are two main reasons to add skip connections: to avoid the problem of vanishing gradients, or to mitigate the Degradation (accuracy saturation) problem; where adding more layers to a suitably deep model leads to higher training error.[1] During training, the weights adapt to mute the upstream layer[clarification needed], and amplify the previously-skipped layer. In the simplest case, only the weights for the adjacent layer's connection are adapted, with no explicit weights for the upstream layer. This works best when a single nonlinear layer is stepped over, or when the intermediate layers are all linear. If not, then an explicit weight matrix should be learned for the skipped connection (a HighwayNet should be used). Skipping effectively simplifies the network, using fewer layers in the initial training stages[clarification needed]. This speeds learning by reducing the impact of vanishing gradients, as there are fewer layers to propagate through. The network then gradually restores the skipped layers as it learns the feature space. Towards the end of training, when all layers are expanded, it stays closer to the manifold[clarification needed] and thus learns faster. A neural network without residual parts explores more of the feature space. This makes it more vulnerable to perturbations that cause it to leave the manifold, and necessitates extra training data to recover.

MobileNet

International Journal For Advanced Research In Science & Technology
A peer reviewed international journal
www.ijarst.in
IJARST
ISSN: 2457-0362

MobileNet uses depthwise separable convolutions. This convolution block was at first introduced by Xception. A depthwise separable convolution is made of two operations: a depthwise convolution and a pointwise convolution.

A standard convolution works on the spatial dimension of the feature maps and on the input and output channels. It has a computational cost of $Df^2 * M * N * Dk^2$; with Df the dimension of the input feature maps, M and N the number of input and output channels, and Dk the kernel size.

A depthwise convolution maps a single convolution on each input channel separately. Therefore its number of ouput channel is the same of the number of input channel. It computational cost is $Df^2 * M * Dk^2$.

5. EXPERIMENTAL RESULTS

Fig 3 Home Page

Fig 4 Signup Page

Fig 5 Signin Page

Fig 6 Upload image

Fig 7 Prediction Result

6. CONCLUSION

Alzheimer's disease is the leading cause of dementia. This paper determines a prospective solution for detecting the disease at an early stage. The models used in this paper have successfully classified the images into the appropriate four classes and indeed provided us with promising results. We observe that DenseNet performs better than ResNet50,Inception V3,Inception Resnet V2&Mobilenet. Further research is required to ensure that this particular model can be implemented in clinical settings, increasing the health care rate against this specific disease. Knowledge should be spread among people regarding this disease, and they should be encouraged to get themselves examined. We are currently working on deploying this model onto a website for better practical usage.

7. FUTURE SCOPE

In the future, this model can also be tested on a larger dataset. We had only 52 and 12 images for training and testing in the current dataset for the 'Moderate Demented' class. The proposed model can help doctors diagnose Alzheimer's Disease more effectively and can be modified to identify other Neurodegenerative Diseases more automatically in the future.

REFERENCES

[1] Suresha, HalebeeduSubbaraya, and SrirangapatnaSampathkumaran Parthasarathy. "Alzheimer Disease Detection Based on Deep Neural Network with Rectified Adam Optimization Technique using MRI Analysis." 2020 Third International Conference on Advances in Electronics, Computers and Communications (ICAECC), pp. 1-6. IEEE, 2020.

[2] Deng, Lan, and Yuanjun Wang. "Hybrid diffusion tensor imaging feature-based AD classification." Journal of X-Ray Science and Technology Preprint, 2020, pp. 1-19.

[3] Khan, Afreen, and Swaleha Zubair. "An Improved Multi-Modal based Machine Learning Approach for the Prognosis of Alzheimer’s disease." Journal of King Saud University-Computer and Information Sciences, 2020.

[4] Khan, Afreen, and Swaleha Zubair. "Usage Of Random Forest Ensemble Classifier Based Imputation And Its Potential In The Diagnosis Of Alzheimer's Disease." Int. J. Sci. Technol. Res. 8, no. 12, 2019, pp. 271- 275.

[5] Asim, Yousra, Basit Raza, Ahmad Kamran Malik, Saima Rathore, Lal Hussain, and Mohammad Aksam Iftikhar. "A multi- modal, multi- atlas-based approach for Alzheimer detection via machine learning." International Journal of Imaging Systems and Technology 28, no. 2, 2018, pp. 113-123.

[6] Alam, Saruar, Goo- Rak Kwon, and Alzheimer's Disease Neuroimaging Initiative. "Alzheimer disease classification using KPCA, LDA, and multi- kernel learning SVM." International Journal of Imaging Systems and Technology 27, no. 2, 2017, pp. 133-143.

[7] Khajehnejad, Moein, ForoughHabibollahiSaatlou, and HodaMohammadzade. "Alzheimer's disease early diagnosis using manifoldbased semi-supervised learning." Brain sciences 7, no. 8, 2017, p. 109.

[8] Lama, Ramesh Kumar, JeonghwanGwak, Jeong-Seon Park, and SangWoong Lee. "Diagnosis of Alzheimer's disease based on structural MRI images using a regularized extreme learning machine and PCA features." Journal of healthcare engineering 2017, 2017.

[9] Bryan, R. Nick. "Machine learning applied to Alzheimer disease." , 2016, pp. 665-668.

[10] Escudero, Javier, Emmanuel Ifeachor, John P. Zajicek, Colin Green, James Shearer, Stephen Pearson, and Alzheimer's Disease Neuroimaging Initiative. "Machine learning-based method for personalized and costeffective detection of Alzheimer's disease." IEEE transactions on biomedical engineering 60, no. 1, 2012, pp. 164-168