

CATEGORIZATION AND COMPARISON OF ALZHEIMER'S DISEASE USING ML & CNN

P. Vineela¹, Shivani², S. Roshni varma³, Mr. M. Kalyan Ram⁴

^{1,2,3}Student, CSE, B. Tech, Aditya Engineering College, Surampalem

⁴Guide, Associate Professor, MCA, M. Tech (PhD), Aditya Engineering College, Surampalem

ABSTRACT

Alzheimer's disease, which falls under the category of dementia, is also known as senile dementia. Memory loss and difficulty performing simple activities cognitively are the most common symptoms. Despite the disease's prevalence, therapy is sometimes delayed because there aren't any obvious biomarkers. There are extremely few patients that undergo a proper diagnosis and obtain the right treatment. Many folks typically acquire a diagnosis when it's too late. Since there is no known treatment for this illness, the only ways to delay its onset are early detection and prevention action. When used to examine the disease's biomarkers, both machine learning and deep neural network algorithms are effective. The likelihood of underlying AD is determined by biomarkers like as plaques and tangles in the grey matter as well as historical trends noticed by the algorithm. Both numerical and neuroimaging data can be used for the analysis. SVM, Random Forest, and Regression are examples of machine learning algorithms that use numerical data to represent the percentage change in the hippocampus and entorhinal cortex. CNN (convolutional neural network), a deep learning algorithm, uses T1W images from MRI scans to classify Alzheimer's disease into mild, moderate, and severe forms.

Keywords – Alzheimer's Disease, ADNI, Machine Learning Algorithms, Convolutional Neural Network

1. INTRODUCTION

The primary diagnosis includes discussing the disease's initial onset with the family, reviewing medical history, cognitive testing that assesses problem-solving skills, routine memory-loss screening, routine MRI, and image-scanning which is used to determine the presence of atypical dementia. The final way to determine whether someone has Alzheimer's disease is by a lumbar puncture or amyloid. Overall, it is a laborious task. However, adopting an ML and CNN model speeds up the process of disease diagnosis and increases the effectiveness of its detection. To get the features, the data needs to be cleaned up and pre-processed using the neuroimaging studies in the dataset.

When these features are combined in the appropriate model, the outcome classifies the severity of AD. The strategy makes use of several of CNN and machine learning methodologies. An effective board to choose the best model to categories the condition is produced by comparing the two models. The models included in CNN and machine learning are VGG16, VGG19, resnet50, and resnet101. Due to its several channels, CNN is strong in categorizing images from one another. The models' individual efficacy is evaluated after fitting. Thus, a result is generated that shows whether the comparison was successful.

2. LITERATURE REVIEW

We investigated 165 papers from the years 2005 to 2019 using different feature extraction and machine learning methods. SVM, ANN, and DL, three significant machine learning algorithms for diagnosing Alzheimer's disease, were investigated by the authors. Additional learning strategies like transfer, ensemble learning, and multi-kernel learning have also been studied. Alzheimer's disease has been successfully treated using SVM-based models, proving its effectiveness. This is because SVM does not suffer from the drawbacks of local minima, unlike methods like ANN. SVM is used frequently because it is more understandable than deep neural networks, which are opaque models. Future research on this issue should concentrate on the ability of deep learning models to be used in clinical settings. Furthermore, researchers have been found to have given the feature extraction stage more thought than the categorization stage. [1] There are two groups of patients in this paper. 1715 participants total, including those from ADNI, make up the initial group. The Parelino Neurodegenerative Disease Biobank of RI Health was used to select the second group. Using the N4 technique to correct the pictures, they are then translated into MNI space with brain masks displaying analogous modifications. The model was constructed using SVM and CNN techniques. On the training set for SVM, the C parameter is utilised five times. The classifier implementation was enabled by Scikit-learn. CNN, on the other hand, employed FCN architecture, which uses two pooling layers similarly to many CNN. The filters are one of seven models that make up the entire network. A 3D convolution layer, dropout, batch normalisation (BN), a ReLu activation function, then another 3D layer, and so on make up the network's sequential layers.. These layers are used in the categorization and comparison of Alzheimer's disease using machine learning and CNN. Another such mechanism that showed a change in the prediction score was saliency maps. The outcome thus obtained demonstrated that the modified GM maps AUC had larger latency than the T1w images AUC. Similar results were seen for CNN, with GM maps

yielding more accuracy than the T1w pictures. [2]A 3D convolution layer, dropout, batch normalisation (BN), a ReLu activation function, then another 3D layer, and so on make up the network's sequential layers. Much more helpful than a group analysis would be a model for categorising each subject separately. The author's goal was to assign each patient to a hierarchically organised diagnosis category using a machine learning-based classification approach and surface-based cortical thickness data. The individual eventually received one of the final clinical designations as a result of this hierarchy being applied across the tree. Each participant was successfully classified into one of five clinical groups with 75.8 percent accuracy, according to classification findings using the entire hierarchical tree. [3]Convolutional neural networks (CNNs) for patch-based deep feature learning were presented by Liu, M., Zhang, J., Nie, D., and Shen (using data to find discriminative anatomical landmarks in MR images). They put out an anatomical landmark-based deep feature learning (LDL) technique for illness prediction. Our suggested method has a general structure made up of four main parts: 1) Finding landmarks; 2) extracting landmark-based patches; 3) learning patch-based features; and 4) applications for disease categorization and image retrieval. In this article, they proposed a deep feature learning method based on landmarks for automatically deriving patch-based representations from MR images for the identification of brain illnesses associated with Alzheimer's disease. This method opens the door for discriminative biomarkers in morphological MR image processing and computer-aided Alzheimer's disease diagnosis. [4]

3. METHODOLOGY

VGG16:For image classification tasks, the VGG16 model is a deep Convolutional Neural Network (CNN) architecture. It consists of 16 layers, including convolutional layers, max pooling layers, and fully connected layers.To use VGG16 for categorization and comparison of Alzheimer's disease using machine learning and CNN, the following steps can be taken:

1. Collect and pre-process the dataset: The first step is to collect and pre-process the dataset. The collection should include both healthy and Alzheimer's disease patient brain pictures. The photos should be normalised and scaled to a standard size.
2. Split the dataset: To create training, validation, and testing sets, divide the dataset. The validation set is used to fine-tune the model's hyperparameters, the testing set is used to assess the model's performance, and the training set is used to train the model.
3. Train the VGG16 model: The VGG16 model must then be trained using the training set. A softmax layer should be added as the last layer of the model to divide the photos into two groups: those with Alzheimer's disease and those without the condition.
4. Fine-tune the VGG16 model: After training the model on the training set, the model should be fine-tuned on the validation set to optimize its performance. This involves adjusting the hyperparameters of the model, such as the learning rate, batch size, and number of epochs.
5. Evaluate the performance of the model: Finally, the model's performance on the testing set needs to be assessed. One can assess the model's performance using the accuracy, precision, recall, and F1-score.

Hardware requirements:

- Hard Disk - 500GB
- Monitor
- RAM - 8GB
- Input Devices - Keyboard, Mouse

Software requirements:

- Google Colab, Visual Studio Code
- Windows 10/Windows 8
- Python
- Libraries Used – Matplotlib,NumPy,Pandas,CV2, etc.

4. ALGORITHMS USED

Convolutional Neural Network

Artificial neural networks called convolutional neural networks (CNN) are modelled after biological systems. Instead of being given feature vectors like traditional machine learning models, CNN is given raw picture pixel values. When we look at an image, our brains process a huge amount of information. Each neuron has a distinct field and is connected to other neurons in a fashion that covers the entire visual spectrum. The receptive field, which is a specific area of the visual field, is where each neuron in the neurobiological system exclusively reacts to information. In a CNN, each neuron only considers information that is present in its receptive area.. Due to features like minimum interconnection across

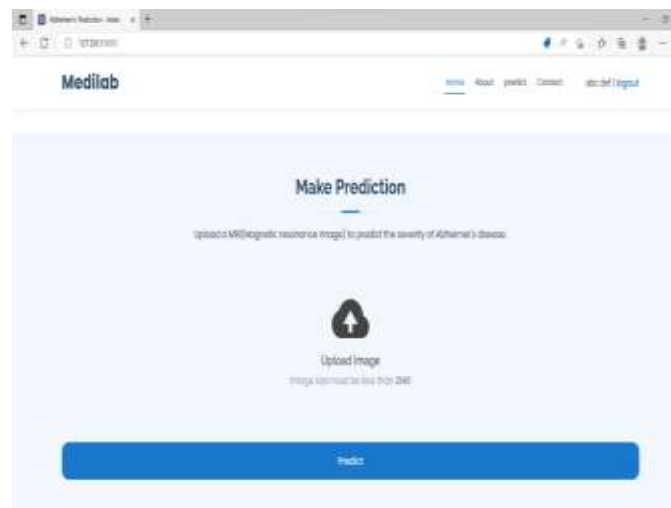
consecutive layers, parametric sharing of weights between adjacent pixels, and similar representation, CNN is more successful than picture detection and classification.

VGG16 -The convolutional neural network (CNN) architecture for this system was developed by the Visual Geometry Group (VGG) at the University of Oxford. It is a deep neural network with 16 layers, hence the name VGG16.

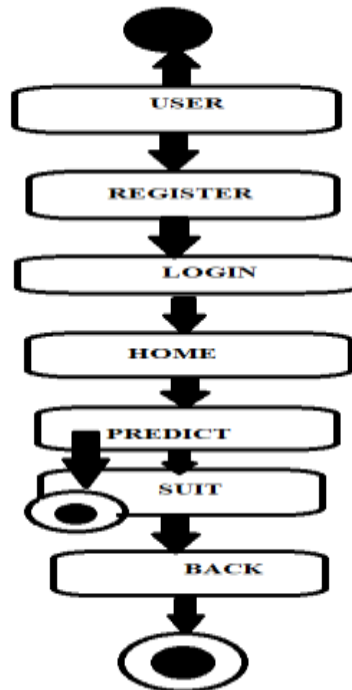
13 convolutional layers make up the VGG16 architecture, which is followed by 3 fully linked layers. The fully connected layers have 4096 neurons apiece and are followed by a max pooling layer for each convolutional layer. The network receives a 224x224x3 input image and produces a probability distribution over 1000 classes. As a benchmark for image classification tasks, VGG16 is well renowned for being straightforward and efficient.

The model was developed using state-of-the-art performance on the ImageNet dataset, which at the time of its release has approximately 1.2 million images and 1000 classes. The ImageNet dataset, which at the time of its publication contained over 1.2 million images and 1000 classifications, served as the basis for the model's development.

5. EXPERIMENTAL RESULTS



6. FLOWCHART



7. ADVANTAGES

- The accuracy of our suggested system outweighs that of others.
- Our suggested methodology will accurately forecast the severity of Alzheimer's, in contrast to other methodologies.
- Identify trends in the CNN algorithm that pinpoint the most effective Alzheimer's treatment for individuals of different ages.
- Comparatively speaking, it minimises the expense, time, and effort.

8. CONCLUSION AND FUTURE SCOPE

Alzheimer's can be discovered more quickly by leveraging databases of medical reports to detect Alzheimer's infection. This forecast might assist patients in taking the necessary proactive steps to assist the patient. Although there is no known cure for this illness, patients can receive treatments to help them combat the illness. Prediction algorithms make it simpler to determine where they are in the process so that treatment may begin. If a treatment for Alzheimer's were to be discovered in the near future, this prediction might assist patients in determining what stage of the disease they are experiencing and whether or not receiving it would be beneficial to them. The accuracy-generated comparison between the capabilities of ML models and CNN models is shown. To make sure the values fit into the ML models without difficulty, they were standardised. After that, SVM, logistic regression, and random forest models were trained using the dataset. The accuracy of SVM is higher than that of other ML models. While the CNN model has generally produced more accurate results. Given this, it can be argued that CNN is the most reliable model for forecasting the onset of Alzheimer's disease's initial phases. 1 in 100 people are affected by this illness. However, the ability to be treated does not ensure that illness can be cured. The system models could be improved in the future by utilising new machine learning models like AdaBoost, KNN, Majority Voting, and Bagging, as well as a larger dataset. This will increase the system's dependability and performance. The ML system can help the general public comprehend the likelihood of dementia in adult patients by simply inputting MRI data.

9. REFERENCES

- [1] Tanveer, M., Richhariya, B., Khan, R. U., Rashid, A. H., Khanna, P., Prasad, M., & Lin, C. T. (2020). Machine learning techniques for the diagnosis of Alzheimer's disease. *ACM Transactions on Multimedia Computing, Communications, and Applications*, 16(1s), 1–35.
- [2] Bron, E. E., Klein, S., Jiskoot, L. C., Linders, J., van Swieten, J. C., van der Flier, W. M., & van der Lugt, A. (2021). Cross-cohort generalizability of deep and conventional machine learning for MRI-based diagnosis and prediction of Alzheimer's disease. *NeuroImage: Clinical*, 31, 102712.
- [3] Kim, J. P., Kim, J., Park, Y. H., Park, S. N., Seo, S. W., & Seong, J.-K. (2019). Machine learning based hierarchical classification of frontotemporal dementia and Alzheimer's disease. *NeuroImage: Clinical*, 23, 101811.

-
- [4] Liu, M., Zhang, J., Nie, D., & Shen, D. (2018). Anatomical landmark based deep feature representation for MR images in brain disease diagnosis. *IEEE Journal of Biomedical and Health Informatics*, 22(5), 1476–1485.
 - [5] Naik, B., Mehta, A., & Shah, M. (2020). Denouements of Machine Learning and Multimodal Diagnostic Classification of Alzheimer's disease. *Visual Computing for Industry, Biomedicine, and Art*, 3(1)
 - [6] Mofrad, S. A., & Lundervold, A. S. (2021). A predictive framework based on brain volume trajectories enabling early detection of Alzheimer's disease. *Computerized Medical Imaging and Graphics*, 90, 101910.