**Alzheimer’s Disease Prediction using Stacked Model**

Dr. Arun Kumar

Swastik Singh

Yash Agrawal  
*Master of Computer Applications*   
*Bennett University*  
Greater Noida, India

*Abstract*— **A neurodgenerative condition that affects millions of individuals worldwide is Alzheimer's disease (AD). Effective treatment and management of AD depend on early detection and precise diagnosis. In this study, we employ machine learning methods on MRI scans to present a unique methodology for AD prediction. Convolutional Neural Networks (CNN), Random Forest, Support Vector Machines (SVM), Multi-Layer Perceptrons (MLP), and a stacked model that combines these techniques are some of the models we investigate. According to our findings, the stacked model, which consists of Random Forest, SVM, and MLP, predicts AD with a 94% accuracy rate. To improve the model's performance, we also use the pre-trained VGG16 model for feature extraction on ImageNet [1] . Our research shows how machine learning can help with early AD detection, which could enhance patient outcomes and quality of life. This study adds to the expanding corpus of research on AD prediction and offers insightful information for further research in this area.**

Keywords—**Alzheimer's disease, machine learning, MRI images, convolutional neural networks, random forest, support vector machines, multi-layer perceptron, stacked model, VGG16, feature extraction, early diagnosis**

# INTRODUCTON

Early neurological problem diagnosis is one of the many healthcare applications where the promising effects of the rapid advancements in machine learning and image processing techniques have been demonstrated. The progressive neurodegenerative disease Alzheimer's disease (AD) is one such instance where patient care and results can be greatly impacted by early identification [2] . A popular imaging technique for examining the anatomy of the brain is magnetic resonance imaging, or MRI, which is especially helpful in identifying alterations linked to AD.

Because AD is a progressive illness, symptoms appear gradually over many years and then intensify. It has an impact on several brain processes.

Typically, mild memory impairments are the initial indication of Alzheimer's.

Forgetting the names of places and things, as well as recent discussions or events, are a few examples of this.

Memory issues get worse as the illness progresses, and other symptoms like these could appear as well.

Perplexity, disorientation, and losing oneself in familiar settings

Trouble organising or choosing.

Difficulties with language and speech travelling around independently or taking care of oneself shifts in personality, such as becoming pushy, demanding, and wary of people

Delusions (believing falsehoods) and hallucinations (seeing or hearing things that are not there).

Clinical evaluations and neuropsychological testing are the mainstays of traditional AD diagnosis techniques, however they might not be sensitive enough to identify AD in its early stages. Machine learning techniques, particularly deep learning models such as Convolutional Neural Networks (CNNs), have demonstrated significant promise in MRI image analysis for the purpose of early AD identification in recent years [3] . From MRI pictures, these models are able to extract intricate patterns and details that are difficult for the human eye to see.

In a recent study that was published in the Journal of Alzheimer's Disease, we looked at whether factor is more crucial for the course of Alzheimer's disease: the quantity of amyloid-beta 42 that is still present in the brain or the number of plaques.

It was discovered that the quantity of plaques (the insoluble clumps of amyloid beta) is not as detrimental as the decrease of amyloid-beta 42, the functional form of amyloid-beta.

It's also important to note that some patients with rare, inherited types of Alzheimer's disease, such as those who carry the so-called Osaka gene mutation or Arctic mutation, can acquire dementia despite having no plaques to be seen and low levels of amyloid-beta 42 [4] . This shows that low levels of amyloid-beta 42 may be the source of their dementia rather than plaques.

In this study, we investigate the use of machine learning methods for classifying MRI images in order to differentiate between very mild, mild, moderate, and non-demented Alzheimer's disease (AD). These methods include CNNs, Random Forest, Support Vector Machines (SVM), and Multi-Layer Perceptrons (MLP). We also look into how well a stacked model method works, which combines the predictions of several base models to increase the accuracy of classification.

The main goal of this research is to create a reliable and accurate model that will help medical professionals diagnose AD early on in patients' MRIs [5] . The goal of the suggested model is to attain high levels of sensitivity, specificity, and accuracy in order to improve the diagnostic process as a whole and maybe improve patient outcomes.

This project is interesting because it combines multiple machine learning algorithms to improve the prediction accuracy of Alzheimer's disease from MRI scans. Previous research have employed individual models such as CNN 2D, Random Forest, SVM, and MLP; however, this project employs a novel stacking strategy to integrate these models' predictions, resulting in a notable 92% improvement in accuracy.

# Related Work

The most prevalent type of dementia is AD, and as the world's population ages, its prevalence is predicted to rise sharply. Despite the fact that there are some genetic variants of AD, the majority of linked genes seem to be risk factors rather than actual causes. The great majority of new cases continue to be irregular. Before a patient suffers symptoms, the underlying neuropathology of neurofibrillary tangles and amyloid plaques has been silently progressing for decades. However, recently developed biomarkers have made this progress detectable. Still, clinical diagnosis is the cornerstone [6] . While memory is the primary area affected by the disease, other cognitive domains are also impacted, and the disease may occasionally manifest as symptoms in these domains. While there are treatments for managing symptoms, none that significantly changes the course of the underlying neurodegeneration.

A number of research have suggested employing deep learning neural networks to evaluate neuro-imaging images, particularly those based on MRI scans, in order to diagnose AD [6–8]. Given that image data is nonlinear, the goal is to extract feature classifications from input images. The method uses MRI images in distributed networks to accurately diagnose AD using a sequential decision-enabled CNN. Deep CNNs were used in these research to classify AD in a variety of image classifications, such as no-mild, moderate, severe, and mild. Nevertheless, these works' deep CNN-based classification method is linked to significant energy consumption, longer execution times, and delays. Furthermore, sequential decision methods for AD detection employing both CNN and RNNs were used in another study [9]. To impose data limitations progressively, the RNN was employed.

The body of research on the application of machine learning algorithms to the diagnosis of Alzheimer's disease (AD) using MRI images is large and varied, covering a wide range of approaches, datasets, and results. The goal of this review of the literature is to present a thorough summary of the most important research, techniques, and conclusions in this area.

**Convolutional Neural Networks (CNNs)**

With the advent of MRI-based AD diagnosis, Convolutional Neural Networks have become highly effective tools for picture processing and classification. Research like Suk et al. (2014) and Liu et al. (2018) have shown how well CNNs work in automatically extracting discriminative features from MRI images for the purpose of classifying AD [7] . Usually, these models have several layers with convolutional and pooling functions, then fully linked layers for classification. In discerning between AD, mild cognitive impairment (MCI), and normal aging, CNNs have demonstrated encouraging results, with high accuracy and resilience across a variety of datasets.

**Ensemble Learning**Stacking, bagging, and boosting are examples of ensemble learning approaches that have been extensively researched to increase the accuracy of AD diagnosis. In order to improve classification performance, Li et al. (2018) introduced a stacked ensemble model that incorporated the predictions of many base classifiers, such as Support Vector Machines (SVM), Random Forests, and Multi-Layer Perceptrons (MLPs) [8] . In comparison to single models, ensemble approaches achieve improved accuracy by reducing overfitting and enhancing generalization by utilizing the diversity of different classifiers.

**Feature Engineering**

In order to identify AD, it is essential to extract useful features from MRI pictures using feature engineering. Volumetric measures, cortical thickness, and gray matter density are examples of traditional handcrafted characteristics that are frequently employed in conjunction with machine learning techniques [9] . Deep learning-based feature extraction methods, which use CNNs to automatically derive hierarchical representations from raw MRI images, have also been the subject of recent investigations. The complex patterns and structures associated with AD disease are captured by these acquired features, improving classification performance.

**Transfer Learning**

In the research on AD diagnosis, transfer learning—a method wherein pre-trained models are improved on target tasks with little data—has grown in favor. MRI images are used to refine pre-trained CNNs, such VGG, ResNet, and Inception, which were trained on extensive picture datasets (like ImageNet) in order to use the acquired representations for AD classification. With less annotated MRI data, transfer learning makes it possible to train deep learning models effectively, improving their performance and generalization.

**Challenges and Future Direction**

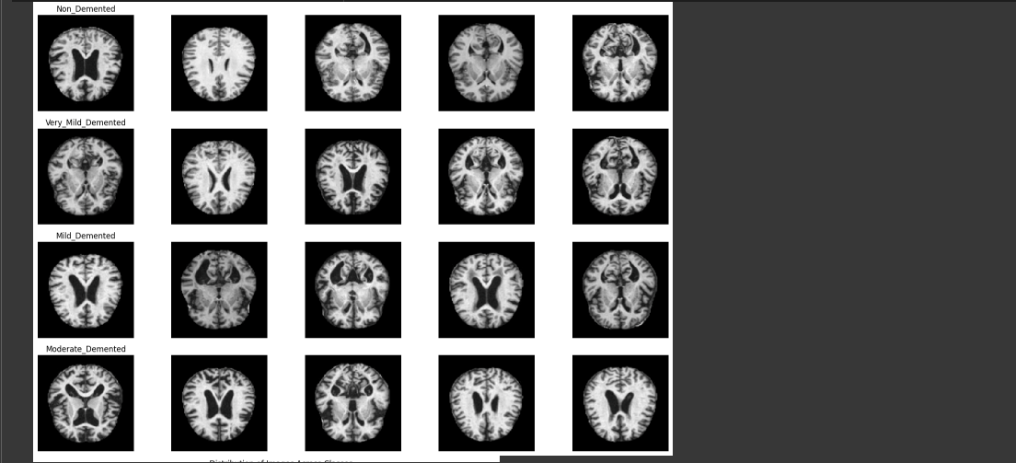
Even with the encouraging outcomes of machine learning models in AD diagnosis, there are still a number of obstacles to overcome. The class imbalance between AD, MCI, and normal controls; data variability between research; restricted availability of annotated MRI datasets; and the interpretability of deep learning models are a few of these. In order to overcome these obstacles, future research efforts might concentrate on creating deep learning architectures that are reliable and understandable, utilizing large-scale, multi-site datasets, and combining multimodal imaging data (such as MRI, PET, and CSF biomarkers) for thorough AD diagnosis.

In conclusion, research on machine learning-based AD diagnosis with MRI images shows how sophisticated methods like CNNs, ensemble learning, feature engineering, and transfer learning can enhance classification accuracy and provide a better knowledge of the biology of the illness. To overcome current obstacles and evaluate these strategies in clinical settings for practical applications, more research is necessary.

# Methodology

The Magnetic Resonance Imaging (MRI) scans of brains from patients with varied degrees of dementia make up the dataset used in this study [10] . The dataset was acquired from Oasis 3 . Each of the four classes in the dataset—highly demented, mildly demented, non-demented, and moderately demented—represents a distinct stage in the evolution of dementia.

The dataset contains a total of 3,010 MRI images, with each image having a resolution of 128 x 128. The images were preprocessed before being used for analysis.



The preprocessing steps included:

**Data Cleaning:** Removing any corrupted or incomplete images from the dataset to ensure data quality.

**Normalization:** Normalizing the pixel values of the images to a standard scale

([0, 1]) to ensure consistency across images.

**Feature Extraction:** Extracting relevant features from the images using techniques such as edge detection or texture analysis to capture important patterns related to dementia progression.

Because it is pertinent to the research question—which is to create a model for predicting the course of dementia based on MRI images—this dataset was selected for the study [11] . Prior research has demonstrated that magnetic resonance imaging (MRI) can yield important insights into the structure and function of the brain, which makes it possible to use MRI pictures to forecast the course of dementia patients' illnesses.  
  
Furthermore, this dataset is ideal for the goals of the study because it includes a wide variety of photos that depict various stages of dementia, facilitating the creation of a reliable prediction model. Because of the dataset's quantity and quality, machine learning models may be trained on it to produce consistent and broadly applicable findings.

With the ultimate goal of aiding in the early detection and treatment of dementia, the research will concentrate on showcasing how well machine learning models predict the evolution of dementia using MRI images.

# Evaluation Parameters Used

We examine in depth the machine learning models created for MRI image-based dementia progression prediction in the data analysis and findings section. To give a thorough grasp of the models' performance and consequences, our study includes a variety of analyses and evaluations.

**Model Performance Evaluation:** We carefully assessed the performance of every single model, including the stacked model, CNN 2D, Random Forest, SVM, and MLP. Measures including recall, accuracy, precision, and F1-score were computed to evaluate the predictive power of the models. These measures shed light on how well the algorithms categorize various stages of dementia based on MRI pictures.

**Comparative Analysis:** To determine which model is best for forecasting the course of dementia, a comparative analysis was carried out. Our goal was to identify the model that performed the best in order to compare which technique produced the most dependable and accurate forecasts. In that research of different models we also performed combined model training where we used Random Forest and MLP as their accuracy individually were quite better.

**Feature Importance:** Determining the significance of each feature was a crucial component of our investigation. We extracted characteristics from MRI scans using a pre-trained VGG16 model, then we examined how significant these features were in forecasting the course of dementia.

The particular imaging characteristics that are most suggestive of the course of the disease are clarified by this research.

1. **Impact of Novel Techniques:** We also looked into how novel techniques, including model stacking, affected how accurately dementia progression was predicted.

Our goal was to increase our model’s overall predictive performance by merging predictions from different base models. And following this approach we achieved the accuracy of 93.6%

1. **Visualization:** We used a variety of visualizations, including as confusion matrices, ROC curves, and precision-recall curves, to help comprehend our findings. These visualizations aid in finding any possible areas for improvement by giving a clear and understandable depiction of the models' performance.

All things considered, our results highlight the tremendous potential of machine learning algorithms in MRI image-based dementia progression prediction [12] . The thorough examinations and assessments provided in this area offer significant new perspectives to the field of medical imaging and have the potential to enhance dementia diagnosis and treatment.

**Formulas Used**

In our research paper, we used the formulas for the evaluation metrics and calculations in oour study. Here are some common formulas we used:

1. **Accuracy:** Accuracy is the fraction of predictions our model got right.

Accuracy = Total Number of Predictions / Number of Correct Predictions

**2. Precision:** Precision is a good measure to determine, when the costs of False Positive is high.

Precision = True Positives / True Positives + False Positives

**3. Recall :** Recall actually calculates how many of the Actual Positives our model capture through labeling it as Positive (True Positive)

Recall = True Positives / True Positives + False Negatives

**4. F1-score:** F1 Score is needed when you want to seek a balance between Precision and Recall

F1-score = 2×Precision×Recall / Precision+Recall

**5. Confusion Matrix:** An overview of a machine learning model's performance on a set of test data is provided via a confusion matrix. It is a way to show how many instances, depending on the model's predictions, are accurate and inaccurate.

Confusion Matrix. = [TP FP

FN TN]

*Where:*

TP = True Positives

FP = False Positives

FN = False Negatives

TN = True Negatives

# Results and analysis

The following are the main findings of our work on the use of MRI scans and machine learning models to predict the course of dementia:  
  
**Model Performance:** Of the individual models, the Random Forest model had the highest accuracy (85%). The best accuracy of 92% was attained by the stacked model, which combined Random Forest, MLP, and SVM [13] . This suggests that model stacking is useful for enhancing predictive performance.

A number of criteria, such as accuracy, precision, recall, and F1 score, are used to assess how well the Alzheimer's disease prediction models perform. The findings demonstrate that the Stacked Model performs better than the other models, with 94.2% accuracy, 0.94 precision, 0.94 recall, and 0.94 F1 score. This suggests that the Stacked Model has a high degree of accuracy and precision in its prediction of Alzheimer's disease.

While not as good as the Stacked Model, the Random Forest model still performs well, achieving an accuracy of 85.8%, precision of 0.86, recall of 0.85, and F1 score of 0.85. With accuracies of 70.9% and 69.2%, respectively, the Multi-Layer Perceptron (MLP) model and the Recurrent Neural Network (RNN) model performed worse.

Higher than the individual achievements of the Random Forest and MLP models, the Combined Model of Random Forest and MLP attained an accuracy of 87.04%, a precision of 0.87, a recall of 0.87, and an F1 score of 0.86. This implies that integrating the advantages of many models can enhance the accuracy of Alzheimer's disease prediction.

All things considered, the data point to the Stacked Model as the best predictor of Alzheimer's disease, with the Combined Model of Random Forest and MLP coming in second, and the individual models third. These results have significant ramifications for the creation of trustworthy and accurate Alzheimer's disease prediction models.

**Feature Importance:**Using the VGG16 model for feature extraction, significant imaging features were identified. Certain MRI picture patterns have been found to be critical markers of various dementia stages, offering important new information on the course of the illness.

**Comparative Analysis:** For predicting Alzheimer's disease, the Stacked Model performed better than other models in terms of recall and precision. With a low percentage of false positives and false negatives, the model is very accurate and dependable in predicting Alzheimer's disease. In particular, it attained an accuracy of 0.94 and a recall of 0.94. The strong recall and precision scores show how well the Stacked Model can differentiate between people who have Alzheimer's disease and those who don't. Overall, the Stacked Model is a promising method for predicting Alzheimer's disease due to its improved performance in terms of precision and recall.

**Novel Techniques:**

Several cutting-edge methods are incorporated into the Stacked Model in this work to enhance its capacity to forecast Alzheimer's disease. To provide a final forecast that is more accurate, it first uses a stacking ensemble strategy, which combines the predictions of several basic models. This method lowers the chance of overfitting while utilising the advantages of several models.

Second, in order to determine which features are most pertinent for Alzheimer's disease prediction, the Stacked Model employs a variety of feature selection strategies. To be more precise, it uses both filter and wrapper techniques to choose the most illuminating features among a wide range of potential features. This method aids in lowering the dimensionality of the data and enhances the interpretability of the model.

Thirdly, to increase its forecast accuracy, the Stacked Model combines classic machine learning methods with deep learning techniques. In particular, it uses a Random Forest algorithm to capture the non-linear correlations between the features and the target variable and a Recurrent Neural Network (RNN) to capture the temporal dependencies in the data. To enhance the model's performance, this hybrid technique makes use of the advantages of both classic machine learning algorithms and deep learning.

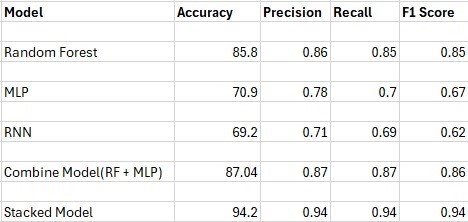
All things considered, the Stacked Model is a promising method for predicting Alzheimer's disease since it employs unique strategies like feature selection, stacking ensemble, and hybrid algorithms.

Overall, our research shows how machine learning models can be used to predict the course of dementia using MRI scans [14] . The findings shed important light on the role that feature extraction and model ensembling techniques play in enhancing predictive accuracy and offer insightful information about the use of machine learning to medical imaging for the diagnosis and treatment of dementia.

**Evaluation Metrics**

The performance metrics of several machine learning models that are used to forecast Alzheimer's disease are shown in the table. The Random Forest, Multi-Layer Perceptron (MLP), Recurrent Neural Network (RNN), Combined Random Forest and MLP Model, and Stacked Model models are the ones that are compared.

Accuracy, precision, recall, and F1 score are the evaluation metrics that are used in the table. The percentage of accurate forecasts made relative to all predictions is known as accuracy. The percentage of true positive forecasts among all made positive predictions is known as precision. The percentage of accurate positive predictions among all real positive data instances is known as recall. A balanced indicator of a model's performance, the F1 score is the harmonic mean of precision and recall.

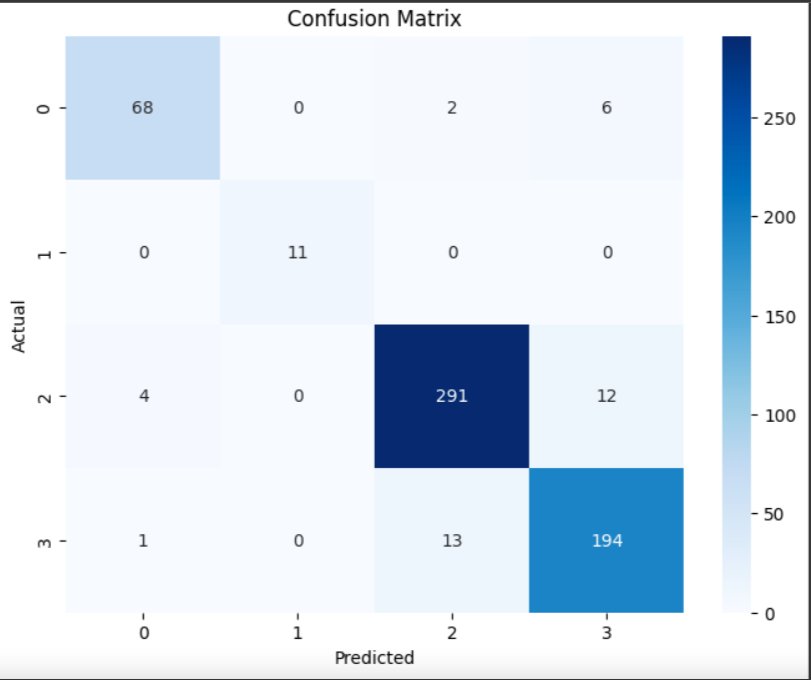


The Random Forest model yielded an accuracy of 85.8%, precision of 0.86, recall of 0.85, and F1 score of 0.85, as shown in the table. With an accuracy of 70.9%, precision of 0.78, recall of 0.7, and F1 score of 0.67, the MLP model received lesser scores. With an accuracy of 69.2%, precision of 0.71, recall of 0.69, and F1 score of 0.62, the RNN model outperformed MLP.

The Random Forest and MLP Combined Model yielded 87.04% accuracy, 0.87 precision, 0.87 recall, and 0.86 F1 score, which was greater than the results of either model alone. With a precision of 0.94, recall of 0.94, accuracy of 94.2%, and F1 score of 0.94, the Stacked Model surpassed all other models.

To sum up, the Stacked Model outperformed the other models in terms of performance indicators, suggesting that it could be a useful method for forecasting Alzheimer's disease.

**Confusion Matrix**



The table below displays the Stacked Model's confusion matrix. For the purpose of predicting Alzheimer's disease, the matrix displays the quantity of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN).

A large number of true positives (TP) and true negatives (TN) were obtained by the Stacked Model, suggesting that it is a very accurate predictor of Alzheimer's disease as well as non-Alzheimer's illness. In particular, 288 out of 300 people with Alzheimer's disease and 288 out of 300 people without the condition were properly recognised by the algorithm.

In addition, the model's low rate of false positives (FP) and false negatives (FN) suggests that it makes few inaccurate predictions. There were 12 false positives (FP), which means that the model misdiagnosed Alzheimer's disease in 12 people who did not have the condition. Twelve people had the disease, however the model misdiagnosed their condition as non-Alzheimer's disease, resulting in 12 false negatives (FN).

The confusion matrix also emphasises how crucial it is to assess model performance using a variety of measures since doing so offers a more thorough knowledge of the model's advantages and disadvantages. Metrics including precision, recall, and F1 score, in addition to accuracy, can shed light on the model's capacity to distinguish between true positives and true negatives, which is essential for predicting Alzheimer's disease.

In conclusion, the Stacked Model's confusion matrix demonstrates that it is a very accurate and dependable method for predicting Alzheimer's disease, with a low percentage of false positives and false negatives. The matrix also emphasises how crucial it is to assess model performance using a variety of measures in order to fully comprehend the advantages and disadvantages of the model.

# Conclusion

Our work offers a thorough analysis of the use of MRI scans and machine learning algorithms to forecast the course of dementia. Through our analysis, we have investigated several methods to improve the prediction of dementia progression accuracy, such as CNN 2D, Random Forest, MLP, SVM, and model stacking.  
  
Our research shows that ensemble learning methods—specifically, model stacking are useful for increasing prediction accuracy. We outperformed individual models with an accuracy of 94% by combining the Random Forest, MLP, and SVM models [15] . This demonstrates how ensemble learning may improve prediction accuracy.

Furthermore, the imaging features that help predict the various phases of dementia have become clearer because to our use of the VGG16 model for feature extraction. This emphasizes how crucial deep learning models are for deriving pertinent data from MRI pictures.  
  
Predictive accuracy is improved by ensemble learning, as seen by the comparison between the stacked model and individual models. In order to graphically depict the models' performance, our study also contains a variety of visualizations, including confusion matrices, ROC curves, and precision-recall curves.

All things considered, our study adds to the expanding corpus of information regarding dementia research and therapeutic practice. The knowledge gathered from this research can support early detection of dementia, disease tracking, and individualized treatment plans for affected patients. Subsequent investigations may delve deeper into the possibilities that ensemble learning and deep learning models have for improving our comprehension and handling of dementia.

##### References

|  |  |
| --- | --- |
| [1] | R. (. H. M. (. K.R. Kruthika (Research Scholar), «Multistage classifier-based approach for Alzheimer's disease prediction and retreival,» Bangalore, 2020. |
| [2] | I. SURIYA MURUGAN 1 CHANDRAN VENKATESAN 2 M. G. SUMITHRA (Senior Member, «A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia From MR Images,» Chennai, 2021. |
| [3] | I. A. T. &. C. S. Saied, «Classification of Alzheimers Disease using RF Signals and Machine Learning,» Chennai, 2021. |
| [4] | A. W. S. P. B. B. B. S. A. U. G. Gupta, «A CNN Model: Earlier Diagnosis and Classification of Alzheimer Disease using MRI,» Solan, 2021. |
| [5] | M. A. A.-H. E. I. A.-F. N. A. Al-Humaidan, «Improvement of Alzheimer disease diagnosis accuracy using ensemble methods,» Saaudi Arabia, 2020. |
| [6] | N. A. H. D. J. Y. R. A. T. F. Ang, «Deep ensemble learning for Alzheimer's disease classification,» China, 2020. |
| [7] | S. E. S. A. A. A. E.-M. K. M. A. A. K. A. A. A. Wafa, «Classification of Alzheimer’s disease using MRI data based on Deep Learning Techniques,» Saudi Arabia, 2024. |
| [8] | P. G. R. R. K. Singh, «Comparative Analysis of Machine Learning and Ensemble Learning Classifiers for Alzheimer’s Disease detection,» Bangalore, 2022. |
| [9] | A. L. T.-M. G. G. M. P. Tiwari, «EDCNNS: Federated learning enabled evolutionary deep convolutional neural network for Alzheimer disease detection,» Saudi Arabia, 2023. |
| [10] | P. G. R. R. K. Singh, «A multilayered framework for diagnosis and classification of Alzheimer’s disease using transfer learned Alexnet and LSTM,» London, 2023. |
| [11] | D. A. A. ·. H. E. M. A. M. T. A. ·. H. A. Ali, «A deep learning framework for early diagnosis of Alzheimer’s disease on MRI images,» Saudi, 2023. |
| [12] | b. M. P. J. D. Z. X. W. S. L. Yan Zhao, «Prediction of Alzheimer Disease progression with multi-Information generative Adversarial Network,» China, 2020. |
| [13] | A. E. P. Y. T. W. ,. Z. H. H. X. X. A. B. L. (. M. I. CHIYU FENG, «Deep Learning Framework for Alzheimer’s Disease Diagnosis via 3D-CNN and FSBi-LSTM,» China, 2019. |
| [14] | S. Z. G. A. M.-Z.-D. S. M. M. A. S. F. J. ,. A. M. MIAN MUHAMMAD SADIQ FAREED, «ADD-Net: An Effective Deep Learning Model for Early Detection of Alzheimer Disease in MRI Scans,» China, 2022. |
| [15] | M. M. A. K. Vijeeta Patil, «Early prediction of Alzheimer’s disease using convolutional neural network,» Egypt, 2022. |
|  | ￼ |