

A Novel Machine Learning and Deep Learning Insight for Alzheimer's Diseases Using Neuroimaging Dataset Analysis

Mamta
Department of CSE, Maharaja Agrasen University

Sushil Kumar Bansal
Department of CSE, Maharaja Agrasen University

<https://hdl.handle.net/2324/7420065>

出版情報 : Evergreen. 13 (2), pp.499-516, 2026-06. 九州大学グリーンテクノロジー研究教育センター
バージョン :
権利関係 : Creative Commons Attribution 4.0 International



A Novel Machine Learning and Deep Learning Insight for Alzheimer's Diseases Using Neuroimaging Dataset Analysis

Mamta^{1,*}, Sushil Kumar Bansal¹

¹Department of CSE, Maharaja Agrasen University, Baddi, District Solan, Himachal Pradesh
174103, India

*Author to whom correspondence should be addressed:
E-mail: mamta.airesearch@gmail.com

(Received August 20, 2025; Revised March 11, 2026; Accepted March 31, 2026)

Abstract: Alzheimer's disease is a progressive neurodegenerative disease that has become one of the most central health issues in the world today. Early detection and proper diagnosis in the initial stages are essential for slowing disease progression and improving patient outcomes. This paper presents an integrated review and a practical evaluation of the most advanced machine-learning (ML) and deep-learning (DL) methods for diagnosing Alzheimer's disease (AD) in its initial stages, using neuroimaging data as input. As there is a possibility of better performance using classifier ensembles, which involve Support Vector Machine (SVM), Random Forest (RF), or Decision Tree (DT), or Logistic Regression (LR), or Naive Bayes (NB), or any combination of ensemble tactics, their diagnostic capabilities have been investigated. At the same time, Convolutional Neural Networks (CNNs), 3-D CNNs, Deep Neural Networks (DNNs), Feedforward Neural Networks (FFNNs), and hybrid CNN- LSTM models (CNN-LSTMs) have been evaluated in detail. The current state-of-the-art pretrained networks for transfer learning include VGG19, ResNet50, DenseNet, EfficientNetB0, MobileNet, Xception, and GapNet. Each model was trained on neuroimaging datasets and evaluated for accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC). The results show that DL-based models improve their ability to identify subtle structural patterns in brain MRI scans compared to traditional ML models. In addition, ensemble ML models offer competitive performance and greater interpretability. This paper provides a comparative discussion of the merits, limitations, and practical challenges of the two methodology families (including overfitting, class imbalance, and generalizability) associated with both methods. The results are a step towards developing effective, automated, and innovative decision support tools for the early detection of Alzheimer's disease.

Keywords: Alzheimer's Disease; Convolutional Neural Networks; Deep Learning; Early Diagnosis; Neuroimaging; Transfer Learning

1. Introduction

The main effects of neurodegenerative diseases, whose biological basis is unknown, are a weakening of thinking, regular activities, and cognitive processes. However, dementia is one of the mental illnesses that encompasses memory problems as well as significant cognitive deficiencies like apraxia, agnosia, and aphasia. It also involves disruptions in daily functioning, such as eating, listening, and other tasks that typically require the brain. The word "dementia" refers to a wide range of illnesses and ailments that result from different brain cells dying or from cells not working normally. Based on studies, AD is the most prevalent type of dementia affecting the majority of people worldwide. Understanding AD and its causes is

essential to learning more about the condition, as these topics are covered below. Alois Alzheimer, a German psychiatrist, was the first to identify this illness 1907¹⁾. The neurofibrillary tangles and amyloid plaques, composed of the proteins A and Tau, in the brain of an AD patient eventually cause neurons to die. The temporal lobe and hippocampus are primarily affected by AD. Because AD causes changes in the brain faster than physical symptoms manifest in patients, the prevention and diagnosis of the illness are challenging.

According to several studies, 1 in 85 persons is expected to have AD by 2050. Furthermore, more than 47 million patients worldwide are affected by it in the current circumstances²⁾. Therefore, a precise consensus on the diagnosis of AD is essential because early treatments

typically slow down the progression of AD, potentially prolonging the quality of life for affected individuals. Our goal is to examine the feature extraction methods used to train classifiers to detect the ailment and establish an appropriate diagnosis. Deep learning models have also demonstrated better performance over conventional machine learning methods in detecting Alzheimer's disease using MRI data in recent years. Nevertheless, it remains difficult to choose appropriate architectures and their relative behaviour for early-stage diagnosis.

In Figure 1, the image compares a healthy brain to one with Alzheimer's disease. In Alzheimer's, the brain doesn't function properly, leading to memory loss, thinking problems, poor communication between brain cells, and a build-up of abnormal proteins. The Figure shows the difference between a healthy neuron and a damaged one in Alzheimer's. In 2012, new guidelines and proposals concerning several types of brain abnormalities connected to Alzheimer's disease and other dementia-related

disorders were made by the Alzheimer's Association and the National Institute on Aging.

Alzheimer's disease has four stages, as shown in Figure 2. The diagram outlines the gradual structural changes in the brain across the preclinical, mild, moderate, and severe stages. The graphic display highlights the progressive cerebral atrophy and accompanying functional dementia that accompanies disease size. Research indicates that out of every five patients with MCI symptoms who check with their doctor during the next three to four years, their condition transforms to AD. But the patient will have the final stage of MCI in case the diagnosis and management are not appropriate during this period, which can lead to loss of multiple memories and issues with various physical and mental problems. The most severe patients with AD cannot perform daily and simple duties, and this can be witnessed during the significant changes in the mood and disposition of people living with dementia³.

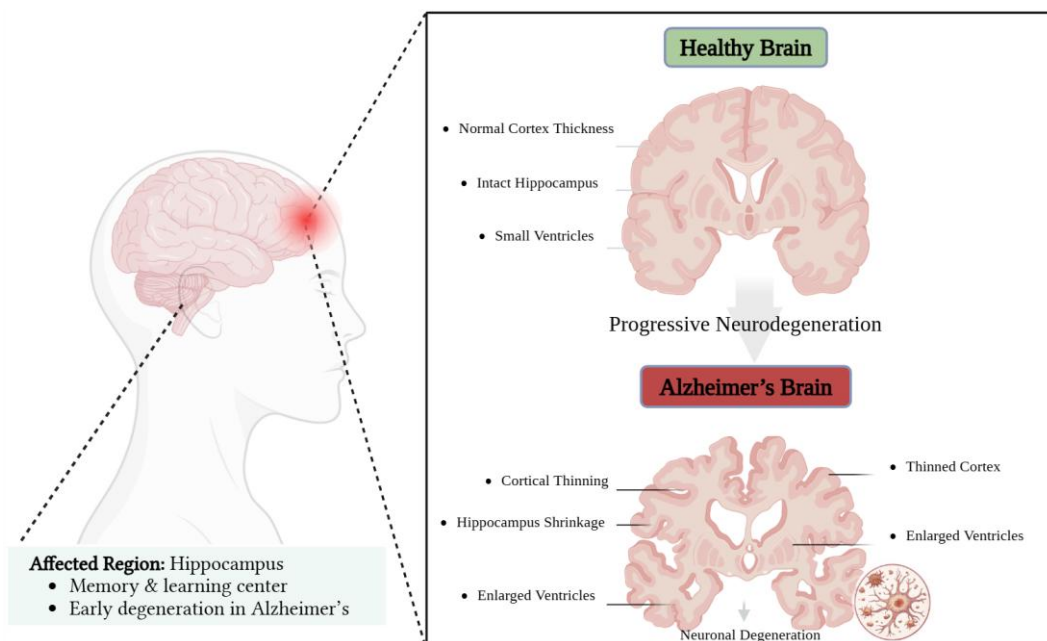


Fig. 1: Healthy vs Alzheimer's Disease Brain

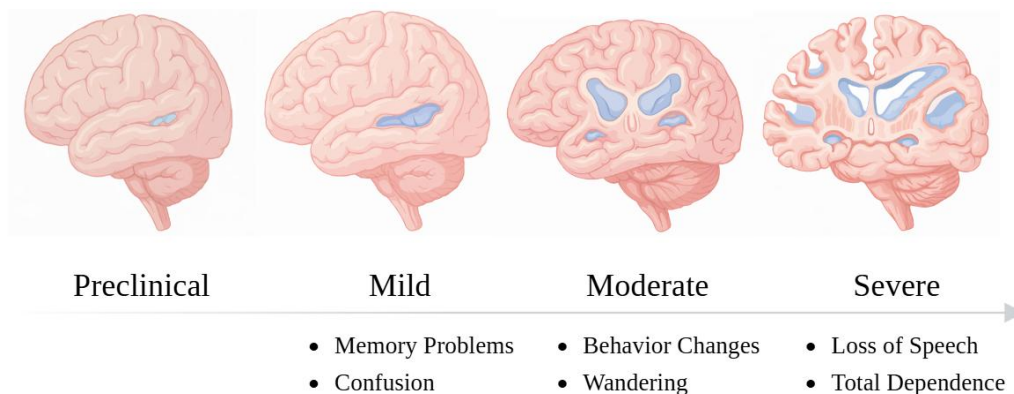


Fig. 2: Stages of Alzheimer's Disease Progression in the Brain

Cite: Mamta, S. Bansal, "A Novel Machine Learning and Deep Learning Insight for Alzheimer's Diseases Using Neuroimaging Dataset Analysis". Evergreen, 13 (02) 499-516 (2026). <https://doi.org/10.5109/7420065>.

Using classifiers across numerous approaches to dementia prediction requires a comprehensive survey of AD diagnosis and its stages. The primary goal of this survey has been to provide participants with information on various feature extraction approaches so they can develop appropriate features and classifiers, which, in turn, enable accurate prediction. According to the report by the Alzheimer's Association (2025), Alzheimer's disease is one of the costliest health care needs in the United States, where the total care costs are projected to reach about 384 billion dollars by the year 2025. Medicare and Medicaid cover a substantial share of these expenditures, totaling approximately 246 billion, or about 64 percent of total spending. The report also indicates that this economic burden is expected to increase significantly as the number of new cells grows within the aging population. The total Medicare and Medicaid expenditure is expected to increase to around 644 billion by 2050, indicating a significant upward trend in healthcare expenditures related to the disease⁴.

Figure 3 shows that Medicare and Medicaid payments from 2025 to 2050 gradually increase as the population ages, highlighting the significant economic pressure of an ageing population. This review analyzes how scientists use MRI scans and advanced computational methods, i.e., machine learning and deep learning, to learn about AD. The datasets used have also been overviewed in the manuscript. The structure of the research includes several parts: Section 2 discusses previous studies in the field of AD; Section 3 outlines how machine-learning models evolve into deep-learning models in AD studies; Section 4 provides an overview of the models and datasets used; Section 5 discusses experimental results; Section 6 presents the discussion; and Section 7 presents concluding remarks. The general goal is to clarify how innovative methods, including intensive learning, can hasten early diagnosis and enhance knowledge about Alzheimer's disease, which has become necessary due to the lack of effective therapeutics and the increasing incidence of the disease.

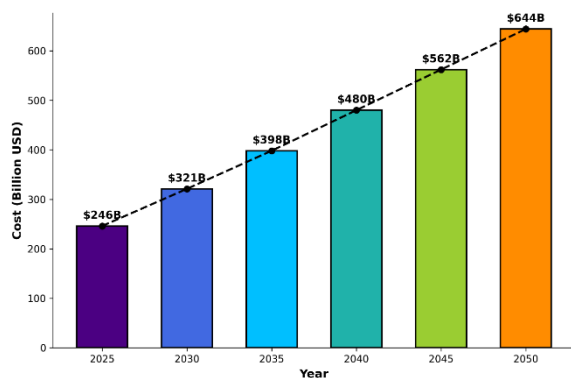


Fig. 3: Projected Alzheimer spending rising to \$246B to \$644B (2025–2050)

1.1. Contribution of the Study

The present paper enhances early-detection studies in AD through artificial intelligence in various unique ways. First, it sequentially runs a set of DL and ML models, such as the SVM, the Random Forest, the XGBoost, the Logistic Regression, the VGG19, the ResNet50, the DenseNet121, and CNN-LSTM, on the OASIS MRI data with the aim of establishing imaging signatures that can possibly be indicative of the incipient stages of AD. A CNN-LSTM model combines spatial features from analysis with sequence learning to represent progressive cerebral changes. The choice of those models was aimed at comparing different learning paradigms on the same MRI data, extending lightweight networks to deep architectures, and learning how to select the best models across different computational environments. Second, it provides a critical comparative analysis of ML and DL approaches, evaluating them using universally recognized classification metrics: accuracy, precision, recall, F1-score, and AUC-ROC. The AUC-ROC measures the discriminative capacity of a given model to distinguish cases of AD from non-AD across serial decision thresholds. As such, this will seriously reflect the limitations and advantages of each method in its applications to clinical practice. Third, the paper proposes a two-pipeline approach that combines handcrafted feature learning methods (e.g., PCA) with ML and end-to-end image-based learning in DL to better examine traditional and modern AI methods. Lastly, it creates a novel hybrid diagnostic pipeline that unites hand-designed feature extraction and ML models with end-to-end DL techniques, providing a robust framework that may eventually enhance the speed and efficiency of the diagnostic process and promote effective integration into clinical practice.

1.2. Objectives of the Study

The primary aim of the present research is to propose and test an effective computerized model of early-phase analysis of Alzheimer's Disease (AD) based on neuroimaging data. The research objectives are the following:

- 1) To apply a set of machine learning (SVM, Random Forest, XGBoost, Logistic Regression) and deep learning (VGG19, ResNet50, DenseNet121, CNN-LSTM) models to the OASIS MRI dataset to determine discriminative patterns in the early stages of Alzheimer's disease.
- 2) To conduct a comparative study of ML and DL within the systematic model-based evaluation with standard performance metrics (accuracy, precision, recall, F1-score, and AUC-ROC) to identify the most clinically-reliable strategy.

2. Literature Review

Various reviews and survey articles have been included in studies of Alzheimer's, with emphasis on the significance of early diagnosis and the future of neuroimaging modalities such as magnetic resonance imaging (MRI). However, the majority of already prepared reviews have already talked about one of the clinical properties of AD, or how the application of MRI or any other scans has assisted in detecting AD at a symptomatic level. New advancements in the early detection of Alzheimer's disease (AD) have prompted a shift from traditional machine learning (ML) to more sophisticated deep learning (DL) models. An extensive overview of the detection of Alzheimer's disease has made the use of artificial intelligence and deep learning methods with various datasets a theme of focus, including the extraction of features, modalities, and validation strategies. The paper also notes several critical challenges, such as the inability to interpret, and it highlights the need to have more explainable and robust models⁵⁾. The deep learning methods have also been proven to perform better as compared to the traditional machine learning techniques in detecting Alzheimer's disease in EEG by automatic feature extraction of complex patterns of raw data, as compared to the conventional method, which used manual feature extraction. The paper also shows that a combination of the two methodologies can empower early diagnosis, but there are still some issues like interpretability, data quality, and model generalization⁶⁾. Deep learning method has proven to perform exceptionally better compared to traditional machine learning algorithms in detecting Alzheimer's disease using MRI, whereby complex patterns are automatically identified based on imaging data (compared to manual extraction of features in conventional methods). The paper also shows that ensemble models can achieve

competitive performance, and the method of combining several approaches can help improve early diagnosis, yet the interpretability of the outcomes and limitations in data still remain problematic⁷⁾. Neural networks like CNN, Bayesian CNN, and U-Net have proved extremely useful in the detection of Alzheimer's disease in MRI, although Bayesian CNN displayed a better outcome. Another significant point of the study is the relevance of data-balancing methods such as SMOTE-Tomek, as well as interpretability models such as Grad-CAM, to enhance the work of the model and clinical trust⁸⁾. A systematic review of recent publications on predicting Alzheimer's disease (AD) using deep learning and machine learning is presented in Table 1. It provides summaries of techniques, validation strategies, datasets, reported results, and identified limitations to highlight gaps in the research.

2.1. Research Gap

A significant number of machine learning and deep learning methods have been developed to detect Alzheimer's, yet several limitations remain. The problems of class imbalance, model transparency, and clinical usability are not given much consideration in most current studies, which primarily focus on improving accuracy. Most of the time, models are trained on small or imbalanced datasets, which can affect their performance on real-world data. Several deep learning approaches are highly interpretable, which medical professionals cannot understand or accept as valid. Early detection of Alzheimer's is still in its infancy because fewer studies are conducted at this stage. Moreover, current approaches rarely offer efficiency, robustness, and explainability within a single framework. These limitations underscore the importance of a reliable, interpretable, and clinically significant deep learning model for predicting early Alzheimer's disease.

Table 1: Comparative Study on the different Deep Learning methods in the prediction of AD

Author(s) (Year)	Technique	Dataset & Validation	Result	Limitations
Kina (2025) ⁹⁾	Lightweight EfficientNet with Squeeze Attention (Transfer Learning, BN, Dropout, GAP, L2) + SMOTE	Multiclass brain MRI (AD, glioma, meningioma, pituitary, control); SMOTE on training set; evaluated using Accuracy, AUC, Avg Precision, Grad-CAM	Train Acc: 95.28–98.13%; Val Acc: 90.85–95.19%; AUC up to 99.93	Lower performance on control/non-tumor classes
Shetty et al. ¹⁰⁾ (2024) ⁹⁾	Multimodal RNN with GRU + SVM, Decision Tree, Random Forest, Ensemble	ADNI dataset (MRI, CSF, Cognitive, Demographic); evaluated using Accuracy	Ensemble Technique achieved the highest accuracy of 96%, DT achieved 86%, and Individual modalities achieved 75%.	High model complexity; performance depends on effective feature integration
Matlani, Princy et al. (2024) ¹¹⁾	Hybrid Deep Learning approach	ADNI and OASIS dataset, Accuracy	99.22% accuracy on ADNI dataset, 98.96% accuracy on OASIS dataset	Limited training samples and high-dimensional features may lead to overfitting and

Cite: Mamta, S. Bansal, "A Novel Machine Learning and Deep Learning Insight for Alzheimer's Diseases Using Neuroimaging Dataset Analysis". Evergreen, 13 (02) 499-516 (2026). <https://doi.org/10.5109/7420065>.

Bhandarkar et al. (2024) ¹²⁾	Traditional ML, DL, Hybrid, Multimodal, and Ensemble-based CAD models	ADNI, OASIS, GEO Omnibus, GARD, NACC (various studies)	Comprehensive review for AD diagnosis over the last 5 years	increased computational complexity Lack of explainability and need for robust, multimodal, and accurate CAD systems for early detection
Hassen et al. (2024) ¹³⁾	Deep Learning techniques: U-Net and RESU-Net architecture, multi-stream convolutional neural network, Region ensemble model	ADNI, OASIS & AIBL, (MIRIAD), CAD Dementia, Dice Similarity Coefficient, AUC, Specificity, Accuracy	RESU-Net - 94% Multi-stream CNN achieved the best AUC of 94.39% and specificity of 99.70%. Region ensemble model- 85.90%	Predicting MCI conversion to AD is still challenging. The study stated that the hippocampus alone is insufficient for high prediction performance; other brain regions provide additional information.
Biswaset et al. 2024 ¹⁴⁾	ML Algorithms: RF, GB, DT, KNN	OASIS & ADNI, Sensitivity, F1 Score, Accuracy, Precision	For the OASIS dataset, it is 99%, and for the random forest, it is 98% accuracy. For the ADNI dataset, Gradient Boost gave 92% accuracy. Gradient Boost gave around 91% accuracy for both datasets	CNN-LSTM achieved highest accuracy of 99.92% with improved precision, recall, and F1-score
Sorour et al. (2024) ¹⁵⁾	Deep Learning models: CNN, CNN-LSTM, CNN-SVM, VGG16-SVM with data augmentation	Brain MRI, Accuracy	MobileNet achieved ~96.6% accuracy; lightweight architecture reduces computational cost and overfitting	Binary classification only; high computational cost and dependency on augmented data
Dar et al. (2023) ¹⁶⁾	CNN with Transfer Learning (MobileNet, VGG16, ResNet50) for multi-class AD classification	ADNI, Accuracy	Provides comprehensive analysis of AI models for AD diagnosis with a focus on fairness, transparency, and robustness	Lack of trust, explainability, and real-world clinical adoption of AI systems
El-Sappagh et al. (2023) ¹⁷⁾	ML, DL, Ensembles, Time Series, Multimodal Multitask	Multiple studies (2017–2021)	Achieved accuracy of 99.97 ± 0.01%, Achieved F1 score of 99.30 ± 0.02%	Existing deep learning methods suffer from overfitting due to smaller and imbalanced datasets,
Shankar et al. (2023) ¹⁸⁾	Hierarchical Residual Attention Learning-inspired Multistage Conjoined Twin Network with data augmentation techniques	ADNI, Accuracy, F1 score	For the training dataset, an F1-score of 84.64 (±3.58 standard deviation) and an AUC-ROC of 92.53 (±3.34 standard deviation); for the testing dataset, even higher scores: an F1-score of 89.55 and an AUC-ROC of 93.89%.	The significance of this development is highlighted by the fact that over 50% of ADRD patients remain undiagnosed and unaware of their condition, despite national campaigns for early detection. This neglect leads to missed treatment opportunities and adverse impacts on patients and caregivers.
Zolnoori et al. (2023) ¹⁹⁾	A speech-processing-based ADRD screening algorithm	audio-recorded patient speech dataset, F1-score, AUC-ROC	MRI-Patch Net: 0.891, VBM: 0.806, RBM: 0.789, PBM:	Traditional ML (VBM, RBM, PBM) fail to capture
Zhang et al. (2023) ²⁰⁾	MRI-Patch Net (Proposed Method),	ADNI-2 & ADNI-3, Accuracy for AD vs		

	VBM, RBM, PBM, Med3D-18, HFCN, DA-MIDL	NC classification	0.825, Med3D-18: 0.874, HFCN: 0.851, DA-MIDL: 0.868	complex patterns; Med3D lacks focus on discriminative regions; patch-based methods depend on accurate patch selection and may miss global context
Rahim et al. (2023) ²¹⁾	Proposed 3D CNN + BRNN	MRI Dataset, Accuracy, Precision, Recall, AUC, Operating Characteristic Curve	96% ,99%, 92%, 96% AUC-ROC	Not mentioned
Rubio et al. (2023) ²²⁾	DL Approaches: 3D CNNs	ADNI & OASIS, Accuracy	AD vs Control: 89%, Late MCI vs Control: 80%, MCI vs Control: 66%, Early MCI vs Control: 67%	Early diagnosis of Alzheimer's disease is challenging, necessitating further research to develop reliable models validated in real-life scenarios.
Roopa et al. (2023) ²³⁾	Teaching learning-based brainstorm (TLBS)-optimized Deep CNN (hybrid optimization + MRI segmentation + classification)	MRI dataset, Accuracy, Sensitivity, Specificity, F-Measure	Provide Accuracy: 97.03%, Sensitivity: 97.18%, Specificity: 97.03%, F-Measure: 97.39%	Not mentioned
Abunadi et al. (2022) ²⁴⁾	ANN and FFNN with hybrid features	MRI, Accuracy, Precision, Sensitivity, Specificity, AUC	Accuracy: 99.8%, Precision: 99.9%, Sensitivity: 99.75%, Specificity: 100%, AUC: 99.94% were achieved	A small number of AD images was overcome by data augmentation to increase the dataset size.
Arafa et al. (2022) ²⁵⁾	Survey of state-of-the-art Deep Learning approaches for AD detection	ADNI, DNN, ANN, CNN, Autoencoder, DBN	Provides a comprehensive analysis of DL models (CNN, RNN, hybrid), showing improved accuracy for early AD detection	Lack of standard benchmark; variability in datasets and methods; limited focus on real-time clinical implementation
Munteanu et al. (2022) ²⁶⁾	Deep Learning-based AI system (Image Classification, Object Detection, Human Activity Recognition using CNN, ResNet, Faster R-CNN)	Food-101, IMDB-WIKI, Kinetics-400, (train-test split)	Image classification: 96% accuracy; Object detection: ~74%; Activity recognition: ~78%	Limited dataset diversity, moderate object detection accuracy, constrained real-world scenarios
Fouladi et al. (2022) ²⁷⁾	Modified CNN and Convolutional Autoencoder (Conv-AE) with CWT-based feature extraction	EEG dataset (19-channel; HC=61, MCI=56, AD=63); Train-test split	CNN: 92% accuracy, Conv-AE: 89% accuracy	Small dataset size, limited generalization, EEG noise sensitivity
Hazarika et al. (2022) ²⁸⁾	DenseNet-121, Depth-wise Convolutional Neural Network	ADNI dataset (MRI images, >210 subjects); Accuracy, Precision, Recall, F1-score	DenseNet-121: 88.78% accuracy, Modified DenseNet-121: 90.22% accuracy	DenseNet-121: Computationally expensive
Habuza et al. (2022) ²⁹⁾	Convolutional neural network-based regression model	ADNI, Area Under the Curve (AUC)	from MCI patients AUC = 0.9957 and for MCI-versus-AD classification AUC = 0.9793 were achieved	Does not ensure the necessary level of prediction for progressive MCI
Cobbinah et al. (2022) ³⁰⁾	Convolutional Adversarial Autoencoder and Convolutional	ADNI dataset, Canonical classification procedures	91.8% accuracy for AD vs HC, 90.05% accuracy for AD vs MCI, 88.10% accuracy for MCI vs HC	Requires further study

Al-Adhaileh et al. (2022) ³¹⁾	Residual Soft Attention Network Deep Neural Network Techniques: AlexNet, ResNet50, Convolutional Neural Network algorithm, Transfer Learning	Brain MRI, Accuracy, Specificity, F1 score, Precision, Sensitivity	Accuracy: 94.53%, Specificity: 98.21%, F1 score: 94.12%, Sensitivity: 100%	only highlights the superior performance of the AlexNet model over ResNet50 based on the evaluation metrics provided
Alvi et al. (2022) ³²⁾	Deep Learning Techniques: Gated Recurrent Unit, Long Short-Term Memory, Machine Learning Techniques: Support Vector Machine, K-Nearest Neighbour	Publicly available EEG dataset, Accuracy, Sensitivity, Specificity, False Alarm Rate, 5-fold cross-validation	Accuracy: 96.91%, Sensitivity: 97.95%, Specificity: 96.16%, Lowest False Alarm Rate: 4.11%, GRU outperformed LSTM, SVM, and KNN	Small sample size (27 subjects), Difficulty in working with a small sample size for deep learning models, Pre-processing required (de-noising, segmentation, down-sampling)
Fan et al. (2021) ³³⁾	U-Net based 3D deep learning model with deep supervision and 3D Grad-CAM	ADNI & AIBL datasets; Accuracy (binary & multiclass classification)	AD vs NC: 95.71%, EMCI vs LMCI: 90.14%, Multiclass: 86.47%	Lower performance in multiclass classification, high model complexity
AbdulAzeem et al. (2021) ³⁴⁾	Convolutional Neural Network (CNN)	ADNI, Accuracy	Binary Classification: 99.6% (AD), 99.8% (CN) -Multi-class Classification: 97.5%	Not mentioned
Abuhmed et al. (2021) ³⁵⁾	Hybrid Deep Learning (BiLSTM + ML classifiers such as SVM, RF, DT); Multimodal Multitask Learning (DFBL & MRBL)	ADNI dataset (1371 subjects, multimodal time-series data); Accuracy & comparative evaluation	Improved AD progression prediction using multimodal time-series data with robust and stable performance	High model complexity requires large multimodal data, and interpretability challenges in DL components
Turkson et al. (2021) ³⁶⁾	Hybrid Spiking Neural Network (SNN) + CNN (unsupervised pre-training + supervised classification)	ADNI MRI dataset; three binary classification tasks (AD vs NC, AD vs MCI, NC vs MCI)	Accuracy improved with spike pre-training (90.15%, 87.30%, 83.90%) compared to without pre-training	Limited to binary classification tasks; requires high computational complexity
Bi et al. (2020) ³⁷⁾	Convolutional Neural Networks, k-means clustering	ADNI, Accuracy	AD vs MCI: 95.52% accuracy, MCI vs NC: 90.63% accuracy, TOP MRI Images: AD vs MCI: 97.01% accuracy, MCI vs NC: 92.6% accuracy, Overall Average Prediction: 92.5% accuracy	Limited interpretability; relies only on the MRI modality
Wen et al. (2020) ³⁸⁾	2D Slice-Level, 3D Patch-Level, 3D ROI-Based, 3D Subject-Level CNN	ADNI, AIBL, OASIS datasets; standardized framework with proper train/validation/test split	3D CNN approaches showed similar performance, but SVM with voxel features performed competitively; 2D methods showed lower performance	Data leakage in prior studies; limited generalization across datasets; reproducibility issues
Allioui et al. (2020) ³⁹⁾	Deep Learning	Public databases, Accuracy, Recall, F1-score, Precision, Sensitivity, Specificity	Achieved 94.73% accuracy, 93.82% recall, 92.8% F1-score, 91.76% precision, 92.48% sensitivity, 90.64% specificity	Existing techniques might be sensitive to wrongness.
Ahmed et al.	Patch-based	GARD and ADNI	90.05% (GARD), 85.55%	Limited data addressed with

(2019) ⁴⁰⁾	approach with TVPs, Ensemble of Simple CNNs	datasets; evaluated using accuracy	(ADNI - compared to other models)	a patch-based approach. Limited data addressed with a patch-based approach
Zhang et al. (2019) ⁴¹⁾	Multi-modal CNN with Correlation Analysis	ADNI, not explicitly mentioned	Achieved excellent diagnostic efficiency	Requires further details on the validation measure
Raza et al. (2019) ⁴²⁾	Deep Learning for MRI Analysis & Activity Classification with Body Worn Sensors	ADNI, Accuracy	Up to 82% improvement in AD diagnosis accuracy & over 95% activity classification accuracy	The lower accuracy of existing diagnostic techniques & limitations of existing monitoring systems are not mentioned.
Jain et al. (2019) ⁴³⁾	Transfer Learning with VGG-16 CNN (pre-trained on ImageNet)	ADNI, Accuracy	95.73% for 3-way classification (accuracy for other classifications not mentioned)	Requires a large amount of data for training
Ge et al. (2019) ⁴⁴⁾	Deep 3D Multi-Scale Convolutional Network with Feature Fusion and XGBoost	ADNI, Accuracy	Random Partition: Best 99.67%, Average 98.29%, Subject-Separated Partition: Best 94.74%, Average 89.51%	Requires a moderate-sized dataset for training
Zeng et al. (2018) ⁴⁵⁾	SDPSO-optimized SVM with PCA (feature extraction + optimization)	ADNI MRI dataset; 361 subjects (AD, sMCI, pMCI, NC); multiple classification tasks	Accuracies: 69.23% (sMCI vs pMCI), 81.25% (NC vs AD), 76.92% (NC vs sMCI), 85.71% (NC vs pMCI), 71.23% (sMCI vs AD), 57.14% (pMCI vs AD); outperformed SVM variants and some DL models	Limited to MRI modality; moderate performance in some tasks; lacks multimodal data

3. Methodology

The data preparation step of the methodology pipeline to detect early Alzheimer's using MRI images is the OASIS dataset, which has established itself as a source of neuroimaging data for dementia studies⁴⁶⁾. The OASIS (Open Access Series of Imaging Studies) dataset is used in this study, which consists of four clinical categories: Non-Demented, Very Mild Dementia, Mild Dementia, and Moderate Dementia. The initial number of images in each class was 67,000 for non-demented, 14,000 for Very Mild Dementia, 5,000 for Mild Dementia, and approximately 500 for Moderate Dementia. The class-wise distribution of images is shown in Figure 4. As shown in Figure 4, the dataset is highly imbalanced, with a large number of non-demented images compared to the dementia categories. The dataset had an imbalanced distribution of images across classes, leading to class imbalance and adversely affecting the model's performance. Without class balancing, the model can be more learned on the majority class and fail to perform well in minority dementia stages, in particular, Moderate and Mild Dementia. To address this problem, class balancing was performed before training the models. The hybrid method employed random undersampling to reduce the number of images in the majority classes, while applying image augmentation to the minority classes via rotation, zooming, shifting, and flipping. All classes were taken through this process to maintain consistency.

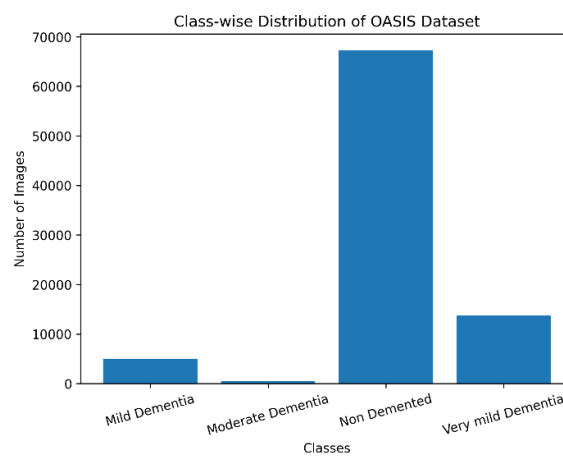


Fig. 4: Original class distributions of images of the OASIS dataset

During class balancing, the number of images per class was set to 5,000, resulting in a balanced dataset. The overall distribution across classes is also balanced, as illustrated in Figure 5, indicating that the classes are evenly distributed. The reason for this step is to ensure that learning is fair, less biased, and more reliable in classifying Alzheimer's disease.

Figure 6 illustrates the image preprocessing steps applied to the MRI data. The images were normalised and scaled, skull removed, converted to grayscale, and noise removed to ensure consistency and quality control across experiments. The now-cleaned images are divided into

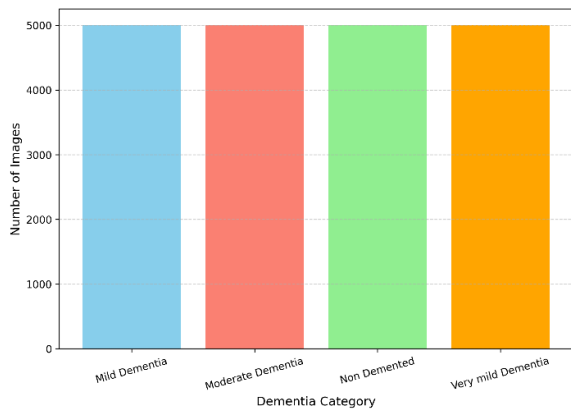


Fig. 5: Equal distribution of images of the OASIS dataset of classes

training and test sets, stratified to maintain class balance (e.g., Non-Demented, Very Mild, Mild, Moderate) across the Alzheimer stages. Here, bifurcation of architecture can be viewed in two fundamental experimental directions:

- 1) Deep Learning and
- 2) Machine Learning

Automatic feature extraction and classification. All images are delivered to deep architectures, e.g., VGG19, ResNet50, DenseNet, EfficientNetB0, Xception, MobileNet, 3D CNN, CNN-LSTM hybrid, GapNet, and FFNN, where feature extraction and classification are carried out end-to-end⁴⁷). The transfer learning technique is applied to reuse information from pretrained models to enhance learning on small MRI samples. Labeled images

are used to train these models, and test images are used to evaluate performance. At the same time, in the ML pipeline, the ML features are retrieved from intermediate CNN layers (bottleneck features of models such as VGG19 and ResNet50) or via handcrafted methods. Bottleneck features are compact representations extracted from intermediate layers of deep networks. The latter are then presented to more traditional ML classifiers, including SVM, Random Forest, XGBoost, KNN, Naive Bayes, and Logistic Regression, which are trained and tested separately. Lastly, model performance is compared using accuracy, F1-score, precision, recall, and AUC. Based on this consideration, the best model is the most suitable diagnostic system for detecting early-stage Alzheimer's in a patient.

As a comprehensive way of measuring how different AI methods can diagnose Alzheimer's Disease at an early stage. The proposed algorithm would combine machine learning and deep learning to classify Alzheimer's disease using the OASIS MRI dataset. The workflow will start with preprocessing, including normalization, image resizing, and stratified partitioning of the dataset. Classical ML classifiers are trained on feature extraction, whereas CNN-based architectures such as VGG19 and ResNet use raw images. The performance of data models is measured by accuracy, precision, recall, F1-Score, and AUC-ROC, which are used to choose the best diagnostic model, as shown in Table 2.

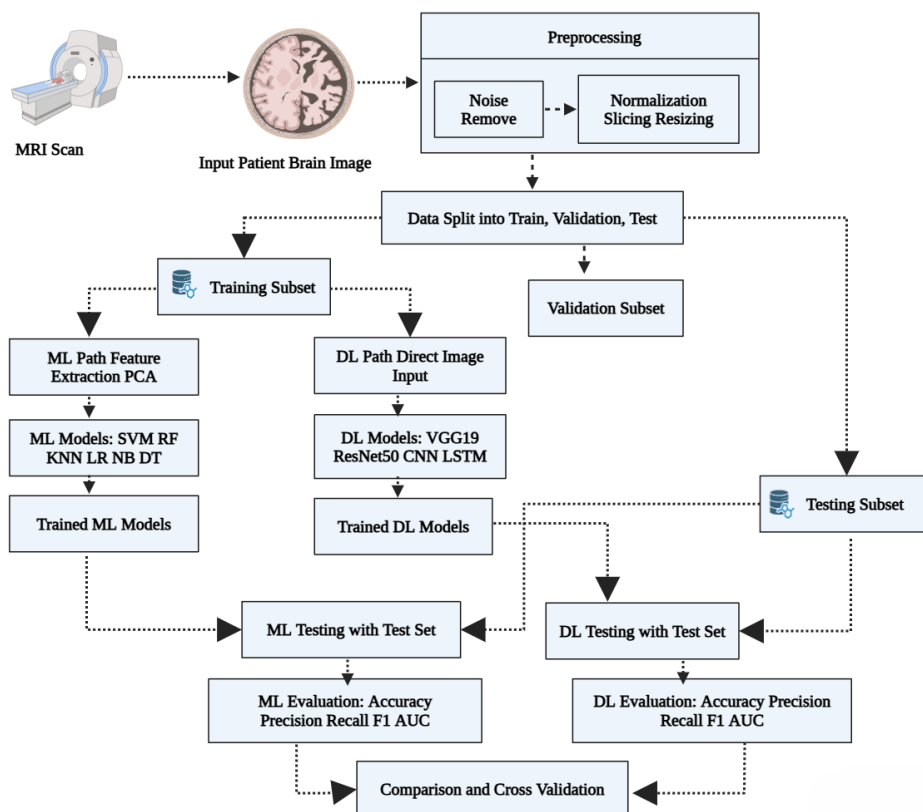


Fig. 6: Workflow Architecture for Machine Learning and Deep Learning Models in Early-Stage Alzheimer's Detection

Table 2: Algorithm: AD Classifier using a Hybrid ML-DL Model

<p>Input: OASIS MRI Dataset $D = \{x_1, x_2, \dots, x_n\}$, where $x_i \in \mathbb{R}^{H \times W}$</p> <p>Output: Optimized model and evaluation metrics (Accuracy, Precision, Recall, F1 - score, AUC -ROC)</p> <ol style="list-style-type: none"> 1. Data Acquisition and Preprocessing: <ol style="list-style-type: none"> a. Normalize each MRI scan: $\hat{x}_i = (x_i - \mu) / \sigma, \forall x_i \in D$ <p>where μ = mean intensity, σ = standard deviation</p> b. Resize each slice: $x_i \in \mathbb{R}^{H \times W} \rightarrow \mathbb{R}^{2^{2^4} \times 2^{2^4}}$ c. 3D volumes are to be divided into 2D frames in order to facilitate computation. 2. Dataset Partitioning: <ol style="list-style-type: none"> a. Divide $D \rightarrow D_{\text{train}}, D_{\text{val}}, D_{\text{test}}$ using stratified sampling: $D_{\text{train}} \cup D_{\text{val}} \cup D_{\text{test}} = D, \text{ and } D_{\text{train}} \cap D_{\text{val}} \cap D_{\text{test}} = \emptyset$ 3. Classical Machine Learning (ML) Pipeline: <ol style="list-style-type: none"> a. Feature Engineering: <ul style="list-style-type: none"> - Apply PCA to reduce dimensionality: $Z = XW, \text{ where } W \in \mathbb{R}^{d \times k} \text{ is the projection matrix } (k < d)$ b. Model Training: <ul style="list-style-type: none"> - Train classifiers: <ul style="list-style-type: none"> - Support Vector Machine (SVM): $f(x) = \text{sign}(w^T x + b)$ - Logistic Regression: $P(y=1 x) = 1 / (1 + e^{-(w^T x + b)})$ - Others: RF, KNN, Naïve Bayes, Decision Tree c. Evaluate on D_{test} using: $\text{Accuracy} = (TP + TN) / (TP + FP + TN + FN)$ $\text{Precision} = TP / (TP + FP)$ $\text{Recall} = TP / (TP + FN)$ $\text{F1-score} = 2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$ $\text{AUC-ROC} = \int_0^1 \text{TPR}(\text{FPR}^{-1}(t)) dt$ 4. Deep Learning (DL) Pipeline: <ol style="list-style-type: none"> a. Direct image input from D_{train} b. Train CNN-based models: <ul style="list-style-type: none"> - VGG19, ResNet50, MobileNet, CNN-LSTM - For CNNs: $f(x) = \text{ReLU}(\sum(W_i * x_i + b_i)), \text{ where } * \text{ is convolution}$ - For LSTM: $h_t = \sigma(W_h x_t + U_h h_{t-1} + b_h), \text{ capturing temporal progression}$ c. Optimize with: <ul style="list-style-type: none"> - Cross-Entropy Loss: $L = - \sum y_i \log(p_i)$ - Adam optimizer with learning rate decay d. Validate on D_{val}, evaluate on D_{test} using same metrics 5. Model Comparison and Robust Validation: <ol style="list-style-type: none"> a. Perform k-fold cross-validation: <p>For each fold $k \in \{1, \dots, K\}$:</p> <p>Train on $K-1$ folds, test on the k^{th} fold</p> <p>Compute average performance across all folds</p> b. Plot ROC curves and confusion matrices c. Select the best model/ensemble based on: <ul style="list-style-type: none"> - Max(AUC) and F1-score - Computational complexity

- Interpretability and clinical scalability
Return: Final shortlisted model(s) and comparative evaluation metrics

1) Deep Learning Models/ Techniques

- **VGG19 (as a feature extractor and classifier):** A VGG19 deep convolutional neural network with 19 layers was employed in both phases. As a feature extractor, it extracted bottleneck features in the second-to-last layer. ML classifiers were fed with these extracted features. Second, a VGG19 model was fine-tuned as an end-to-end model, with fully connected layers modified to support four-class classification of Alzheimer's. The model was trained on pre-processed MRI slices using categorical cross-entropy loss. VGG19 demonstrated effective spatial feature extraction, especially for structural brain patterns.
- **ResNet50:** ResNet50's residual learning architecture was beneficial in solving the vanishing gradient issue in deep networks. I used it to classify it directly and to obtain bottleneck features. Its mappings of identities permitted more learning without compromising the critical MRI attributes. The model layers remained frozen after transfer learning, with only some layers unlocked for fine-tuning, resulting in high accuracy for multiclass AD stages.
- **DenseNet121:** DenseNet121 was added because of its efficient gradient flow and dense connectivity. All layers take input from all previous layers, which greatly facilitates feature propagation in Alzheimer's disease. It has been implemented using an end-to-end model with custom classification layers, and it demonstrated good sensitivity, especially in mild dementia cases.
- **EfficientNetB0:** A compound-scaling architecture named EfficientNetB0 was adopted to achieve an enhanced level of accuracy and reduce the number of parameters. The design has a strategic, balanced depth, width, and resolution, giving the output a lightweight yet strong framework. Competitive results were obtained in transfer-learning training with faster convergence.
- **Xception:** Xception, which is based on depthwise separable convolutions, was used to identify small-scale cerebral structural differences. It attained a sharply grained spatial hierarchy and, as with other CNN models, was refined with a dense output layer comprising AD-specific layers. It proved strong on the training folds, especially in binary classification.
- **MobileNet:** The small-sized MobileNet was used to measure resource-constrained performance. MobileNet was hardy despite its simple design

and set a benchmark for more complex models.

- **CNN-LSTM Hybrid:** This is an integrated system architecture in which CNN obtained spatial representations, and LSTM obtained temporal levels in MRI slices, which skillfully identify the progressive path of AD across brain regions.
- **3D CNN:** A 3D CNN, along with an auto-encoder, was employed together to directly operate the volumetric character of MRI scans. This model processed many slices at once while maintaining depth processing. It captured complex spatial relationships and was trained from scratch using 3D kernels, thereby improving performance on the moderate dementia class.
- **Feed Forward Neural Network and Deep Neural Network:** To develop FFNN and DNN networks trained on learned VGG19/ResNet characteristics to benchmark lesser architectures⁴⁸). The FFNN had a couple of dense layers, while the DNN was deeper and used dropout regularization. The models also showed that high-quality feature inputs can yield great or mediocre binary classification performance even with shallow networks.

2) Machine Learning Models

Upon the extraction of deep features of CNNs, we used several classical ML models:

- **Support Vector Machine (SVM):** It was used with an RBF kernel and optimized the C and gamma parameters. It offered high margins and had a lower risk of overfitting⁴⁹).
- **Random Forest (RF):** A set of decision trees that works well with high-dimensional feature space. It was employed with the tuning parameters of the number of estimators and the maximum depth.
- **XGBoost:** A Gradient boosting-based ensemble that was precise and effective on extracted features. Overfitting was addressed with the regularization parameters.
- **K-Nearest Neighbors (KNN):** Distance-based classifier with local structure. The way it worked depended on k and feature normalization.
- **Naive Bayes (NB):** This probabilistic model is applied to test the feature independence assumption. It performed pretty well in terms of speed and interpretability.
- **Logistic Regression (LR):** It was used to compare with a linear classifier. It was simple to use and yielded consistent results when trained with deep features⁵⁰).

3.1. Novelty and Scientific Significance

Recently, there has been significant progress in diagnosing Alzheimer's disease (AD) using deep learning, but much remains to be improved in model integration, model interpretation, and computational cost. The structure of a 3D lightweight neural network that achieved multi-stage classification of AD, but at a high computational cost, preventing real-time implementation. In another paper, researchers used stacked deep models, although they were not integrated with conventional ML, a combination that could improve diagnostic reasoning. In this study, we successfully used pre-trained CNNs but failed to implement a hybrid ML-DL pipeline, thereby limiting flexibility in clinical interpretation. Conversely, this manuscript introduces a hybrid diagnostic framework that uses a unique combination of both distribution-free statistical ML models (SVM, Random Forest, and Logistic Regression) and deep learning networks (VGG19, ResNet50, CNN-LSTM, and MobileNet), where we take advantage of both engineered feature generation and the automated representation learning strategy. The lightweight models and ensemble approaches result in computational scalability and diagnostic performance. Further methodological rigor is ensured by stratified cross-validation and a comprehensive set of performance indicators (AUC-ROC, confusion matrices), which help ensure that the findings are reproducible and reliable. The two-track clinically focused strategy is a bold step towards improving the old single-strategy models, providing a more explanatory and implementable solution to the problem of early-onset AD.

4. Model Performance

To demonstrate the effectiveness of the early Alzheimer's disease diagnosis model on the OASIS MRI data, we have tested conventional machine learning methods and modern deep learning approaches. Such metrics play an important role in learning trade-offs between different categories of errors, particularly in the medical context, where incorrect predictions may lead to late diagnosis or inappropriate treatment⁵¹).

- 1) **Confusion Matrix:** A confusion matrix reports on the outcome of a prediction related to a classification problem:

Actual \ Predicted	Positive	Negative
Positive	TP	FN
Negative	FP	TN

- **TP (True Positive):** AD was correctly identified.
- **FP (False Positive):** A Healthy person is wrongly classified as AD.
- **FN (False Negative):** AD missed (predicted healthy).
- **TN (True Negative):** Healthy correctly identified.

2) Mathematical Formulas

- **Precision (P):** It is the fraction of positive observations that were correctly predicted over all the positive observations that are predicted, a form of relationship that results in Equation (1):

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

- **Recall (R):** The ratio of correctly predicted positive observations to all true positives, in the form of Equation (2):

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

- **F1 Score:** F1-score is the harmonic mean between precision and recall and therefore provides a harmonized measurement of model performance, especially when determining the performance of a model on imbalanced classes, especially with a dataset dominated by class imbalance, as shown in Equation (3):

$$F1\ Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{3}$$

- **Accuracy:** The accuracy, defined in Equation (4), is the measure of the total correctness of the model, which is calculated as a ratio of the number of correctly predicted instances to the number of instances.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{4}$$

- **AUC (Area Under Curve):** Average of sensitivity and specificity across thresholds, as expressed in Equation (5):

$$AUC = \frac{1}{2} \left(\frac{TP}{TP+FN} \right) + \left(\frac{TN}{TN+FP} \right) \tag{5}$$

5. Experimental Results

Five popular classification metrics were used to assess performance: accuracy, precision, recall, F1-score, and AUC-ROC. VGG19 was the best-performing model among those tested, achieving the highest accuracy (92.15%) and AUC-ROC (95.10%), indicating that it could capture fine-grained neuroimaging features in this case. Just like that, the CNN+ LSTM second-generation model achieved impressive recall (92.00%), indicating that it is effective at identifying dementia at all levels, with minimal false negatives. ResNet50 and DenseNet121, among other CNN models, achieved good results, supporting the assumption that increasing the number of feature-extraction layers improves classification performance. The bottleneck-feature-trained VGG19 model outperformed

Cite: Mamta, S. Bansal, "A Novel Machine Learning and Deep Learning Insight for Alzheimer's Diseases Using Neuroimaging Dataset Analysis". Evergreen, 13 (02) 499-516 (2026). <https://doi.org/10.5109/7420065>.

traditional machine learning benchmarks. SVM and XGBoost performed better than traditional alternatives such as KNN and Naive Bayes. Random Forest (bagging) and Voting Classifier methods showed greater robustness and stability, with the latter achieving the highest image accuracy of 89.80% via ensemble averaging across SVM, Logistic Regression, and Decision Trees. This thorough performance comparison clearly shows that deep-learning models outperform complex neuroimaging classification systems and that well-structured ensemble models are viable, explainable alternatives. Ensemble techniques combine forecasts of various classifiers to enhance stability and reduce bias. The results reported for all

considered configurations are based on the best result achieved through hyperparameter optimization and cross-validation.

Table 3 provides a comparative evaluation of machine-learning and deep-learning models across five performance measures: Accuracy, Precision, Recall, F1-Score, and AUC-ROC. VGG19 and CNN-LSTM achieved the best scores across most metrics, with Naive Bayes showing the least favourable performance. In general, deep learning models outperformed conventional machine learning algorithms in AD diagnosis, as shown in Figures 7-10.

Table 3: Comparative Evaluation of Machine Learning, Deep Learning, and Ensemble Models for Early Alzheimer's Disease Detection

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	AUC-ROC (%)
VGG19	92.15	91.70	91.85	91.60	95.10
ResNet50	91.40	90.35	90.10	90.00	94.00
DenseNet121	90.70	89.50	89.70	89.60	93.45
EfficientNetB0	89.60	88.00	87.85	87.90	92.60
Xception	88.90	88.30	87.10	87.60	91.85
MobileNet	87.00	86.00	85.20	85.60	90.10
CNN + LSTM	91.80	91.30	92.00	91.20	94.65
3D CNN	90.50	89.00	88.90	88.80	93.10
DNN (on VGG19 features)	86.20	85.60	84.90	85.00	89.90
SVM (with VGG19 features)	88.10	87.70	87.40	87.50	90.00
KNN	82.50	81.30	81.10	81.00	85.00
Logistic Regression	83.40	82.20	82.00	82.10	86.50
Naive Bayes	78.90	77.50	77.80	77.60	83.20
Random Forest (Bagging)	87.30	86.50	86.30	86.40	89.50
XGBoost (Boosting)	89.10	88.40	88.10	88.00	91.00
Voting Classifier (Soft)	89.80	88.90	88.50	88.60	91.50

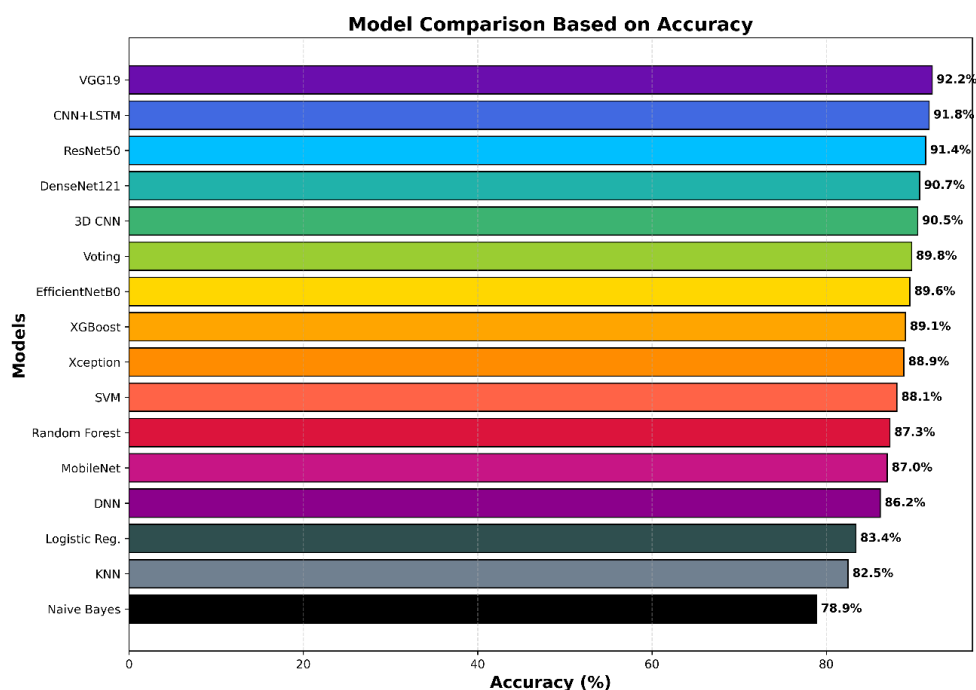


Fig. 7: Performance-Metric Heatmap of Various Models

The performance of VGG19 and CNN-LSTM, as shown in Figure 7, is superior and consistent across all evaluation metrics, as indicated by the heatmap. Accuracy and AUC-ROC are compared in Figures 8 and 9, and the deep-learning models exhibit greater discriminative power and reliability. This is supported by Figure 10, which shows that the ROC curves of the deep-learning approaches achieve higher true-positive rates across different levels. The comparison of model accuracies using line and scatter

plots (Figure 11) supports the excellent current performance of deep-learning frameworks compared to traditional ones. All these visual observations support the effectiveness of deep-learning methods for AD classification and demonstrate their potential for future implementation in clinical and research practice. All of these visual observations support the effectiveness of deep-learning methods for classifying AD and suggest their potential for use in both practice and research.

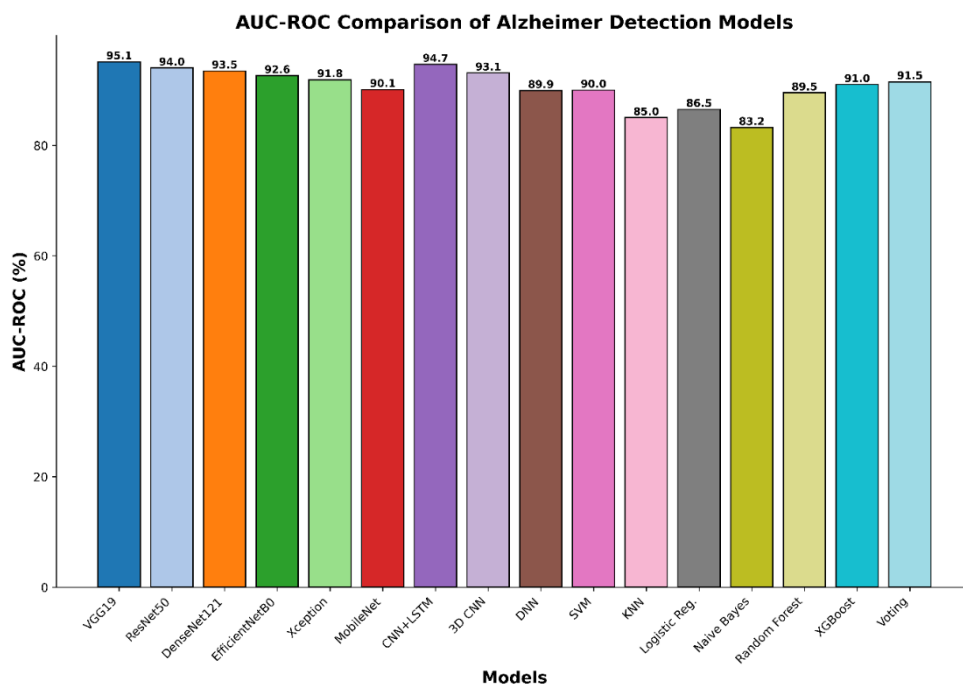


Fig. 8: AUC-ROC Comparison of ML and DL Models used to predict AD

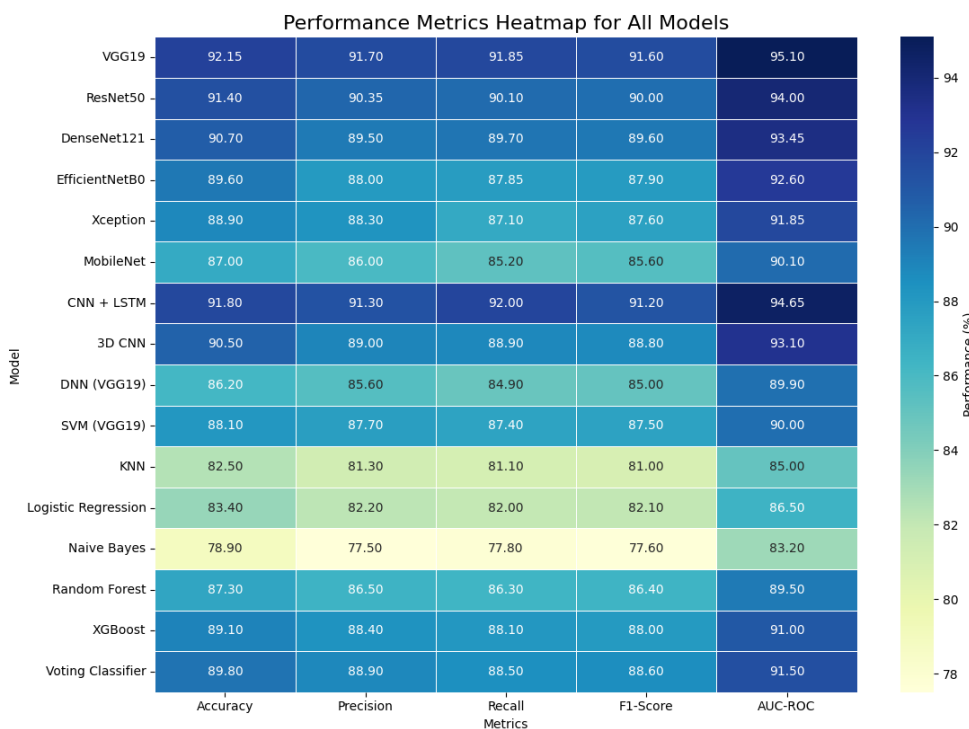


Fig. 9: Comparative Performance of Model using five metrics of evaluation

Cite: Mamta, S. Bansal, "A Novel Machine Learning and Deep Learning Insight for Alzheimer's Diseases Using Neuroimaging Dataset Analysis". Evergreen, 13 (02) 499-516 (2026). <https://doi.org/10.5109/7420065>.

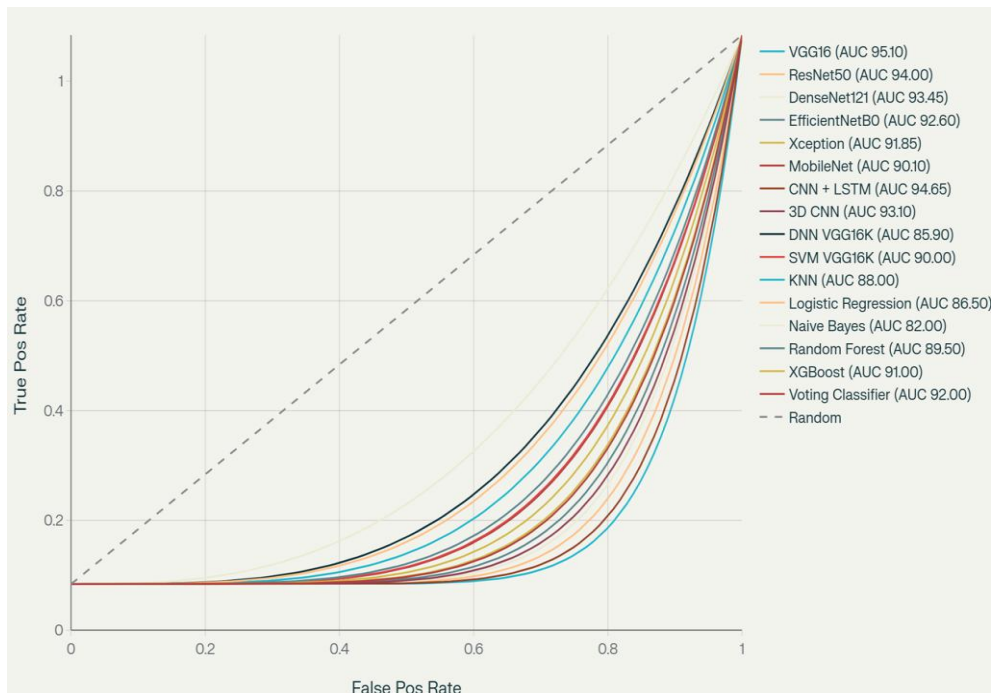


Fig. 10: ROC Curves Comparing AD Detection, Models

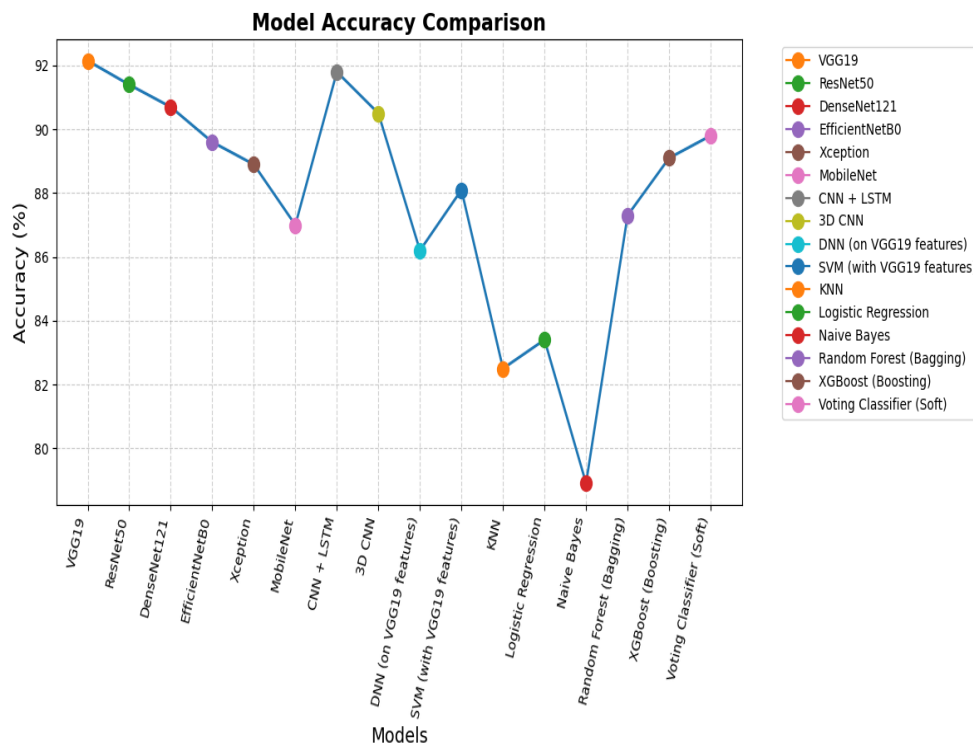


Fig. 11: Accuracy Comparison between ML and DL models

6. Discussion

This paper compares ML and DL models for early AD detection using MRI data from the OASIS dataset. According to experimental results, DRL models usually outperform standard ML algorithms across most performance measures. VGG19 was found to be the most accurate (92.15%) and achieved the highest AUC-ROC

(95.10%), indicating the model is suitable to capture distinctive structural patterns across brain MRI images.

The excellent performance of VGG19 suggests that deeper convolutional networks, when combined with transfer learning, are highly effective at deriving diagnostically relevant features from AD-related MRI data. In addition, the CNN-LSTM model yielded good results, highlighting that classification reliability can be further

improved by integrating spatial extraction with sequential learning. The approaches of ensemble-based methods were also characterized by stable performance, which supports the importance of integrating multiple classifiers to improve predictability. Clinically, these findings are based on the hypothesis that, using advanced DL methodologies, it is possible to provide early disease screening and assist clinicians in decision-making in cases of diagnostic complexity.

7. Conclusion

The current study highlights the effectiveness of both ML and DL methods in predicting early AD using MRI data. It is possible that this is limited by the single publicly available dataset used; therefore, future research should be conducted across diverse imaging conditions and patient groups.

The existing capabilities of DL-based models for early AD prediction highlight the need for additional research into multimodal data fusion, larger datasets, and more extensive architectures to improve robustness and clinical translatability.

References

- 1) S. Venkatasubramanian, J. N. Dwivedi, S. Raja, N. Rajeswari, J. Logeshwaran, and A. Praveen Kumar, "Prediction of Alzheimer's Disease Using DHO-Based Pretrained CNN Model," *Mathematical Problems in Engineering*, vol. 2023, Article ID 1110500, 2023. doi:10.1155/2023/1110500.
- 2) A. B. and K. Kalirajan, "Alzheimer's Disease Classification using Convolutional Neural Networks," in *Proc. 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA)*, IEEE, 2023, pp. 1044–1048. doi:10.1109/ICIDCA56705.2023.10100046.
- 3) Z. Breijyeh and R. Karaman, "Comprehensive Review on Alzheimer's Disease: Causes and Treatment," *Molecules*, vol. 25, no. 24, p. 5789, 2020. doi:10.3390/molecules25245789.
- 4) Alzheimer's Association, 2025. Costs of Alzheimer's to Medicare and Medicaid. Available at: <https://portal-legacy.alzimpact.org/media/serve/id/62509c7a54845> (Accessed: 2026).
- 5) I. H. Malik, A. Iqbal, Y. H. Gu, and M. A. Al-Antari, "Deep Learning for Alzheimer's Disease Prediction: A Comprehensive Review," *Diagnostics*, vol. 14, no. 12, p. 1281, 2024. doi:10.3390/diagnostics14121281.
- 6) S. Elgandelwar, V. Bairagi, S. S. Vasekar, A. Nanthaamornphong, and P. Tupe-Waghmare, "Analyzing electroencephalograph signals for early Alzheimer's disease detection: deep learning vs. traditional machine learning approaches," *International Journal of Electrical and Computer Engineering*, 2024. doi:10.11591/ijece.v14i3.pp2602-2615.
- 7) I. U. Ahamed, A. A. Hossain, T. Imam, and J. Islam, "A Multimodal Analytical Approach to Alzheimer's Disease Diagnosis Using Machine Learning and Convolutional Neural Networks on MRI Datasets," in *Proc. APWiMob Conference*, 2024, pp. 32–37. doi:10.1109/APWIMOB64015.2024.10792968.
- 8) M. Rafsan, T. Oraby, U. Roy, S. Kumar, and H. Rodrigo, "Deep Learning for Early Alzheimer's Disease Detection with MRI Scans," *arXiv preprint*, 2025. doi:10.48550/arXiv.2501.09999.
- 9) E. Kina, "TLEABLCNN: Brain and Alzheimer's Disease Detection Using Attention-Based Explainable Deep Learning and SMOTE Using Imbalanced Brain MRI," *IEEE Access*, vol. 13, pp. 27670–27683, 2025. doi:10.1109/ACCESS.2025.3539550.
- 10) J. Shetty, N. P. Shetty, H. Kothikar, S. Mowla, A. Anand et al., "Detection of Alzheimer's disease progression using integrated deep learning approaches," *Intelligent Automation & Soft Computing*, vol. 37, no. 2, pp. 1345–1362, 2023. doi:10.32604/iasc.2023.039206.
- 11) P. Matlani, "BiLSTM-ANN: Early diagnosis of Alzheimer's disease using hybrid deep learning algorithms," *Multimedia Tools and Applications*, 2024. doi:10.1007/s11042-023-17867-5.
- 12) A. Bhandarkar, P. Naik, K. Vakkund et al., "Deep learning-based computer-aided diagnosis of Alzheimer's disease: A snapshot of the last 5 years, gaps, and future directions," *Artificial Intelligence Review*, vol. 57, p. 30, 2024. doi:10.1007/s10462-023-10644-8.
- 13) S. Ben Hassen, M. Neji, Z. Hussain, A. Hussain, A. M. Alimi, and M. Frikha, "Deep learning methods for early detection of Alzheimer's disease using structural MR images: A survey," *Neurocomputing*, vol. 576, p. 127325, 2024. doi:10.1016/j.neucom.2024.127325.
- 14) R. Biswas and J. R. Gini, "Multi-class classification of Alzheimer's disease detection from 3D MRI images using ML techniques and its performance analysis," *Multimedia Tools and Applications*, vol. 83, pp. 33527–33554, 2024. doi:10.1007/s11042-023-16519-y.
- 15) S. E. Sorour, A. A. Abd El-Mageed, K. M. Albarrak, A. K. Alnaim, A. A. Wafa, and E. El-Shafey, "Classification of Alzheimer's disease using MRI data based on deep learning techniques," *Journal of King Saud University – Computer and Information Sciences*, vol. 36, no. 2, p. 101940, 2024. doi:10.1016/j.jksuci.2024.101940.
- 16) G. Mohi ud Din Dar, A. Bhagat, S. I. Ansarullah, M.

- T. B. Othman, Y. Hamid, H. K. Alkahtani, I. Ullah, and H. Hamam, "A novel framework for classification of different Alzheimer's disease stages using CNN model," *Electronics*, vol. 12, p. 469, 2023. doi:10.3390/electronics12020469.
- 17) S. El-Sappagh, J. M. Alonso-Moral, T. Abuhmed et al., "Trustworthy artificial intelligence in Alzheimer's disease: State of the art, opportunities, and challenges," *Artificial Intelligence Review*, vol. 56, pp. 11149–11296, 2023. doi:10.1007/s10462-023-10415-5.
 - 18) G. Shankar, D. Sisodia, and P. Chandrakar, "An intelligent hierarchical residual attention learning-based conjoined twin neural network for Alzheimer's stage detection and prediction," *Computational Intelligence*, 2023. doi:10.1111/coin.12594.
 - 19) M. Zolnoori, A. Zolnour, and M. Topaz, "ADscreen: A speech processing-based screening system for automatic identification of patients with Alzheimer's disease and related dementia," *Artificial Intelligence in Medicine*, vol. 143, p. 102624, 2023. doi:10.1016/j.artmed.2023.102624.
 - 20) X. Zhang, L. Han, L. Han, H. Chen, D. Dancey, and D. Zhang, "sMRI-PatchNet: A Novel Efficient Explainable Patch-Based Deep Learning Network for Alzheimer's Disease Diagnosis With Structural MRI," *IEEE Access*, vol. 11, pp. 108603–108616, 2023. doi:10.1109/ACCESS.2023.3321220.
 - 21) N. Rahim, S. El-Sappagh, S. Ali, K. Muhammad, J. Del Ser, and T. Abuhmed, "Prediction of Alzheimer's progression based on multimodal deep-learning-based fusion and visual explainability of time-series data," *Information Fusion*, vol. 92, pp. 363–388, 2023. doi:10.1016/j.inffus.2022.11.028.
 - 22) A. Mora-Rubio, M. A. Bravo-Ortíz, S. Quiñones Arredondo, J. M. Saborit Torres, G. A. Ruz, and R. Tabares-Soto, "Classification of Alzheimer's disease stages from magnetic resonance images using deep learning," *PeerJ Computer Science*, vol. 9, p. e1490, 2023. doi:10.7717/peerj-cs.1490.
 - 23) Y. M. Roopa, B. B. Reddy, M. R. Babu et al., "Teaching learning-based brain storm optimization tuned Deep-CNN for Alzheimer's disease classification," *Multimedia Tools and Applications*, vol. 82, pp. 33333–33356, 2023. doi:10.1007/s11042-023-14815-1.
 - 24) I. Abunadi, "Deep and hybrid learning of MRI diagnosis for early detection of the progression stages in Alzheimer's disease," *Connection Science*, vol. 34, no. 1, pp. 2395–2430, 2022. doi:10.1080/09540091.2022.2123450.
 - 25) D. A. Arafa, H. E. D. Moustafa, A. M. T. Ali-Eldin et al., "Early detection of Alzheimer's disease based on the state-of-the-art deep learning approach: A comprehensive survey," *Multimedia Tools and Applications*, vol. 81, pp. 23735–23776, 2022. doi:10.1007/s11042-022-11925-0.
 - 26) D. Munteanu, C. Bejan, N. Munteanu, C. Zamfir, M. Vasić, S.-M. Petrea, and D. Cristea, "Deep-Learning-Based System for Assisting People with Alzheimer's Disease," *Electronics*, vol. 11, p. 3229, 2022. doi:10.3390/electronics11193229.
 - 27) S. Fouladi, A. A. Safaei, N. Mammone et al., "Efficient deep neural networks for classification of Alzheimer's disease and mild cognitive impairment from scalp EEG recordings," *Cognitive Computation*, vol. 14, pp. 1247–1268, 2022. doi:10.1007/s12559-022-10033-3.
 - 28) R. A. Hazarika, D. Kandar, and A. K. Maji, "An experimental analysis of different deep learning-based models for Alzheimer's disease classification using brain magnetic resonance images," *Journal of King Saud University – Computer and Information Sciences*, vol. 34, no. 10, pp. 8576–8598, 2022. doi:10.1016/j.jksuci.2021.09.003.
 - 29) T. Habuza, N. Zaki, E. A. Mohamed, and Y. Statsenko, "Deviation From Model of Normal Aging in Alzheimer's Disease: Application of Deep Learning to Structural MRI Data and Cognitive Tests," *IEEE Access*, vol. 10, pp. 53234–53249, 2022. doi:10.1109/ACCESS.2022.3174601.
 - 30) B. M. Cobbinah, C. Sorg, Q. Yang, A. Ternblom, C. Zheng, W. Han, L. Che, and J. Shao, "Reducing variations in multi-center Alzheimer's disease classification with convolutional adversarial autoencoder," *Medical Image Analysis*, vol. 82, p. 102585, 2022. doi:10.1016/j.media.2022.102585.
 - 31) M. H. Al-Adhaileh, "Diagnosis and classification of Alzheimer's disease using a convolutional neural network algorithm," *Soft Computing*, vol. 26, pp. 7751–7762, 2022. doi:10.1007/s00500-022-06762-0.
 - 32) A. Alvi, S. Siuly, H. Wang, K. Wang, and F. Whittaker, "A deep learning based framework for the diagnosis of mild cognitive impairment," *Knowledge-Based Systems*, vol. 248, p. 108815, 2022. doi:10.1016/j.knosys.2022.108815.
 - 33) Z. Fan, J. Li, L. Zhang et al., "U-net based analysis of MRI for Alzheimer's disease diagnosis," *Neural Computing and Applications*, vol. 33, pp. 13587–13599, 2021. doi:10.1007/s00521-021-05983-y.
 - 34) Y. AbdulAzeem, W. M. Bahgat, and M. Badawy, "A CNN-based framework for classification of Alzheimer's disease," *Neural Computing and Applications*, vol. 33, pp. 10415–10428, 2021. doi:10.1007/s00521-021-05799-w.
 - 35) T. Abuhmed, S. El-Sappagh, and J. M. Alonso, "Robust hybrid deep learning models for Alzheimer's progression detection," *Knowledge-Based Systems*, vol. 213, p. 106688, 2021. doi:10.1016/j.knosys.2020.106688.

- 36) R. E. Turkson, H. Qu, C. B. Mawuli et al., "Classification of Alzheimer's disease using a deep convolutional spiking neural network," *Neural Processing Letters*, vol. 53, pp. 2649–2663, 2021. doi:10.1007/s11063-021-10514-w.
- 37) X. Bi, S. Li, B. Xiao, Y. Li, G. Wang, and X. Ma, "Computer-aided Alzheimer's disease diagnosis by an unsupervised deep learning technology," *Neurocomputing*, vol. 392, pp. 296–304, 2020. doi:10.1016/j.neucom.2018.11.111.
- 38) J. Wen, E. Thibeau-Sutre, M. Diaz-Melo, J. Samper-González, A. Routier, S. Bottani, D. Dormont, S. Durrleman, N. Burgos, and O. Colliot, "Convolutional neural networks for classification of Alzheimer's disease: Overview and reproducible evaluation," *Medical Image Analysis*, vol. 63, p. 101694, 2020. doi:10.1016/j.media.2020.101694.
- 39) H. Alloui, M. Sadgal, and A. Elfazziki, "Utilization of a convolutional method for Alzheimer's disease diagnosis," *Machine Vision and Applications*, vol. 31, p. 25, 2020. doi:10.1007/s00138-020-01074-5.
- 40) S. Ahmed et al., "Ensembles of Patch-Based Classifiers for Diagnosis of Alzheimer's Diseases," *IEEE Access*, vol. 7, pp. 73373–73383, 2019. doi:10.1109/ACCESS.2019.2920011.
- 41) F. Zhang, Z. Li, B. Zhang, H. Du, B. Wang, and X. Zhang, "Multi-modal deep learning model for auxiliary diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 361, pp. 185–195, 2019. doi:10.1016/j.neucom.2019.04.093.
- 42) M. Raza, M. Awais, W. Ellahi, N. Aslam, H. X. Nguyen, and H. Le-Minh, "Diagnosis and monitoring of Alzheimer's patients using classical and deep learning techniques," *Expert Systems with Applications*, vol. 136, pp. 353–364, 2019. doi:10.1016/j.eswa.2019.06.038.
- 43) R. Jain, N. Jain, A. Aggarwal, and D. J. Hemant, "Convolutional neural network-based Alzheimer's disease classification from magnetic resonance brain images," *Cognitive Systems Research*, vol. 57, pp. 147–159, 2019. doi:10.1016/j.cogsys.2018.12.015.
- 44) C. Ge, Q. Qu, I. Y.-H. Gu, and A. S. Jakola, "Multi-stream multi-scale deep convolutional networks for Alzheimer's disease detection using MR images," *Neurocomputing*, vol. 350, pp. 60–69, 2019. doi:10.1016/j.neucom.2019.04.023.
- 45) N. Zeng, H. Qiu, Z. Wang, W. Liu, H. Zhang, and Y. Li, "A new switching-delayed-PSO-based optimized SVM algorithm for diagnosis of Alzheimer's disease," *Neurocomputing*, vol. 320, pp. 195–202, 2018. doi:10.1016/j.neucom.2018.09.001.
- 46) R. K. Shrivastava, S. P. Singh, and G. Kaur, "Machine learning models for Alzheimer's disease detection using OASIS data," in *Cognitive Technologies*, Springer Nature Singapore, 2023, pp. 111–126. doi:10.1007/978-981-99-2154-6_6.
- 47) N. B. Balakrishnan, S. P. S., and J. J. Panackal, "Alzheimer's disease diagnosis using machine learning: A review," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 71, no. 3, pp. 120–129, 2022. doi:10.14445/22315381/IJETT-V71I3P213.
- 48) T. J. Saleem, S. R. Zahra, F. Wu, A. Alwakeel, M. Alwakeel, F. Jeribi, and M. Hijji, "Deep learning-based diagnosis of Alzheimer's disease," *Journal of Personalized Medicine*, vol. 12, no. 5, p. 815, 2022. doi:10.3390/jpm12050815.
- 49) M. Zubair, A. Jaffar, S. Hussain, and S. Akram, "Alzheimer's disease classification using a hybrid deep learning approach with multi-layer U-net segmentation and XAI-driven analysis," *PLOS ONE*, vol. 20, no. 9, p. e0332572, 2025. doi:10.1371/journal.pone.0332572.
- 50) A. K. Mondol, S. M. H. S. Iqbal, P. K. Biswas, M. H. Islam, S. S. Arjon, and T. Yasmin, "Logistic regression-based machine learning model for early Alzheimer's detection using the OASIS dataset," in *Data Science, AI and Applications*, S. Palaiahnakote et al., Eds., *Communications in Computer and Information Science*, vol. 2683, Springer, Cham, 2025. doi:10.1007/978-3-032-11355-9_15.
- 51) I. Thayumanasamy and K. Ramamurthy, "Performance analysis of machine learning and deep learning models for classification of Alzheimer's disease from brain MRI," *Traitement du Signal*, vol. 39, no. 6, pp. 1961–1970, 2022. doi:10.18280/ts.390608.