

Hybrid Approach for Alzheimer's Disease Prediction Using MRI: Deep Feature Extraction and Machine Learning Models

Himanshu Pant^{1*}, Garima Joshi² and Bhupesh Rawat¹

¹School of Computing, Graphic Era Hill University, Bhimtal Campus, India.

²Department of Zoology, Kumaun University, Nainital, India.

*Corresponding Author Email: himpant7@gmail.com

<https://dx.doi.org/10.13005/bpj/3408>

(Received: 13 May 2025; accepted: 03 July 2025)

Since Alzheimer's disease (AD) is a neurodegenerative brain ailments characterized by a progressive loss of cognitive abilities, early identification is essential to providing appropriate treatment. Transfer learning on MRI scans is investigated in this study in order to organize AD into four stages: no impairment, very mild, mild, and moderate impairment. To extract significant visual features, pre-trained convolutional neural networks (VGG-16, VGG-19, InceptionNet, and ResNet50) were used. Four distinct machine learning methods were subsequently employed to classify these features: Random Forest, Gradient Boosting, Neural Networks, and Logistic Regression. Metrics like F1-score, AUC, recall, accuracy, and precision were used to assess the predictive models. Neural Networks consistently delivered superior results across all feature sets, with Gradient Boosting performing nearly as well. Logistic Regression and Random Forest also yielded reliable outcomes, though slightly less effective. Among all evaluated models and classifiers, VGG16 combined with Gradient Boosting achieves the highest classification accuracy of 95.7%, along with strong AUC (0.987) and F1-score (0.957). While ResNet50 also performs well, VGG16 with Gradient Boosting emerges as the best overall in terms of raw accuracy and consistency.

Keywords: Alzheimer's Disease Prediction; Feature Extraction; MRI Image Classification; Neurological Disorder; Transfer Learning.

Memory impairment, cognitive decline, and altered behavior are among the major indications for Alzheimer's disease (AD), a chronic, progressive neurological illness that predominantly affects older persons.¹ As the leading cause of dementia globally, AD places a substantial emotional and financial strain on patients, their caregivers, and healthcare infrastructures. The accumulation of tau protein tangles and amyloid-beta plaques in the brain, which disrupt typical neuronal activity and eventually cause cell death, is a molecular signature of the illness. Alzheimer's

typically unfolds in progressive stages; ranging from No Impairment (NI) to Very Mild (VMD), Mild, and finally Moderate Impairment (Mod)² highlighting the importance of early and precise diagnosis for effective disease management and improved patient care.

Conventional approaches to diagnosing AD rely heavily on cognitive assessments, clinical history, and neuroimaging techniques. While widely used, these methods are often subjective and inconsistent, particularly in the early stages when symptoms are subtle or overlap with other

neurological disorders. Artificial intelligence (AI), especially when used with neuroimaging techniques like Magnetic Resonance Imaging (MRI), has proven beneficial for improving diagnosis precision in recent years.³ MRI provides detailed insight into brain anatomy and functional alterations associated with neurodegenerative diseases, making it a powerful modality for detecting the structural changes characteristic of AD.

The rise of deep learning (DL) and machine learning (ML) has ushered in transformative changes across medical diagnostics. Using pre-trained Convolutional Neural Networks (CNNs), one such technique, transfer learning, has been extensively utilized to extract intricate data from medical visuals. These networks may be optimized for particular tasks after being first trained on enormous datasets such as ImageNet. This enables optimal feature extraction even in scenarios when there is a scarcity of labeled medical data.⁴ This approach not only speeds up the training process but also helps achieve higher accuracy and reduces the risk of overfitting—critical factors when working with sensitive health data.

This study proposes a reliable and efficient classification framework that integrates CNN-based feature extraction with traditional ML classifiers to categorize Alzheimer's stages using MRI scans. Popular pre-trained models—VGG-16, VGG-19, InceptionV3, and ResNet50—are employed to derive deep features from the image data.⁵ In order to determine each classifier's ability to discriminate between the four AD phases, the features retrieved are then processed using Random Forest (RF), Neural Network (NN), Logistic Regression (LR), and Gradient Boosting (GB).

CNN architectures such as VGG, ResNet, and InceptionV3 have demonstrated strong performance in various image classification tasks due to their ability to identify and encode hierarchical patterns in images.⁶ VGG models rely on small filters for detailed feature extraction, while ResNet's use of residual connections facilitates deeper model training without performance degradation. InceptionV3, on the other hand, utilizes parallel convolutional layers to capture multi-scale features, enhancing its representational power.

The classification task involves sorting MRI samples into four diagnostic categories—NI, VMD, Mild, and Mod—each representing a different severity level of Alzheimer's. Achieving accurate classification across these categories is crucial for timely therapeutic intervention tailored to the patient's condition.⁷

Prior research has shown promising outcomes when combining CNN-based features with classical machine learning algorithms. For instance, studies have applied similar approaches to detect diseases like lung cancer and diabetic retinopathy, leveraging CNNs to extract disease-specific patterns from medical images.⁸ In the context of Alzheimer's, CNNs have been successful in identifying key biomarkers such as hippocampal shrinkage and cortical thinning. Integrating CNN-derived features with classifiers like Random Forest or Gradient Boosting has been found to enhance model performance by uniting deep learning's pattern recognition strength with traditional ML's decision-making capabilities.⁹

Transfer learning proves particularly advantageous when data availability is constrained, a common issue in medical domains due to ethical and logistical limitations. By reapplying knowledge from broader datasets, pre-trained models can deliver reliable results even with fewer medical samples. This study utilizes this benefit by training models on a limited MRI dataset of Alzheimer's patients and extracting rich, high-level features for subsequent classification.

Each ML algorithm included in this study offers unique advantages. Neural Networks can model complex nonlinear interactions in high-dimensional datasets, while Random Forests excel in handling imbalanced data with minimal tuning. Gradient Boosting enhances prediction accuracy through iterative learning, and Logistic Regression, despite its simplicity, remains a staple in healthcare analytics due to its transparency and ease of interpretation.¹⁰

By combining deep learning and machine learning methodologies, the proposed approach presents a scalable and automated solution for Alzheimer's diagnosis. This dual-method strategy minimizes the need for manual feature engineering while ensuring adaptability across different clinical environments. The work aligns with ongoing

efforts to develop intelligent diagnostic support systems for neurological disorders, demonstrating how AI can assist clinicians in making more accurate and efficient decisions.

In conclusion, the research introduces a comprehensive methodology for classifying stages of Alzheimer's disease using MRI scans. Through the fusion of deep feature extraction and machine learning classification, the study identifies optimal model combinations that yield high diagnostic accuracy. This framework not only advances the development of AI-based tools for early detection but also addresses a pressing global health issue by facilitating timely medical intervention and improving patient care outcomes.

Related work

The tools of artificially intelligent (AI), machine learning, and deep learning have evolved considerably. Deep learning and artificial intelligence have grown increasingly important in the healthcare sector due to their constant advancement, particularly in the identification, prognosis, and tracking of neurodegenerative illnesses like Alzheimer's. To aid in the early and precise diagnosis of Alzheimer's, numerous studies have employed an assortment of datasets, including medical imaging data from MRI and PET scans. Researchers have explored a range of clinical markers associated with Alzheimer's disease, including cognitive decline, memory impairment, structural brain changes, and behavioral symptoms, to develop automated systems capable of identifying and classifying different stages of the disease. These AI-driven approaches have shown considerable promise in enhancing diagnostic precision and enabling timely interventions for individuals at risk of or suffering from Alzheimer's disease.

For the purpose of early Alzheimer's detection using MRI, Sharma *et al.*¹¹ recommended a CNN model combining a neural network classifier with feature extraction based on VGG16. Utilizing two datasets (6400 and 6330 pictures), it exhibited 90.4% and 71.1% accuracy in classifying AD stages, respectively. Despite outperforming prior methods, the model has trouble recognizing early neurological alterations and mild-stage detection.

The HEMRDTL model, which incorporates fused CT-MRI and EEG data for Alzheimer's detection, has been proposed by

Leela *et al.*¹² Implementing feature fusion and deep learning-based classification, it leverages VGG-19 for image feature extraction and M-RPCA for EEG data. Though the model boosts stage-wise accuracy, its high processing requirements and reliance on multimodal inputs make it less viable for clinical settings.

Mamun *et al.*¹³ utilized 6,219 MRI scans from multiple phases of dementia to provide a method of deep learning for Alzheimer's recognition. The customized CNN achieved the best results, with 97.6% accuracy, 97% recall, and 99.26% AUC, out of the assessment of four models: CNN, ResNet101, DenseNet121, and VGG16. Notwithstanding excellent performance, drawbacks include the use of just one modality and a small dataset.

El-Geneedy *et al.*¹⁴ leveraged 2D T1-weighted MRI images to generate a shallow CNN-based framework for the non-invasive diagnosis of Alzheimer's disease. It permits both detailed MCI classification (VMD, MD, MoD) and general classification (normal, MCI, AD). It obtained 99.68% accuracy, outperforming models such as ResNet50 and EfficientNetB7. Limitations, however, include a lack of external validation and a restricted variety of datasets.

Using MRI scans from the OASIS dataset, Hussain *et al.*¹⁵ demonstrated a 12-layer CNN for the binary classification of Alzheimer's disease. Having an accuracy of 97.75%, the custom model surpassed pre-trained models such as InceptionV3 and MobileNetV2. Its binary concentration and lack of assessment on external or multi-class datasets are shortcomings, despite its impressive findings.

A collaborative deep learning model for early Alzheimer's detection was designed by Fathi *et al.*¹⁶ utilizing MRI data from the ADNI dataset. The ensemble of the top six CNN classifiers, which underwent training on multiple CNNs, attained up to 99.83% accuracy for LMCI/AD and 93.92% accuracy for three-class tasks. Limited external validation and some performance variations are among the constraints, despite the impressive findings.

Lanjewar *et al.*¹⁷ used 6,400 MRI images from four classes to present a CNN-KNN hybrid model for Alzheimer's recognition. CNN was liable for feature extraction, while KNN was in

possession of classification. The model surpassed ResNet50 and other models with high results, such as 99.43% F1-score and 99.58% accuracy.

An extensive analysis of both traditional machine learning and deep learning techniques for Alzheimer's diagnosis using MRI is offered by Zhao *et al.*¹⁸ They look at input formats and feature extraction methods, as well as models like SVM, Random Forest, CNNs, autoencoders, and transformers. The paper emphasizes issues including data leaking and class imbalance, emphasizing the need of trustworthy datasets and standard procedures.

Using structural MRI, EHR, and SNP data, Nilanjana Pradhan *et al.*¹⁹ created deep learning models for Alzheimer's disease identification. Techniques for feature selection were used to improve specificity, sensitivity, and accuracy. Despite the models' good validation results, early diagnosis of mild cognitive impairment remained challenging, despite their efficient models on standard datasets.

With an emphasis on early detection using AI-based image processing, Suganyadevi Sellappan *et al.*²⁰ surveyed deep learning methods for MRI scan-based Alzheimer's disease diagnosis. Standardized methods are recommended and advanced designs are highlighted in the study. The study's drawbacks, despite its transparency, include the need for broader clinical validation and practical implementation of the procedures assessed.

Using MRI data, Alejandro Mora-Rubio *et al.*²¹ demonstrated a deep learning method for Alzheimer's stage identification. They used models involving EfficientNet, DenseNet, Siamese networks, and Vision Transformers together with data augmentation and preprocessing based on FreeSurfer. The model's performance dropped to 66–67% for early MCI, although reaching 89% accuracy for AD compared to control, suggesting difficulties in early-stage identification.

MATERIALS AND METHODS

This section outlines the data, preprocessing steps, and modeling techniques adopted to develop the proposed Alzheimer's Disease (AD) detection system. The main objective of the study is to improve the accuracy of AD stage

identification by applying an image classification approach. Using well-known convolutional neural network designs as VGG16, VGG19, ResNet50, and InceptionV3, this approach extracts profound characteristics.²² These pre-trained models provide attributes that form the basis for further research. Traditional machine learning classifiers, such as logistic regression, neural networks, gradient boosting, and random forest, receive these deep characteristics rather than relying entirely on end-to-end deep learning.

Dataset Descriptions

Obtaining high-quality and well-annotated data is a crucial yet challenging component in deep learning research, especially for neurodegenerative diseases like Alzheimer's. For this study, an open-source and publicly accessible dataset of brain MRI scans was utilized. The dataset is structured into four clearly defined categories that reflect the severity of cognitive decline: No Impairment, Very Mild Impairment, Mild Impairment, and Moderate Impairment.²³ The dataset includes a total of 10,250 grayscale MRI images used for model training, with 1,279 additional images reserved for testing.

Data Pre-Processing

The raw dataset included grayscale MRI scans of varying sizes, initially lacking consistent labeling and formatting. To prepare the data for model training, images were labeled according to the four diagnostic categories. Since the images varied in resolution, preprocessing steps were essential to bring uniformity. All images were resized to a fixed dimension of 255×255 pixels,²⁴ ensuring consistency and compatibility with the CNN input requirements.

This standardization step was performed using Python libraries, specifically with the Keras framework backed by TensorFlow. Once the images were resized and prepared, deep feature extraction was carried out using pre-trained models—VGG16, VGG19, ResNet50, and InceptionV3—as illustrated in Figure 1. This preprocessing pipeline improved the quality of input data and positively influenced the downstream classification process.

Projected Framework

The core objective of the proposed model is to classify MRI brain scans into four categories reflecting stages of Alzheimer's disease. The classification process is built upon transfer learning principles, utilizing pre-trained CNN architectures

as feature extractors. VGG16, VGG19, ResNet50, and InceptionV3, originally trained on the ImageNet dataset, are fine-tuned for the multi-class AD classification task.

For each CNN model, the convolutional base is retained to capitalize on its learned features, while the top layers are modified. These layers are adapted to the current classification objective by incorporating fully connected layers followed by dropout for regularization. The final output layer uses the softmax activation function to predict the appropriate cognitive stage. The block diagram for AD severity classification is shown in Figure-2.

For multi-class classification problems, the models are trained using the categorical cross-entropy loss function after having been built using the Adam optimizer. To minimize overfitting and improve generalization, techniques including early halting and model check pointing are used.

Model performance is measured using multiple evaluation metrics, including accuracy, precision, recall, F1-score, and confusion matrices. Additionally, ROC-AUC curves are plotted to assess the classification effectiveness across all four categories. This comprehensive modeling strategy facilitates reliable and early-stage detection of Alzheimer's disease, offering potential support in clinical decision-making and monitoring.²⁵

To assess the effectiveness of deep learning methods on the Alzheimer's MRI dataset, several convolutional neural networks—VGG16, VGG19, ResNet50, and InceptionV3—were utilized for both feature extraction and classification. These networks were trained to identify and categorize brain MRI scans into four cognitive categories: No Impairment, Very Mild Impairment, Mild Impairment, and Moderate Impairment. The approach is centered around a deep learning-based classification system, which involves preparing the input data through image preprocessing and segmentation, thereby enabling the networks to recognize and learn intricate brain structure patterns.²⁶

Each model utilizes 3x3 convolutional layers to extract spatial features, followed by max-pooling layers for reducing the dimensionality while retaining important details. Padding is applied to preserve image dimensions across layers, and ReLU activation introduces non-linearity to enhance learning. For validation, Keras' test iterator is employed, and training is done over several epochs with calculated steps per epoch—determined by dividing the total image count by batch size—to ensure consistent learning.

To improve model generalization, a variety of data augmentation techniques are applied

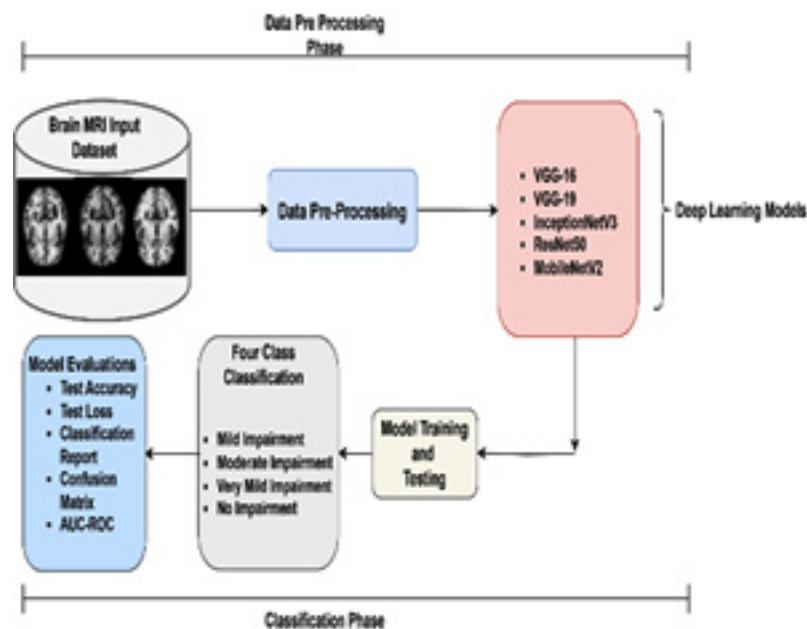


Fig. 1. Workflow of the Proposed Alzheimer's Disease Classification System

to the MRI scans, including cropping, padding, rotating, scaling, translating, and flipping. These augmentations create a richer dataset by simulating natural variations, reducing overfitting risks.

Model performance is analyzed using a confusion matrix, offering a clear view of predicted vs. actual labels. In addition, a classification report is generated to summarize evaluation metrics like accuracy, precision, recall, and F1-score.²⁷ These metrics provide insights into how effectively the models can distinguish between the four AD stages on both the test and validation datasets. The chosen CNN architectures, such as VGG16 and ResNet50, offer deep layers that are well-suited for extracting the nuanced features necessary for accurate diagnosis.

Feature extraction models

Extracting relevant features from MRI scans is vital for developing an accurate

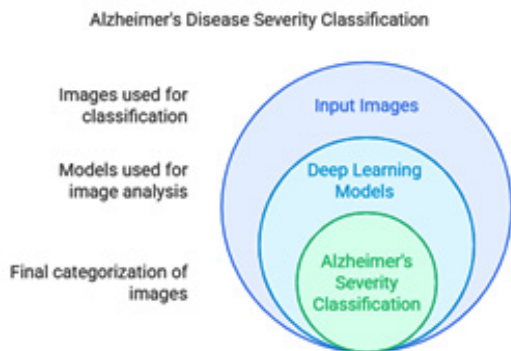


Fig. 2. Block Diagram for AD Severity classification

classification system for Alzheimer’s detection. This study employs high-performing, pre-trained CNN models—VGG16, VGG19, ResNet50, and InceptionV3—widely used in medical image analysis for their robustness in pattern recognition. These models, initially trained on large-scale datasets, are either fine-tuned or used as fixed feature extractors to detect features linked to the different stages of AD.²⁷

The extracted deep features are then passed to classical machine learning algorithms such as Logistic Regression, Random Forest, Gradient Boosting, and Multilayer Perceptron (MLP) for classification. This hybrid strategy combines the representational power of CNNs with the interpretability of traditional classifiers, offering a balanced approach that enhances the model’s diagnostic performance. This approach has proven effective in separating various cognitive states, which is crucial for timely and accurate Alzheimer’s diagnosis.

Deep learning approaches

Transfer learning is leveraged in this study to maximize the utility of pre-trained deep learning models. Instead of training models from scratch, pre-existing weights from large datasets are repurposed for new classification tasks, saving time and computational resources. This process involves two key approaches: feature extraction and fine-tuning.²⁸ In feature extraction, the final classification layers of the pre-trained model are removed, and the remaining layers are used to

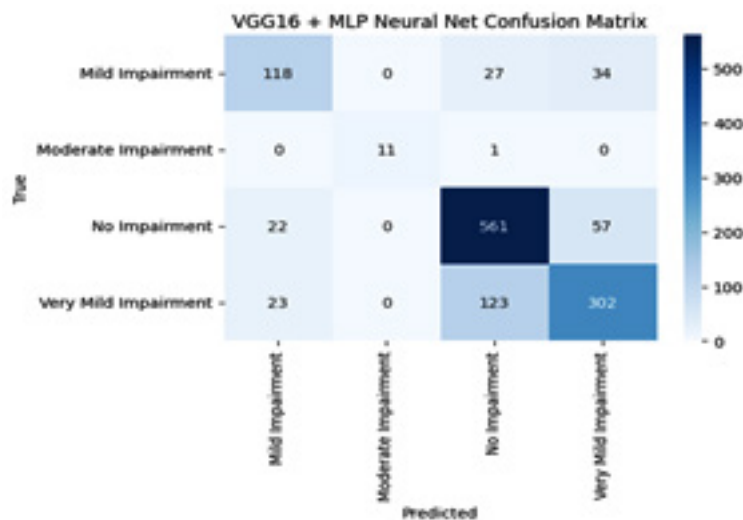


Fig. 3. VGG16 with MLP Neural Net Confusion Matrix

extract relevant features from the new dataset. These features are then classified using new custom classifiers. Fine-tuning, on the other hand, involves modifying and retraining the model's upper layers along with the classification head to adapt more specifically to the new dataset. This results in more tailored features for improved performance.

VGG 16

Introduced by the Visual Geometry Group at Oxford, VGG16 consists of 13 convolutional layers and 3 dense layers. It accepts 224×224×3 input images and uses 3x3 convolution filters throughout.²⁹ Max-pooling is used after certain convolutional blocks to reduce dimensionality, with padding to maintain spatial integrity. ReLU

activations help introduce non-linearity, improving the model's learning capacity.

In this study, VGG16's convolutional layers serve as feature extractors for brain MRIs. These features are passed to a classifier built with two dense layers (256 and 128 units) followed by a softmax output layer for multi-class prediction. This setup has shown strong performance in distinguishing between various AD stages. The confusion matrix for VGG16 with MLP Neural Net as shown in Figure-3.

VGG19

VGG19 is an extended version of VGG16, featuring 16 convolutional layers and 3 fully connected layers. Its deeper architecture enables

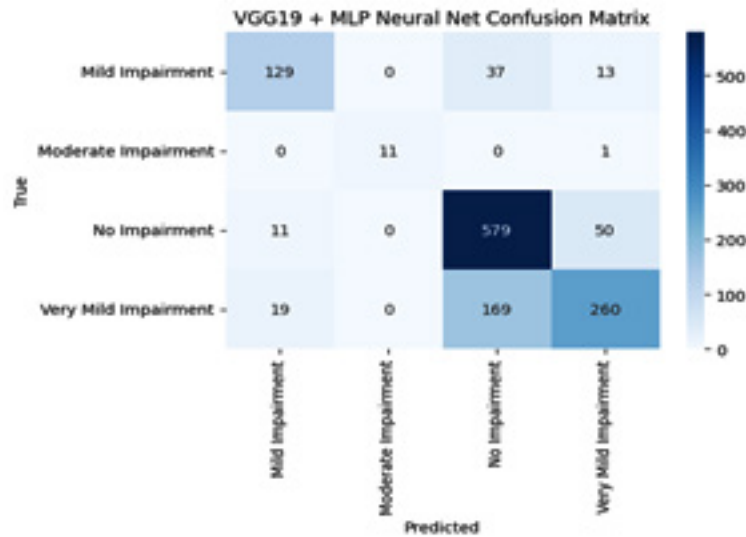


Fig. 4. VGG19 with MLP Neural Net Confusion Matrix

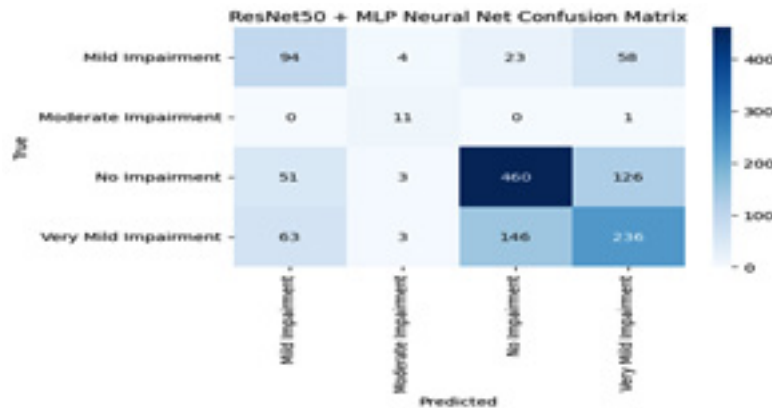


Fig. 5. Resnet50 with MLP Neural Net Confusion Matrix

it to learn more detailed patterns from medical images.³⁰ Like VGG16, it uses 3x3 convolutions, ReLU activations, and pooling layers. To mitigate overfitting, dropout is incorporated in the dense layers, and the model uses categorical cross-entropy loss for training on multiclass data. The confusion matrix for VGG19 with MLP Neural Net as shown in Figure-4.

ResNet50

ResNet50 is a 50-layer deep neural network that utilizes residual connections to address the vanishing gradient issue common in deep models. These skip connections enable better gradient flow, making training more stable.³¹

In the context of AD detection, ResNet50 efficiently extracts robust features from brain scans. These features are then input into classical classifiers, aiding in the identification of different

stages of cognitive decline. Its strong generalization ability and high accuracy make it well-suited for medical diagnostic tasks. The confusion matrix for ResNet50 with MLP Neural Net as shown in Figure-5.

InceptionV3

InceptionV3 is a highly optimized CNN architecture that utilizes asymmetric and factorized convolutions (e.g., 1x7 followed by 7x1) to minimize computational overhead. It processes multiple filter sizes in parallel, enabling it to capture features at multiple scales—an ideal trait for analyzing brain structures in MRI scans.³²

This architecture extracts a wide range of abstract features that are then classified using traditional machine learning methods. InceptionV3 is known for its rapid training convergence and high accuracy, even when trained on relatively

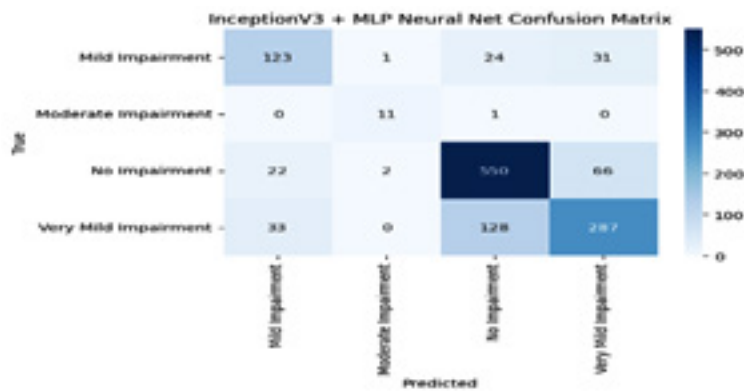


Fig. 6. InceptionV3 with MLP Neural Net Confusion Matrix

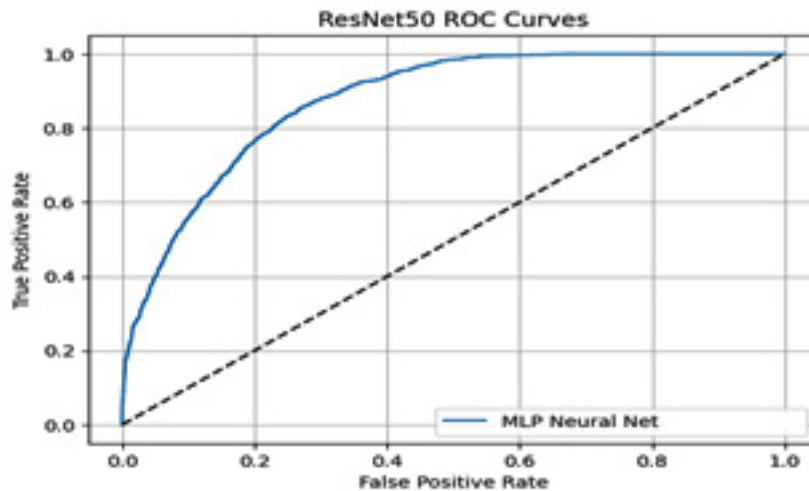


Fig. 7. ROC Curve for AD Predicted for VGG16 Model

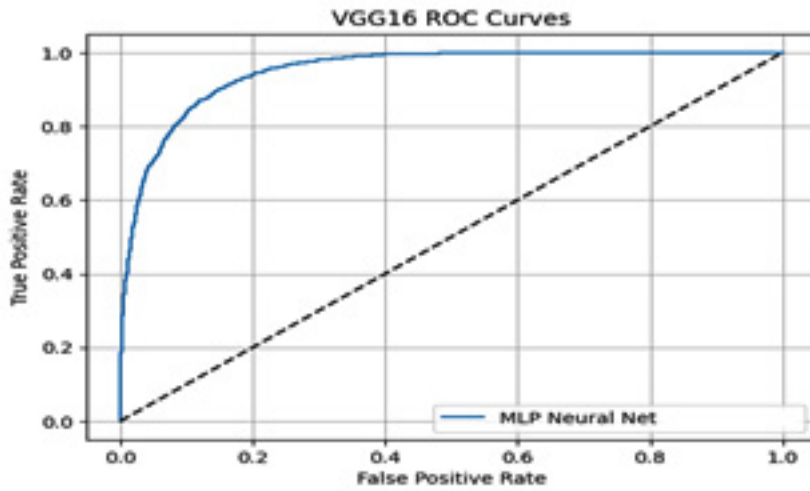


Fig. 8. ROC Curve for AD Predicted for VGG19

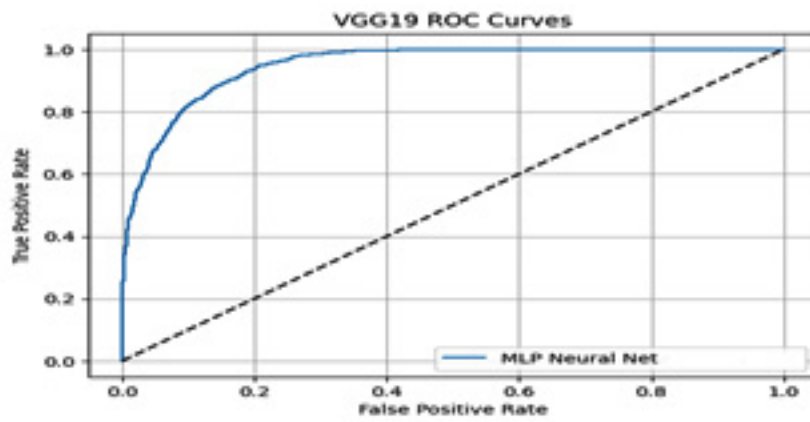


Fig. 9. ROC Curve for AD Predicted for Resnet50

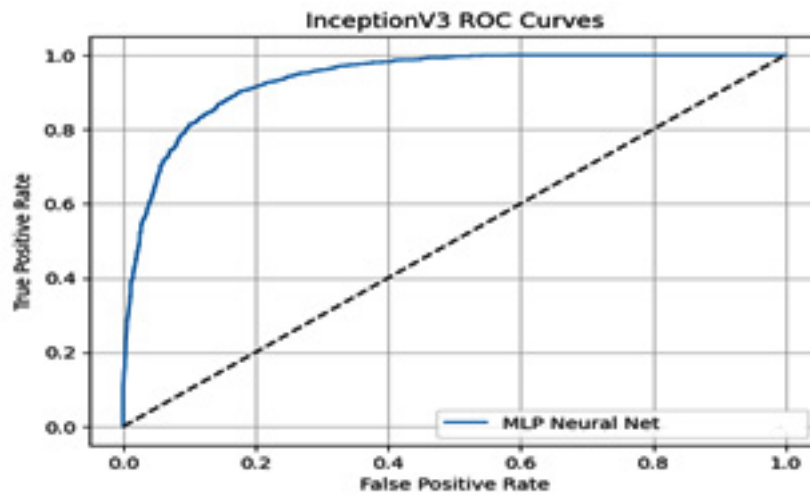


Fig. 10. ROC for AD Predicted for InceptionV3

Table 1. Comparison of different models in terms of performance

Feature Extraction Models	Classifiers	AUC	CA	F1	Precession	Recall
VGG-16	Gradient Boosting	0.987	0.957	0.957	0.957	0.957
	Logistic Regression	0.980	0.921	0.921	0.921	0.921
	Neural Network	0.986	0.953	0.973	0.973	0.973
	Random Forest	0.984	0.923	0.923	0.923	0.923
VGG-19	Gradient Boosting	0.950	0.863	0.862	0.865	0.863
	Logistic Regression	0.952	0.863	0.862	0.864	0.863
	Neural Network	0.952	0.856	0.855	0.857	0.856
	Random Forest	0.932	0.852	0.851	0.856	0.852
ResNet50	Gradient Boosting	0.985	0.953	0.953	0.953	0.953
	Logistic Regression	0.971	0.951	0.951	0.951	0.951
	Neural Network	0.986	0.943	0.943	0.943	0.943
	Random Forest	0.966	0.933	0.933	0.933	0.933
InceptionNet	Gradient Boosting	0.910	0.640	0.660	0.710	0.660
	Logistic Regression	0.920	0.710	0.710	0.710	0.710
	Neural Network (MLP)	0.940	0.760	0.760	0.760	0.760
	Random Forest	0.930	0.730	0.730	0.730	0.730

small medical datasets. The confusion matrix for InceptionV3 with MLP Neural Net as shown in Figure-6.

RESULTS

This section presents the experimental outcomes derived from the proposed framework for Alzheimer's Disease (AD) prediction. The framework integrates deep learning-based feature extraction using pre-trained convolutional models with classical machine learning classifiers. Specifically, we evaluated the performance of VGG-16, VGG-19, ResNet50, and InceptionV3 models in combination with four classifiers—Logistic Regression (LR), Multilayer Perceptron (MLP), Gradient Boosting (GB), and Random Forest (RF). The models were assessed using several performance metrics: Area Under the ROC Curve (AUC), Classification Accuracy (CA), F1-score, Precision, and Recall.³³ ROC Curve for AD Predicted for multiple models is shown in Figure-7 to Figure-10.

The ROC (Receiver Operating Characteristic) curves offer a visual representation of the performance of an MLP (Multi-Layer Perceptron) neural network classifier utilizing deep features extracted from four pre-trained

convolutional neural networks: VGG16, VGG19, InceptionNet, and ResNet50. These curves plot the True Positive Rate (sensitivity) against the False Positive Rate at various classification thresholds, with the Area Under the Curve (AUC) serving as a key metric to evaluate the classifier's ability to distinguish between classes.³⁴

Based on the results, VGG16 and ResNet50 emerged as the top-performing models when paired with the MLP classifier, each achieving an AUC of 0.986. This indicates a very high discriminative ability and strong classification power using features extracted from these architectures. VGG19 demonstrated moderate performance with an AUC of 0.952, suggesting it is capable of effective classification but slightly less discriminative than VGG16 and ResNet50. InceptionNet, while still maintaining acceptable performance, recorded a lower AUC of 0.940, making it the weakest among the four models in terms of class separability.³⁵

Overall, the findings highlight that VGG16 and ResNet50 provide the most robust and reliable feature representations for Alzheimer's disease classification using MRI images when combined with an MLP classifier. VGG19 also performs well, whereas InceptionNet, although still effective, offers comparatively less discriminative

power. These insights are vital for selecting the most appropriate base model for transfer learning in medical imaging applications.

DISCUSSION

Based on the evaluation metrics for different deep learning feature extraction models (VGG16, VGG19, ResNet50, and InceptionNet) combined with various classifiers, the following detailed performance analysis can be made as shown in Table 1.

The VGG16 model continues to demonstrate strong performance across all classifiers. When combined with a Neural Network, it achieves the highest overall performance with an F1-score, precision, and recall of 0.973, an AUC of 0.986, and a classification accuracy (CA) of 95.3%. Gradient Boosting is also very effective with AUC = 0.987 and CA = 95.7%, indicating excellent discriminative ability.³⁶ Logistic Regression and Random Forest provide slightly lower but still strong performance, with AUCs of 0.980 and 0.984, and classification accuracies of 92.1% and 92.3%, respectively.

VGG19 shows moderate performance across all classifiers. Gradient Boosting and Logistic Regression yield identical CAs of 86.3% and very close AUCs of 0.950 and 0.952, respectively. The Neural Network gives a slightly lower CA of 85.6%, while Random Forest performs the lowest among the four with an AUC of 0.932 and CA of 85.2%. Overall, while VGG19 performs decently, it is consistently behind VGG16.

The ResNet50 model maintains its reliability, particularly with Gradient Boosting (AUC = 0.985, CA = 95.3%) and Neural Network (AUC = 0.986, CA = 94.3%). Logistic Regression and Random Forest also perform competitively, achieving CAs of 95.1% and 93.3%, respectively. Across all classifiers, ResNet50 demonstrates high and consistent precision, recall, and F1-scores, establishing it as one of the top-performing feature extractors in this evaluation.

In contrast, InceptionNet shows the weakest performance. Even with its best configuration using a Neural Network (MLP), it achieves a comparatively lower AUC of 0.940 and CA of 76.0%. Gradient Boosting has the

poorest result overall with CA = 64.0% and AUC = 0.910, despite relatively higher recall (0.660–0.710).³⁷ Logistic Regression and Random Forest fare slightly better with CAs of 71.0% and 73.0%, respectively. These metrics suggest that InceptionNet is not as effective in extracting discriminative features for this task as the other models.

CONCLUSION

This study demonstrates the effectiveness of a hybrid deep learning and machine learning approach for early detection of Alzheimer's Disease (AD) using MRI scans. By leveraging transfer learning with pre-trained CNNs—specifically VGG16, VGG19, ResNet50, and InceptionNet—deep features were efficiently extracted and classified using traditional classifiers including Logistic Regression, Neural Network (MLP), Gradient Boosting, and Random Forest. The experimental analysis reveals that VGG16 and ResNet50, when paired with Neural Network classifiers, consistently outperform other combinations, achieving the highest values in AUC, accuracy, precision, recall, and F1-score. These findings suggest that deep feature representations from VGG16 and ResNet50 are particularly well-suited for capturing complex neuroimaging patterns associated with AD.

On the other hand, InceptionNet and VGG19, especially when used with Random Forest, exhibited relatively weaker performance, indicating their limited ability to generalize well to this dataset. Nevertheless, the overall low loss and high precision metrics across most models underscore the reliability of the proposed framework.

In the conclusion, VGG16 paired with Gradient Boosting outperforms all other model-classifier combinations, achieving the highest classification accuracy of 95.7%, along with an impressive AUC of 0.987 and F1-score of 0.957. Although ResNet50 also delivers competitive results, VGG16 with Gradient Boosting stands out for its superior accuracy and reliability. This approach holds significant potential for supporting clinical decision-making and enhancing early intervention strategies in Alzheimer's care.

ACKNOWLEDGEMENT

The author would like to thank Graphic Era Hill University for providing the necessary resources, facilities, and a conducive environment for completing the research work.

Funding source

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Conflict of interest

The author(s) do not have any conflict of interest.

Data availability

This statement does not apply to this article.

Ethics statement

This research did not involve human participants, animal subjects, or any material that requires ethical approval.

Informed consent statement

This study did not involve human participants, and therefore, informed consent was not required.

Clinical Trial Registration

This research does not involve any clinical trials.

Permission to reproduce material from other sources

Not Applicable.

Authors' Contribution

Himanshu Pant: Conceived and designed the research study, developed the methodology, Writing the original Draft; Garima Joshi: Data Collection, conducted data preprocessing, assisted with visualization; Data Preprocessing, Results interpretation; Data Analysis, Domain Expert; Bhupesh Rawat: Worked on deep learning and machine learning models and assisted in fine-tuning the algorithms.

REFERENCES

- Singh A, Kumar R. Brain MRI Image Analysis for Alzheimer's Disease (AD) Prediction Using Deep Learning Approaches. *SN Comput Sci.* 2024;5(1):160. doi:10.1007/s42979-023-02461-1
- Hazarika RA, Kandar D, Maji AK. A novel machine learning based technique for classification of early-stage Alzheimer's disease using brain images. *Multimed Tools Appl.* 2023;83(8):24277-24299. doi:10.1007/s11042-023-16379-6
- Vanaja T, Shanmugavadivel K, Subramanian M, Kanimozhiselvi CS. Advancing Alzheimer's detection: integrative approaches in MRI analysis with traditional and deep learning models. *Neural Comput Appl.* 2025;37(14):8527-8546. doi:10.1007/s00521-025-10993-1
- Kumar PN, Maguluri LP. Deep Learning-based Classification of MRI Images for Early Detection and Staging of Alzheimer's Disease. *International Journal of Advanced Computer Science and Applications.* 2024;15(5). doi:10.14569/IJACSA.2024.0150545
- Alshammari M, Mezher M. A Modified Convolutional Neural Networks For MRI-based Images For Detection and Stage Classification Of Alzheimer Disease. In: *2021 National Computing Colleges Conference (NCCC).* IEEE; 2021:1-7. doi:10.1109/NCCC49330.2021.9428810
- Lokesh K, Challa NP, Satwik AS, Kiran JC, Kumar Rao N, Naseeba B. Early Alzheimer's Disease Detection Using Deep Learning. *EAI Endorsed Trans Pervasive Health Technol.* 2023;9. doi:10.4108/eetpht.9.3966
- Sisodia PS, Ameta GK, Kumar Y, Chaplot N. A Review of Deep Transfer Learning Approaches for Class-Wise Prediction of Alzheimer's Disease Using MRI Images. *Archives of Computational Methods in Engineering.* 2023;30(4):2409-2429. doi:10.1007/s11831-022-09870-0
- Murugan S, Venkatesan C, Sumithra MG, et al. DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia From MR Images. *IEEE Access.* 2021;9:90319-90329. doi:10.1109/ACCESS.2021.3090474
- Yao Z, Wang H, Yan W, et al. Artificial intelligence-based diagnosis of Alzheimer's disease with brain MRI images. *Eur J Radiol.* 2023;165:110934. doi:10.1016/j.ejrad.2023.110934
- Vanaja T, Shanmugavadivel K, Subramanian M, Kanimozhiselvi CS. Advancing Alzheimer's detection: integrative approaches in MRI analysis with traditional and deep learning models. *Neural Comput Appl.* 2025;37(14):8527-8546. doi:10.1007/s00521-025-10993-1
- Sharma S, Guleria K, Tiwari S, Kumar S. A deep learning based convolutional neural network model with VGG16 feature extractor for the detection of Alzheimer Disease using MRI scans. *Measurement: Sensors.* 2022;24:100506. doi:10.1016/j.measen.2022.100506
- Leela M, Helenprabha K, Sharmila L. Prediction and classification of Alzheimer Disease categories using Integrated Deep

- Transfer Learning Approach. *Measurement: Sensors*. 2023;27:100749. doi:10.1016/j.measen.2023.100749
13. Mamun M, Bin Shawkat S, Ahammed MS, Uddin MM, Mahmud MI, Islam AM. Deep Learning Based Model for Alzheimer's Disease Detection Using Brain MRI Images. In: *2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*. IEEE; 2022:0510-0516. doi:10.1109/UEMCON54665.2022.9965730
 14. EL-Geneedy M, Moustafa HED, Khalifa F, Khater H, Abdelhalim E. An MRI-based deep learning approach for accurate detection of Alzheimer's disease. *Alexandria Engineering Journal*. 2023;63:211-221. doi:10.1016/j.aej.2022.07.062
 15. Hussain E, Hasan M, Hassan SZ, Hassan Azmi T, Rahman MA, Zavid Parvez M. Deep Learning Based Binary Classification for Alzheimer's Disease Detection using Brain MRI Images. In: *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*. IEEE; 2020:1115-1120. doi:10.1109/ICIEA48937.2020.9248213
 16. Fathi S, Ahmadi A, Dehnad A, Almasi-Dooghaee M, Sadegh M. A Deep Learning-Based Ensemble Method for Early Diagnosis of Alzheimer's Disease using MRI Images. *Neuroinformatics*. 2023;22(1):89-105. doi:10.1007/s12021-023-09646-2
 17. Lanjewar MG, Parab JS, Shaikh AY. Development of framework by combining CNN with KNN to detect Alzheimer's disease using MRI images. *Multimed Tools Appl*. 2023;82(8):12699-12717. doi:10.1007/S11042-022-13935-4/METRICS
 18. Zhao Z, Chuah JH, Lai KW, et al. Conventional machine learning and deep learning in Alzheimer's disease diagnosis using neuroimaging: A review. *Front Comput Neurosci*. 2023;17. doi:10.3389/fincom.2023.1038636
 19. Pradhan N, Sagar S, Singh AS. Analysis of MRI image data for Alzheimer disease detection using deep learning techniques. *Multimed Tools Appl*. 2024;83(6):17729-17752. doi:10.1007/S11042-023-16256-2/METRICS
 20. Suganyadevi S, Pershiya AS, Balasamy K, Seethalakshmi V, Bala S, Arora K. Deep Learning Based Alzheimer Disease Diagnosis: A Comprehensive Review. *SN Comput Sci*. 2024;5(4):391. doi:10.1007/s42979-024-02743-2
 21. Mora-Rubio A, Bravo-Ortíz MA, Quiñones Arredondo S, Saborit Torres JM, Ruz GA, Tabares-Soto R. Classification of Alzheimer's disease stages from magnetic resonance images using deep learning. *PeerJ Comput Sci*. 2023;9:e1490. doi:10.7717/peerj-cs.1490
 22. Bhimavarapu U. A hybrid approach for Alzheimer's disease diagnosis: image segmentation and deep learning classification. *Neural Comput Appl*. Published online May 5, 2025. doi:10.1007/s00521-025-11249-8
 23. Lim BY, Lai KW, Haiskin K, et al. Deep Learning Model for Prediction of Progressive Mild Cognitive Impairment to Alzheimer's Disease Using Structural MRI. *Front Aging Neurosci*. 2022;14. doi:10.3389/fnagi.2022.876202
 24. Al-Rawashdeh HS, Usman A, Dutta AK, Sait ARW. Hybrid Feature Extraction Technique-based Alzheimer's Disease Detection Model Using MRI Images. *Journal of Disability Research*. 2024;3(6). doi:10.57197/JDR-2024-0073
 25. Pant H, Lohani MC, Pant J, Petshali P. GUI-Based Alzheimer's Disease Screening System Using Deep Convolutional Neural Network. In: ; 2021:259-272. doi:10.1007/978-981-33-6862-0_22
 26. Salahudin A, Qadeer E, Ahmed A, Fahad LG. Enhanced Classification of Alzheimer's Disease Using Deep Learning and Machine Learning Models with MRI Scans. In: *2024 International Conference on Frontiers of Information Technology (FIT)*. IEEE; 2024:1-6. doi:10.1109/FIT63703.2024.10838463
 27. Jenber Belay A, Walle YM, Haile MB. Deep Ensemble learning and quantum machine learning approach for Alzheimer's disease detection. *Sci Rep*. 2024;14(1):14196. doi:10.1038/s41598-024-61452-1
 28. Ishfaq QUA, Bibi R, Ali A, et al. Automatic smart brain tumor classification and prediction system using deep learning. *Sci Rep*. 2025;15(1):14876. doi:10.1038/s41598-025-95803-3
 29. L S, V S, Ravi V, E.A G, K.P S. Deep learning-based approach for multi-stage diagnosis of Alzheimer's disease. *Multimed Tools Appl*. 2023;83(6):16799-16822. doi:10.1007/s11042-023-16026-0
 30. Elazab A, Wang C, Abdelaziz M, et al. Alzheimer's disease diagnosis from single and multimodal data using machine and deep learning models: Achievements and future directions. *Expert Syst Appl*. 2024;255:124780. doi:10.1016/j.eswa.2024.124780
 31. Heckkel W, Helali A. Unveiling Alzheimer's Disease Early: A Comprehensive Review of Machine Learning and Imaging Techniques. *Archives of Computational Methods in Engineering*. 2025;32(1):471-484. doi:10.1007/s11831-024-10179-3

32. Bhandarkar A, Naik P, Vakkund K, Junjappanavar S, Bakare S, Pattar S. Deep learning based computer aided diagnosis of Alzheimer's disease: a snapshot of last 5 years, gaps, and future directions. *Artif Intell Rev.* 2024;57(2):30. doi:10.1007/s10462-023-10644-8
33. Sathyabhama B, Kannan M. An effective deep learning-based automatic prediction and classification of Alzheimer's disease using EGELU-SZN technique. *Neural Comput Appl.* 2025;37(9):6915-6932. doi:10.1007/s00521-025-10994-0
34. Heenaye-Mamode Khan M, Reesaul P, Auzine MM, Taylor A. Detection of Alzheimer's disease using pre-trained deep learning models through transfer learning: a review. *Artif Intell Rev.* 2024;57(10):275. doi:10.1007/s10462-024-10914-z
35. A N S, Kattennanavar G, K S A, S A A, H M A. Deep Learning-based Alzheimer's Detection: A Novel Approach. In: *2025 International Conference on Computing for Sustainability and Intelligent Future (COMP-SIF)*. IEEE; 2025:1-7. doi:10.1109/COMP-SIF65618.2025.10969870
36. Shinde S, Badwaik D, Mane S. Comparative Analysis of Machine Learning and Deep Learning Approaches for Multiclass Identification and Classification of Alzheimer's Disease. In: *2024 International Conference on Emerging Smart Computing and Informatics (ESCI)*. IEEE; 2024:1-6. doi:10.1109/ESCI59607.2024.10497347
37. Alsubaie MG, Luo S, Shaukat K. Alzheimer's Disease Detection Using Deep Learning on Neuroimaging: A Systematic Review. *Mach Learn Knowl Extr.* 2024;6(1):464-505. doi:10.3390/make6010024