

Research Article

A Vertical Stacking-Based Ensemble Deep Learning Model for Early Diagnosis of Alzheimer's Disease Using Multimodal MRI Scans

Parvatham Niranjana Kumar and Lakshmana Phaneendra Maguluri

Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation Green Fields, Vaddeswaram, Guntur, Andhra Pradesh, India

Article history

Received: 14-01-2025

Revised: 09-04-2025

Accepted: 10-04-2025

Corresponding Author:

Lakshmana Phaneendra Maguluri
Department of Computer Science
and Engineering, Koneru
Lakshmaiah Education Foundation
Green Fields, Vaddeswaram,
Guntur, Andhra Pradesh, India
Email: niranjan.1216@gmail.com

Abstract: Early detection of Alzheimer's Disease (AD) is crucial for timely interventions that can slow disease progression, enhance quality of life, and assist with future planning. Convolutional Neural Networks (CNNs) are an efficient method for processing image-based data. In this work, we used CNN-based deep learning models to extract structural information from structural MRI (sMRI) and brain neuron connectivity patterns from functional MRI (fMRI) data. In this study, a stacking-based ensemble multimodal framework was proposed by integrating both texture features and brain neuron connectivity patterns using Deep Learning (DL) models such as GoogLeNet, DenseNet-121, GNN, and U-Net. The prediction probabilities were combined using a vertical stacking approach to create a meta-feature matrix, which was utilized by the Meta model and trained using the Random Forest classification algorithm to generate the final predictions. This approach leveraged the complementary strengths of structural and functional data, thereby improving classification accuracy and generalization. The proposed method demonstrated remarkable accuracy of 95.18%, reflecting its exceptional performance and minimal error rates. It surpassed the effectiveness of existing state-of-the-art methods, showing high precision in early AD detection and highlighting its potential for neurodegenerative disease research.

Keywords: Alzheimer's Disease, GoogLeNet, DenseNet-121, Graph Neural Network, U-Net, Meta Model

Introduction

The leading cause of AD related dementia is progressive degeneration of brain neuron, predominantly affecting older adults. It shows up abnormal behavioral changes, memory loss and decreased reasoning, all of which have a significant influence on a person's day-to-day activities. The disease disrupts communication between brain cells, resulting in their degeneration and eventual death. AD advances through stages, starting with mild cognitive impairment and eventually leading to severe disability, where individuals struggle with basic activities (Alzheimers and Dementia, 2023).

According to the 2023 report on Alzheimer's disease, the condition significantly impacts public health. More than 6 million Americans aged 65 and older are currently living with Alzheimer's, a figure expected to increase as the population continues to age (Xiao *et al.*, 2023). The report also underscores that Alzheimer's ranks as the sixth-leading cause of death in the United States,

highlighting the significant societal impact of the disease (Ciurea *et al.*, 2023).

Although a definitive cure for AD has not been found, early detection is crucial for managing its progression. Early diagnosis enables interventions with existing treatments that may alleviate symptoms and improve cognitive function (Patra *et al.*, 2023). Furthermore, early-stage diagnosis enables individuals to participate in clinical trials, adopt novel therapeutic approaches and make lifestyle changes to support cognitive and emotional well-being (Gunawan *et al.*, 2024). For families and caregivers, early diagnosis provides the opportunity to prepare for the long-term implications of the disease, thereby improving family dynamics and reducing caregiver stress. Imaging modalities like structural MRI and functional MRI play a pivotal role in diagnosing Alzheimer's disease with enhanced accuracy. Structural MRI provides high-resolution images of brain anatomy, identifying physical changes caused by Alzheimer's, such as cortical thinning

and hippocampal atrophy. Meanwhile, fMRI evaluates how structural changes affect brain activity and neuron connectivity over time, providing understanding of the functional consequences of neurodegeneration (Wang *et al.*, 2021). By combining sMRI and fMRI data, clinicians can create comprehensive profiles that account for both structural and functional brain changes. Research shows that this dual approach improves predictions of cognitive decline and facilitates timely interventions (Abrol *et al.*, 2019).

Deep learning technologies have revolutionized the analysis of sMRI and fMRI data for Alzheimer's diagnosis. Convolutional Neural Networks (CNNs) like GoogLeNet and DenseNet have demonstrated remarkable effectiveness in sMRI analysis. GoogLeNet, with its Inception modules, processes multiple filter sizes simultaneously, capturing a diverse range of features, such as subtle anatomical changes associated with AD. Its efficient design supports deep networks while minimizing computational costs, making it highly effective for image classification tasks (Talha *et al.*, 2024). DenseNet, characterized by its densely connected layers, promotes feature reuse and mitigates the vanishing gradient problem. This architecture has shown exceptional accuracy in distinguishing between healthy patients, those with mild cognitive impairment and Alzheimer's patients (Zia-Ur-Rehman *et al.*, 2024).

Functional MRI data, which captures complex brain interactions, is particularly well-suited for analysis using Graph Neural Networks (GNN). Graph Neural Networks (GNNs) represent the brain like a graph, with nodes denoting brain areas and edges indicating functional connections. This structure allows the network to learn spatial and temporal dynamics, revealing patterns of disrupted neural connectivity typical in Alzheimer's patients (Zhang *et al.*, 2023). U-Net, another deep learning model, excels in fMRI-based segmentation tasks. Its encoder-decoder structure with skip connections preserves spatial information while capturing abstract features, making it ideal for identifying abnormal activity in specific brain regions (Bhosale *et al.*, 2023).

The adoption of hybrid models combining deep learning techniques (Saxena *et al.*, 2023) represents a significant advancement in Alzheimer's classification. By integrating models such as GoogLeNet, DenseNet, GNN and U-Net, Hybrid systems utilize the distinct advantages of each methodology to enhance diagnostic efficacy. For instance, while one model might excel at detecting structural changes, another might be more adept at analyzing functional abnormalities. This ensemble approach reduces model-specific biases, enhances classification accuracy and ensures more reliable predictions in clinical practice. The use of hybrid deep learning models is particularly advantageous in addressing the complexity of Alzheimer's pathology.

These models facilitate a multi-faceted analysis by combining insights from structural and functional imaging. This comprehensive understanding of the condition allows healthcare providers to make accurate decisions, perhaps resulting in improved patient outcomes and more successful management of Alzheimer's disease. The ensemble method emphasizes the significance of leveraging diverse computational tools to advance diagnostic precision and improve clinical interventions.

The rest of the paper is organized to cover the review of related literature, followed by a detailed explanation of the materials and methods used in the study. This is succeeded by a presentation of the results along with a comprehensive discussion. The paper concludes with a summary of key findings and final remarks.

Literature Survey

Jiao *et al.* (2024) introduced a Framework (MFASN) to integrate structural (sMRI), functional (fMRI) imaging data and genetic information for early AD diagnosis and biomarker discovery. The method employs a deep auto-encoder for non-linear feature extraction and a sparse self-representation module for subspace clustering, capturing meaningful data relationships. Experimental results on ADNI datasets demonstrate the framework's effectiveness in identifying AD-related biomarkers and SNP associations. Despite its success, challenges such as high-dimensional data, limited sample size and complex non-linear relationships persist.

In the study by Yu *et al.* (2024) GSCANet, a novel method designed for early AD diagnosis by leveraging multimodal MRI data was proposed. This includes Haralick features, functional brain neuron connectivity patterns and neuropsychological scores. In order to capture long-term contextual interactions between spatial and channel information, GSCANet incorporates a coordinate attention mechanism in addition to a group self-calibrated module to improve spatial characteristics. The model demonstrated significant classification performance: 78.70% accuracy for four-class classification (AD, early MCI, late MCI and normal controls), 83.33% for three-class (AD, MCI and NC) and over 92% for binary classifications. These results underscore its effectiveness in differentiating AD stages. Key contributions include the integration of multimodal data for comprehensive analysis, the development of an innovative architecture for feature extraction and superior classification across AD stages. While the study does not explicitly address its limitations, potential challenges include training complexity and the need for larger datasets to ensure generalization.

Dolci *et al.* (2024) investigated the use of multimodal MRI to identify amyloid- β (A β) positivity, a biomarker of AD, in unbalanced cohorts across the spectrum.

Combining sMRI, fMRI and diffusion MRI allowed for the capture of complimentary structural and connectivity alterations associated with A β deposition. This multimodal approach achieved a classification accuracy of 76.2%, with key regions identified as critical for A β detection through explainability analysis. The research contributes by demonstrating the effectiveness of non-invasive MRI techniques in detecting A β positivity, identifying brain regions strongly associated with amyloid accumulation and providing an alternative to PET scans and cerebrospinal fluid tests for AD diagnosis. However, the study notes challenges, including potential biases from the unbalanced cohort and difficulties integrating high-dimensional multimodal data.

Gupta *et al.* (2020) suggested a multimodal neuroimaging approach to group Alzheimer's patients according to the disease's prodrome. Combining structural MRI, rs-fMRI, AV45-PET, FDG-PET, DTI and APOE genotype data, it employed advanced feature extraction techniques like brain parcellation, voxel-wise and graphical analysis. Classification was performed using kernel-based methods and the EasyMKL classifier, with a one-at-a-time exclusion cross-validation strategy ensuring robust performance evaluation. Results demonstrated improved classification accuracy, sensitivity and specificity with multimodal approaches, particularly highlighting the effectiveness of FDG-PET, AV45-PET and rs-fMRI. Limitations included a small sample size, dataset variability, potential overfitting, reliance on selected features and the lack of longitudinal data, limiting generalizability.

Meng *et al.* (2022) introduces a multi-modal LassoNet framework that combines resting-state fMRI and DTI to classify Alzheimer's Disease (AD). The framework demonstrated high accuracy, especially in differentiating Alzheimer's disease from healthy controls and early mild cognitive impairment. Results showed the highest accuracy in the AD-HC group, reaching 92.79%. It outperformed other models and exhibited stability in classification. However, the deterministic fiber tracking method used in the study has limitations, potentially introducing biases.

Massalimova & Varol (2021) utilizes deep learning models trained on structural MRI (T1W) and Diffusion Tensor Imaging (DTI) from the Open Access Series of Imaging Studies (OASIS)-3 dataset. The study introduces a unique input-agnostic architecture, enabling the model to use either or both modalities, distinguishing it from prior multi-modal approaches. The results demonstrate that combining MRI and DTI scans yields the best performance, achieving an accuracy of up to 0.97. Comparisons with similar studies indicate superior outcomes, suggesting that whole-brain feature extraction and advanced network architectures contribute to the improved results.

Vaithianathan *et al.* (2024) focuses on resting-state fMRI (rs-fMRI) analysis for classifying Alzheimer's disease. It employs a seven-stage preprocessing pipeline and extracts features using wavelets, shearlets and scattering transforms. An automated system integrates heterogeneous features for classification. The results indicate that wavelet features achieve the highest classification accuracy. While the model outperforms others, challenges remain, including overfitting and the need for external validation. Future work aims to incorporate additional imaging modalities.

Khatri & Kwon (2022) highlights Alzheimer's Disease (AD) diagnosis using structural MRI and resting-state fMRI data. For classification, it leverages functional network features along with the volume of amygdala and hippocampus subfields. Feature selection techniques such as LASSO, SVM-RFE and JMI enhance classification accuracy. The results indicate that combining structural and functional features yields high classification performance, with SVM combined with JMI outperforming other methods. The study underscores the significance of specific brain regions in distinguishing different stages of AD.

Pan *et al.* (2021) presents a multimodality framework by integrating features from different brain modalities. It extracts grey matter volume, surface area and cortical thickness from pre-processed sMRI data. After feature extraction using a 3DCNN-SE module, dimensionality is decreased using an indicator selection strategy. The Multi-Attention-Fusion Module (MAFM) fuses the extracted features for classification. Evaluating 596 patients from the ADNI dataset, the model achieves 88% accuracy in distinguishing AD, MCI and cognitively normal individuals, outperforming existing methods.

Pamarapa *et al.* (2024) developed an SVM classification model to differentiate between Cognitive Normal (CN), Early Mild Cognitive Impairment (EMCI) and Late Mild Cognitive Impairment (LMCI) in people 65–75 years old using T1-weighted MR and F-18 FDG PET data. The methodology involved preprocessing, registering images to standardized templates and using SVMs with various feature sets for classification. Results showed that combined PET/MR imaging improved accuracy, particularly for CN vs. LMCI classification, while PET alone was most effective for EMCI vs. LMCI.

In order to diagnose Alzheimer's Disease (AD), the study Wang *et al.* (2024) presents a deep joint learning model and a multimodal feature fusion technique known as "MRI-p value." It combines ResNet for feature extraction and attention mechanisms for location information, enhancing classification accuracy. Using data from the ADNI, the method achieved high accuracy and AUC scores across different classifications. Additionally, six novel genes were identified, providing

new targets for potential AD treatments. Possible enhancements include incorporating more pre-trained models, improving the integration of genetic and imaging data and addressing increasingly complex multi-classification task.

Rama Lakshmi and Radhika (2024b) focuses on early Alzheimer's disease diagnosis using sophisticated Machine Learning (ML) methods. It employs a dataset with biomarkers and cognitive traits and introduces an adaptive filtering process that combines Gaussian and bilateral filters for noise reduction. This is followed by feature extraction using a customized ResNet-53 model. Additionally, the study fine-tunes SVM models for both binary and multiclass classification. To optimize performance, Bayesian optimization is applied for Hyperparameter tuning. The integration of multimodal data enhances the model's robustness, making it a reliable tool for early diagnosis.

Using multimodal neuroimaging data from the ADNI dataset, Desai *et al.* (2024) investigate the application of Deep Learning (DL) in the diagnosis of Alzheimer's disease (AD). Various Convolutional Neural Networks (CNNs) were tested, with DenseNet achieving the highest accuracy (90%) and the lowest validation loss (0.17). The research highlights the significance of multimodal fusion, combining MRI, PET scans and other data to enhance diagnostic precision. It also underscores the importance of ethical considerations, including data privacy and fairness, in the deployment of DL models for AD diagnosis. This study can be further enhanced by fully utilizing these models to advance analysis and treatment, emphasizing the need for interdisciplinary collaboration.

A Pareto-optimal cosine color map is presented in the study by Odusami *et al.* (2024b) to enhance fused image classification and visual clarity. For feature extraction and classification, a mobile vision transformer (ViT) model with a Swish activation function is employed. The fused images from the ADNI, AANLIB and OASIS datasets are used to train the model, while non-fused images are used for evaluation. The proposed method achieves high classification accuracy, with notable results such as 99.25% accuracy for distinguishing AD from healthy images in ADNI and 99.50% precision for AD versus CN in OASIS. Evaluation metrics demonstrate the model's strong performance.

Thanh *et al.* (2024) introduces a novel method called Tensor Kernel Learning (TKL) for Alzheimer's disease assessment. By integrating data from MRI, PET, CSF and SNP using tensor and kernel learning, to enhance AD prediction. The method uses a supervised SVM classifier after fusing the data using CP/PARAFAC decomposition and graph diffusion. Applied to the ADNI dataset, TKL outperforms single-modality methods, achieving higher classification accuracies for distinguishing between cognitive stages. Additionally, Clearer patterns in the

data are made possible by TKL, which facilitates the comprehension of intricate interactions between modalities.

Odusami *et al.* (2024a) proposes a novel method for Alzheimer's Disease (AD) diagnosis using fused sMRI and FDG-PET images, enhanced with Gaussian Laplacian pyramid and Static Pulse-Coupled Neural Network (SPCNN) techniques. The fused images are processed by a Pareto-optimized Mobile Vision Transformer (MViT) for classification. The approach achieves high accuracy in distinguishing AD from CN and MCI stages, with precision rates of 94.73, 92.98 and 89.36%, respectively. The fusion method outperforms traditional techniques, ensuring detailed image quality.

In order to categorize Alzheimer's Disease (AD) and other cognitive illnesses using 2D MRI images, Shah *et al.* (2024) suggest the Bi-Vision Transformer (BiViT) design. For improved feature learning, the model integrates two new modules: Parallel Coupled Encoding Strategy (PCES) and Mutual Latent Fusion (MLF). According to the study, performance may be enhanced by addressing data imbalance and scarcity.

Kamal & Nimmy (2024) use a multi-modality strategy that integrates text and visual data to improve Alzheimer's disease detection. The method involves pre-processing MRI images with a U-net segmentation technique to isolate Regions of Interest (ROIs). Vision transformers (ViT) and BERT are then used to process the pre-processed data, addressing the complexities of multi-modal datasets. To improve interpretability, Explainable AI techniques, such as LIME and LRP, are incorporated, providing understanding of the model's decision-making process. The study uses medical demographic and image data from Kaggle, achieving an accuracy of 86%, surpassing other methods.

By combining multi-modal neuroimaging data, Liu *et al.* (2024) suggest a Hierarchical Attention-based Multi-task Multi-modal Fusion (HAMMF) model to improve the diagnosis of Alzheimer's Disease (AD). The model simultaneously performs AD classification, cognitive score regression and age regression using MRI and PET scans from the ADNI dataset. Using channel and spatial attention mechanisms, a Contextual Hierarchical Attention Module (CHAM) is shown to capture fine-grained data. The Transformer model efficiently integrates features from the two modalities. The approach achieves 93.15% accuracy in AD/NC recognition, showcasing strong pathological feature recognition.

A novel adversarial learning-based technique for fusing MRI and PET scans for early Alzheimer's Disease (AD) diagnosis is presented in the study by Choudhury *et al.* (2024). It proposes preprocessing steps to align and normalize these images, making them suitable for fusion. The core of the method is a Coupled GAN (CGAN) architecture that uses dual convolutional auto encoders

and discriminators to extract and fuse structural (MRI) and metabolic (PET) features into a shared latent space, which are then classified for AD stage identification. The model is tested on the ADNI dataset and shows superior performance compared to existing methods. This approach allows for more accurate AD diagnosis by leveraging both MRI and PET data.

Duenias *et al.* (2024) introduces a novel framework based on hypernetworks to effectively Merge Medical Imaging (MRI) with Electronic Health Record (EHR) data for enhanced healthcare diagnostics. By conditioning image processing on the patient's EHR data, the approach improves the accuracy of brain age prediction and Alzheimer's disease classification. The framework outperforms both single-modality models and existing MRI-tabular fusion methods, offering superior data utilization and flexibility. Validation on two clinical tasks demonstrates its robustness and generalizability.

Using both 3D MRI and amyloid PET imaging, Castellano *et al.* (2024) investigate the creation of multi-modal diagnostic models for Alzheimer's Disease (AD). The research demonstrates that models using volumetric data (3D MRI and PET) offer better performance than those relying on 2D images alone. Integrating these modalities significantly enhances model accuracy. The use of Grad-CAM for explainability highlights key brain regions linked to AD. Limitations include the use of only 50 slices from the axial plane and the loss of temporal resolution in PET scans.

Mahmud Joy *et al.* (2025) present ViTAD, a Vision Transformer-based model for classifying five stages of Alzheimer's Disease (AD) using 1,296 brain MRI images from the ADNI dataset. The model improves upon Google's ViT by fine-tuning hyper parameters and adding layers for enhanced performance. Preprocessing steps include grayscale-to-RGB conversion, cropping and Laplacian sharpening, while data augmentation techniques such as flipping, zooming and rotation ensure robustness. The model was trained for 20 epochs with a learning rate of 0.0001, achieving 99.98% accuracy, 100% precision and an F1-score of 1.00. It outperforms CNN-based models like DenseNet and EfficientNet, reaching optimal accuracy in just 8 epochs. Despite its high accuracy, the model's reliance on a relatively small dataset raises concerns about generalizability and potential overfitting.

Khan *et al.* (2024c) proposes Dual-3DM3-AD, a multi-modal fusion model for early Alzheimer's diagnosis using MRI and PET scans. Preprocessing includes noise reduction (QNLN), skull stripping (Morphology function) and 3D conversion (BDM). A Mixed-Transformer with Furthered U-Net is employed for segmentation, followed by multi-scale feature extraction and fusion using DCFAM. A multi-head attention mechanism enhances feature selection before

classification with a softmax layer. The model achieves 98% accuracy, outperforming existing approaches. Potential limitations include computational complexity and dataset generalizability concerns.

Aghdam *et al.* (2023) proposes PVTAD, a novel approach for Alzheimer's disease classification using Pyramid Vision Transformer (PVT) and White Matter (WM) features from T1-weighted sMRI scans. The method integrates CNN and ViT to extract both local and global patterns, enhancing biomarker identification. Experiments on the ADNI dataset achieved 97.7% accuracy and a 97.6% F1-score, surpassing traditional CNN and ViT models. While effective, the approach may be computationally intensive and dataset-dependent.

Mora-Rubio *et al.* (2023) utilizes MRI scans from the ADNI and OASIS datasets to classify different stages of Alzheimer's Disease (AD). It employs FreeSurfer for preprocessing and data augmentation techniques like rotation, flipping and zooming. The models used include 3D CNNs (EfficientNet, DenseNet, a Siamese network) and a Vision Transformer architecture. The best accuracy achieved was 89% for AD vs. Control but dropped to 66-67% for early-stage detection, highlighting the challenge of identifying mild cognitive impairment. The study emphasizes the need for improved detection techniques for early AD diagnosis.

The existing literature highlights several data fusion techniques and their limitations, such as small sample sizes and challenges in integrating features from different modalities, which often lead to inaccurate analysis of critical features. The Vision Transformer (ViT) techniques discussed in the literature survey are primarily applied to small datasets and face generalizability issues across different datasets, affecting their reliability in medical imaging. Additionally, ViTs require large datasets to prevent overfitting and struggle with early-stage disease detection due to their limited inductive biases.

In order to tackle these problems, we suggest a new strategy that makes use of proven deep learning. Specifically, GoogLeNet and DenseNet-121 are employed for sMRI feature extraction and classification, while GNN and U-Net are utilized for fMRI dataset feature extraction and classification. Rather than merging extracted features, which can lead to high dimensionality, redundant or overlapping features, varying feature vector sizes and reduced interpretability, we combined the prediction probabilities from these models to create a meta-model. This meta-model is designed to diagnose Alzheimer's disease with greater precision by mitigating overfitting and addressing the limitations of feature merging. The meta-model effectively generalizes these issues, as it utilizes prediction probabilities as feature matrix, ensuring sufficient input for robust learning. Comparative results, as listed in Table (1), the proposed

meta-model outperforms state-of-the-art techniques. The key contribution of this research are:

- Novel stacking ensemble method: The research introduces a ensemble learning technique using "stacking" to combine predictions from multiple deep learning models trained on both sMRI and fMRI data.
- Multimodal data integration: This approach effectively integrates information from different imaging modalities (sMRI and fMRI), leveraging their complementary strengths for improved diagnostic accuracy
- Addressing feature merging challenges: Instead of directly merging extracted features, which can lead to high dimensionality, redundancy and interpretability issues, the research utilizes prediction probabilities from individual models as input for the meta-model
- Enhanced accuracy: The suggested meta-model outperformed current state-of-the-art techniques in the early diagnosis of Alzheimer's disease, achieving a noticeably higher accuracy of 95.18%

Table 1: Shows comparison of proposed method with state of art techniques

| S.No | Author | Dataset | Methodology | Accuracy |
|------|-----------------------------------|-----------------------|--|----------|
| 1 | Yu <i>et al.</i> (2024) | ADNI | Group self-calibrated coordinate attention network GSCANet | 83.33% |
| 2 | Dolci <i>et al.</i> (2024) | ADNI | Deep Neural Network | 76.2%, |
| 3 | Jiao <i>et al.</i> (2024) | ADNI | Multi-modal data fusion framework (MFASN) | --- |
| 4 | Wang <i>et al.</i> (2024) | ADNI | deep joint learning diagnostic model | 93.44% |
| 5 | Desai <i>et al.</i> (2024) | ADNI | DenseNet | 90% |
| 6 | Thanh <i>et al.</i> (2024) | ADNI | Tensor kernel learning (TKL) | 91.31% |
| 7 | Liu <i>et al.</i> (2024) | ADNI | Hierarchical attention-based multi-task and multi-modal fusion model | 93.15% |
| 8 | Long <i>et al.</i> (2023) | Nanfang Hospital data | SVM and ANN | 80.36% |
| 9 | Meng <i>et al.</i> (2022) | ADNI | multi-modal LassoNet framework | 92.79%. |
| 10 | Khatri and Kwon (2022) | ADNI | LASSO, SVM-RFE, and JMI(Joint Mutual Information) | 90.35% |
| 11 | Pan <i>et al.</i> (2021) | ADNI | Multi-Attention-Fusion Module (MAFM) | 88% |
| 12 | Mora-Rubio <i>et al.</i> , (2023) | ADNI | Vision Transformer | 89% |
| 13 | Aghdam <i>et al.</i> , 2023 | ADNI | CNN and ViT | 97.70% |
| 14 | Proposed System | ADNI | Stacking based Meta model using Random Forest | 95.18% |

Materials

We utilized datasets that are publicly accessible on the Alzheimer's disease Neuroimaging Initiative (ADNI) website (Weiner *et al.*, 2017). This study integrates structural Magnetic Resonance Imaging (sMRI) and functional Magnetic Resonance Imaging (fMRI) data from the ADNI repository to categorize Alzheimer's Disease (AD) using multimodal imaging. The characteristics of dataset are described in Table (2).The dataset is divided into three categories: Normal Controls (NC), Early Mild Cognitive Impairment (EMCI) and Mild Cognitive Impairment (MCI). The fMRI dataset includes resting-state fMRI image, each fMRI 3D image contains 9800 2D slices and we calculate histograms for pairs of images and measure the differences between them using the Chi-Square distance. Images are then selected based on the degree of variation observed. We collected 980 images from each 3D images based on content variation score. The fMRI images are distributed as follows: 8,316 for EMCI, 9,256 for MCI and 9,849 for NC. The sMRI dataset, collected from the 55 individuals, also includes 2 to 5 visits per subject and contains 7,200 EMCI, 7,376 MCI and 7,430 NC images. By combining these two types of imaging, the study investigates both structural and functional brain changes to better classify the progression of Alzheimer's disease. The longitudinal nature of the data facilitates the examination of changes over time.

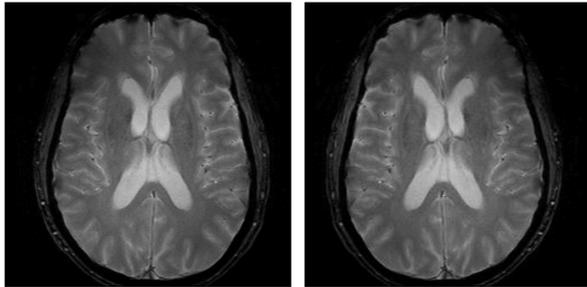
Table 2: Dataset characteristics

| Characteristic | sMRI Dataset | fMRI Dataset |
|----------------------------|--------------|--------------------------|
| Age(years) range | 60-90 | 60-90 |
| Number of Subjects | 55 | 55 |
| Total Number of MRI images | 5069 | 21600 |
| Number of Visits | 5-Feb | 5-Feb |
| Type of modality | sMRI | Axial rsfMRI (Eyes Open) |

Preprocessing

To prepare the data for effective analysis, preprocessing techniques were employed to improve its quality and ensure balanced label distribution. Horizontal flipping and a CycleGAN-based data augmentation method (Parvatham and Maguluri, 2024) were applied to sMRI dataset as shown in Figs. (1-2). These techniques helped to address class imbalance by generating additional data samples, thereby improving the robustness and reliability of the model. Additionally, all images were resized to a uniform resolution, ensuring compatibility across the dataset and enabling seamless processing by the model. Image normalization was also performed to scale pixel values to a consistent range, reducing variability and improving the model's training efficiency.

To further refine the data, an adaptive median filter was employed to mitigate the impact of salt-and-pepper noise as shown in Figure (3), which is a common issue in image datasets. This filter not only removed noise but also preserved critical details in the images, leading to improved overall image quality. By integrating augmentation, resizing, normalization and noise reduction, the preprocessing stage ensured the dataset was well-structured and ready for subsequent feature extraction and classification tasks.



(a) Original image (b) Augmented image

Fig. 1: Horizontal fragmentation

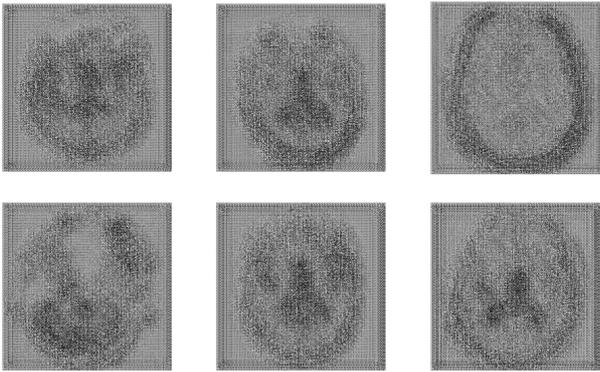
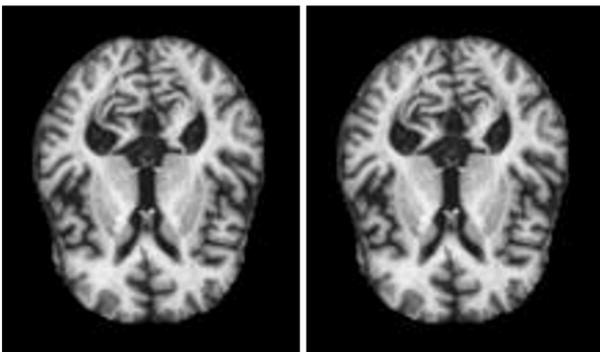


Fig. 2: Cycle GAN generated images



(a) Before noise removal (b) After noise removal

Fig. 3: Illustrating noise removal using adaptive median filter

Dataset Partitioning for Optimized Processing

The dataset consists of a total of 43,606 images, which presents significant computational challenges for

direct processing due to its size and complexity. To address these challenges and enhance the efficiency of preprocessing and analysis, the dataset was divided into four manageable partitions. Specifically, the sMRI dataset was split into two equal portions, designated as Dataset-SM1 and Dataset-SM2 shown in Table (3), ensuring a balanced distribution of samples across both partitions, similarly the fMRI dataset was divided into two separate parts, named Dataset-FM1 and Dataset-FM2 as shown in Table (4). This partitioning approach not only reduced computational demands but also allowed for parallel processing and efficient utilization of available resources.

Table 3: Dataset-SM1 and Dataset-SM2 (sMRI Dataset)

| Label Name | Number of Images |
|------------|------------------|
| CN | 3715 |
| EMCI | 3600 |
| MCI | 3688 |
| Total | 11003 |

Table 4: Dataset-FM1 and Dataset-FM2 (fMRI Dataset)

| Label Name | Number of Images |
|------------|------------------|
| CN | 3600 |
| EMCI | 3600 |
| MCI | 3600 |
| Total | 10800 |

Methods

In this study, we introduce a stacking-based ensemble approach that integrates predictions from multiple models, including GoogLeNet, DenseNet-121, GNN and U-Net, to improve the classification of Alzheimer’s disease (Acharya *et al.*, 2021) using sMRI and fMRI datasets. Initially, we started experimenting with various state-of-the-art techniques such as GoogLeNet, VGG-16, ResNet, DenseNet and EfficientNet for sMRI image analysis. However, GoogLeNet and DenseNet performed best with the meta-model, so we selected them for sMRI image analysis. For fMRI image analysis, GNN and U-Net were used. GNN effectively captures the spatial and temporal relationships in brain connectivity graphs, while U-Net was included for its ability to preserve fine spatial details, which is crucial in detecting subtle changes in brain structures. The combination of GNN and U-Net produced good results when we trained the meta-model using extracted features from these models.

GoogLeNet and DenseNet extract spatial patterns from sMRI, GNNs analyze functional connectivity in fMRI and U-Net identifies localized features. The predictions from these models are combined into a meta-feature matrix, which is then processed by a second-level classifier, such as logistic regression, to produce the final output. This meta-model learns to integrate the strengths of individual models for final predictions. By leveraging the complementary nature of structural and functional data, stacking improves classification accuracy and

generalization, making it a powerful tool for multi-modal analysis in neurodegenerative disease research.

Structural MRI Image Analysis

GoogLeNet

GoogLeNet, a prominent deep learning architecture in medical imaging, excels at identifying features in brain MRI data. Its effectiveness stems from its 22-layer structure and Inception modules, which capture fine details in MRI scans. Figure (4) illustrates the pipeline for Alzheimer’s disease classification using GoogLeNet (Inception V3) and sMRI images. The dataset undergoes preprocessing and is split into training, validation and test sets. A pre-trained Inception V3 model serves as the base feature extractor, initially with frozen layers to retain learned knowledge. Feature extraction is performed using Global Average Pooling, followed by a dense layer (128 neurons, ReLU activation) and a Dropout layer (0.5) to reduce overfitting. The model is then compiled with the Adam optimizer and trained using categorical cross-entropy loss. After initial training, the base model layers are unfrozen for fine-tuning. Finally, the model is evaluated on test data to determine its classification accuracy (Oh *et al.*, 2019).

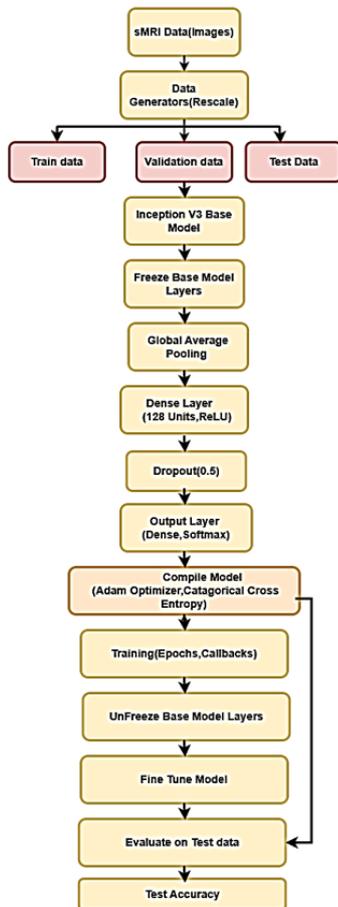


Fig. 4: GoogleNet Architecture for AD Classification

Equations (1-3) are equations to compute the outputs of convolution, activation and pooling operations respectively. The model also utilizes global average pooling and auxiliary classifiers to improve performance and ensure training stability (Parvatham and Maguluri, 2024). Finally the softmax function converts scores into probabilities as shown in (4).

$$Y [i, j, k] = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X [i + m, j + n] \cdot W [m.n.k] \quad (1)$$

where:

X is the input image

W is the filter (kernel)

Y is the output feature map

$Y [i, j, k]$ is the output feature map value at position (i, j) in the k^{th} channel after applying the convolution operation

M and N are height and width of the filter

$$f(x) = \max(0, x) \quad (2)$$

where:

x is the input value

$f(x)$ is the output after applying ReLU:

$$Y [i, j, k] = \max(X [i : i + P, j : j + P, k]) \quad (3)$$

where:

P is the pooling size

X is the input feature map

Y is the output of the pooling layer:

$$P(y = j|x) = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad (4)$$

where:

$P(y = j|x)$ is the probability that the input x belongs to class j

z_j is the j^{th} element of the output vector z

K is the total number of classes

DenseNet-121

Figure (5) represents an Alzheimer’s disease classification pipeline using Dataset-SM2 partition and DenseNet-121. The Dataset-SM2 partition of the sMRI dataset, which is divided into 70% training, 20% testing and 10% validation. Its dense layer connections enable efficient feature reuse, improving accuracy in detecting subtle brain structure changes linked to AD (Solano-Rojas *et al.*, 2020; Hazarika *et al.*, 2023). The model undergoes preprocessing, feature extraction using Global Average Pooling and classification with a dense layer and Dropout (0.5). Initially, the pre-trained model’s layers are frozen, then fine-tuned after initial training. The model is compiled using Adam optimizer and

categorical cross-entropy loss and evaluated on test data to determine its classification accuracy. Feature extraction using DenseNet-121 involves applying convolution operations to capture spatial details, followed by batch normalization shown in (5) and ReLU for regularization and non-linearity shown in (2), with dense blocks concatenating previous layer outputs to improve feature propagation as shown in (6). The process ends with global average pooling as in (7) and softmax in (4) for classification.

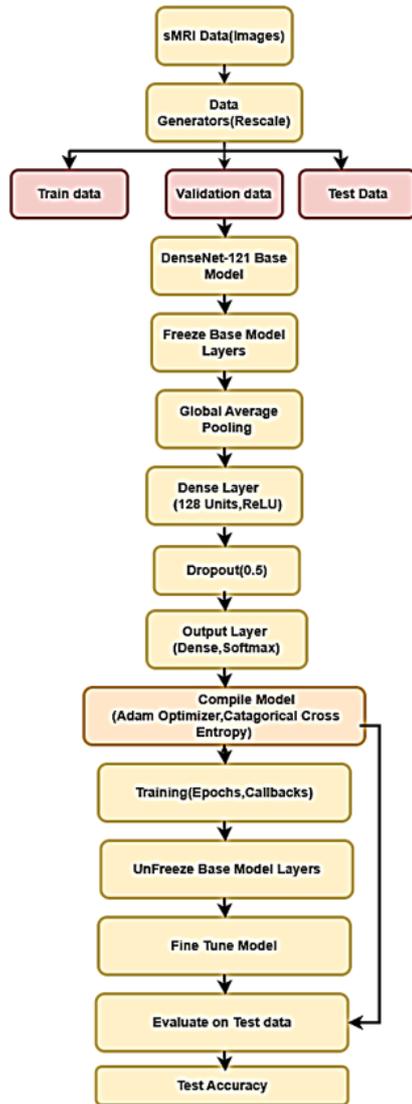


Fig. 5: DenseNet-121 Architecture for AD Classification

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}, y_i = \gamma \hat{x}_i + \beta \quad (5)$$

where:

x_i is the input value for the i^{th} feature in the batch,

μ_B is the mean of the input values in the current batch

m is the number of inputs in the batch

σ_B^2 is the variance of the input values in the current batch
 ϵ is the small constant added for numerical stability to prevent division by zero,

\hat{x}_i is the normalized value of x_i

γ is a learnable scaling parameter applied to the normalized value

β is the learnable shift parameter that adjusts the mean of the normalized values,

y_i The final output of the batch normalization:

$$H_l = [X_0, X_1, \dots, X_{l-1}, X_l] \quad (6)$$

where:

H_l is input to the l^{th} layer, which is the result of concatenating all the feature maps produced by the preceding layers up to the l^{th} layer X_0, X_1, X_2, \dots

X_{l-1}, X_l represent the feature maps generated by each of the previous layers from the 0^{th} layer up to the l^{th} layer. Each X_i is a set of feature maps from the i^{th} layer

$$Y[k] = \frac{1}{H \times W} \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} X[i, j, k] \quad (7)$$

where:

$Y[k]$ is the output value for the k^{th} feature map after applying Global Average Pooling,

$X[i, j, k]$ is the input feature map

i and j index the spatial dimensions of the feature map

k indexes the channel or depth dimension

H is the height of the input feature map

W is the width of the input feature map

Functional MRI Image Analysis

Graph Neural Networks (GNNs)

Graph Neural Networks (GNNs) are employed for Alzheimer's disease classification by analyzing resting-state fMRI data, which captures interactions between different brain regions. These functional connectivity patterns, measured through blood flow variations, differ between healthy individuals and Alzheimer's Disease (AD) patients (Han *et al.*, 2024). The flowchart illustrated in Figure (6) represents an Alzheimer's disease classification pipeline using Dataset-FM1 and a Graph Neural Network (GNN). The fMRI images are first converted to grayscale and then transformed into graph representations, where nodes represent brain regions and edges define their connectivity patterns (Tong *et al.*, 2023; Zhou *et al.*, 2024; Gao *et al.*, 2023). Two Graph Convolutional Network (GCN) layers extract spatial features by aggregating information from neighboring

nodes. A Global Mean Pooling layer reduces feature dimensions while preserving crucial information. Finally, a fully connected layer classifies brain scans into Cognitively Normal (CN), Early Mild Cognitive Impairment (EMCI), or Mild Cognitive Impairment (MCI). This graph-based method improves classification accuracy by leveraging the structural relationships in brain connectivity, making it a valuable tool for early Alzheimer's detection.

In this study, every fMRI image is converted into a graph where the nodes correspond to individual pixels and edges link each pixel to its adjacent right and lower neighbors. The Graph $G(V, E)$ is constructed as shown in (8) and (9), where V is the set of nodes and E is the set of edges. It employs graph convolutional layers to propagate features, followed by global mean pooling to aggregate node features into a graph-level representation as shown in (10), (11) & (12). Finally, a fully connected layer classifies the graph based on the extracted features:

$$V = \{v_i | i \in [0, H \times W - 1]\} \quad (8)$$

where:

Each pixel value is treated as a node feature $x_i, x_i \in R$

The feature matrix X for all nodes is shown in Eq. 10:

$$E = \{v_i, v_j\} | i f j = i + 1 (\text{rightneighbor}), = i + w (\text{lowerneighbor}) \quad (9)$$

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_N \end{bmatrix} \in R^{N \times 1} \quad (10)$$

where: $N = H \times W$

The GNN uses Graph Convolutional Network (GCN) layers to propagate and aggregate information across the graph. The graph convolution operation is defined as:

$$H^{l+1} = \sigma \left(\hat{D}^{-\frac{1}{2}} \hat{A} \hat{D}^{-\frac{1}{2}} H^{(l)} W^{(l)} \right) \quad (11)$$

where:

$H(l)$ is the node feature matrix at layer l

$\hat{A} = A + I$ is the adjacency matrix with added self-loops

\hat{D} is the degree matrix of \hat{A}

W^l is the learnable weight matrix at layer l , σ is an activation function

$$h_G = \frac{1}{N} \sum_{i=1}^N h_i \quad (12)$$

where:

h_i is the features of the i^{th} node after the final GCN layer.

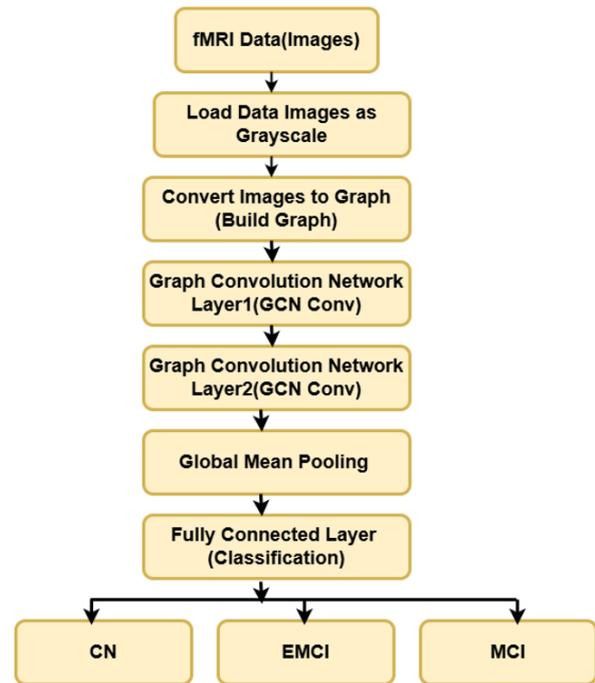


Fig. 6: GNN Architecture for classification of AD

U-Net Architecture

The U-Net architecture is popular in medical image analysis, especially for fMRI-based Alzheimer's disease classification, due to its effective pixel-level segmentation. It includes a contracting path for capturing context and an expanding path for precise feature localization, preserving spatial details in high-resolution images (Asiri *et al.*, 2023). The U-Net architecture excels in medical imaging, particularly for identifying structures and abnormalities in brain scans and is effective in segmenting brain tumors and regions of interest in fMRI data (Pang *et al.*, 2023). The flowchart in Figure (7) represents an Alzheimer's disease classification pipeline using dataset-FM2 and a U-Net-based CNN. The model extracts spatial and hierarchical features through multiple convolutional layers with ReLU activation and MaxPooling, progressively increasing the number of feature channels (64, 128, 256, 512). A Global Average Pooling (GAP) layer compresses feature representations before passing them to a fully connected classification layer. The model classifies brain scans into Cognitively Normal (CN), Early Mild Cognitive Impairment (EMCI) and Mild Cognitive Impairment (MCI).

For Alzheimer's disease classification, U-Net aids in delineating subtle changes in brain activity by segmenting and analyzing regions associated with cognitive functions. Its capacity to process high-dimensional fMRI data efficiently and detect nuanced variations supports both clinical diagnosis and research (Fujita *et al.*, 2023). The U-Net model applies

convolution, ReLU activation, max pooling and adaptive average pooling for feature extraction as shown in (13), (2), (14) & (15) respectively, followed by a dense layer for classification, using cross-entropy loss for optimization and the Adam optimizer.

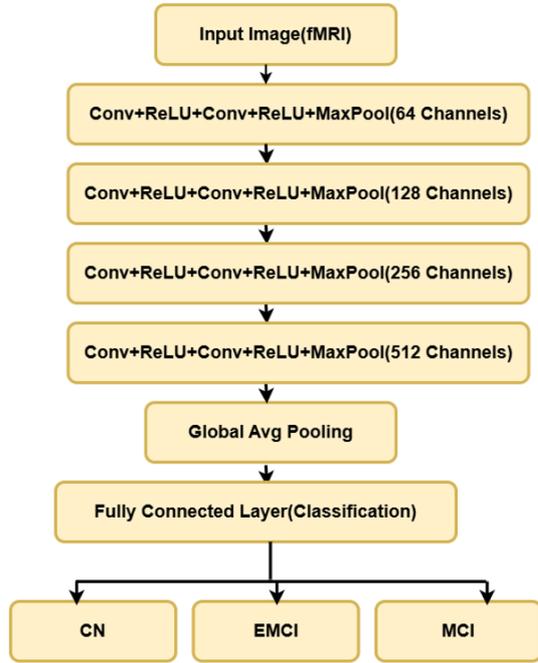


Fig. 7: U-Net architecture for classification of AD

$$Conv2d(x, W) = (x * W + b) \quad (13)$$

where:

x is the input image or feature map

W represents the convolution filters

b is the bias term

$*$ denotes the convolution operation:

$$MaxPool2d(x) = \max_{i,j \in window} x_{i,j} \quad (14)$$

where:

$x_{i,j}$ represents the pixel values in pooling window

$$AdaptiveAvgPool2d(x) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W x_{ij} \quad (15)$$

where:

H and W are the height and width of the feature map

x_{ij} is the pixel value at position (i,j)

Proposed System Architecture

The proposed Alzheimer's Disease Classification System aims to categorize patients into three stages—Normal Control, Early Mild Cognitive Impairment and Mild Cognitive Impairment using a multimodal approach with sMRI and fMRI datasets. The architecture of proposed system shown in Figure (8), the system begins

by acquiring and preprocessing the imaging data, employing techniques like augmentation, resizing, normalization and noise reduction to enhance quality and address class imbalance. Preprocessed data is divided into four datasets, each tailored for a specific model: GoogLeNet and DenseNet-121 for sMRI and fMRI classification, respectively, GNN for analyzing relationships in sMRI data and U-Net for localized fMRI feature extraction. Each model independently generates predictions and their prediction probabilities are combined using a vertical stacking ensemble approach (Chen *et al.*, 2022) to create a feature matrix as shown in Eq. 16. In this matrix, each row represents a sample and the columns consist of the concatenated prediction probabilities from all models across all classes. This resulting feature set (8722, 3) is then used for training the meta-model with various classifiers. We split the resultant feature set into two partitions:

$$Stacked\ Probabilities = \begin{bmatrix} GNN\ Pred.\ Prbabilities \\ GoogLeNet\ Pred.\ Prbabilities \\ U - Net\ Pred.\ Prbabilities \\ DenseNet - 121\ Pred.\ Prbabilities \end{bmatrix} \quad (16)$$

80% for training and 20% for testing the model. This meta-classifier delivers the final classification into NC, EMCI, or MCI. Performance is assessed using metrics such as accuracy, sensitivity, specificity, precision and F1-score. The proposed system enhances the reliability and precision of Alzheimer's disease classification, offering a robust tool for neurodegenerative disease research.

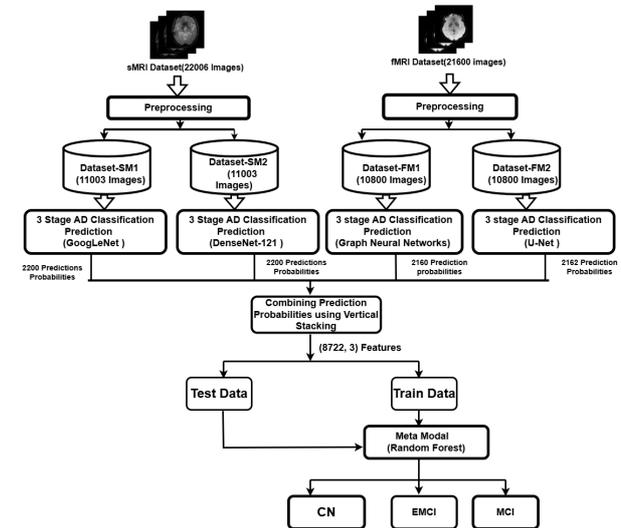


Fig. 8: Architecture of proposed system

Results

This research explored four deep learning architectures: Graph Neural Networks and U-Net for analyzing functional MRI data, alongside GoogLeNet and DenseNet-121 for static MRI analysis. The

prediction probabilities generated by these models were integrated into a meta-model to enhance the accuracy of the final predictions. Given the substantial size of the dataset, the experiments were performed on Google Colab, utilizing CUDA—NVIDIA's parallel computing framework designed to harness GPU cores for accelerated computation. Google Colab offered free access to high-performance GPUs, such as the NVIDIA Tesla, along with a pre-configured Python environment. This setup facilitated efficient dataset handling and significantly faster training and inference processes compared to traditional CPU-based systems.

GNN model Using Dataset-FM1

The model attains a total accuracy of 57.82%, demonstrating its strongest performance in the CN class with an accuracy of 80.05%, while its weakest performance is observed in the EMCI class, achieving an accuracy of 65.70%. It shows effectiveness in identifying non-CN cases, reflected by a specificity of 0.8691 and achieves a balanced F1-score of 0.6812 for CN. However, for EMCI, lower precision (0.4810) and recall (0.3870) suggest substantial potential for enhancement. The MCI class demonstrates moderate performance, achieving an accuracy of 70.10% and an F1-score of 0.6120. It performs better than EMCI but does not surpass CN.

U-Net Model Using Dataset-FM2

The U-Net model achieves an overall accuracy of 99.90%, demonstrating an outstanding ability to classify cases with minimal errors. For the CN class, it achieves nearly perfect accuracy (99.91%) with high precision, recall and F1-score values (0.9986 each), reflecting exceptional reliability. The EMCI class is classified flawlessly with a 100% accuracy and all metrics at 1.00. Similarly, the MCI class shows near-perfect accuracy (99.91%) with excellent specificity, precision, recall and F1-score (0.9986 each). Overall, the model delivers superior performance across all classes, with particularly remarkable results in EMCI classification.

GoogLeNet Model Using Dataset-SM1

The GoogLeNet model utilizes the Dataset-SM1 partition of the sMRI dataset, which is further split into three subsets: Train, test and validation. For model training, testing and validation, 70% of the images in each category are allocated for training, 20% for testing and the remaining 10% for validation. The GoogLeNet architecture delivers outstanding results, achieving a total accuracy of 99.54%, which signifies a minimal error rate. For the CN class, the model attains an accuracy of 99.73%, accompanied by high specificity (0.9979), precision (0.9960), recall (0.9960) and an F1-score of 0.9960. The EMCI class also performs exceptionally well, with an accuracy of 99.68%, a specificity of 0.9966

and balanced metrics such as precision (0.9930), recall (0.9972) and an F1-score of 0.9950. Similarly, the MCI class achieves an accuracy of 99.68%, with specificity (0.9986), precision (0.9972), recall (0.9932) and an F1-score of 0.9951. These findings highlight the model's strong generalization ability and robustness when applied to unseen data.

DenseNet Model Using Dataset-SM2

The DenseNet-121 model achieves impressive performance with an accuracy of 97.86%, reflecting a remarkably low error rate. It achieves perfect accuracy in the CN class (100%) with perfect precision, recall, specificity and F1-score (all 1.00). For the EMCI class, the model reaches 97.90% accuracy, with high precision (0.9490), recall (0.9870) and a strong F1-score of 0.9680. The MCI class also shows near-perfect performance with 98.31% accuracy, high specificity (0.9939), precision (0.9873) and recall (0.9488), leading to an F1-score of 0.9687. Overall, the model demonstrates outstanding performance across all classes and is well-suited for the classification task, likely to generalize well to new data. The Table (5) shows the results of deep learning models using single modality images.

Table 5: Results of single modality with various deep learning models

| Label | Metric | Dataset-FM1 (GNN) | Dataset-FM2 (U-Net) | Dataset-SM1 (GoogLeNet) | Dataset-SM2 (DenseNet-121) |
|---------|-------------|-------------------|---------------------|-------------------------|----------------------------|
| CN | Accuracy | 80.05 | 99.91 | 99.73 | 100 |
| | Specificity | 0.8691 | 0.9993 | 0.9979 | 1 |
| | Precision | 0.7098 | 0.9986 | 0.996 | 1 |
| | Recall | 0.6546 | 0.9986 | 0.996 | 1 |
| | F1-Score | 0.6812 | 0.9986 | 0.996 | 1 |
| EMCI | Accuracy | 65.7 | 100 | 99.68 | 97.9 |
| | Specificity | 0.792 | 1 | 0.9966 | 0.974 |
| | Precision | 0.481 | 1 | 0.993 | 0.949 |
| | Recall | 0.387 | 1 | 0.9972 | 0.987 |
| | F1-Score | 0.428 | 1 | 0.995 | 0.968 |
| MCI | Accuracy | 0.701 | 99.91 | 99.68 | 98.31 |
| | Specificity | 0.705 | 0.9993 | 0.9986 | 0.9939 |
| | Precision | 0.546 | 0.9986 | 0.9972 | 0.9873 |
| | Recall | 0.692 | 0.9986 | 0.9932 | 0.9488 |
| | F1-Score | 0.612 | 0.9986 | 0.9951 | 0.9687 |
| Overall | Accuracy | 57.82 | 99.9 | 99.54 | 97.86 |

Figure 9(a) presents a comparison of accuracy among four deep learning models: U-Net, GoogLeNet, DenseNet-121 and Graph Neural Network (GNN). U-Net demonstrates the highest accuracy of 99.90%, while GoogLeNet follows closely with 99.54%. In contrast, the Graph Neural Network achieves a notably lower accuracy of 57.82% and DenseNet-121 attains an accuracy of 97.86%. The results emphasize the superior accuracy performance of U-Net and GoogLeNet over DenseNet-121 and the Graph Neural Network. Figures 9(b-d) illustrate a comparison between specificity (TNR)

and sensitivity (TPR), which are critical metrics in medical diagnosis. The results are analyzed for each Alzheimer's disease (AD) stage, including Cognitively Normal (CN), Early Mild Cognitive Impairment (EMCI) and Mild Cognitive Impairment (MCI). DenseNet demonstrates exceptional accuracy in classifying CN, while U-Net excels in accurately classifying EMCI and MCI stages.

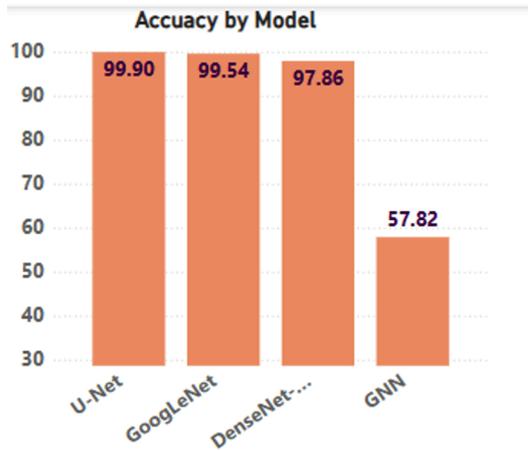


Fig. 9(a): Accuracy comparison of deep learning models using single modality

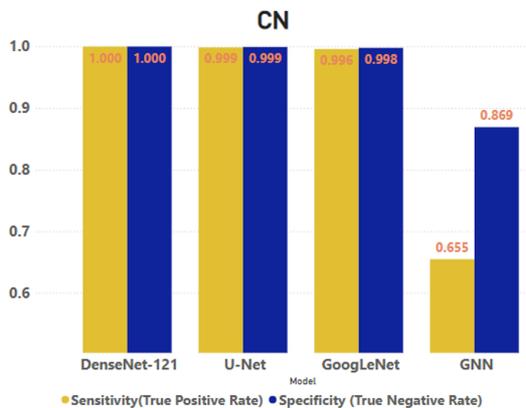


Fig. 9(b): Sensitivity and Specificity comparison of CN stage

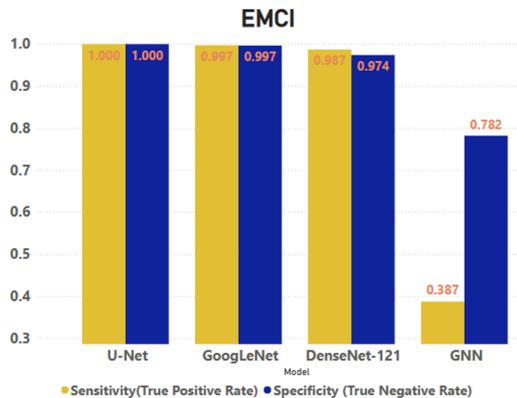


Fig. 9(c): Sensitivity and Specificity comparison of EMCI stage

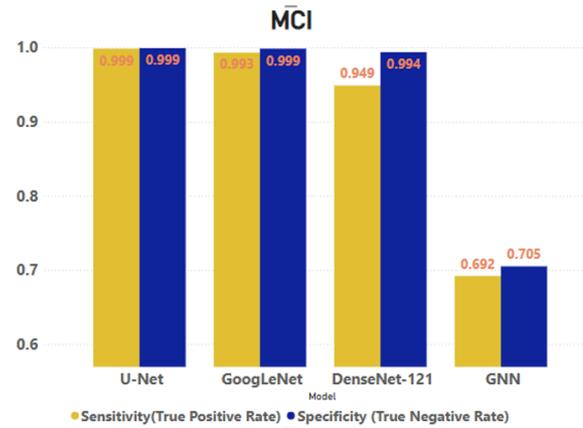


Fig. 9(d): Sensitivity and Specificity comparison of MCI stage

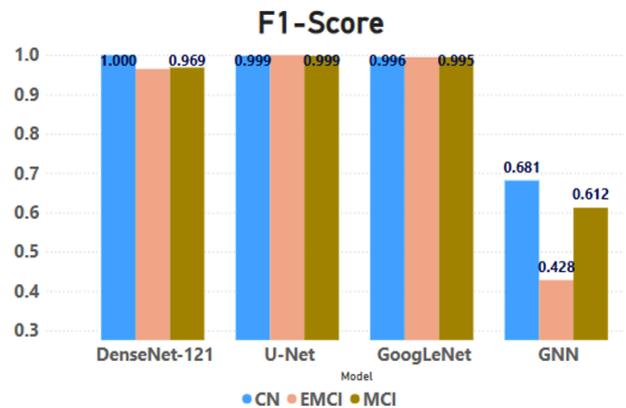


Fig. 9(e): F1-score comparison by model

Figure 9(e) illustrates an F1-score comparison among four deep learning models, which presents a balanced evaluation of their performance across three AD stages. U-Net and GoogLeNet consistently achieve high F1-scores across all stages, while DenseNet-121 also performs well. In contrast, the GNN demonstrates significantly poorer performance, with a notably lower F1-score compared to the other models.

GoogLeNet and DenseNet-121 outperformed GNN and U-Net primarily because the latter models were trained on fMRI data, which presents unique challenges. Functional MRI (fMRI) captures temporal brain activity and connectivity, requiring graph-based representations to model inter-region relationships. Unlike sMRI images, which have a well-defined spatial structure, fMRI data must first be converted into a graph, where brain regions are represented as nodes and functional connectivity defines the edges. One of the major challenges with this approach is that fMRI connectivity patterns vary across individuals, making it difficult for GNNs to generalize effectively. In contrast, CNNs like GoogLeNet and DenseNet work with spatially consistent sMRI images, allowing them to learn more stable and robust features.

Additionally, optimizing GNN architectures is inherently more complex than training CNNs. A potential strategy to enhance performance is by integrating CNN-based spatial features from sMRI with GNN-based functional connectivity features from fMRI, creating a more comprehensive classification model that leverages both structural and functional information.

Multimodality Meta modal

We created various meta-models, including FFNN, Logistic Regression, Random Forest, XGBoost, LightGBM and SVM, for the classification of Alzheimer’s disease by combining two modalities (sMRI and fMRI) and evaluated them to identify the optimal model that yields the most accurate final predictions.

Feed Forward Neural Network

The FFNN model achieves an overall accuracy of 90.09%, demonstrating good performance. For the CN class, it achieves 95.0% accuracy with high specificity (0.96), precision (0.93), recall (0.93) and an F1-score of 0.93. The EMCI class shows 92.0% accuracy, with specificity of 0.94, precision of 0.89, recall of 0.86 and an F1-score of 0.88. For the MCI class, the model attains 93.0% accuracy, specificity of 0.94, precision of 0.88, recall of 0.91 and an F1-score of 0.90. Overall, the FFNN model performs well across all classes, making it a suitable choice for classifying the dataset.

Logistic Regression

The Logistic Regression (LR) classification algorithm achieves an overall accuracy of 89.91%, indicating solid performance. For the CN class, it has 95.0% accuracy, with high specificity (0.96), precision (0.92), recall (0.93) and an F1-score of 0.93. In the EMCI class, the model reaches 91.0% accuracy, with specificity of 0.94, precision of 0.88, recall of 0.87 and an F1-score of 0.88. For the MCI class, the accuracy is 93.0%, with specificity of 0.95, precision of 0.90, recall of 0.89 and an F1-score of 0.90. Overall, the model performs well across all classes, making it suitable for classifying the dataset.

XGBoost

The XGBoost model achieves an overall accuracy of 94.15%, demonstrating excellent performance. For the CN class, it achieves 97.0% accuracy, with high specificity (0.99), precision (0.97) and a solid recall (0.94), resulting in an F1-score of 0.96. The EMCI class shows 95.0% accuracy, with good specificity (0.96), precision (0.92) and recall (0.95), leading to an F1-score of 0.93. In the MCI class, the model performs with 96.0% accuracy, high specificity (0.97) and balanced precision and recall (both 0.94), achieving an F1-score of 0.94. Overall, the model excels across all classes and is likely to generalize well to new data.

Light GBM

The Light GBM model achieves an overall accuracy of 94.72%, indicating strong performance with low error. For the CN class, it reaches 97.0% accuracy, with high specificity (0.99), precision (0.97), recall (0.94) and an F1-score of 0.96. The model performs similarly for the EMCI class, with 96.0% accuracy, a specificity of 0.96 and a precision of 0.92, achieving an F1-score of 0.94. For the MCI class, the model demonstrates 97.0% accuracy, with specificity of 0.98 and precision of 0.95, resulting in an F1-score of 0.95. Overall, the model performs consistently well across all classes and is expected to generalize effectively to new data.

SVM

The SVM model achieves an overall accuracy of 89.62%, indicating solid performance. For the CN class, it performs well with 95.0% accuracy, high specificity (0.96) and good precision (0.93), recall (0.93) and an F1-score of 0.95. In the EMCI class, the model has 91.0% accuracy, with specificity of 0.94, precision of 0.88, recall of 0.86 and an F1-score of 0.87. For the MCI class, the model achieves 92.0% accuracy, specificity of 0.94, precision of 0.88, recall of 0.89 and an F1-score of 0.89. Overall, the SVM model performs well across all classes and is suitable for classifying the dataset.

Table 6: Results of proposed Meta model and other classifiers

| Label | Metric | FFNN | Logistic Regression | XGBoost | Light GBM | SVM | Proposed Meta model (Random Forest) |
|------------------|-------------|-------|---------------------|---------|-----------|-------|-------------------------------------|
| CN | Accuracy | 95 | 95 | 97 | 97 | 95 | 98 |
| | Specificity | 0.96 | 0.96 | 0.99 | 0.99 | 0.96 | 0.99 |
| | Precision | 0.93 | 0.92 | 0.97 | 0.97 | 0.93 | 0.97 |
| | Recall | 0.93 | 0.93 | 0.94 | 0.94 | 0.93 | 0.95 |
| | F1-Score | 0.93 | 0.93 | 0.96 | 0.96 | 0.95 | 0.96 |
| EMCI | Accuracy | 92 | 91 | 95 | 96 | 91 | 96 |
| | Specificity | 0.94 | 0.94 | 0.96 | 0.96 | 0.94 | 0.97 |
| | Precision | 0.89 | 0.88 | 0.92 | 0.92 | 0.88 | 0.93 |
| | Recall | 0.86 | 0.87 | 0.95 | 0.96 | 0.86 | 0.96 |
| | F1-Score | 0.88 | 0.88 | 0.93 | 0.94 | 0.87 | 0.94 |
| MCI | Accuracy | 93 | 93 | 96 | 97 | 92 | 97 |
| | Specificity | 0.94 | 0.95 | 0.97 | 0.98 | 0.94 | 0.98 |
| | Precision | 0.88 | 0.9 | 0.94 | 0.95 | 0.88 | 0.95 |
| | Recall | 0.91 | 0.89 | 0.94 | 0.94 | 0.89 | 0.95 |
| | F1-Score | 0.9 | 0.9 | 0.94 | 0.95 | 0.89 | 0.95 |
| Overall accuracy | | 90.09 | 89.91 | 94.15 | 94.72 | 89.62 | 95.18 |

Random Forest

The Proposed Meta Model using Random Forest achieves an overall accuracy of 95.18%, indicating high performance and low error rates. For the CN class, it has an accuracy of 98.0%, with high specificity (0.99), precision (0.97) and recall (0.95), resulting in an F1-score of 0.96. The EMCI class shows 96.0% accuracy,

specificity of 0.97, precision of 0.93 and recall of 0.96, with an F1-score of 0.94. For the MCI class, the model achieves 97.0% accuracy, specificity of 0.98, precision of 0.95 and recall of 0.95, leading to an F1-score of 0.95. Overall, the model performs excellently across all classes as shown Table (6), suggesting it is well-suited for classification tasks and generalizes well to new data.

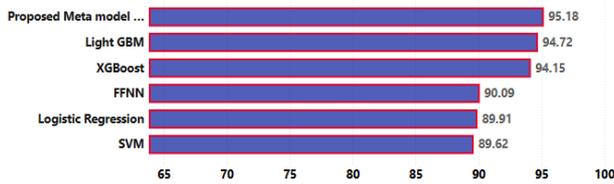


Fig. 10(a): Illustrates the performance comparison of various classifiers

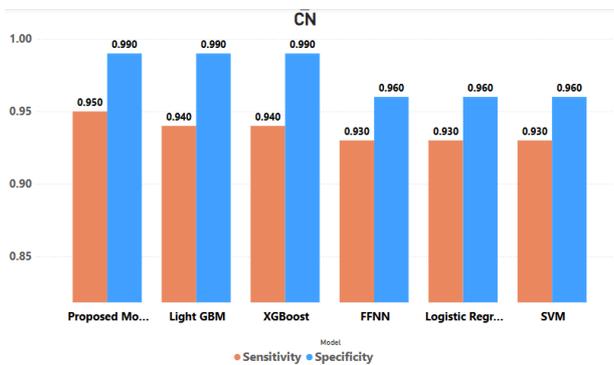


Fig. 10(b): Sensitivity and Specificity comparison of CN stage

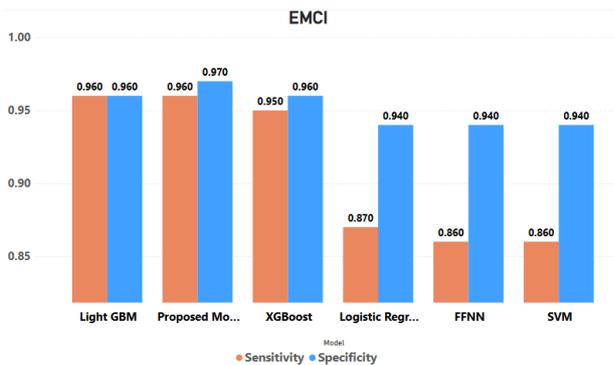


Fig. 10(c): Sensitivity and Specificity comparison of EMCI

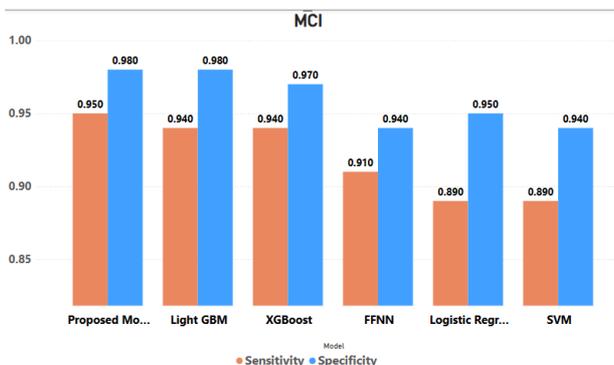


Fig. 10(d): Sensitivity and Specificity comparison of MCI

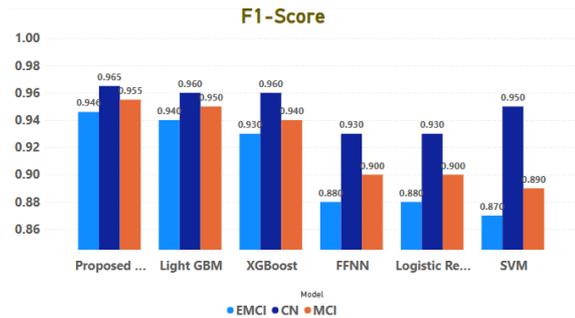


Fig. 10(e): F1-score comparison by model

Figure 10(a) illustrates the performance comparison of various meta classifiers, the proposed meta model with the Random Forest classifier leads with an accuracy of 95.18, demonstrating superior effectiveness. Light GBM and XGBoost follow closely with scores of 94.72 and 94.15, respectively, indicating strong performance. The Feed-Forward Neural Network (FFNN) achieves a score of 90.09, performing reasonably well but lagging behind the tree-based models. Traditional methods like Logistic Regression and SVM have the lowest scores, at 89.91 and 89.62, respectively, showing their limitations compared to more advanced techniques.

Figure 10(b) evaluates the sensitivity and specificity of different classifiers in classifying the AD CN stage. The proposed model stands out with a sensitivity of 0.95 and specificity of 0.99, showcasing outstanding performance. Light GBM and XGBoost also perform robustly, with sensitivity and specificity values of 0.94 and 0.99, respectively. The Feed-Forward Neural Network (FFNN), Logistic Regression and SVM show slightly lower performance, with sensitivity and specificity values of 0.93 and 0.96.

Figure 10(c) showcases the sensitivity and specificity of different classifiers for EMCI classification. The proposed model exhibits robust and consistent performance, achieving sensitivity and specificity values of 0.96 and 0.97, respectively. Light GBM and XGBoost also show competitive performance with similar metrics. In contrast, models such as Logistic

Regression, FFNN and SVM tend to underperform compared to these more advanced methods. Figure 10(d) evaluates the sensitivity and specificity of different classifiers for MCI classification. The proposed model, listed first, demonstrates strong performance with values of 0.95 and 0.98. Light GBM and XGBoost also show competitive and reliable results. In contrast, traditional models such as FFNN, Logistic Regression and SVM perform less effectively compared to the advanced techniques.

Figure 10(e) compares the F1-scores of different classifiers across EMCI, CN and MCI stages. The proposed model excels with F1-scores of 0.965, 0.946 and 0.955 for CN, EMCI and MCI, respectively, showcasing outstanding performance. Light GBM and

XGBoost also deliver strong and consistent results, closely trailing the proposed model. While FFNN performs reasonably well, it falls short compared to the top models. Logistic Regression and SVM achieve the lowest scores, reflecting their limited effectiveness in this context. The Proposed Random Forest model consistently surpasses other models such as FNN, Logistic Regression, XGBoost, Light GBM and SVM in terms of accuracy and various evaluation metrics.

K – fold Cross-Validation is used to evaluate the performance of a proposed model by dividing the dataset into *k* equal subsets (folds). The model is trained on *k – 1* folds and tested on the remaining fold, repeating the process *k* times so that each fold serves as the test set once. The final performance is determined by averaging the results across all folds, ensuring a more reliable evaluation. The results of 15-Fold Cross-Validation are presented in the Table (7), showing the accuracy obtained in each fold. The model demonstrates consistent performance across folds, with an overall average accuracy of 95.13%. Its strong performance demonstrates its ability to generalize effectively to unseen data, making it suitable choice for AD early prediction.

Table 7: Results of K-fold Cross validation

| Fold No | ACC | Fold No | ACC | Fold No | ACC |
|---------------|-------|---------|-------|---------|-------|
| 1 | 95.7 | 6 | 93.33 | 11 | 95.69 |
| 2 | 95.7 | 7 | 95.69 | 12 | 95.48 |
| 3 | 92.9 | 8 | 93.11 | 13 | 96.98 |
| 4 | 93.97 | 9 | 94.62 | 14 | 95.48 |
| 5 | 96.12 | 10 | 96.34 | 15 | 95.69 |
| Mean Accuracy | | | 95.13 | | |

ACC: Accuracy

Discussion

Alzheimer’s disease (AD) has no definitive cure, but early detection is critical for managing its progression. Accurate identification, particularly in very mild and mild cases, is challenging with a single modality like sMRI. Combining multi-modal imaging, such as sMRI for structural analysis and fMRI for functional analysis, provides a more comprehensive approach. This study utilized a dataset of 43,600 images, including both sMRI and fMRI data, divided into four subsets. Four deep learning models—GNN, U-Net, GoogLeNet and DenseNet-121—were developed using transfer learning. The prediction probabilities from these models were combined through vertical stacking to create a feature matrix, which was used to train a meta-model. This meta-model effectively reduced overfitting issues associated with single-modality models, offering more accurate and reliable results, particularly for early detection. Figure (11) compares deep learning techniques using single modality with a stacking-based ensemble approach with multi-modality. The multi-modality approach achieves an average accuracy of 92.27%,

significantly higher than the single modality's 89.29%. Additionally, the multi-modality method shows superior average specificity (93.98%) and sensitivity (88.77%) compared to the single modality approach. This highlights that multimodal ensemble techniques are more effective for achieving high accuracy and reliability.

Figure (12) shows a comparison of various meta-classifiers based on their confusion matrices and reveals that the proposed Random Forest (RF) model achieves the highest accuracy, correctly classifying 1,661 out of 1,745 cases with only 84 misclassifications, demonstrating superior reliability and robustness. FFNN and XGBoost also perform well, particularly in identifying CN and EMCI stages, though they exhibit some confusion between EMCI and MCI. Logistic Regression (LR) and SVM maintain strong accuracy for CN but encounter challenges in clearly distinguishing EMCI from MCI. Light GBM, although effective in detecting CN and EMCI, struggles with MCI classification due to notable misclassifications.

Figure (13) presents the AUROC analysis, highlighting the superior performance of the proposed system, which achieves high scores across all classes (CN: 0.97, EMCI: 0.96, MCI: 0.96), indicating strong and balanced classification capabilities. Light GBM closely matches this performance, while XGBoost also performs well with slightly lower AUC values. In contrast, Logistic Regression, SVM and FFNN exhibit moderate effectiveness, particularly struggling with EMCI classification, which consistently yields lower AUROC scores across these models. This suggests that EMCI is the most challenging stage to distinguish. Overall, the proposed model demonstrates higher reliability and precision, especially in addressing the complexities of early stage cognitive impairment.

Figure (14) shows the Precision-Recall curve of proposed system, which demonstrates strong performance across all three classes: CN, EMCI and MCI, with Average Precision (AP) scores of 0.94, 0.91 and 0.92 respectively. The model maintains high precision and recall, indicating it correctly identifies most cases with few false positives or negatives. Cognitively Normal (CN) shows the highest AP, suggesting it is the easiest for the model to classify accurately. EMCI has the lowest AP, which is expected due to its subtle symptoms making it harder to detect.

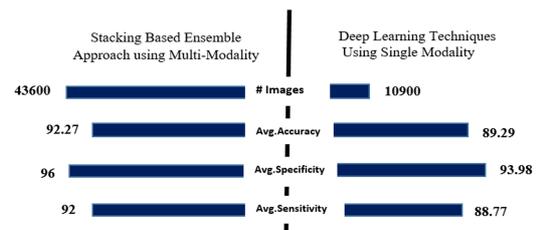


Fig. 11: Performance Comparison between single modality models with multi-modality models

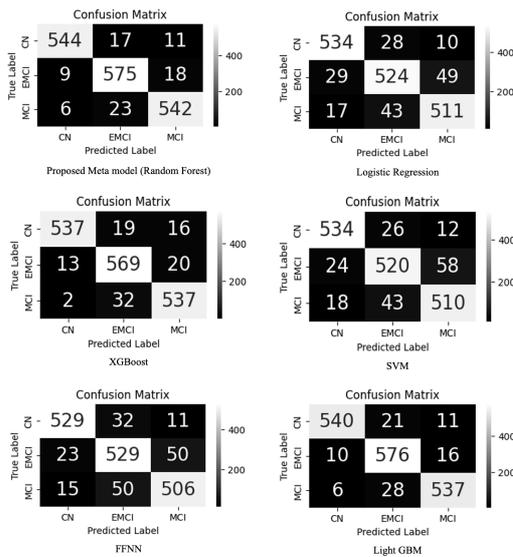


Fig. 12: Confusion matrix of proposed system and other classifiers

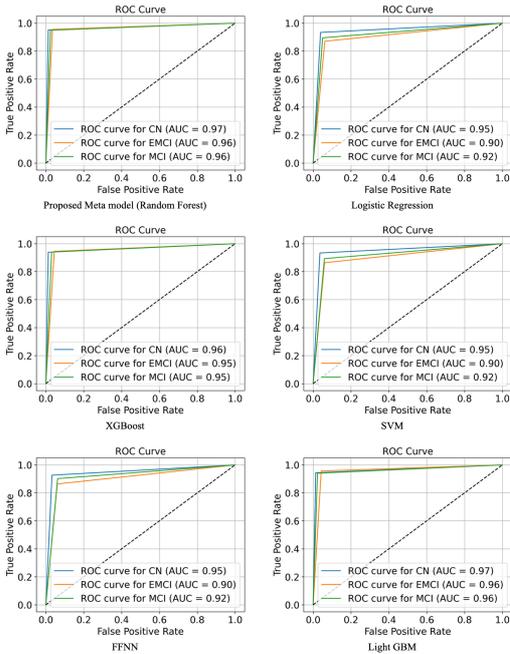


Fig. 13: Comparison of Proposed System AUROC Curve with other classifier

A sharp drop at high recall values is observed, which is common when the model starts to misclassify more to achieve full recall.

Figure (15) shows the stage-wise confusion matrices for the proposed system. These matrices demonstrate the robustness of the model in classifying CN, EMCI and MCI stages. For the CN class, the model shows excellent precision and recall, indicating a strong capability to distinguish normal cases with minimal error. The EMCI classification also performs well, although a slightly higher number of false negatives suggests room for

improved sensitivity. The MCI results show balanced performance, with consistent true and false prediction counts, reflecting reliable detection. Overall, the model exhibits high classification accuracy across all stages, with further refinement potentially enhancing EMCI detection. Figure (16) shows comparison of proposed method with state-of-the-art techniques and it is revealed its superior accuracy across all categories (CN, EMCI, MCI), highlighting its potential for improving treatment outcomes.

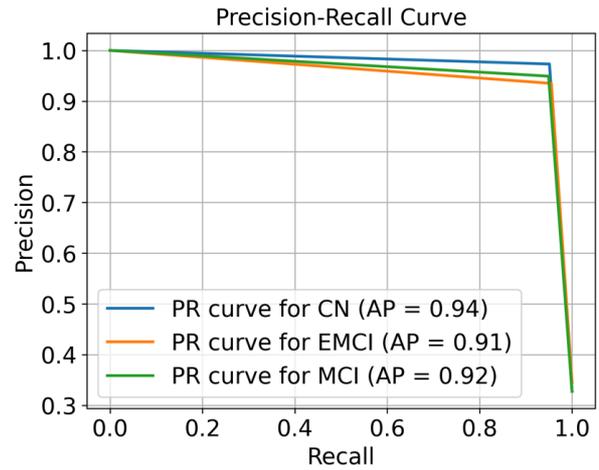


Fig. 14: Precision-Recall Curve of proposed model

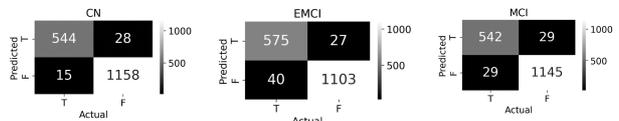


Fig. 15: Stage-wise confusion matrix of proposed meta model

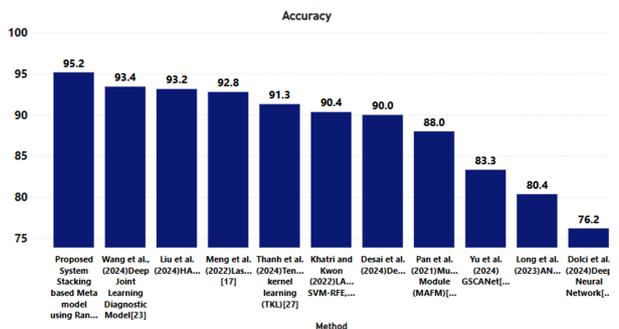


Fig. 16: Comparing proposed method with state-art-techniques

Conclusion

This study introduces a novel ensemble method based on stacking for early Alzheimer's disease diagnosis using multimodal neuroimaging data. By combining predictions from deep learning models trained on both sMRI and fMRI data, the proposed approach capitalizes on the complementary strengths of these different modalities. The meta-model achieved an impressive accuracy of 95.18%, surpassing current state-of-the-art

methods. This reliable and precise system has significant potential to enhance early AD detection, enable timely interventions and advance research in neurodegenerative diseases. Future improvements could include incorporating additional modalities and geographical data to better predict AD progression. Additionally, vision transformer techniques can be used with large datasets to enhance the proposed system.

Acknowledgment

We extend our sincere thanks to the publisher for their support in the publication of this research article. We are grateful for the resources and platform they provided, which allowed us to share our findings with a broader audience. We also appreciate the editorial team's efforts in reviewing and refining our work and we are thankful for the opportunity to contribute to the advancement of research through this publication.

Funding Information

The authors have not received any financial support or funding to report.

Author's Contributions

Niranjana Kumar Parvatham: The research scholar, Data collection, processing, core implementation, experimentation and analysis of the proposed methodology.

Lakshmana Phaneendra Maguluri: The supervisor, provided overall guidance, technical direction and critical revisions throughout the research.

Ethics

This article is original and contains unpublished material. The corresponding author confirms that all of the other authors have read and approved the manuscript and no ethical issues involved.

References

Acharya, H., Mehta, R., & Kumar Singh, D. (2021). Alzheimer Disease Classification Using Transfer Learning. *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, 1503-1508.
<https://doi.org/10.1109/iccmc51019.2021.9418294>

Aghdam, M. A., Bozdog, S., & Saeed, F. (2023). Pvtad: Alzheimer's Disease Diagnosis Using Pyramid Vision Transformer Applied to White Matter of T1-Weighted Structural Mri Data. *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, 1-4.
<https://doi.org/10.1109/ISBI56570.2024.10635541>

Alzheimers & Dementia. (2023). 2023 Alzheimer's disease facts and figures. *Alzheimers & Dementia*, 19(4), 1598-1695.
<https://doi.org/10.1002/alz.13016>

Asiri, A. A., Shaf, A., Ali, T., Aamir, M., Irfan, M., Alqahtani, S., Mehdar, K. M., Halawani, H. T., Alghamdi, A. H., Alshamrani, A. F. A., & Alqhtani, S. M. (2023). Brain Tumor Detection and Classification Using Fine-Tuned CNN with ResNet50 and U-Net Model: A Study on TCGA-LGG and TCIA Dataset for MRI Applications. *Life*, 13(7), 1449.
<https://doi.org/10.3390/life13071449>

Bhosale, T., Gulame, M., Shendkar, B., Kadam, P., More, R., & Mali, R. (2023). Alzheimer's Disease MRI Image Segmentation Based on the Enhanced U-Net. *2023 IEEE International Conference on ICT in Business Industry & Government (ICTBIG)*, 1-5.
<https://doi.org/10.1109/ictbig59752.2023.10455753>

Castellano, G., Esposito, A., Lella, E., Montanaro, G., & Vessio, G. (2024). Automated detection of Alzheimer's disease: a multi-modal approach with 3D MRI and amyloid PET. *Scientific Reports*, 14(1), 5210.
<https://doi.org/10.1038/s41598-024-56001-9>

Chen, D., Yi, F., Qin, Y., Zhang, J., Ge, X., Han, H., Cui, J., Bai, W., Wu, Y., & Yu, H. (2022). A Stacking Framework for Multi-Classification of Alzheimer's Disease Using Neuroimaging and Clinical Features. *Journal of Alzheimer's Disease*, 87(4), 1627-1636. <https://doi.org/10.3233/jad-215654>

Choudhury, C., Goel, T., & Tanveer, M. (2024). A coupled-GAN architecture to fuse MRI and PET image features for multi-stage classification of Alzheimer's disease. *Information Fusion*, 109, 102415.
<https://doi.org/10.1016/j.inffus.2024.102415>

Ciurea, V. A., Covache-Busuioc, R.-A., Mohan, A. G., Costin, H. P., & Voicu, V. (2023). Alzheimer's disease: 120 years of research and progress. *Journal of Medicine and Life*, 16(2), 173-177.
<https://doi.org/10.25122/jml-2022-0111>

Desai, M. B., Kumar, Y., & Pandey, S. (2024). Efficient Approach for Diagnosis and Detection of Alzheimer Diseases Using Deep Learning. *2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET)*, 1-5.
<https://doi.org/10.1109/acroset62108.2024.10743886>

Dolci, G., Ellis, C. A., Cruciani, F., Brusini, L., Abrol, A., Galazzo, I. B., Menegaz, G., & Calhoun, V. D. (2025). Multimodal MRI accurately identifies amyloid status in unbalanced cohorts in Alzheimer's disease continuum. *Network Neuroscience*, 9(1), 259-279.
https://doi.org/10.1162/netn_a_00423

Duenias, D., Nichyporuk, B., Arbel, T., & Riklin Raviv, T. (2025). Hyperfusion: A hypernetwork approach to multimodal integration of tabular and medical imaging data for predictive modeling. *Medical Image Analysis*, 102, 103503.
<https://doi.org/10.1016/j.media.2025.103503>

- Fujita, Y., Tanaka, T., Hori, T., & Hamamoto, Y. (2023). Classification Model Based on U-Net for Crack Detection from Asphalt Pavement Images. *Journal of Image and Graphics, 11*(2), 121-126. <https://doi.org/10.18178/joig.11.2.121-126>
- Gao, J., Liu, J., Xu, Y., Peng, D., & Wang, Z. (2023). Brain age prediction using the graph neural network based on resting-state functional MRI in Alzheimer's disease. *Frontiers in Neuroscience, 17*, 1222751. <https://doi.org/10.3389/fnins.2023.1222751>
- Gunawan, D., Zuama, R. A., & Ghani, M. A. (2024). Analysis of Machine Learning Algorithms for Early Detection of Alzheimer's Disease: A Comparative Study. *Journal of Artificial Intelligence and Engineering Applications (JAIEA), 3*(3), 898-902. <https://doi.org/10.59934/jaiea.v3i3.579>
- Gupta, Y., Kim, J.-I., Kim, B. C., & Kwon, G.-R. (2020). Classification and Graphical Analysis of Alzheimer's Disease and Its Prodromal Stage Using Multimodal Features From Structural, Diffusion, and Functional Neuroimaging Data and the APOE Genotype. *Frontiers in Aging Neuroscience, 12*, 238. <https://doi.org/10.3389/fnagi.2020.00238>
- Han, S., Sun, Z., Zhao, K., Duan, F., Caiafa, C. F., Zhang, Y., & Solé-Casals, J. (2024). Early prediction of dementia using fMRI data with a graph convolutional network approach. *Journal of Neural Engineering, 21*(1), 016013. <https://doi.org/10.1088/1741-2552/ad1e22>
- Hazarika, R. A., Maji, A. K., Kandar, D., Jasinska, E., Krejci, P., Leonowicz, Z., & Jasinski, M. (2023). An Approach for Classification of Alzheimer's Disease Using Deep Neural Network and Brain Magnetic Resonance Imaging (MRI). *Electronics, 12*(3), 676. <https://doi.org/10.3390/electronics12030676>
- Jiao, C.-N., Gao, Y.-L., Ge, D.-H., Shang, J., & Liu, J.-X. (2024). Multi-modal imaging genetics data fusion by deep auto-encoder and self-representation network for Alzheimer's disease diagnosis and biomarkers extraction. *Engineering Applications of Artificial Intelligence, 130*, 107782. <https://doi.org/10.1016/j.engappai.2023.107782>
- Kamal, M. S., & Farhana Nimmy, S. (2024). Interpretable Transformers for Alzheimer Disease Diagnosis on Multi-modal Data. *2024 International Joint Conference on Neural Networks (IJCNN), 1-8*. <https://doi.org/10.1109/ijcnn60899.2024.10651416>
- Khan, A. A., Mahendran, R. K., Perumal, K., & Faheem, M. (2024). Dual-3DM3AD: Mixed Transformer Based Semantic Segmentation and Triplet Pre-Processing for Early Multi-Class Alzheimer's Diagnosis. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 32*, 696-707. <https://doi.org/10.1109/tnsre.2024.3357723>
- Khatri, U., & Kwon, G.-R. (2022). Alzheimer's Disease Diagnosis and Biomarker Analysis Using Resting-State Functional MRI Functional Brain Network With Multi-Measures Features and Hippocampal Subfield and Amygdala Volume of Structural MRI. *Frontiers in Aging Neuroscience, 14*, 818871. <https://doi.org/10.3389/fnagi.2022.818871>
- Kumar, P. N., & Maguluri, L. P. (2024). Deep Learning-based Classification of MRI Images for Early Detection and Staging of Alzheimer's Disease. *International Journal of Advanced Computer Science and Applications, 15*(5). <https://doi.org/10.14569/ijacsa.2024.0150545>
- Liu, X., Li, W., Miao, S., Liu, F., Han, K., & Bezabih, T. T. (2024). HAMMF: Hierarchical attention-based multi-task and multi-modal fusion model for computer-aided diagnosis of Alzheimer's disease. *Computers in Biology and Medicine, 176*, 108564. <https://doi.org/10.1016/j.compbiomed.2024.108564>
- Long, Z., Li, J., Fan, J., Li, B., Du, Y., Qiu, S., Miao, J., Chen, J., Yin, J., & Jing, B. (2023). Identifying Alzheimer's disease and mild cognitive impairment with atlas-based multi-modal metrics. *Frontiers in Aging Neuroscience, 15*, 1212275. <https://doi.org/10.3389/fnagi.2023.1212275>
- Mahmud Joy, Md. A., Nasrin, S., Siddiqua, A., & Farid, D. Md. (2025). ViTAD: Leveraging modified vision transformer for Alzheimer's disease multi-stage classification from brain MRI scans. *Brain Research, 1847*, 149302. <https://doi.org/10.1016/j.brainres.2024.149302>
- Massalimova, A., & Varol, H. A. (2021). Input Agnostic Deep Learning for Alzheimer's Disease Classification Using Multimodal MRI Images. *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2875-2878*. <https://doi.org/10.1109/embc46164.2021.9629807>
- Meng, X., Liu, J., Fan, X., Bian, C., Wei, Q., Wang, Z., Liu, W., & Jiao, Z. (2022). Multi-Modal Neuroimaging Neural Network-Based Feature Detection for Diagnosis of Alzheimer's Disease. *Frontiers in Aging Neuroscience, 14*, 911220. <https://doi.org/10.3389/fnagi.2022.911220>
- Mora-Rubio, A., Bravo-Ortiz, M. A., Quiñones Arredondo, S., Saborit Torres, J. M., Ruz, G. A., & Tabares-Soto, R. (2023). Classification of Alzheimer's disease stages from magnetic resonance images using deep learning. *PeerJ Computer Science, 9*, e1490. <https://doi.org/10.7717/peerj-cs.1490>
- Oduşami, M., Damaševičius, R., Milieškaitė-Belousovienė, E., & Maskeliūnas, R. (2024). Alzheimer's disease stage recognition from MRI and PET imaging data using Pareto-optimal quantum dynamic optimization. *Heliyon, 10*(15), e34402. <https://doi.org/10.1016/j.heliyon.2024.e34402>

- Odusami, M., Damaševičius, R., Milieškaitė-Belousovienė, E., & Maskeliūnas, R. (2024). Multimodal Neuroimaging Fusion for Alzheimer's Disease: An Image Colorization Approach With Mobile Vision Transformer. *International Journal of Imaging Systems and Technology*, 34(5), e23158. <https://doi.org/10.1002/ima.23158>
- Oh, K., Chung, Y.-C., Kim, K. W., Kim, W.-S., & Oh, I.-S. (2019). Classification and Visualization of Alzheimer's Disease using Volumetric Convolutional Neural Network and Transfer Learning. *Scientific Reports*, 9(1), 18150. <https://doi.org/10.1038/s41598-019-54548-6>
- Pamarapa, C., Keeratitivattayut, R., Ekjeen, T., Shoombuatong, W., & Vichianin, Y. (2024). Multimodal SVM Classification for Early-stage Alzheimer's Disease Diagnosis Using T1-weighted MR and F-18 FDG PET Imaging. *Journal of Medical Imaging and Radiation Sciences*, 55(3), 101539. <https://doi.org/10.1016/j.jmir.2024.101539>
- Pan, Q., Ding, K., & Chen, D. (2021). Multi-Classification Prediction of Alzheimer's Disease based on Fusing Multi-modal Features. *2021 IEEE International Conference on Data Mining (ICDM)*, 1270-1275. <https://doi.org/10.1109/icdm51629.2021.00156>
- Pang, J., Zhang, Y., Jiang, H., Li, M., Wang, L., Yu, P., Liu, J., & Hu, L. (2023). An Exploration of the accuracy of TOF Automatic Segmentation Algorithm based on U-net model. *2023 16th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, 1-6. <https://doi.org/10.1109/cisp-bmei60920.2023.10373230>
- Parvatham, N. K., & Maguluri, L. P. (2024). Cycle GAN-Based MRI Augmentation and Hybrid Deep Learning Approach (GDP) for Binary Classification of Alzheimer's Disease: Differentiating Normal and Very Mild Dementia. *International Journal of Intelligent Engineering and Systems*, 17(6), 1281-1299. <https://doi.org/10.22266/ijies2024.1231.93>
- Patra, B., Maity, N. P., & Charan Sutar, B. (2023). Early Detection of Alzheimer's Disease using Feed Forward Neural Network. *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1-4. <https://doi.org/10.1109/icccnt56998.2023.10307151>
- Rama Lakshmi, B., & Radhika, Y. (2024). Resnet-53 for Alzheimer's Disease Detection from MRI Images and Analysis with SVM Tuning with Hyper Optimization Technique. *2024 4th International Conference on Sustainable Expert Systems (ICSES)*, 1065-1072. <https://doi.org/10.1109/icses63445.2024.10763183>
- Saxena, S., Prasad, S. N., & Murthy T S, D. (2023). Utilizing deep learning techniques to diagnose nodules in lung computed tomography (CT) scan images. *IAENG International Journal of Computer Science*, 50(2), 537-552.
- Shah, S. M. A. H., Khan, M. Q., Rizwan, A., Jan, S. U., Samee, N. A., & Jamjoom, M. M. (2024). Computer-aided diagnosis of Alzheimer's disease and neurocognitive disorders with multimodal Bi-Vision Transformer (BiViT). *Pattern Analysis and Applications*, 27(3), 76. <https://doi.org/10.1007/s10044-024-01297-6>
- Solano-Rojas, B., Villalón-Fonseca, R., & Marín-Raventós, G. (2020). Alzheimer's Disease Early Detection Using a Low Cost Three-Dimensional Densenet-121 Architecture. *The Impact of Digital Technologies on Public Health in Developed and Developing Countries*, 3-15. https://doi.org/10.1007/978-3-030-51517-1_1
- Talha, A., Dhanasree, Ch., Divya, E., Prabhas, K. S., & Syed Abudhagir, U. (2024). Performance Evaluation of Deep Learning Models for Alzheimer's Disease Detection. *2024 10th International Conference on Communication and Signal Processing (ICCSPP)*, 317-322. <https://doi.org/10.1109/iccsp60870.2024.10543787>
- Thanh, V. D., Trung Le, T., Tuan, P. M., Linh Trung, N., Abed-Meraim, K., Adel, M., Dung, N. V., Thanh Trung, N., Long, D. D., & Chén, O. Y. (2024). Tensor Kernel Learning for Classification of Alzheimer's Conditions using Multimodal Data. *2024 International Conference on Multimedia Analysis and Pattern Recognition (MAPR)*, 1-6. <https://doi.org/10.1109/mapr63514.2024.10661014>
- Tong, W., Li, Y.-X., Zhao, X.-Y., Chen, Q.-Q., Gao, Y.-B., Li, P., & Wu, E. Q. (2023). fMRI-Based Brain Disease Diagnosis: A Graph Network Approach. *IEEE Transactions on Medical Robotics and Bionics*, 5(2), 312-322. <https://doi.org/10.1109/tmrb.2023.3270481>
- Vaithianathan, K., Pernabas, J. B., Parthiban, L., Rashid, M., & Alshamrani, S. S. (2024). Normalized group activations based feature extraction technique using heterogeneous data for Alzheimer's disease classification. *PeerJ Computer Science*, 10, e2502. <https://doi.org/10.7717/peerj-cs.2502>
- Vergara, V. M., Abrol, A., Espinoza, F. A., & Calhoun, V. D. (2019). Selection of Efficient Clustering Index to Estimate the Number of Dynamic Brain States from Functional Network Connectivity. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. 2019 41st Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Berlin, Germany. <https://doi.org/10.1109/embc.2019.8856284>

- Wang, J., Wen, S., Liu, W., Meng, X., & Jiao, Z. (2024). Deep joint learning diagnosis of Alzheimer's disease based on multimodal feature fusion. *BioData Mining*, 17(1), 48.
<https://doi.org/10.1186/s13040-024-00395-9>
- Wang, Y., Liu, X., & Yu, C. (2021). Assisted Diagnosis of Alzheimer's Disease Based on Deep Learning and Multimodal Feature Fusion. *Complexity*, 2021(1), 6626728.
<https://doi.org/10.1155/2021/6626728>
- Weiner, M. W., Veitch, D. P., Aisen, P. S., Beckett, L. A., Cairns, N. J., Green, R. C., Harvey, D., Jack, C. R., Jagust, W., Morris, J. C., Petersen, R. C., Salazar, J., Saykin, A. J., Shaw, L. M., Toga, A. W., & Trojanowski, J. Q. (2017). The Alzheimer's Disease Neuroimaging Initiative 3: Continued innovation for clinical trial improvement. *Alzheimer's & Dementia*, 13(5), 561-571.
<https://doi.org/10.1016/j.jalz.2016.10.006>
- Xiao, J., Li, J., Wang, J., Zhang, X., Wang, C., Peng, G., Hu, H., Liu, H., Liu, J., Shen, L., Zhang, N., Yan, N., Ma, Q., Xu, W., Liao, Z., Ren, R., Wang, M., Yu, E., Tian, J., ... Wang, G. (2023). 2023 China Alzheimer's disease: facts and figures. *Human Brain*, 2(3), 1-13.
<https://doi.org/10.37819/hb.3.1771>
- Yu, X., Liu, J., Lu, Y., Funahashi, S., Murai, T., Wu, J., Li, Q., & Zhang, Z. (2024). Early diagnosis of Alzheimer's disease using a group self-calibrated coordinate attention network based on multimodal MRI. *Scientific Reports*, 14(1), 24210.
<https://doi.org/10.1038/s41598-024-74508-z>
- Zhang, Y., He, X., Chan, Y. H., Teng, Q., & Rajapakse, J. C. (2023). Multi-modal graph neural network for early diagnosis of Alzheimer's disease from sMRI and PET scans. *Computers in Biology and Medicine*, 164, 107328.
<https://doi.org/10.1016/j.compbiomed.2023.107328>
- Zhou, Z., Wang, Q., An, X., Chen, S., Sun, Y., Wang, G., & Yan, G. (2024). A novel graph neural network method for Alzheimer's disease classification. *Computers in Biology and Medicine*, 180, 108869.
<https://doi.org/10.1016/j.compbiomed.2024.108869>
- Zia-ur-Rehman, Awang, M. K., Rashid, J., Ali, G., Hamid, M., Mahmoud, S. F., Saleh, D. I., & Ahmad, H. I. (2024). Classification of Alzheimer disease using DenseNet-201 based on deep transfer learning technique. *PLOS ONE*, 19(9), e0304995.
<https://doi.org/10.1371/journal.pone.0304995>