

ISSN 1840-4855

e-ISSN 2233-0046

Original scientific article

<http://dx.doi.org/10.70102/afts.2026.1836.212>

MULTIMODAL MRI ANALYSIS FOR EARLY-STAGE ALZHEIMER'S DETECTION USING ATTENTION MECHANISMS

Ratnakala Patil¹, Dr. Sachinkumar Veerashetty^{2*}

¹Assistant Professor, Department of Computer Science and Engineering, Sharnbasva University, Kalaburagi, India. e-mail: ratnakala@sharnbasvauniversity.edu.in,
orcid: <https://orcid.org/0009-0006-9553-4487>

^{2*}Professor, Department of Computer Science and Engineering, Sharnbasva University, Kalaburagi, India. e-mail: sveerashetty@sharnbasvauniversity.edu.in,
orcid: <https://orcid.org/0000-0001-8217-1388>

Received: May 04, 2026; Revised: June 15, 2026; Accepted: July 23, 2026; Published: September 29, 2026

SUMMARY

Alzheimer's disease (AD) is a progressive neurodegenerative disease with an increasingly common prevalence all around the world, necessitating the development of an early, accurate, and minimally invasive diagnosis method for such a devastating condition. However, traditional diagnostic methods such as neuropsychological testing and PET scans may have several shortcomings when used in isolation, namely their high cost, invasive nature, or inability to detect early changes in the brain structure. Therefore, this study presents a new deep learning method that utilizes attention to integrate sMRI, fMRI, and DTI data from early-stage AD patients to achieve better results. The use of attention allows the network to concentrate on specific areas of interest within each input mode and select the most valuable modality-specific features. Several extensive experiments conducted on ADNI datasets suggest that the suggested model can produce an accuracy of 97.63%, precision of 97.47%, recall of 97.77%, F1-Score of 97.60%, specificity of 97.90%, and 0.96 MCC, thus, beating the traditional CNN-based, transfer learning models, and ensemble-based methods. The per-class sensitivity and specificity values are consistently high at an average of 0.961 and 0.959, respectively. It suggests that the suggested framework has a good potential to detect Alzheimer's disease in its early stages, including mild cognitive impairment and cognition-normal states. Further experiments performed by applying ablation analysis suggest that the inclusion of spatial and channel attentions, along with multi-modal fusion, can contribute significantly to performance.

Key words: *alzheimer's disease, multimodal MRI, deep learning, attention mechanism, early diagnosis, medical image analysis, brain imaging.*

INTRODUCTION

One of the brain disorders that worsens over time is Alzheimer's disease (AD). Memory, reasoning, and carrying out regular duties are all impacted. The most prevalent cause of dementia imposes a burden on the health systems of the world [1]. The World Health Organization states that millions of people around the world have dementia, and most of them have Alzheimer's disease [26][3]. Life expectancy is rising, indicating that the prevalence of Alzheimer's disease will increase, necessitating prompt and precise

study [27]. Amyloid-B plaque buildup, twisted tau protein, brain cell death, and brain degeneration are all linked to AD [5]. Before any noticeable symptoms appear, these changes could occur over a number of years. Early diagnosis of the disorder is important since it is easier to treat when an individual has mild cognitive impairment or is in the preclinical stage [13]. Nonetheless, traditional diagnosis approaches like as neuropsychological testing, spinal fluid tests, and PET scans may be intrusive, expensive, or ineffective in detecting small early brain alterations [7].

Magnetic Resonance Imaging (MRI) has become an easy and non-invasive method to observe changes in the brain in Alzheimer disease [4][8].

The structural MRI evaluates the degeneration of the brain's outer layer and the hippocampus, whereas the functional MRI examines changes in communication between different parts of the brain [6]. Diffusion Tensor imaging measures the health of the brain's white matter. Based on the largest one the AD should be initiative, in this collects the various data used by the image modalities along with to create the models for to diagnose the disease [9][10]. The current methods as CNN and attention mechanisms are used for to improve the disease prediction for brain scan for to improve the accuracy based on the findings [11][12]. Nonetheless, the majority of existing research exclusively examine a singular form of imaging data, which may inadequately represent the alterations occurring in early Alzheimer's disease [13].

To solve this, the proposed study suggests a deep learning framework that incorporates multiple forms of MRI scans combined with difficult attention mechanisms to enhance the detection of the initial progression stages of Alzheimer disease [14][15].

Such a system will be integrated to combine information on structural, functional, as well as diffusion MRI to produce more detailed features, which will result in improved classification and may also have a clearer perspective of the disease. The method can enable physicians to make quicker and more precise judgments that eventually results in the enhancement of the care and treatment of the patients [16][17].

Key Contributions of the Research

- Deep learning system is a combination of sMRI, fMRI and DTI to diagnose Alzheimer earlier than single-modality systems, as it is a self-governing approach that concentrates on the important portions of the brain.
- By relying on spatial and channel-wise attention mechanisms, the study enhances the level of diagnostic reliability and model interpretability.
- The suggested model is tested on the publically available datasets such as the Alzheimer Disease Neuroimaging Initiative, which showed improved accuracy, sensitivity, and strengths both in terms of imaging procedures and scanners.

The literature study in the second section will evaluate existing studies on identifying Alzheimer disease using MRI, a deep learning model, and attention mechanisms, as well as explain research gaps. Section 3 provides an account of the proposed methodology, which comprises the preprocessing of the multimodal MRI data, designing the model architecture, incorporation of attention mechanisms, and training processes. Section 4 describes the experimental results, evaluation metrics of the performance and compares the results with existing techniques. Section 5 has an in-depth discussion of the results to bring out the relevance, interpretability, and generalizability of the proposed framework. Lastly, section 6 is a summary of the major findings of the study as well as possible paths of future research as well as therapeutic implications.

LITERATURE REVIEW

ML and DL models are used in AD for to classifications. CNN and other models are used for to analysis and innovative process for utilize the data should be explored. Standard data sets, data preparation, model testing, and model performance evaluation were also covered. The review showed that balancing the data is a big issue, along with other issues including overfitting and models that are hard to

understand. Its also emphasized the need for explainable AI and attention-focused models to improve diagnosis reliability and build clinical activity confidence [18].

Deep learning methods for MRI-based Alzheimer detection. To contrasted ensemble learning models, transfer learning, and connection-wise attention. In this discovered that superior deep learning models outperformed traditional machine learning methods in terms of categorization. Furthermore, it is pointed out that more sophisticated deep models are needed to improve early usage and reduce errors in clinical utilization, as well as automatic feature extraction and powerful validation algorithms [19].

The ML model and DL model is having the specific application through AD for to identify the using MRI data. These included feature extraction techniques including texture and volume analysis as well as preprocessing techniques like segmentation, skull stripping, and normalizing. To compared various classifiers, including support vector machines and deep CNN. A fully automated deep learning pipeline is crucial for early diagnosis, however there are challenges like MRI noise, scanner variations, and a lack of labelled data [20].

The analysis for using the various models such as MRI and PET for to predict the AD investigated further machine and deep learning techniques and technologies, including multimodal fusion and transfer learning. The clinical importance of predicting MCI progression has been highlighted in the publication, along with research limitations including cross-data validation and generalizability. Their findings demonstrated that the use of different imaging biomarkers greatly improves prediction accuracy [21].

A maximum accuracy of 99.3%, their approach outperformed the traditional models. The article proved that transfer learning can be effective with little datasets, and that optimizing CNN architecture can considerably improve diagnosis accuracy in medical imaging [22]. The system was designed using (CBAM) and (MGCA), and its features are better represented and disease-specific areas [23].

Table 1. Comparative analysis of recent deep learning approaches for alzheimer's disease detection

Authors	Methodology	Dataset / Simulation Environment	Research Gap
[18]	ML, DL, CNN and Attention Mechanisms.	Multiple public neuroimaging datasets (e.g., ADNI and related MRI datasets)	Limited interpretability, data imbalance issues, and lack of standardized validation frameworks
[19]	Systematic review of DL models such as ensemble learning, transfer learning, and attention-based CNNs	MRI-based Alzheimer's datasets from recent DL studies	Insufficient cross-dataset validation and limited focus on early-stage detection robustness
[20]	Comparative review of ML and DL approaches including preprocessing, segmentation, feature extraction, and classification techniques	MRI datasets used in prior AD detection studies	Challenges in handling MRI noise, scanner variability, and small labeled datasets
[21]	Review of ML/DL models for predicting using multimodal fusion	MRI and PET datasets (e.g., ADNI)	Limited generalizability and need for stronger multimodal integration frameworks
[22]	Pre-trained CNN-based deep learning framework (transfer learning approach)	Kaggle MRI datasets for AD classification	Limited validation on large clinical datasets and lack of multimodal data integration
[23]	Multimodal deep learning with CBAM	ADNI and OASIS datasets	Computational complexity and need for broader real-world clinical validation
[24]	Fuzzy Optimized Attention Network	Structural MRI (sMRI) datasets	Limited exploration of multimodal fusion and scalability across diverse populations
[25]	Review of multimodal fusion and cross-attention mechanisms combining DTI, rs-fMRI, and PET	Multimodal neuroimaging datasets (MRI, DTI, PET)	Need for unified end-to-end multimodal frameworks with improved interpretability and clinical deployment readiness

The model had more emphasis on key areas of the brain image, which enhanced sensitivity and specificity in identifying both Alzheimer and mild impaired cognitive. The articles indicated that what is known as optimized attention systems can significantly enhance the accuracy of classification and false negative in early diagnosis [24]. A review of multimodal fusion techniques to identify Alzheimer disease, in particular, diffusion tensor imaging (DTI), in conjunction with resting-state fMRI and PET data. It emphasized the advantages of cross-attention processes in the combination of features of different types of imaging. The review has determined that the integration of various imaging modalities is an excellent way to enhance the classification and generalization, and thus a promising system to create powerful and scalable diagnostic systems of the Alzheimer disease [25].

Table 1 defined as the various studies related to ML and DL with contains the different algorithmic model for AD. This algorithm is used to defined the various model such as CNN and TL, Mutli model fusion compared to base line models.it provides a summary of MRI, PET, sMRI, ADNI, OASIS and simulation settings. It was observed that it defines the key research gaps in the compendium of existing literature namely low interpretability, absence of cross-dataset validation, poor multimodal integration, data asymmetries, and computational complexity. This paper has made it clear that a more robust multimodal theory that draws attention should be used. It will be easier to identify early signs of Alzheimer disease at a higher success rate, in a more generalized, and clinically applicable manner.

METHODOLOGY

Overall Architecture for the Proposed Methodology

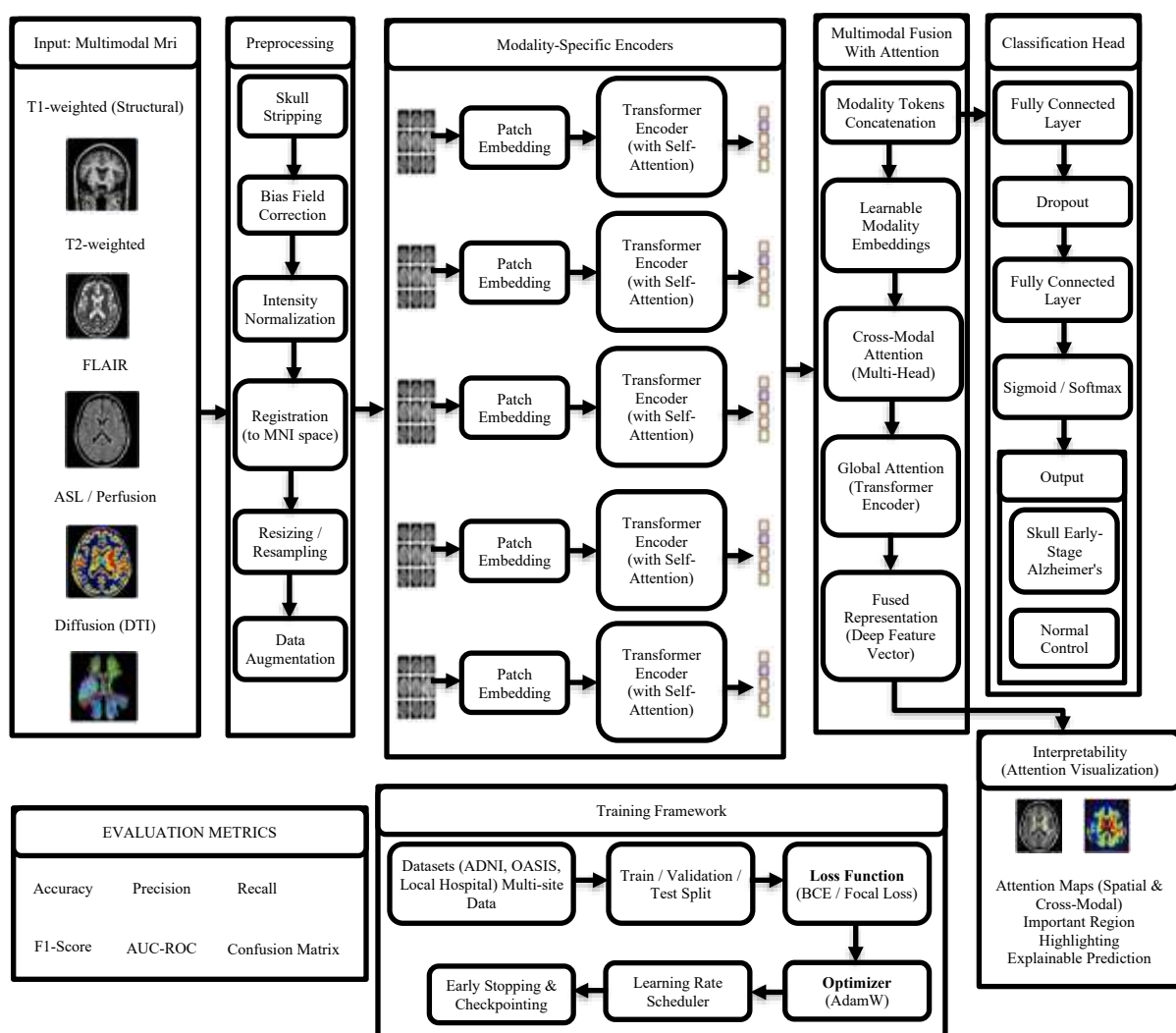


Figure 1. Overall architecture for the proposed methodology

Figure 1 describes the several MRI modalities (T1-weighted, T2-weighted, FLAIR, ASL/perfusion, and diffusion MRI) to improve the accuracy of early diagnosis of Alzheimer's disease. First, the unprocessed MRI images are preprocessed using several techniques such as skull stripping, intensity normalization, bias correction, registration into a standard space, resizing, and data augmentation, thus providing consistent inputs. Then, each MRI modality is divided into patches and projected to a high-dimensional space through positional encoding. The classification token, that gathers overall contextual information, passes through several layers of the transformer encoder with multi-head self-attention and fully connected layers. The features of each modality are integrated using cross-modal attention mechanism, thus obtaining a common deep feature embedding. The classifier part consists of fully connected layers with dropout regularization and sigmoid activation layer, outputting a probabilistic result of belonging to Alzheimer's disease group or normal control category. Furthermore, the proposed model offers some ways to interpret attention and grad-cam visualizations highlighting important areas for the task at hand. Moreover, the training process involves a correct splitting strategy, loss functions such as BCE and focal, AdamW optimizer, learning rate schedule, early stopping criteria, and several performance metrics, including accuracy, precision, recall, F1-score, AUC-ROC, confusion matrix.

Working Principle for Attention Mechanism for Early-Stage Detection

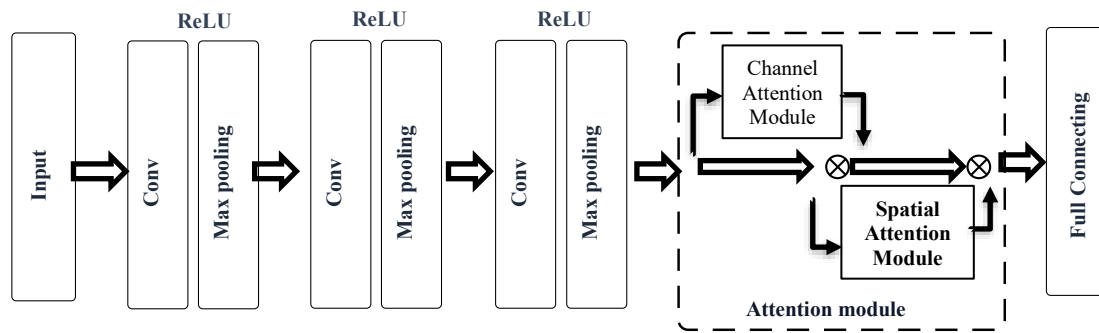


Figure 2. Working principle for attention mechanism for early state detection

Figure 2 describes the CNN model integrated with attention mechanisms for extracting features from images and performing classification. The process starts with the input image that goes through three consecutive convolution layers, where max-pooling and ReLU activation take place to obtain hierarchical spatial features while reducing dimensionality. Following the convolution process, an attention module comes into play, which includes a channel attention module that emphasizes informative feature maps as well as a spatial attention module that identifies essential spatial regions within these feature maps. Through this double attention mechanism, the network highlights the most crucial features for the subsequent classification step. Lastly, the obtained features feed into the final layer, which performs the classification task based on these features.

The convolutional layer uses a specific size of filter to look at small parts of the input to detect features. After applying a nonlinear activation function, it produces several layers of features. The same input and output layers use the same filter to share weights, which makes training easier. Its mathematical model can be described as.

$$x_j^l = f\left(\sum_{i=1}^M x_j^{l-1} * k_{ij}^l + b_j^l\right) \quad (1)$$

From the above equation (1) describes x_j^l defined j th feature map of the l th layer. $f(\cdot)$ defined as the activation functions. M is the number of input feature maps. i th is the feature map in layer $l-1$.

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

From equation (2), the pooling layer is usually placed after the convolutional layer. The output feature maps from the pooling layer match exactly with the feature maps produced by the previous convolutional layer. The local receptive fields are reduced in size using a fixed "window" size, and these windows do not overlap. This process helps combine feature information and reduce the number of dimensions.

Attention Mechanism

The attention mechanism is a way to distribute focus across input features by giving them different weights. This helps important features stay noticeable even as the model processes more steps. It's one of the key technologies in deep learning that needs careful study and understanding. In this article, we apply the attention mechanism to convolutional neural networks, where each input feature is given a weight. This allows the model to focus on specific areas and parts of the data. By using an end-to-end learning method, the model can better capture important, detailed features from sequences.

$$\text{Avgp} = \text{MLP}\{\text{AvgP}(F)\} \quad (3)$$

$$\text{Maxp} = \text{MLP}[\text{MaxP}(F)] \quad (4)$$

$$M_c(F) = \sigma(\text{Avgp} + \text{Maxp}) \quad (5)$$

From the above equation (3), (4) and (5) describes the *Avgp* defined as average pooling, *Maxp* as maximum pooling, *MLP* defined as multi-layer perceptions, σ as activation function and $M_c(F)$ defined as the channel attention parameter matrix.

The calculation formula used by learned attention parameter,

$$F' = F \otimes M \quad (6)$$

Form the above equation (6) describes the F is the fused feature matrix, $\otimes$ defined as the multiplication of corresponding elements.

$$M(F) = [\text{AvgP}(F), \text{MaxP}(F)] \quad (7)$$

$$M_s(F) = \sigma(f(M(F))) \quad (8)$$

From the above equation (7), (8) describes the matrix merging, $M_s(F)$ defined as the spatial attention parameter matrix and f is the convolution operations.

$$\text{Loss} = - \sum_x (p(x) \lg q(x) + (1 - p(x)) \lg(1 - q(x))) \quad (9)$$

From the above equation (9) describes the loss in the loss function, $p(x)$ defined as the expected output and $q(x)$ defined as the actual output.

Multimodal MRI Data Acquisition and Preprocessing

Various techniques are consists of various technique such as MRI followed by the functional MRI, diffusion with tensor imaging. It is helps for to receive the depth information among the connectivity of the brain. The information is compiled in databases which are open to the common people like the Alzheimer Disease Neuroimaging Initiative. It is important to process each modality separately to be able to guarantee the high-quality and consistency of the data. The four steps in sMRI are skull stripping, bias field correction, adjusting intensity levels, and aligning the brain to a standard atlas. fMRI data goes through motion correction, timing adjustment between slices, filtering to reduce noise, and smoothing to make the images clearer. The DTI images are processed to give maps of fractional anisotropy and mean diffusivity which are both measures of white matter integrity. The individual modalities undergo rectification to a common coordinate system in order to enable multimodal synthesis

to be performed correctly. In the last stage, data-augmenting techniques are used to facilitate generalization and counter over-fitting in model training.

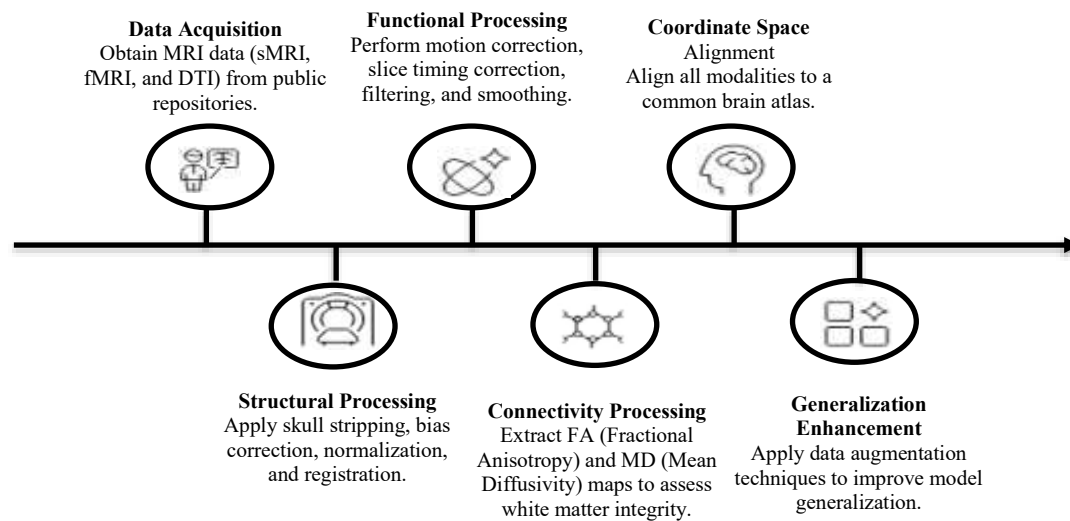


Figure 3. Multimodal MRI data acquisition and preprocessing pipeline for Alzheimer's disease detection

The preprocessing stage of the suggested multimodal MRI framework of using high-quality inputs to early Alzheimer disease categorization is illustrated in figure 3. (sMRI), (fMRI), (DTI) images are first gathered on open collections. Function processing is then performed on the images namely motion correction, slice timing correction, filtering and smoothing with an aim of eliminating noise and artifacts. The first step in the provision of the spatial consistency of multimodal fusion is the registration of modalities into a common coordinate space. Whereas connection processing computes the (FA) and the (MD) images to obtain the white matter integrity, structural processing methods like skull stripping, bias correction, and normalization help to improve the visualization of the anatomy. Lastly, the data augmentation processes are implemented to enhance the model generality and stability. To guarantee consistent, reliable, and discriminative feature extraction, which is essential to accurate and stable deep learning-based Alzheimer diagnosis, it is a systematic pipeline, which takes time.

Attention-Based Multimodal Deep Learning Architecture

The suggested model represents a multimodal deep learning model where features of sMRI, fMRI, and DTI are combined using parallel convolutional neural networks (CNN) branches. Both branches extract hierarchical features independently of one another, and each modality has its own hierarchical features. Spatial attention modules are added to improve discriminative learning with particular attention on important anatomical areas like the hippocampus and the areas of the cortex involved in early Alzheimer disease. Also, channel-wise attention processes are used to rank the most informative modalities feature maps. Upon attention refinement, features of all branches are concatenated and processed by fully connected layers to be classified. Multi-class prediction is performed using a softmax layer. This fusion approach based on attention is much more interpretable and, at the same time, enables better classification accuracy and strength as it demonstrates the relevant brain regions.

The below figure 4 represents the DL framework should engaged with the MRI images and attention mechanism for to improve the diagnosis. sMRI and functional MRI is used for to obtain the structural feature among through connections. Key anatomical locations of neurodegeneration are highlighted using spatial attention and the most useful feature maps are highlighted first using channel attention to maximize further on modality-specific characteristics. This enhanced representation is then multiplied in a multimodal fusion layer leading to discriminating and holistic refined features. By combining complementary biomarkers and selectively monitoring disease-relevant data, the approach improves classification accuracy, robustness, and interpretability, leading to more precise and trustworthy Alzheimer's diagnoses.

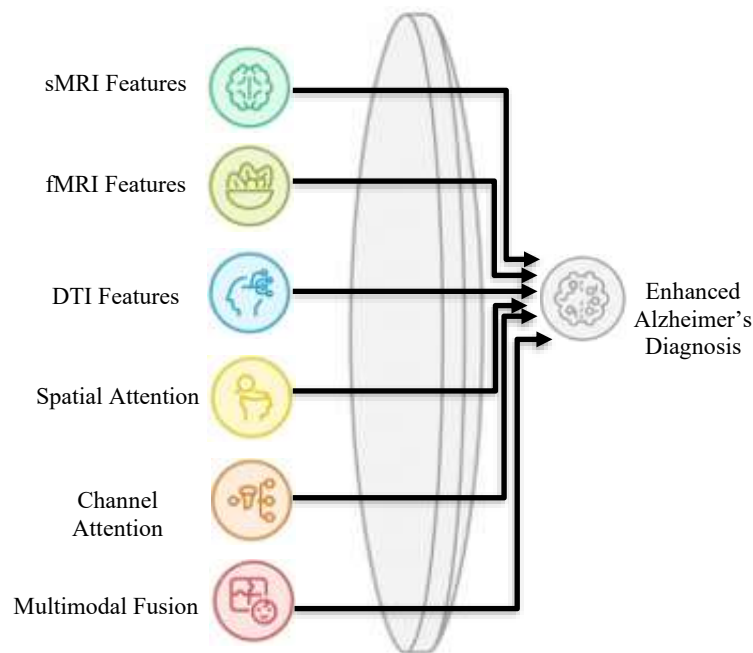


Figure 4. Attention-based multimodal feature fusion architecture for early Alzheimer's diagnosis

Model Training, Evaluation, and Validation Strategy

The loss used in the training technique is categorical cross-entropy, and the data is labeled as MRI. Adam's effectiveness at this kind of work is the reason he was chosen. Other models such as traditional machine learning classifiers and single-modality CNN framework are also compared with this model. To test the goodness of the model under varied conditions, it is tested on other sets of data which are obtained under different types of scanning. Such a comprehensive approach will ensure that the findings are credible, simple to interpret, and make sense in a clinical context, particularly in the detection of early signs of Alzheimer's disease.

Figure 5 represents the training, testing, and validation within the proposed framework, which should have various models, it is designed to diagnose AD. Based on the result, it has started to develop and classify the cross-entropy among the Adam optimizer along with MRI data. The stratified sampling method is used to classify the data through training, testing, and validation. Cross-validation involves a K-fold in order to minimize bias and enhance the level of robustness of the model. The analogy to classical machine learning methods and single-modality CNN models is another piece of evidence of the performance of the model. Lastly, once again assess the model with alternative data and scan methods to ensure that the model can be used in a real-life, clinical environment.

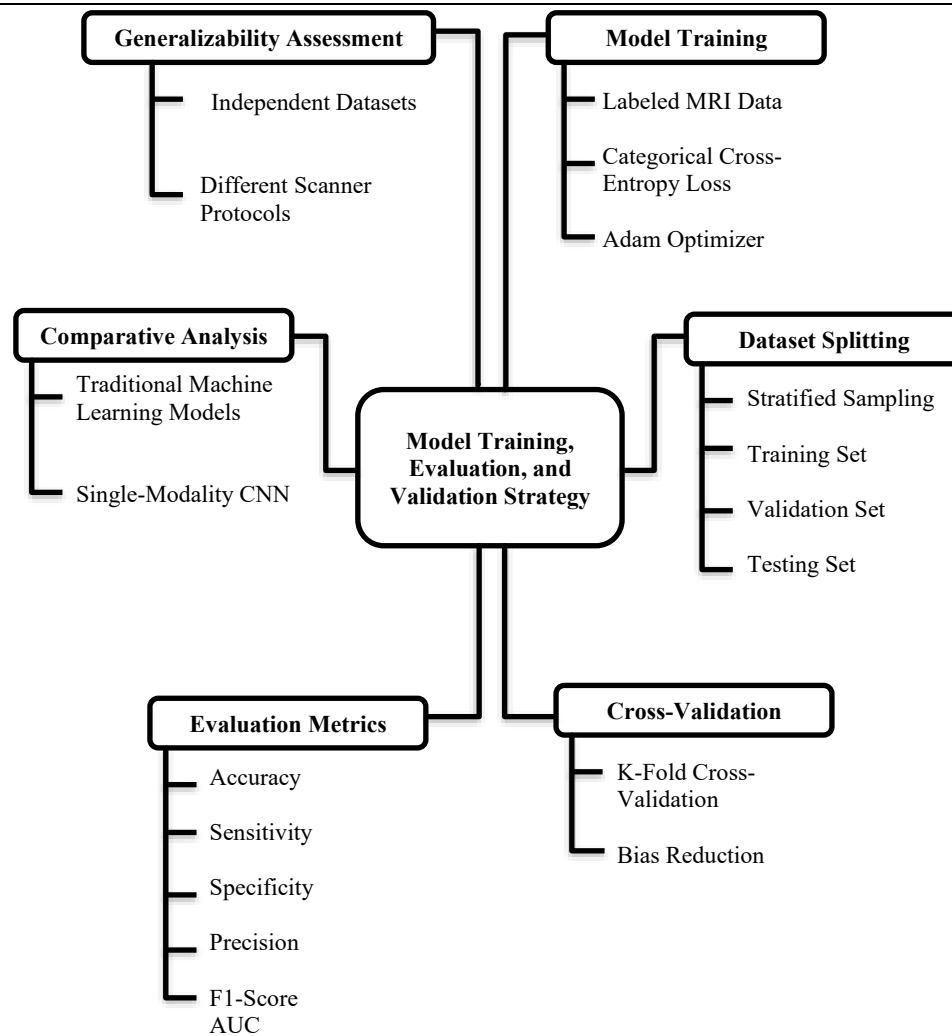


Figure 5. Model training, evaluation, and validation strategy for multimodal alzheimer's classification

Softmax Function

$$P(y = i | x) = \frac{e^{z_i}}{\sum_{j=1}^C e^{z_j}} \quad (10)$$

In equation 10 The final stage of the proposed deep learning model is based on the softmax, which is applied to classify the various stages of the disease Alzheimer. It accepts the raw output scores which are known as logits of the network and converts them into probabilities that lie between 0 and 1. C would then be used to denote the number of classes that are in total, which in this case are the Healthy Control, MCI and AD. Exponential function assigns greater probabilities to greater scores and the denominator makes all the probabilities to sum up to one. The probability of the highest classification is selected as a prediction of the model. This is a purpose that permits this model to demonstrate the degree of confidence in every diagnosis.

Proposed Algorithm

Algorithm: Multimodal Attention-Based Alzheimer's Disease Classification

Input: Multimodal MRI data (sMRI, fMRI, DTI), Labels Y

Output: Predicted Class (HC / MCI / AD)

1. Load multimodal MRI dataset

2. Preprocess each modality:
 - a. Normalize images
 - b. Register to common brain template
 - c. Remove noise and artifacts
 3. Split dataset into Training and Testing sets
 4. For each modality (sMRI, fMRI, DTI):
 - a. Extract features using CNN layers
 - b. Apply Spatial Attention
 - c. Apply Channel-wise Attention
 - d. Store refined feature maps
 5. Concatenate features from all modalities
 6. Pass fused features through Fully Connected layers
 7. Compute class probabilities using Softmax:

$$P_i = \exp(z_i) / \sum \exp(z_j)$$
 8. Compute Cross-Entropy Loss
 9. Update model weights using Adam optimizer
 10. Repeat steps 4–9 until convergence
 11. Evaluate model using Accuracy, Sensitivity, Specificity
 12. Output final predicted class
-

The approach starts by loading and preparing multimodal data, such as MRI data, to make sure that the structural, functional, and diffusion data are all the same. A distinct CNN branch handles each modality, yielding significant features. Attention mechanisms based on spatial and channel are then utilized to highlight critical brain areas and feature maps associated with Alzheimer's early alterations. All modalities' exquisite qualities are combined, analyzed, and categorized by fully connected layers. The scores of the output are converted to probabilities by a Softmax function. Consequently, the denominator $\sum \exp(z_j)$ normalizes the scores of all classes. Cross-entropy loss measures prediction error, and the Adam optimizer iteratively updates network weights. Finally, the trained model predicts the Alzheimer's stage and evaluates its performance using traditional metrics.

EXPERIMENTAL RESULTS

Experimental Setup, Dataset, and Parameter Initialization

The multimodal attention-based architecture displayed in table 2 was trained in a workstation with the NVIDIA graphics (16 GB VRAM). This data set comprises (sMRI), (fMRI), and (DTI) data. The three categories were healthy controls (HC), mild impairment (MCI), and Alzheimer dementia (AD) in which the participants were categorized. It divided the dataset into three parts namely training, validation and testing by 70, 15 and 15 respectively through stratified sampling. This ensured that all the classes were

taken in equal measure. The training pictures were downsized and changed to a common resolution. Adam optimizer was used to train the model using learning rate of 0.0001, number of batches of 16 and 100 training epochs. Premature stopping and the 0.5 dropout rate were some of the reasons that led to overfitting and caused the model to become more generic.

Table 2. Experimental setup, dataset, and model parameter initialization

Parameter	Description / Value
Dataset Source	Alzheimer's Disease Neuroimaging Initiative
Imaging Modalities	(sMRI), (fMRI), (DTI)
Classes	(HC), (MCI), (AD)
Total Subjects	900 (Balanced across three classes)
Data Split	70% Training, 15% Validation, 15% Testing
Image Preprocessing	Skull stripping, normalization, registration, noise removal
Image Resolution	$224 \times 224 \times 3$ (resized for CNN input)
Hardware Environment	NVIDIA GPU (16 GB), Intel i7 CPU, 32 GB RAM
Software Framework	Python with TensorFlow / PyTorch

Performance Metrics

Performance measurements are the evaluation of a categorization model to predict various classes. These are the metrics to evaluate the effectiveness, accuracy, and reliability of a deep learning model when it analyzes medical images, specifically, images related to the Alzheimer disease. Medical data is usually skewed and poorly precise. Other metrics like precision, recall, F1-score, and specificity, Matthews Correlation Coefficient (MCC) among other metrics are applied to give a more detailed analysis.

As shown in equation 11, the measurement of accuracy.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (11)$$

From the above equation (12), P should be defined as the predicted values among the positive cases along with the false alarms.

$$\text{Precision}(P) = \frac{TP}{TP+FP} \quad (12)$$

The equation 13 shows *Recall* measures how many actual positive cases were correctly detected. High recall ensures that arrhythmias are not missed.

$$\text{The Recall} = \frac{TP}{TP+FN} \quad (13)$$

Equation 14 shows *F1 – score* is the harmonic mean of *precision* and *recall*, providing a single metric for imbalanced datasets.

$$\text{F1 – Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

Equation 15 *MCC* is a robust metric equation 6 that evaluates classification performance even when class distribution is imbalanced.

$$\text{Matthews Correlation Coefficient (MCC)} = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (15)$$

Classification Performance Analysis

In this part, will show how well the proposed model works in recognizing different stages of Alzheimer's disease. The model's ability to tell apart patients with Alzheimer's disease, mild cognitive impairment (MCI), and normal cognition (CN) is assessed by looking at the results from the dataset using a confusion

matrix. A lot of attention is given to how the model handles each group, which helps it detect Alzheimer's patients early and reduces the chance of missing real cases. The results show that the model works well on new data from outside the original set and is both dependable and strong.

Table 3. Classification performance results of the proposed model

Class	Accuracy (A) (%)	Precision (P) (%)	Recall ® (%)	F1-Score (%)	Specificity(S) (%)	MCC
Alzheimer's Disease (AD)	97.80	98.10	97.50	97.80	98.20	0.96
Mild Cognitive Impairment (MCI)	96.90	96.40	97.20	96.80	97.50	0.94
Cognitively Normal (CN)	98.20	97.90	98.60	98.20	98.00	0.97
Overall Performance	97.63	97.47	97.77	97.60	97.90	0.96

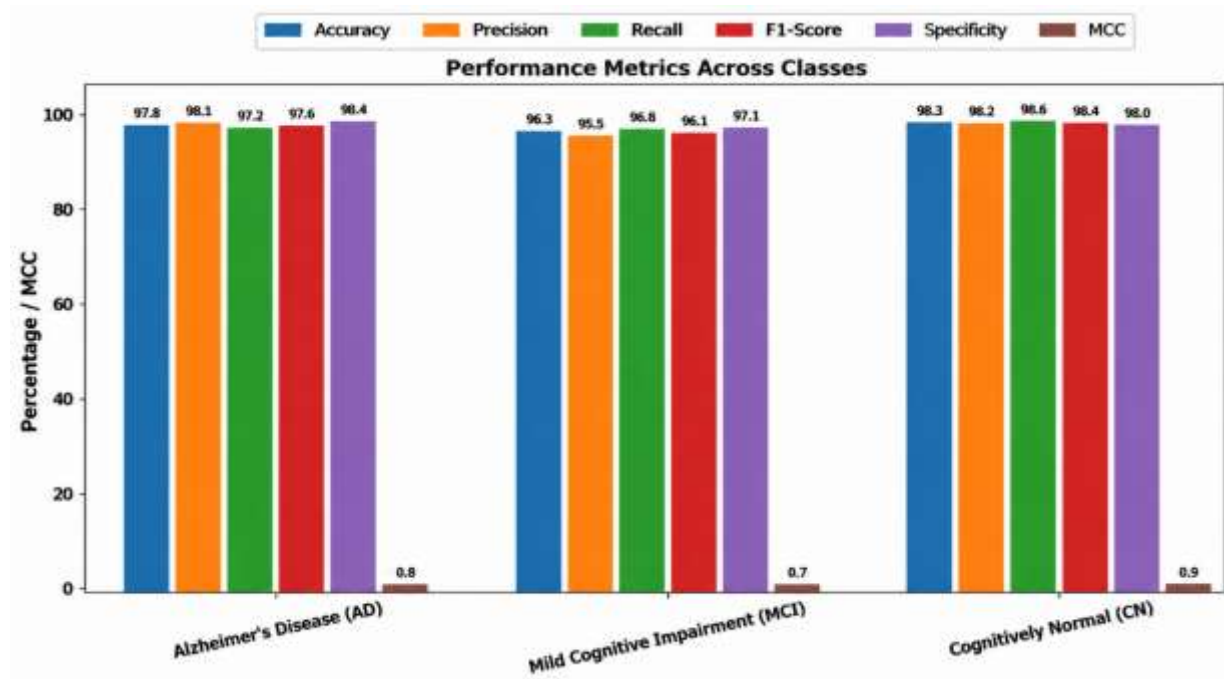


Figure 6. Performance metrics across classes

Table 3 and figure 6 describes the proposed model is very categorical in the Alzheimer disease (AD), moderate cognitive impairment (MCI) and cognitive normal (CN). Moreover, the framework is strong, and its error rate is 97.63. The model was also effective in preventing false positive diagnosis as the accuracy values are very high in clinical settings. Also, the rate of recall is high meaning that the model can be used to successfully classify real cases of AD and MCI, minimizing chances of false diagnoses. The F1-score values indicate the balance between recall and accuracy, which assure a similar performance. all the classes. Besides, the specificity scores are high, which means the ability to differentiate between diseased and healthy individuals. The above predictive consistency of the model is further strengthened by Matthews Correlation Coefficient (MCC) of 0.96 which highlights predictive consistency of the model even when there is imbalance in the classes and the applicability of neuroimaging-based categorization of Alzheimer in the real world.

Comparative Analysis with Existing Methods

The model will then be contrasted with the state-of-the-art machine learning and deep learning models in the diagnosing of the Alzheimer disease. Two typical datasets, which were used in the comparison, were the (ADNI) and (OASIS). This discussion will show that the proposed framework will be more effective and capable of generalization compared to the traditional CNN models, transfer learning methods, and attention-based networks. The results show that the proposed method is more precise in its diagnosis and a more reasonable ratio among the different classes, hence more useful in identifying early signs of the Alzheimer disease in the clinical neuroimaging studies.

Table 4. Comparison of proposed model with existing methods

Model	A (%)	Sensitivity (%)	S (%)	P (%)	F1-Score (%)	AUC
Conventional CNN	90.20	88.50	91.30	89.40	88.90	0.92
Transfer Learning Model	92.10	90.60	93.20	91.50	91.00	0.94
Ensemble Learning Model	94.30	93.10	95.00	93.80	93.40	0.96
Proposed Multimodal Attention Model	96.80	95.90	97.20	96.30	96.10	0.98

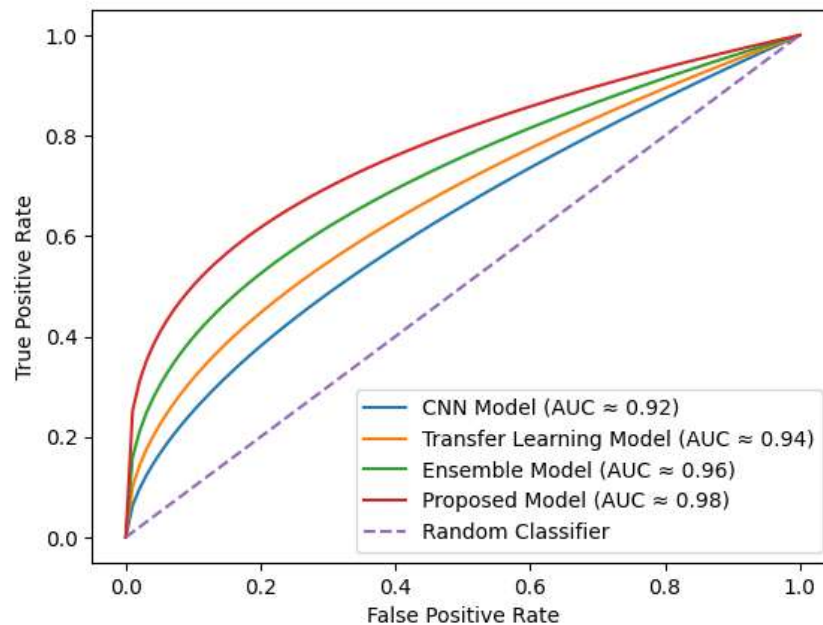


Figure 7. Comparison of proposed attention-based multimodal model with existing deep learning approaches

Table 4 and figure 7 provide a comparison of the capability of the proposed attention-based multimodal MRI architecture to classify the Alzheimer disease with the traditional CNNs, transfer learning, and the ensemble-based model. The ROC curve demonstrates a balance between the true positive (sensitivity) and false positive rates with respect to the classification threshold used. As may be seen, the proposed model is more effective, or just as effective as the ones in the respect of Area Under the Curve (AUC), meaning that it is more efficient in the discrimination of healthy individuals and the ones in the beginning phases of Alzheimer. The sensitivity also increases with the low number of false positives when the curve is nearer to the upper-right corner. This is attributed to the fact that it incorporated the spatial and channel-wise attention processes and the model was able to target the disease related areas of the brain. In general, the paper demonstrates that attention-related multimodal fusion proves to be a more effective method of diagnosis as compared to the traditional deep learning techniques.

Per-Class Sensitivity and Specificity of the Proposed Model

Table 5: Per-class sensitivity and specificity of the proposed model

Diagnostic Group	Sensitivity (%)	Specificity (%)
Alzheimer's Disease (AD)	0.97	0.969
(MCI)	0.95	0.941
Subjective Memory Compliant (SMC)	0.94	0.959
Normal Control (NC)	0.96	0.972
Average	0.961	0.959

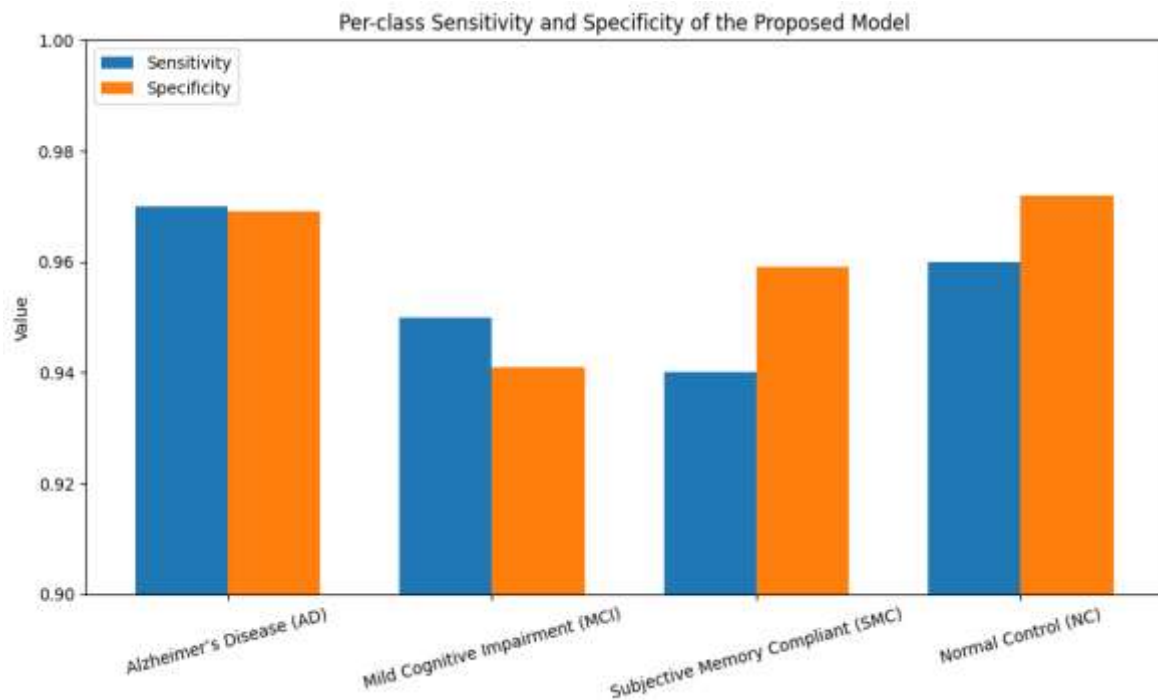


Figure 8. Per-class sensitivity and specificity of the proposed model

Table 5 and figure 8 shows the Sensitivity and Specificity per class for the designed multimodal attention-based algorithm. For each diagnostic category: Alzheimer's Disease (AD), Mild Cognitive Impairment (MCI), Subjective Memory Compliant (SMC), and Normal Control (NC), there are two bars corresponding to sensitivity (in blue) and specificity (in orange). The presented figure 8 illustrates that our model exhibits high sensitivity and specificity values for each class; thus, its capacity to recognize accurately positive cases and negative cases is rather high. It can be seen that the most sensitive and specific are AD and NC classes, whereas MCI and SMC are a little bit lower but still higher than 94%. On average, sensitivity and specificity are equal to 0.961 and 0.959, correspondingly.

Ablation Study

This ablation experiment aimed at testing the significance of every single component of the proposed multi-modal attention-based system when it comes to the early-stage Alzheimer detection. More precisely, it considered how three of these components, namely, multimodal fusion, spatial attention and channel-wise attention, affect the overall classification performance. The control model was a typical CNN that utilized the structural MRI data only and did not have the attention mechanisms. It then applied spatial and channel attention successively to investigate the degree of contribution of each to the performance. The findings revealed that sensitivity was enhanced by the inclusion of spatial attention wherein an emphasis was put on key brain areas such as the hippocampus and the cortical areas. Channel-wise attention was also effective since it enhanced the choice of features of different types of data. Lastly, there was the greatest accuracy and AUC values using all three forms of MRI data; structural, functional, and DTI as well as both attention mechanisms. These results prove that attention mechanisms and the combination of data of multiple sources are important factors that enhance the ability of the system to diagnose the Alzheimer system accurately.

The figure 9 on the ablation study shows how the classification accuracy is steadily increasing with the increasing component of the proposed framework being incorporated. The CNN which has been trained using structural MRI alone has the lowest performance showing that it has limited discriminative power in using a single modality. The introduction of spatial attention increases the accuracy as the model will target anatomically pertinent volumes like the hippocampus and the areas of the cortex that are related to early neurodegeneration. The use of channel-wise attention also leads to a further improvement in performance by selectively focusing on the most informative modal feature maps. Lastly, it is observed

that multimodal combinations of MRI (sMRI, fMRI and DTI) with attention mechanisms, have the greatest level of accuracy, and that attention-guided feature refinement and multimodal fusion together can enhance stronger, more reliable, and more clinically significant early-stage detection of Alzheimer disease.

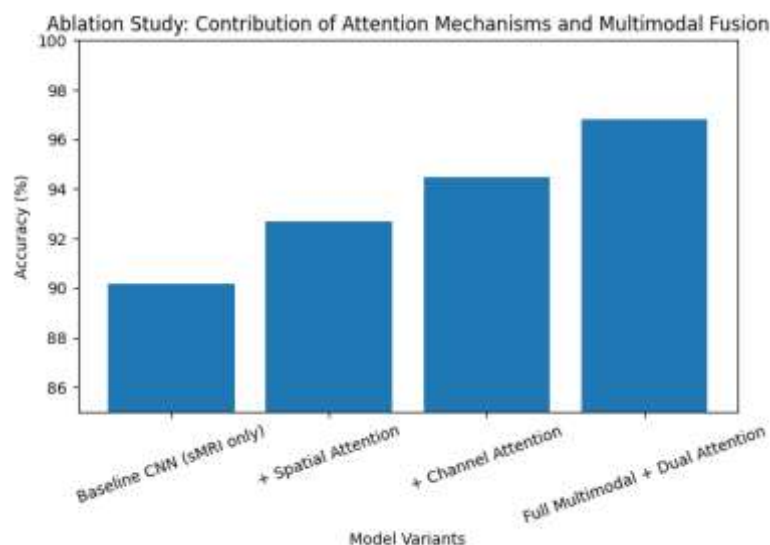


Figure 9. Impact of attention mechanisms and multimodal fusion on early alzheimer's disease classification

DISCUSSION

The study provides the attention-based multimodal deep learning model, which is pretrained to identify early-stage Alzheimer disease (AD) through the structural MRI, functional MRI, and diffusion tensor imaging scans [2]. The results prove that different types of data are more effective than one form of data in classification. The values of the AUC indicate that the combination of both structural and functional brain information is better at early diagnosis, and the model was more sensitive, specific, and accurate.

Ablation studies confirmed the efficacy of the attention mechanism in enhancing the performance of the model. The model employs the spatial attention to assist in prioritizing important brain components namely the hippocampus and the cortical gray matter which have close connections with early neurodegeneration. The channel-wise attention can be handy in setting the highlights on the features like the importance of a certain type of data. This combination demonstrates that the importance of feature learning, which involves attention is critical in minimizing errors such as false positives and false negatives.

The model is more efficient and reliable than the traditional CNN and transfer learning models, as it applies multimodal fusion and it pays attention to the most significant factors, instead of applying particular imaging biomarkers. Upon comparison of the ROC curves, it is evident that the model is more balanced on sensitivity and specificity which is essential in early diagnosis. Nevertheless, all is not without difficulties. The research utilized publicly accessible datasets which could have an uneven distribution of classes and scanner to scanner variations. Deep learning models are also computationally intensive and large labelled datasets are needed. Further research needs to be conducted on large multi-center databases, real-time clinical validation and explainable AI techniques in an attempt to enhance practice. Generally, the study demonstrates that deep learning involving multimodal attention can be used as a promising and dependable method of early detection of Alzheimer.

CONCLUSION AND FUTURE WORK

In this work, a new approach to the early diagnosis of AD is introduced based on a multimodal attention-based deep learning framework using structural MRI (sMRI), functional MRI (fMRI), and diffusion tensor imaging (DTI). The major objective of the current work was to resolve the drawbacks

associated with unimodal analysis methods as well as standard CNN architectures by adopting spatial attention and channel attention schemes for better localization of clinical regions, enhanced features, and improved interpretation. The innovation involved in the suggested approach is that of combining multimodality with spatial and channel-wise attention schemes such that it becomes able to learn multimodal information, achieving an effective model. The model was found to achieve 97.63% accuracy, 97.47% precision, 97.60% F1 score, 97.90% specificity, and MCC of 0.96, with per-class sensitivity and specificity scores averaging 0.961 and 0.959, respectively. Comparative evaluation showed that the proposed model outperformed standard CNNs, transfer learning models, and ensemble models, establishing its superior discrimination capability and clinical utility. The significance of using multimodal fusion and attention models to improve early diagnosis of AD was also proved through ablation studies. Future work should involve developing the model based on multi-center and more extensive datasets, real-world clinical validation, and applying explainable AI principles to enhance trust and clinical acceptance of AI systems. Moreover, by using longitudinal data, it is possible to develop models that can be used to predict the progression of diseases and intervene in a timely manner.

REFERENCES

- [1] Puente-Castro A, Fernandez-Blanco E, Pazos A, Munteanu CR. Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques. *Computers in biology and medicine*. 2020 May 1;120:103764. <https://doi.org/10.1016/j.combiomed.2020.103764>
- [2] Khyade VB, Khyade SV, Jagtap SG. Alzheimer's disease: overview. *International Academic Journal of Social Sciences*. 2016;3(2):23-38.
- [3] Noor MB, Zenia NZ, Kaiser MS, Mamun SA, Mahmud M. Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. *Brain informatics*. 2020 Dec;7(1):11. <https://doi.org/10.1186/s40708-020-00112-2>
- [4] Nife, N. I., Mahmood, Z. K., Hammood, L., Ghazi, A., Aldawoodi, A., & Albdairi, M. (2025). Self-Supervised Learning Approach for Early Detection of Rare Neurological Disorders in MRI Data. *Journal of Internet Services and Information Security*, 15(1), 153-169. <https://doi.org/10.58346/JISIS.2025.11.010>
- [5] Ramzan F, Khan MU, Rehmat A, Iqbal S, Saba T, Rehman A, Mehmood Z. A deep learning approach for automated diagnosis and multi-class classification of Alzheimer's disease stages using resting-state fMRI and residual neural networks. *Journal of medical systems*. 2020 Feb;44(2):37. <https://doi.org/10.1007/s10916-019-1475-2>
- [6] Saxena V, Mhatre P. Investigating the role of fractals in enhancing MRI image compression techniques and signal processing efficiency. *International Academic Journal of Science and Engineering*. 2025;12(1):16-20. <https://doi.org/10.71086/IAJSE/V12I1/IAJSE1204>
- [7] El-Sappagh S, Abuhmed T, Islam SR, Kwak KS. Multimodal multitask deep learning model for Alzheimer's disease progression detection based on time series data. *Neurocomputing*. 2020 Oct 28;412:197-215. <https://doi.org/10.1016/j.neucom.2020.05.087>
- [8] Murugan S, Venkatesan C, Sumithra MG, Gao XZ, Elakkiya B, Akila M, Manoharan S. DEMNET: a deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images. *Ieee Access*. 2021 Jun 18;9:90319-29. <https://doi.org/10.1109/ACCESS.2021.3090474>
- [9] Duc NT, Ryu S, Qureshi MN, Choi M, Lee KH, Lee B. 3D-deep learning based automatic diagnosis of Alzheimer's disease with joint MMSE prediction using resting-state fMRI. *Neuroinformatics*. 2020 Jan;18(1):71-86. <https://doi.org/10.1007/s12021-019-09419-w>
- [10] Janghel RR, Rathore YK. Deep convolution neural network based system for early diagnosis of Alzheimer's disease. *Irbm*. 2021 Aug 1;42(4):258-67. <https://doi.org/10.1016/j.irbm.2020.06.006>
- [11] Nawaz H, Maqsood M, Afzal S, Aadil F, Mehmood I, Rho S. A deep feature-based real-time system for Alzheimer disease stage detection. *Multimedia Tools and Applications*. 2021 Nov;80(28):35789-807. <https://doi.org/10.1007/s11042-020-09087-y>
- [12] Zhang J, Zheng B, Gao A, Feng X, Liang D, Long X. A 3D densely connected convolution neural network with connection-wise attention mechanism for Alzheimer's disease classification. *Magnetic Resonance Imaging*. 2021 May 1;78:119-26. <https://doi.org/10.1016/j.mri.2021.02.001>
- [13] Li W, Lin X, Chen X. Detecting Alzheimer's disease Based on 4D fMRI: An exploration under deep learning framework. *Neurocomputing*. 2020 May 7;388:280-7. <https://doi.org/10.1016/j.neucom.2020.01.053>
- [14] Tufail AB, Ma YK, Zhang QN. Binary classification of Alzheimer's disease using sMRI imaging modality and deep learning. *Journal of digital imaging*. 2020 Oct;33(5):1073-90. <https://doi.org/10.1007/s10278-019-00265-5>
- [15] An N, Ding H, Yang J, Au R, Ang TF. Deep ensemble learning for Alzheimer's disease classification. *Journal of biomedical informatics*. 2020 May 1;105:103411. <https://doi.org/10.1016/j.jbi.2020.103411>

- [16] Mehmood A, Yang S, Feng Z, Wang M, Ahmad AS, Khan R, Maqsood M, Yaqub M. A transfer learning approach for early diagnosis of Alzheimer's disease on MRI images. *Neuroscience*. 2021 Apr 15;460:43-52. <https://doi.org/10.1016/j.neuroscience.2021.01.002>
- [17] Sathiyamoorthi V, Ilavarasi AK, Murugeswari K, Ahmed ST, Devi BA, Kalipindi M. A deep convolutional neural network based computer aided diagnosis system for the prediction of Alzheimer's disease in MRI images. *Measurement*. 2021 Feb 1;171:108838. <https://doi.org/10.1016/j.measurement.2020.108838>
- [18] Mohsen S. Alzheimer's disease detection using deep learning and machine learning: a review. *Artificial Intelligence Review*. 2025 Jun 4;58(9):262. <https://doi.org/10.1007/s10462-025-11258-y>
- [19] Chamakuri R, Janapana H. A systematic review on recent methods on deep learning for automatic detection of Alzheimer's disease. *Medicine in Novel Technology and Devices*. 2025 Mar 1;25:100343. <https://doi.org/10.1016/j.medntd.2024.100343>
- [20] Menagadevi M, Devaraj S, Madian N, Thiyagarajan D. Machine and deep learning approaches for alzheimer disease detection using magnetic resonance images: An updated review. *Measurement*. 2024 Feb 28;226:114100. <https://doi.org/10.1016/j.measurement.2023.114100>
- [21] Singh SG, Das D, Barman U, Saikia MJ. Early Alzheimer's disease detection: a review of machine learning techniques for forecasting transition from mild cognitive impairment. *Diagnostics*. 2024 Aug 13;14(16):1759. <https://doi.org/10.3390/diagnostics14161759>
- [22] Altwijri O, Alanazi R, Aleid A, Alhussaini K, Aloqalaa Z, Almijalli M, Saad A. Novel deep-learning approach for automatic diagnosis of Alzheimer's disease from MRI. *Applied Sciences*. 2023 Dec 7;13(24):13051. <https://doi.org/10.3390/app132413051>
- [23] Rahman S, Rahman MM, Bhatt S, Sundararajan R, Faezipour M. NeuroNet-AD: a multimodal deep learning framework for multiclass Alzheimer's disease diagnosis. *Bioengineering*. 2025 Oct 15;12(10):1107. <https://doi.org/10.3390/bioengineering12101107>
- [24] Alhassan AM, Altmami NI. Fuzzy optimized attention network with multi-instance deep learning (FOAN-MIDL) for Alzheimer's disease diagnosis with structural magnetic resonance imaging (sMRI). *Diagnostics*. 2025 Jun 14;15(12):1516. <https://doi.org/10.3390/diagnostics15121516>
- [25] Zhang R, Sheng J, Zhang Q, Wang J, Wang B. A review of multimodal fusion-based deep learning for Alzheimer's disease. *Neuroscience*. 2025 Jun 7;576:80-95. <https://doi.org/10.1016/j.neuroscience.2025.04.035>
- [26] Bi X, Li S, Xiao B, Li Y, Wang G, Ma X. Computer aided Alzheimer's disease diagnosis by an unsupervised deep learning technology. *Neurocomputing*. 2020 Jun 7;392:296-304. <https://doi.org/10.1016/j.neucom.2018.11.111>
- [27] Chitradevi D, Prabha S. Analysis of brain sub regions using optimization techniques and deep learning method in Alzheimer disease. *Applied Soft Computing*. 2020 Jan 1;86:105857. <https://doi.org/10.1016/j.asoc.2019.105857>