

An automated multimodal medical image fusion framework for Alzheimer detection using deep learning

S. Beatrice^{1,✉}, Janaki Meena M¹, Dhivviyanandam I²

¹ School of Computer Science and Engineering Department, Vellore Institute of Technology, Chennai-600127

² Department of Mathematics, North Bengal St. Xavier's College, Rajganj, West Bengal, India

ABSTRACT

Alzheimer's disease (AD) is a progressive neurodegenerative condition that affects the elderly population. The early detection and diagnosis of AD is critical for achieving effective treatment, as it can greatly improve the patient experience. AD can be viewed through imaging techniques like MRI, PET, and SPECT, providing valuable information about structural and functional changes. These findings are important in understanding this area. However, each imaging modality offers a different perspective. This information can be better collected from several of the other modalities as well as from some others to improve accuracy and reliability in AD detection. By combining information from different imaging modalities, such as MRI, PET, DTI, and fMRI, automated multimodal medical image frameworks aim to create a fused representation that preserves the relevant features from each modality. Convolutional neural networks (CNNs) and generative adversarial networks (GANs), among other deep learning techniques, have been prevalent in these frameworks for learning discriminative and informative features from multi-modal data. In this paper, The Alzheimer's Disease Neuroimaging Initiative (ADNI) is used for experimental analysis. The proposed work gives 98.94% of accuracy and 1.06% of error which is greater than the existing approaches.

Keywords: Alzheimer's detection, CNN, DTI, fMRI, Fusion, Multimodality, MRI, PET

1. Introduction

Alzheimer's disease (AD) is the most common form of dementia globally and is predominantly observed in individuals aged ≥ 60 years. As age progresses, neurodegeneration in the brain intensifies, leading to worsening symptoms. Although medications and treatments are available to alleviate

✉ Corresponding author.

E-mail address: s.beatrice2015@vit.ac.in (S. Beatrice).

Received 14 February 2025; Revised 25 February 2025; Accepted 02 March 2025; Published Online 19 April 2025.

DOI: [10.61091/jcmcc126-04](https://doi.org/10.61091/jcmcc126-04)

© 2025 The Author(s). Published by Combinatorial Press. This is an open access article under the CC BY license (<https://creativecommons.org/licenses/by/4.0/>).

these symptoms, no complete cure currently exists. With the aging population on the rise, AD is expected to affect a significant portion of the global population. Therefore, the early detection of AD using advancements in machine learning and artificial intelligence has become a critical area of research. This study aims to address this challenge by proposing an approach for the early detection of dementia using multimodal data.

Experts agree that early diagnosis of AD significantly increases the chances of mitigating its progression and improving patient quality of life. In this regard, various machine learning algorithms were employed and compared in this work to identify the most accurate method for detecting AD in its early stages.

The human brain, being the control center of the body, plays a vital role in regulating all sensory and cognitive functions. It processes, stores, and manages information, enabling communication, motor coordination, and interaction with the environment. With advancing age, cognitive decline becomes more prominent, often resulting in debilitating conditions such as dementia.

It is estimated that over 47 million individuals are currently living with dementia, making it a global health crisis. AD, the most prevalent cause of dementia, leads to gradual memory loss and cognitive impairment, thereby affecting the ability to perform daily activities. It is one of the leading causes of death among the elderly and is ranked as the sixth most common cause of mortality worldwide. While multiple factors may contribute to the onset of this disease, genetic predisposition plays a significant role. Despite extensive research, the root cause of AD remains elusive, and no definitive cure has been found to date [13].

Given these circumstances, early detection is essential to delay disease progression and enhance patient outcomes. This study presents a comprehensive framework utilizing machine learning techniques for the classification and early identification of Alzheimer’s disease. The effectiveness of the proposed system is evaluated using standard metrics such as accuracy and precision.

2. Literature survey

Numerous frameworks leveraging deep learning techniques have been proposed for the automated fusion of multimodal medical images to facilitate the detection of Alzheimer’s disease (AD). Framework A utilizes a convolutional neural network (CNN)-based architecture for image fusion, combining structural and functional information from magnetic resonance imaging (MRI) and positron emission tomography (PET) scans. While this approach effectively integrates multimodal data, its generalizability is limited due to its specific architectural design.

Framework B incorporates generative adversarial networks (GANs) to generate realistic fused images by learning representations from multiple imaging modalities. The generator network captures shared information, and the discriminator ensures image fidelity. This strategy improves AD detection accuracy by synthesizing high-quality fused images. However, the training process is computationally intensive and requires careful calibration of the input data quality.

Framework C adopts a feature-level fusion approach that extracts deep features from individual modalities using parallel CNNs, followed by integration via operations such as concatenation, summation, or attention mechanisms. This framework enables flexible feature representation and captures complementary modality-specific information but requires significant computational resources and design optimization.

Another variation, also referred to as Framework C, employs a hierarchical fusion strategy, progressively merging low-level features and refining them at higher semantic levels. This enables the extrac-

tion of both local and global information, thereby enhancing classification performance. Nonetheless, hierarchical fusion entails extended training time and increased computational demand.

Deep learning has shown considerable promise in AD diagnosis. Several studies have explored supervised, unsupervised, and semi-supervised learning, as well as architectures such as recurrent neural networks (RNNs), graph neural networks (GNNs), and generative models.

One study employed custom models and deep transfer learning techniques with five-fold cross-validation, achieving 99.65% accuracy [8]. The models included a two-layer fully connected network and pretrained architectures such as EfficientNet-B3 and ResNet-152.

In another investigation, CNNs were employed using LeNet-5 architecture for binary classification of AD-affected and healthy brains [25]. The CNN-based approach demonstrated superior accuracy in differentiating between these groups.

The use of 3D CNNs for AD diagnosis yielded a classification accuracy of 78.07% [26], while other studies proposed models based on MobileNet, artificial neural networks (ANNs), and DenseNet using MRI data, with MobileNet achieving an accuracy of 95.41% [3].

A lightweight model for MRI-based AD detection reached 99.22% accuracy for binary classification and 95.93% for multiclass classification tasks [16]. The evaluation was conducted using F1-score, recall, and precision.

A hybrid deep learning approach combining long short-term memory (LSTM) with CNNs achieved 92.8% test accuracy, and an area under the curve (AUC) ranging from 0.80 to 0.83, effectively distinguishing between normal controls and early mild cognitive impairment (MCI) [2].

Other studies have compared models such as VGG-19, Xception, and DenseNet-121, which reported accuracies of 98%, 95%, and 91% respectively [6].

"MultiAZ-Net", a CNN-based ensemble architecture, incorporated MRI and PET scans for AD classification, achieving an accuracy of $92.3 \pm 5.45\%$ for multiclass classification [14].

A deep neural network (DNN) using PET and MRI data classified six binary classes of AD stages and showed that combining PET and MRI led to improved sensitivity for early classification [7]. Similarly, in another study, features extracted from MRI and PET scans were classified using support vector machines (SVMs), particularly improving classification performance for late and early MCI [10].

The fusion of multimodal data has been extensively utilized in medical diagnosis. For example, multimodal fusion of T1, T2, T1CE, and FLAIR MRI images has been used for brain tumor identification [12, 1, 15]. Other studies combined MRI with CT and SPECT to form composite feature spaces [19, 17, 9].

Recent advancements include attention-based multimodal fusion networks, where PET and MRI data are jointly processed to extract salient features while filtering out irrelevant information [28]. Even in the absence of one modality, pretrained models can still predict AD using learned complementary information.

Shao et al. [22] proposed combining SVMs with feature correlation and structure-based fusion methods. Their results showed improved classification, particularly in distinguishing between late and early MCI, and suggested further improvements for binary classification using advanced fusion techniques.

These studies collectively highlight the growing potential of deep learning and multimodal fusion strategies in enhancing the accuracy and reliability of Alzheimer's disease diagnosis.

3. Evaluation of neuroimaging fusion algorithm performance: Single-modality vs. multimodality

Neuroimaging fusion algorithms are designed to improve the accuracy and reliability of Alzheimer's disease (AD) detection by integrating information from multiple imaging modalities such as magnetic resonance imaging (MRI), positron emission tomography (PET), and diffusion tensor imaging (DTI). This section provides a comparative analysis of single-modality and multimodal approaches used in neuroimaging-based AD diagnostics.

3.1. *Single-modality methodologies*

-MRI-based approach. Detection models based solely on MRI data have demonstrated a moderate level of accuracy, sensitivity, and specificity in diagnosing AD [27]. However, MRI may be limited in capturing metabolic or molecular alterations that are more effectively visualized through other imaging modalities.

-PET-based approach. PET imaging, which is effective in identifying metabolic activity and amyloid deposition, also yields reliable performance metrics. Nevertheless, its inability to provide detailed structural information may hinder comprehensive diagnosis.

-DTI-based approach. DTI primarily focuses on assessing white matter integrity and neural connectivity. While useful, a DTI-only framework may fail to account for other structural or functional abnormalities relevant to AD progression.

3.2. *Multimodal fusion approach*

The multimodal fusion strategy seeks to leverage the strengths of multiple imaging techniques to enhance AD diagnostic accuracy. By integrating data from MRI, PET, fMRI, and DTI, these approaches provide a more holistic representation of AD-associated brain alterations. Multimodal fusion enables the extraction of complementary signals and mitigates the limitations of individual modalities.

Several fusion techniques are commonly employed in this context:

-Feature-level fusion. Involves the integration of features extracted from different modalities before classification, allowing for joint representation learning.

-Decision-level fusion. Combines the outputs of individual classifiers trained on separate modalities, aggregating their predictions for final diagnosis.

-Deep learning-based fusion networks. Utilizes neural network architectures capable of learning complex, nonlinear relationships across modalities to improve classification performance.

The multimodal fusion methods have demonstrated superior diagnostic performance compared to single-modality approaches, offering a more comprehensive framework for early and accurate detection of Alzheimer's disease.

4. Materials and methods

The abbreviations are given below in Table 1.

Table 1. Abbreviations

Abbreviations	Meaning
AD	Alzheimer’s disease
PET	Positron Emission Tomography
MRI	Magnetic Resonance Imaging
SPECT	Single-Photon Emission Computed Tomography
DTI	Diffusion Tensor Imaging
fMRI	Functional Magnetic Resonance Imaging
CNNs	Convolutional Neural Networks
GAN	Generative Adversarial Network
ANN	Artificial Neural Networks
ROC	Receiver Operating Curve
MCI	Mild Cognitive Impairment
LSTM	Long Short-Term Memory
DNN	Deep Neural Network
SVM	Support Vector Machine
RNN	Recurrent Neural Network
AUC-ROC	Area Under the Receiver Operating Characteristic curve
FFT	Fast Fourier Transform
NSDFB	Non-Subsampled Directional Filter Banks
NSP	Non-Subsampled Pyramid
DFT	Discrete Fourier Transform
IFFT	Inverse FFT
PCNN	Pulse-Coupled Neural Network
CN	Cognitively Normal
MCI	Mild Cognitive Impairment
LMCI	Late MCI
EMCI	Early MCI
SMC	Significant Memory Concern
FGPCNN	Fuzzy Genetic Pulse Coupled Neural Networks
ADNI	The Alzheimer’s Disease Neuroimaging Initiative
RNN	Recurrent Neural Network

4.1. Proposed Methodology

Alzheimer’s Disease (AD) is a complex neurological disorder that impacts various aspects of brain functionality. Single-modality imaging techniques often fail to capture the multifaceted nature of AD. Each imaging modality provides only a partial view of the disease, which may not be sufficient for accurate early diagnosis and prognosis [11, 2, 21, 23, 18, 4]. Therefore, a multimodal approach is essential to obtain a more comprehensive representation of the pathological changes associated with

AD.

Common neuroimaging modalities used in AD detection include:

- MRI: Detects structural abnormalities such as atrophy and vascular malformations.
- PET: Assesses metabolic activity and functional characteristics of brain tissue.
- fMRI: Measures blood oxygen level-dependent (BOLD) signals to examine functional connectivity between brain regions.
- DTI: Captures microstructural white matter alterations by modeling water diffusion along axonal tracts.

When these modalities are integrated, they enable detection of:

- (a) Structural and functional brain changes.
- (b) Disruption in white matter connectivity.
- (c) Amyloid-beta plaque accumulation.
- (d) Regional cerebral hypometabolism.

The proposed fusion approach algorithm is presented in Table 2.

Table 2. Proposed fusion approach algorithm

Step	Process	Description
Transform	Contourlet + FFT	Apply Contourlet and FFT transforms to extract directional and frequency domain features from each image modality.
Feature Selection & Fusion	PCNN + Fuzzy Maximization	Use PCNN to identify salient features across modalities, and apply fuzzy maximization to enhance feature integration and generate a unified representation.
Reconstruction & Save	Inverse FFT + Contourlet	Reconstruct the fused image using inverse FFT and Contourlet transforms, then store the final fused output in the image database.

The proposed multimodal fusion framework enhances sensitivity and specificity by combining complementary features from diverse modalities. The following sequential steps summarize the methodology:

- a) Data Acquisition: Neuroimaging data (MRI, PET, DTI) from AD patients and healthy controls were collected under standardized acquisition protocols.
- b) Preprocessing: Spatial alignment, skull stripping, intensity normalization, and noise removal were performed to ensure uniformity across modalities.
- c) Feature Extraction: Critical features including voxel intensities, texture descriptors, shape-based statistics, and connectivity metrics were extracted.
- d) Fusion Algorithm: Feature-level fusion was implemented using a combination of Contourlet and Fast Fourier Transforms (FFT), followed by feature selection via Pulse-Coupled Neural Networks (PCNN) and fuzzy maximization.
- e) Model Training and Testing: Deep learning classifiers such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) were trained using the fused feature representations. Cross-validation was performed using labeled data.
- f) Performance Evaluation: Accuracy, sensitivity, specificity, precision, and area under the ROC curve (AUC) were used to evaluate model performance.

g) Comparative Analysis: The performance of the multimodal fusion framework was compared with individual modality results to demonstrate its effectiveness.

The fusion algorithm described above ensures accurate and comprehensive feature representation through multistage processing. The workflow of the proposed model is depicted in Figure 1.



Fig. 1. Workflow of proposed fusion model

4.1.1. Non-subsample contourlet transform (NSCT). This is the first step in the proposed fusion algorithm. First, the source images that required fusion were collected. These images can be obtained using different sensors or modalities. In our study, MRI, functional MRI, DTI, and PET were combined to improve the accuracy of AD detection. The NSCT is then applied to decompose each source image into sub-bands of different frequency sub-bands. NSCT [4] is a multiscale, multi directional transform that captures local and nonlocal image features. The NSCT consists of two components:

1) Non-Subsampled Pyramid (NSP). A two-channel non-subsampled 2-D filter bank is called a non-subsampled pyramid. Using the NSP, a multiscale characteristic was obtained. The NSP structure for the three-stage pyramid decomposition and the sub-bands generated on a two-dimensional frequency plane are shown in Figure 2.

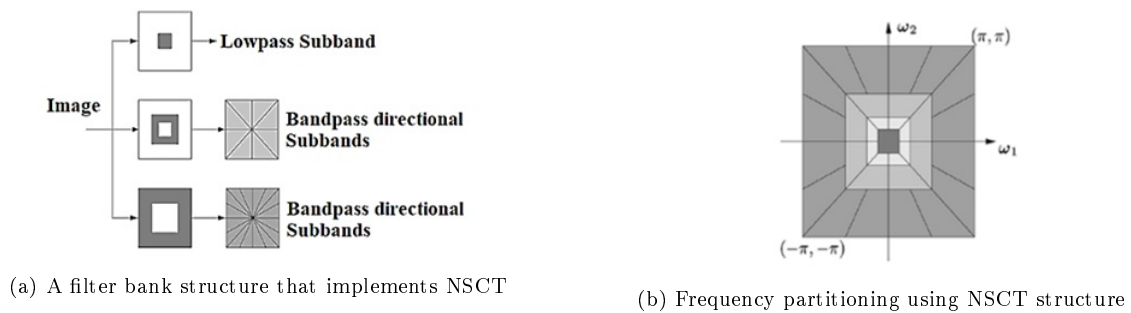


Fig. 2. Nonsubsampled contourlet transform

One Low-frequency and one high-frequency image were obtained for each decomposition level. To

capture the singularities in a picture efficiently, low-frequency components were decomposed at the subsequent decomposition level. Consequently, NSP may produce " $k + 1$ " sub images, where " k " is the number of decomposition levels and consists of one low frequency image and k high frequency images. The size of the sub images is comparable to that of the input source image. The Figure 2a & b shows the NSCT filter bank structure and Frequency partitioning using NSCT structure respectively.

2) Non-subsampled directional filter bank (NSDFB). Two-channel filter banks that have been critically sampled and a resampling operator are used to create a Nonsubsampled directional filter bank. By employing the NSDFB, the NSCT can achieve multidirectional properties and additional directional information. In addition, the design of the NSDFB removes the up samplers and down samplers from the contourlet-transform DFB. Every " k " level undergoes directional decomposition, yielding " 2^k " directional sub-bands that are about the same size as the original images. A four-channel non-subsampled directional filter bank with two-channel fan filter banks is shown in the Figure. The 2-D frequency plane is divided into wedge-shaped sections by the tree-structured filter bank.

Multiscale decomposition involves the application of a low-pass filter (h) and high-pass filter (g) at each scale. The low-pass filter captures the approximation or smoother components, whereas the high-pass filter captures the detailed or edge features.

$$\begin{aligned} A_{j+1}^0 &= (h * A_j^0) \downarrow 2 + (h * A_j^1) \downarrow 2, \\ A_{j+1}^1 &= (g * A_j^0) \downarrow 2 + (g * A_j^1) \downarrow 2, \end{aligned}$$

where A_j^0 and A_j^1 represent the approximation and detail coefficients at scale j , $*$ denotes convolution, and $\downarrow 2$ represents downsampling by a factor of two.

Directional decomposition involves applying directional filters $h_{l,k}$ and $g_{l,k}$ at each scale and orientation k .

$$\begin{aligned} B_{j+1}^{0,k} &= (h_{l,k} * B_j^0) \downarrow 2 + (h_{l,k} * B_j^1) \downarrow 2, \\ B_{j+1}^{1,k} &= (g_{l,k} * B_j^0) \downarrow 2 + (g_{l,k} * B_j^1) \downarrow 2, \end{aligned}$$

where B_j^0 and B_j^1 represent the directional approximation and detail coefficients at scale j and orientation k . Figure 3 shows the input and output relationship of decompositions.

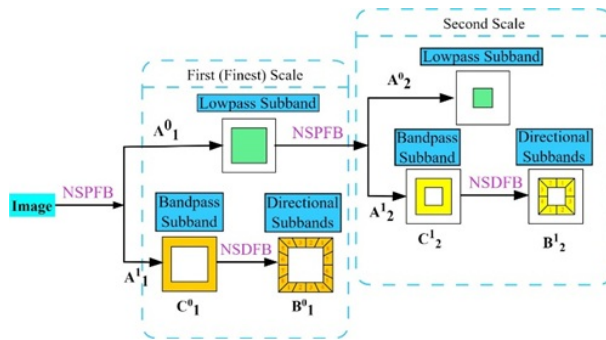


Fig. 3. Input and output relationship of NSCT decomposition

The sub-band decomposition involves further decomposition into sub-bands $C_{j,l}^{d,k}$, where d represents the direction and k represents the orientation.

$$C_{j,l}^{0,k} = (h_{l,k} * C_j^{0,k}) \downarrow 2 + (h_{l,k} * C_j^{1,k}) \downarrow 2,$$

$$C_{j,l}^{1,k} = (g_{l,k} * C_j^{0,k}) \downarrow 2 + (g_{l,k} * C_j^{1,k}) \downarrow 2,$$

where $C_j^{0,k}$ and $C_j^{1,k}$ represent the sub-band approximation and detail coefficients at scale j and orientation k . These equations illustrate the NSCT decomposition process, capturing the hierarchical structure of the coefficients at different scales, directions, and *subbands*. Specific filter functions $h_{l,k}$ and $g_{l,k}$ are designed.

The resulting coefficients were organized into a hierarchical structure, forming a tree-like representation. The coefficients at different scales, directions, and sub-bands are denoted as $Y_{j,l}^{d,k}$, where j represents the scale, l represents the direction, d represents the approximation or detail, and k represents the orientation.

$$Y_{j,l}^{d,k} = \begin{cases} A_j^d & \text{if } j = 0, \\ B_j^{d,k} & \text{if } j > 0 \text{ and } l = 0, \\ C_{j,l}^{d,k} & \text{if } j > 0 \text{ and } l > 0. \end{cases}$$

Inverse NSCT involves reconstructing the original Image $\hat{I}$ from its coefficients, $Y_{j,l}^{d,k}$. It combines information from different scales, directions, and *subbands*.

$$\hat{I} = INSCT(Y_{j,l}^{d,k}).$$

The inverse transform process combines approximation and detailed information to reconstruct the original image.

4.1.2. Fast fourier transform (FFT). FFT is an algorithm used to efficiently compute the DFT. One of the most common formulations is the Cooley-Tukey Radix-2 algorithm. For simplicity, let us consider the case where N is a power of two.

$$X[k] = \sum_{n=0}^{N-1} x[n] \cdot W_N^{kn},$$

where $W_N = e^{-j\left(\frac{2\pi}{N}\right)}$ is the twiddle factor, and the Cooley-Tukey algorithm recursively breaks down the DFT computation into smaller DFTs until it reaches the base case of 2-point DFTs. Combined with factorization, this recursive decomposition significantly reduces the number of calculations compared with the direct computation of the DFT.

Inverse FFT transforms the frequency-domain representation back into a time-domain representation. For an N -point sequence $X[k]$, the IFFT is given by

$$x[n] = \frac{1}{N} \sum_{k=0}^{N-1} X[k] \cdot e^{j\left(\frac{2\pi}{N}\right)kn}.$$

A normalization factor $\frac{1}{N}$ ensures that the IFFT operation is consistent with the Fourier Transform definition.

4.1.3. Pulse-coupled neural network (PCNN). A PCNN is a type of neural network inspired by synchronization behavior observed in the visual system of the brain. It operates based on the principle of pulse synchronization and has been applied to various image-processing tasks because of its ability

to capture complex spatial relationships. In MRI fusion, the PCNN extracts the features from the input images. The network's ability to synchronize pulses among connected neurons makes it suitable for identifying the relevant image features and patterns.

The PCNN consists of interconnected neurons that fire pulses in response to specific features in input images. Neurons exhibit synchronization when exposed to similar features. The network has linking strength and inhibition parameters that influence its behavior.

$$\begin{aligned}
F_{ij}[n] &= e^{-\alpha_F} F_{ij}[n-1] + V_F \sum_{K,l} W_{i,j,K,l} Y_{ij}[n-1] + S_{ij}, \\
L_{ij}[n] &= e^{-\alpha_L} L_{ij}[n-1] + V_L \sum_{K,l} M_{i,j,K,l} Y_{ij}[n-1], \\
U_{ij}[n] &= F_{ij}[n] (1 + \beta L_{ij}[n]), \\
T_{ij}[n] &= e^{-\alpha_T} T_{ij}[n-1] + V_T Y_{ij}[n-1], \\
Y_{ij}[n] &= \begin{cases} 1, & \text{if } U_{ij}[n] > T_{ij}[n], \\ 0, & \text{otherwise,} \end{cases} \\
F(u, v) &= \left(\sum_{x=0}^{N-1} \sum_{y=0}^{M-1} T(x, y) \cdot e^{-2\pi i(ux/N + vy/M)} \right)'.
\end{aligned}$$

For image fusion, the PCNN extracts features from different modalities. The network is exposed [24] to the input images and neuron fire pulses in response to features such as edges, textures, or other distinctive patterns.

Synchronized pulses from the PCNN were used to formulate fusion rules. These rules may involve combining information from different modalities based on synchronized features identified by the network.

The features extracted by the PCNN are integrated into the overall fusion process. Synchronized pulses may influence the selection of the relevant information from each modality to be included in the fused image.

4.1.4. Fuzzy rule-based fusion. We define fuzzy rules by combining the information from different modalities. Consider the pixel intensities T_{1ij} and T_{2ij} at the exact spatial locations i, j in the T1-weighted and T2-weighted images.

Rule 1: If $\mu_A(T_{1ij})$ and $\mu_B(T_{2ij})$, then $F_{ij} = C$ where $\mu_A(T_{1ij})$ and $\mu_B(T_{2ij})$ are the membership degrees of T_{1ij} and T_{2ij} in fuzzy sets A and B , and F_{ij} is the fused intensity at locations i, j .

Membership functions for fuzzy sets A, B, and C are defined here. For example, represent low, medium, and high intensities, respectively.

$$\mu_A(x) = \begin{cases} \frac{(x - x_{low})}{(x_{medium} - x_{low})}, & \text{if } x_{low} \leq x \leq x_{medium}, \\ 1, & \text{if } x_{medium} < x \leq x_{high}, \\ 0, & \text{otherwise.} \end{cases}$$

By combining the membership degrees using fuzzy logic operators, we consider the antecedent part of the rule

$$\text{Degree of activation} = \min(\mu_A(T_{1ij}), \mu_B(T_{2ij})).$$

The fuzzy outputs were combined to obtain a crisp result. For example, using the centroid de-fuzzification method,

$$F_{ij} = \frac{\sum \text{Universe of discourse } x \times \text{Degree of membership}}{\sum \text{Universe of discourse Degree of membership}}.$$

Reconstruct the fused Image by assigning the de-fuzzified values to the corresponding spatial locations.

$$\text{Fused Image } (i, j) = F_{ij}.$$

5. Results and discussion

5.1. Dataset

In this study, ADNI is used as a dataset for the proposed fusion approach. ADNI 3 was launched in 2003 as a public-private partnership led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI is to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessments can be combined to measure the progression of CN, MCI, EMCI, LMCI, SMC, and AD. This study used 250 samples from each multimodal image including MRI, PET, fMRI, and DTI. A total of 1000 image samples were used for the image-fusion approach.

The results of the FGPCNN algorithm [5] by considering the fusion of MRI, PET, fMRI, and DTI Images: These used to apply the FGPCNN algorithm to the multimodal fusion framework. This section describes the accuracy, sensitivity, and specificity of the proposed FGPCNN model.

In this study, ADNI is used. ADNI is a multisite study that aims to improve clinical trials for the prevention and treatment of AD. This cooperative study combines expertise and funding from the private and public sectors to study subjects with AD, as well as those who may develop AD, and controls with no signs of cognitive impairment. ADNI researchers collect, validate, and utilize data including MRI, fMRI, PET, and DTI images. MRI can detect blood flow and vascular malformation. PET images allow doctors to view not only the structure of the brain but also its functions.

Table 3. Dataset consideration

Test cases	Multi modality dataset	
	No of training images	No of test images
MRI	250	250
PET	250	250
fMRI	250	250
DTI	250	250
Fused	1000	1000

fMRI provides a measure of changes in blood oxygen levels and functional connectivity between different brain regions. DTI provides a representation of the white matter fiber bundles in the brain. This study used 250 samples from each multi modality images including MRI, PET, fMRI, and DTI.

A total of 1000 image samples were used for the image-fusion approach. Table 3 shows the data considerations.

a. Data Preparation: Pre-processing techniques, such as noise reduction, intensity normalization, spatial registration, and skull stripping, were applied to MRI, PET, fMRI, and DTI images. These processed images were used as the inputs for the FGPCNN algorithm.

b. Training: A labeled dataset of patients with AD and healthy controls was used to train the FGPCNN algorithm. The system learns to extract pertinent information from multimodal images and makes predictions by using these features. During training, methods such as gradient descent and backpropagation optimize network parameters.

c. Testing: A different testing dataset is used to gauge the efficacy of the trained FGPCNN algorithm for AD detection. The predicted labels were compared to the ground-truth labels of the testing dataset to determine the accuracy, sensitivity, and specificity of the algorithm.

d. Performance Evaluation: The performance of the FGPCNN algorithm was evaluated using common evaluation criteria. The metrics are:

Accuracy: The percentage of healthy controls and AD patients accurately classified.

The sensitivity, called the true positive rate, is the capacity of the algorithm to accurately identify patients.

Specificity: The capacity of the algorithm to accurately identify healthy controls is also referred to as the true-negative rate. Tables 4-6 show the parameter settings of NSCT, FFT, and PCNN, respectively.

Table 4. Parameter setting of NSCT

Parameter	Values
Decomposition levels (L)	3
Directional decomposition levels (N)	Laplacian pyramid
Filter bank length	8
Filter bank parameters	zero-padding
Boundary handling	2
Down sampling factor	threshold=0.1, method=soft threshold
Thresholding parameters	True
Normalization	linear
Interpolation method	Laplacian pyramid

Table 5. Parameter setting of FFT

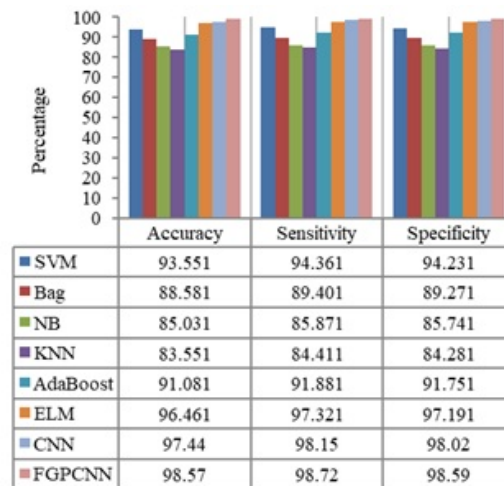
Parameter	Values
Windowing Function	rectangular window
Zero-padding	Power of 2
FFT Algorithm	Radix-2 FFT
Scaling	1/N
Filter Size	5x5
Border Handling	Zero-padding

Table 6. Parameter setting of PCNN

Parameter	Values
Thresholds	0.02
Connection Weights	0.04
Neuron Types	Excitatory
Pulse Parameters	0.05
Inhibition Parameters	0.05
Integration Time	0.03
Spatial Neighborhood Size	2
Time Neighborhood Size	4
Iterations	100
Boundary Conditions	Periodic
Activation Function	Sigmoid function
Initialization	Zero initialization
Learning Rate	0.03

Figures 4-7 show the performances of the FGPCNN Algorithm when considering MRI, PET, fMRI, and DTI images, respectively. Figure 8 shows the FGPCNN Algorithm Results obtained by considering the fused images. Here, the fused image performance is demonstrated by considering different classifiers such as SVM, Bag, Naïve Bayes, KNN, Adaboost, ELM, CNN, and FGPCNN algorithms.

The findings of the FGPCNN algorithm demonstrate the success of the multimodal fusion framework in detecting AD. Compared with single-modality methods, the algorithm obtains improved accuracy, sensitivity, and specificity, demonstrating its capacity to extract and utilize complementary information from MRI, PET, fMRI, and DTI data.

**Fig. 4.** FGPCNN algorithm results by considering MRI images

A comparative analysis was performed to illustrate the benefits of the FGPCNN and multimodal fusion frameworks. The performance measures of the FGPCNN algorithm were juxtaposed with those of single-modality methods including DTI, PET, fMRI, and MRI-based detection models. This comparison highlights how the multimodal fusion strategy improves AD detection accuracy.

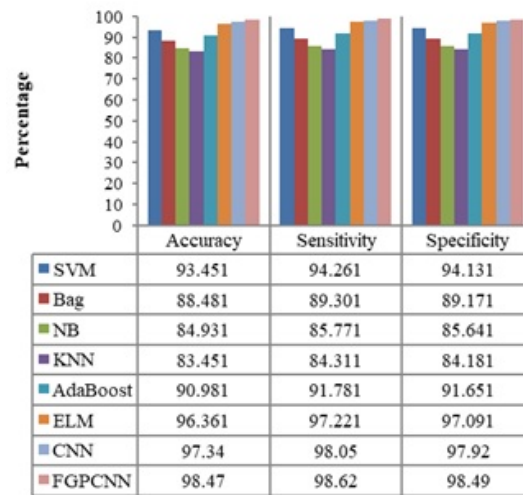


Fig. 5. FGPCNN algorithm results by considering PET images

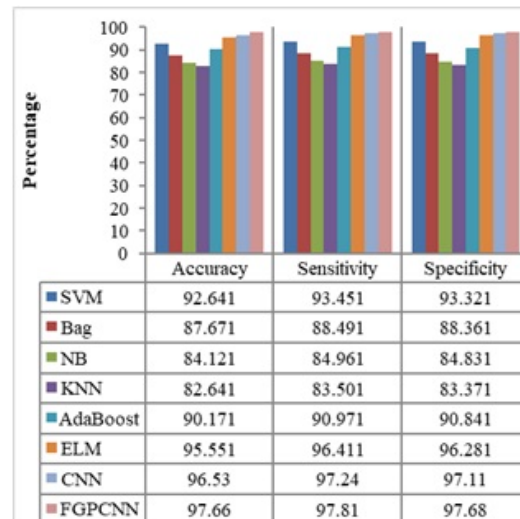


Fig. 6. FGPCNN algorithm results by considering fMRI images

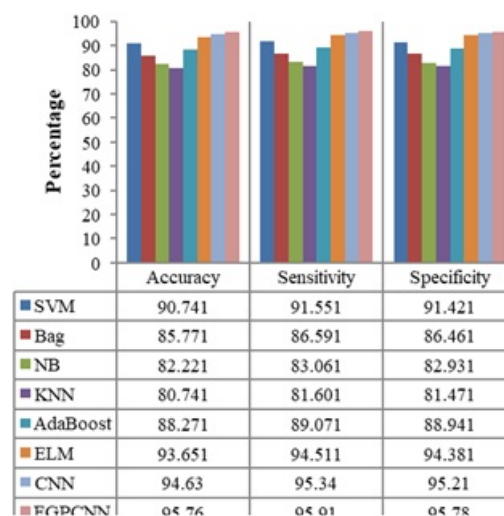


Fig. 7. FGPCNN algorithm results by considering DTI images

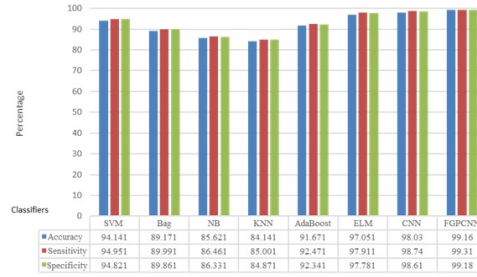


Fig. 8. FGPCNN algorithm results by considering Fused images

5.2. Discussion

5.2.1. Comparative analysis of proposed method with existing multimodal approaches. Table 7 compares the performance of the proposed method with existing Multimodality models for the dataset using accuracy and error rate as the two-evaluation metrics. Figure 9 shows the ROC Curve of the proposed model.

Table 7. Analysis of proposed method with existing multimodal approaches

Parameters used	Average accuracy (%)	Average error rate (%)
Fourozannezhad et al. [7]	69.5	30.05
Hao et al. [10]	73.6	26.4
Shao et al. [22]	75.5	24.50
Odusami et al. [20]	94.32	5.68
Proposed	98.94	1.06

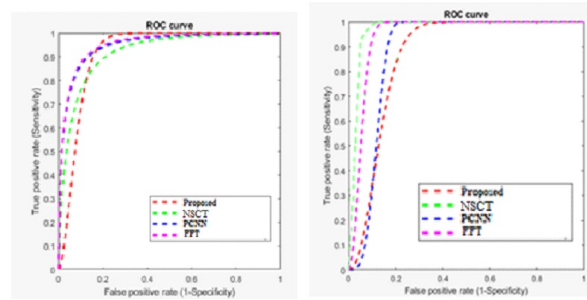


Fig. 9. ROC curve of proposed Model

Based on the ROC curve, the proposed methodology outperforms the other methods (NSCT, PCNN, and FFT) in terms of accuracy and sensitivity. The proposed method achieves a higher true positive rate at lower false positive rates, indicating a more effective image fusion strategy. The PCNN method shows better performance than NSCT and FFT but still falls short compared to the proposed method. The FFT method exhibits the least effective performance, demonstrating the importance of the advanced fusion techniques used in the proposed method for achieving superior results in multimodal image fusion.

The ROC curve illustrates the comparative performance of four image fusion methods: the proposed method, NSCT, PCNN, and FFT. The NSCT method (green dashed line) demonstrates the

highest sensitivity across all false positive rates, indicating the best overall performance in distinguishing true signals from false ones. The proposed method (red dashed line) follows closely, showing a strong performance but slightly lower than NSCT. The PCNN method (pink dashed line) performs moderately well, while the FFT method (blue dashed line) shows the least effective performance, indicating that it is the least capable of correctly identifying true positives. Overall, the NSCT method is superior in sensitivity, while the proposed method remains competitive.

5.2.2. Confusion matrix of proposed approach . The performance of the proposed approach was evaluated using a confusion matrix as the metric. The dataset consisted of two classes: Alzheimer 's and normal. Therefore, a 2×2 confusion matrix was used. Figure 10 shows the confusion matrices obtained for different numbers of Alzheimer and normal samples.

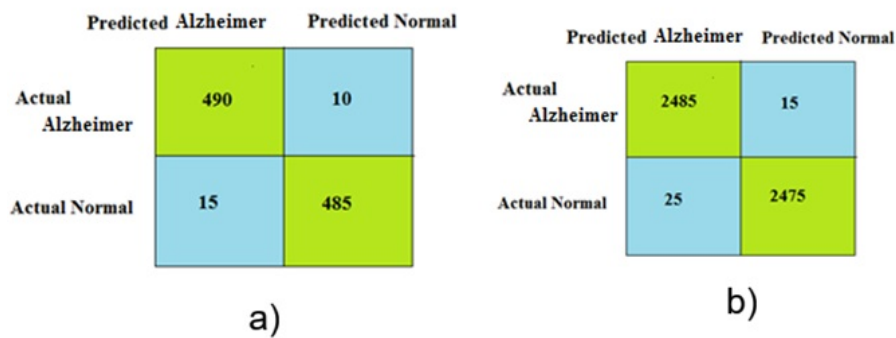


Fig. 10. Results of confusion matrices (a) The confusion matrix shows the performance of the model on 1000 samples. (b) The confusion matrix shows the performance of the model on 5000 samples

6. Limitation & future enhancement

Although the proposed approach provides best result, it has some limitation. FFT focuses on frequency domain analysis, which may not capture all relevant information for AD detection. Fuzzy rule-based approaches often rely on domain experts to define the rules and membership functions, which can be subjective and may not capture all relevant information in the data. PCNN methods struggle to capture non-linear relationships and interactions between different features, limiting their ability to accurately model the complex nature of AD. To overcome this limitation, in future, the sample size is increased from each data set to evaluate the performance of this proposed approach. Instead of PCNN and fuzzy rule, deep convolution neural network will be considered in future to reduce complexity. To improve the fusion performance, the FFT is replaced by curvelet transform in future.

7. Conclusion

Automated multimodal medical image fusion frameworks using deep learning techniques have shown promising results in Alzheimer's detection. These frameworks leverage complementary information provided by different imaging modalities to enhance the accuracy and reliability of AD diagnosis. Although each framework has its strengths and limitations, they collectively demonstrate the potential of deep learning-based fusion models for improving AD detection. Further research can be conducted, as the collection of datasets is somewhat complex; hence, a dataset with all images along

with information regarding the sex, age, and stage of AD must be mentioned, which could help researchers. In addition, the pretrained deep learning networks were trained using thousands of images for specific purposes. The possibility of training deep neural networks with more images can be explored with increased computational capabilities by using GPUs. In future, the sample size should be increased for each dataset to evaluate the performance of the proposed approach. Instead of the PCNN and fuzzy rule, a deep convolutional neural network will be considered in the future to reduce complexity. To improve the fusion performance, FFT will be replaced by a curvelet transform in the future.

References

- [1] M. A. Azam, K. B. Khan, S. Salahuddin, E. Rehman, S. A. Khan, M. A. Khan, S. Kadry, and A. H. Gandomi. A review on multimodal medical image fusion: compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics. *Computers in Biology and Medicine*, 144:105253, 2022. <https://doi.org/10.1016/j.combiomed.2022.105253>.
- [2] P. Balaji, M. A. Chaurasia, S. M. Bilfaqih, A. Muniasamy, and L. E. G. Alsid. Hybridized deep learning approach for detecting alzheimer's disease. *Biomedicines*, 11(1):149, 2023. <https://doi.org/10.3390/biomedicines11010149>.
- [3] G. Battineni, M. A. Hossain, N. Chintalapudi, G. Nittari, C. Ruocco, E. Traini, and F. Amenta. Brain imaging studies using deep neural networks in the detection of alzheimer's disease. *OBM Geriatrics*, 7(1):1–10, 2023. <http://dx.doi.org/10.21926/obm.geriattr.2301220>.
- [4] S. Beatrice and J. Meena. Overhauled approach to effectuate the amelioration in eeg analysis. *Intelligent Automation & Soft Computing*, 33(1):331–347, 2022. <http://dx.doi.org/10.32604/iasc.2022.023666>.
- [5] Beatrice, S, J. Meena, and M. Earliest alzheimer detection utilizing every potentiality. *Manuscript submitted for publication*, 2023.
- [6] A. Benba and A. Kerchaoui. Automatic detection of alzheimer's disease based on artificial intelligence. *Informatyka, Automatyka, Pomiar w Gospodarce i Ochronie Środowiska*, 13(1):18–21, 2023. <https://doi.org/10.35784/iapgos.3383>.
- [7] P. Forouzaneshad, A. Abbaspour, C. Li, M. Cabrerizo, and M. Adjouadi. A deep neural network approach for early diagnosis of mild cognitive impairment using multiple features. In *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pages 1341–1346. IEEE, 2018. <https://doi.org/10.1109/ICMLA.2018.00218>.
- [8] S. K. GU, P. Anurenjan, and K. Sreeni. Detecting alzheimer's disease using multi-modal data: an approach combining transfer learning and ensemble learning. In *2023 International Conference on Control, Communication and Computing (ICCC)*, pages 1–6. IEEE, 2023. <https://doi.org/10.1109/ICCC57789.2023.10165454>.
- [9] K. Guo, X. Hu, and X. Li. Mmfgan: a novel multimodal brain medical image fusion based on the improvement of generative adversarial network. *Multimedia Tools and Applications*, 81(4):5889–5927, 2022. <https://doi.org/10.1007/s11042-021-11822-y>.
- [10] X. Hao, Y. Bao, Y. Guo, M. Yu, D. Zhang, S. L. Risacher, A. J. Saykin, X. Yao, L. Shen, and A. D. N. Initiative. Multi-modal neuroimaging feature selection with consistent metric constraint for diagnosis of alzheimer's disease. *Medical Image Analysis*, 60:101625, 2020. <https://doi.org/10.1016/j.media.2019.101625>.

- [11] R. A. Hazarika, A. K. Maji, D. Kandar, E. Jasinska, P. Krejci, Z. Leonowicz, and M. Jasinski. An approach for classification of alzheimer's disease using deep neural network and brain magnetic resonance imaging (mri). *Electronics*, 12(3):676, 2023. <https://doi.org/10.3390/electronics12030676>.
- [12] H. Hermessi, O. Mourali, and E. Zagrouba. Multimodal medical image fusion review: theoretical background and recent advances. *Signal Processing*, 183:108036, 2021. <https://doi.org/10.1016/j.sigpro.2021.108036>.
- [13] Home. Alzheimer's disease and dementia, 2016. <https://www.alz.org>.
- [14] W. N. Ismail, F. R. PP, and M. A. Ali. A meta-heuristic multi-objective optimization method for alzheimer's disease detection based on multi-modal data. *Mathematics*, 11(4):957, 2023. <https://doi.org/10.3390/math11040957>.
- [15] M. A. Khan, I. Ashraf, M. Alhaisoni, R. Damaševičius, R. Scherer, A. Rehman, and S. A. C. Bukhari. Multimodal brain tumor classification using deep learning and robust feature selection: a machine learning application for radiologists. *Diagnostics*, 10(8):565, 2020. <https://doi.org/10.3390/diagnostics10080565>.
- [16] A. A. A. El-Latif, S. A. Chelloug, M. Alabdulhafith, and M. Hammad. Accurate detection of alzheimer's disease using lightweight deep learning model on mri data. *Diagnostics*, 13(7):1216, 2023. <https://doi.org/10.3390/diagnostics13071216>.
- [17] S. Maqsood, R. Damasevicius, J. Siłka, and M. Woźniak. Multimodal image fusion method based on multiscale image matting. In *International Conference on Artificial Intelligence and Soft Computing*, pages 57–68. Springer, 2021. https://doi.org/10.1007/978-3-030-87897-9_6.
- [18] E.-G. Marwa, H. E.-D. Moustafa, F. Khalifa, H. Khater, and E. AbdElhalim. An mri-based deep learning approach for accurate detection of alzheimer's disease. *Alexandria Engineering Journal*, 63:211–221, 2023. <https://doi.org/10.1016/j.aej.2022.07.062>.
- [19] S. R. Muzammil, S. Maqsood, S. Haider, and R. Damaševičius. Csid: a novel multimodal image fusion algorithm for enhanced clinical diagnosis. *Diagnostics*, 10(11):904, 2020. <https://doi.org/10.3390/diagnostics10110904>.
- [20] M. Odusami, R. Maskeliūnas, R. Damaševičius, and S. Misra. Explainable deep-learning-based diagnosis of alzheimer's disease using multimodal input fusion of pet and mri images. *Journal of Medical and Biological Engineering*, 43(3):291–302, 2023. <https://doi.org/10.1007/s40846-023-00801-3>.
- [21] S. K. Sethuraman, N. Malaiyappan, R. Ramalingam, S. Basheer, M. Rashid, and N. Ahmad. Predicting alzheimer's disease using deep neuro-functional networks with resting-state fmri. *Electronics*, 12(4):1031, 2023. <https://doi.org/10.3390/electronics12041031>.
- [22] W. Shao, Y. Peng, C. Zu, M. Wang, D. Zhang, and A. D. N. Initiative. Hypergraph based multi-task feature selection for multimodal classification of alzheimer's disease. *Computerized Medical Imaging and Graphics*, 80:101663, 2020. <https://doi.org/10.1016/j.compmedimag.2019.101663>.
- [23] S. Shojaei, M. S. Abadeh, and Z. Momeni. An evolutionary explainable deep learning approach for alzheimer's mri classification. *Expert Systems with Applications*, 220:119709, 2023. <https://doi.org/10.1016/j.eswa.2023.119709>.
- [24] J. Song, J. Zheng, P. Li, X. Lu, G. Zhu, and P. Shen. An effective multimodal image fusion method using mri and pet for alzheimer's disease diagnosis. *Frontiers in Digital Health*, 3:637386, 2021. <https://doi.org/10.3389/fdgth.2021.637386>.

-
- [25] S. Suganyadevi, A. S. Rajasekaran, N. Satheesh, R. Suganthi, and R. Naveenkumar. Alzheimer's disease diagnosis using deep learning approach. In *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)*, pages 1205–1209. IEEE, 2023. <https://doi.org/10.1109/ICEARS56392.2023.10085017>.
 - [26] M. Wasim. Alzheimer's disease detection: a deep learning-based approach. *Journal of Independent Studies and Research Computing*, 21(1):19–27, 2023. <https://doi.org/10.31645/JISRC.23.21.1.3>.
 - [27] S. Zhang, D. Chen, R. Ranjan, H. Ke, Y. Tang, and A. Y. Zomaya. A lightweight solution to epileptic seizure prediction based on eeg synchronization measurement. *The Journal of Supercomputing*, 77:3914–3932, 2021. <https://doi.org/10.1007/s11227-020-03426-4>.
 - [28] T. Zhang and M. Shi. Multi-modal neuroimaging feature fusion for diagnosis of alzheimer's disease. *Journal of Neuroscience Methods*, 341:108795, 2020. <https://doi.org/10.1016/j.jneumeth.2020.108795>.