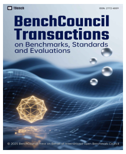


BenchCouncil Transactions on Benchmarks, Standards and Evaluations, 2026

DOI: https://doi.org/10.66834/798kjf21
Research Article

RESEARCH ARTICLE

Design and Evaluation of an Interpretable Multimodal Deep Learning Framework for Early Alzheimer's Disease Detection

Shehu Mohammed,[1,*] Neha Malhotra[1] and Anmol Singh Rai[2]

[1]School of Computer Applications, Lovely Professional University, Phagwara, India and [2]Shrimann Superspeciality Hospitals, Jalandhar, India
[*]Corresponding author. mohammedshehumafara@gmail.com
Received on 22 December 2025; Accepted on 24 June 2026

Abstract

Alzheimer's disease is a progressive neurodegenerative disorder that significantly impairs memory and cognitive functions and affects over 55 million people worldwide. The successful management and planning require early and accurate diagnosis. Conventional radiological assessment is often subjective and time-consuming, which highlights the need for automated and reliable diagnostic solutions. Most deep learning models show promise for classifying neuroimaging data, but they tend to be less computationally efficient and less interpretable, and they cannot be integrated into patient-centric processes. The gap between developing diagnostic algorithms with high accuracy and implementing them in a supportive framework that includes patients and caregivers is large. This paper introduces a comprehensive, hybrid framework that addresses these gaps. We present a dual-modality diagnostic system: a deep learning pipeline using EfficientNetV2-S for CT scan classification, complemented by a Feedforward Neural Network (FNN) that analyses structured clinical data for holistic patient assessment. This diagnostic core is integrated into a user-friendly graphical user interface (GUI) and supplemented by "NeuroBot," an AI-powered chatbot that provides domain-specific information and support. The two models have been trained using the transfer learning method on a curated dataset of 30,000 brain CT slices. The EfficientNetV2-S model achieved an accuracy of 98.19%. After hyperparameter tuning, the FNN model achieved an optimised accuracy of 87.21%. The importance of the features addressed by the models was proved with the help of the statistical t-tests of the corresponding clinical data. The integrated system enables a scalable, translatable, and patient-centered system to improve the early analysis and treatment of Alzheimer's disease.

Key words: Alzheimer's Disease, CT Scan Classification, EfficientNetV2-S, Deep Learning, AI Chatbot, Medical Imaging, Transfer Learning

1. Introduction

Alzheimer's disease (AD) is the most prevalent form of dementia and primarily affects the aging population, accounting for approximately 60-70% of dementia cases worldwide. It is a neurodegenerative, advanced, and progressive disorder that slowly affects the cognitive functions, starting with the memory and then with the dysfunction of judgment, language, and behaviour. The disease has a significant social impact, placing substantial emotional and financial burdens on patients, families, and healthcare systems worldwide. Early and accurate diagnosis is therefore critical important. The timely diagnosis will enable patients and the individuals who provide care to investigate the treatment interventions, develop useful long-term care plans, and eventually improve the quality of life of patients as the disease advances[1, 2].

The neuroimaging procedures are the core of the AD diagnostics process, and they allow clinicians to observe structural changes of the brain typical of the disease, e.g., cortical atrophy and ventricular enlargement. Among these techniques, Computed Tomography (CT) scans serve as a vital tool. CT imaging is particularly valuable due to its rapid acquisition time, widespread availability, and relative cost-effectiveness compared to other modalities like Magnetic Resonance Imaging (MRI), making it a cornerstone of initial neurological workups, especially in resource-limited settings[3]. However, the conventional analysis of these scans is not without its challenges. The diagnostic process is a subjective one and depends on the radiologist's interpretation, and this cannot be adequately performed in a short time, besides being prone to inter-rater inconsistency and human error, especially in detecting the changes at the initial stages of AD.

These diagnostic limitations have been a primary driver for the development of computer-aided diagnostic (CAD) systems, a field that has been revolutionized by the advent of deep learning. As extensively documented in systematic reviews, both machine learning and deep learning techniques have demonstrated profound success in the automated analysis of medical images[4]. CNNs have been exceptionally skilled at this task, in particular. Their architecture allows them to automatically learn and identify complex, hierarchical patterns within visual data, enabling the detection of subtle pathological indicators in MRI and brain CT slices that may be imperceptible to the human eye[5, 6]. This capability has spurred the creation of numerous advanced models, including various hybridized deep learning approaches, all aimed at pushing the boundaries of diagnostic accuracy and reliability[7].

Despite this remarkable progress, a significant gap persists between the development of high-accuracy algorithms in a research setting and their practical deployment in clinical workflows. Many state-of-the-art CNN models face considerable hurdles, including the need for substantial computational resources, which limits their real-time application. Moreover, the inherent "black-box" nature of many deep learning models can be a barrier to clinical adoption; for a diagnosis to be trusted, clinicians must be able to understand and verify the reasoning behind it, a critical factor in differential diagnosis[8]. Beyond the technical challenges, most existing research has focused almost exclusively on the diagnostic algorithm itself. This narrow focus often neglects the broader clinical ecosystem and the critical need for an integrated, user-friendly system that not only provides a diagnosis but also supports patients and caregivers with accessible, contextual information[9].

To overcome these complex issues, this paper proposes a new and unique hybrid framework that combines the high-performance diagnostic engine with an interactive patient support system that is based and comprehensive. Our primary contribution is the development of a dual-model deep learning pipeline that leverages a powerful and efficient EfficientNetV2-S architecture for the classification of brain CT slices with exceptional accuracy and computational efficiency[10]. The core novelty of our work lies in embedding these performant and interpretable models, which utilize Grad-CAM for visual explanations, within a complete, end-to-end ecosystem. This system includes an intuitive graphical user interface (GUI) for seamless interaction by clinicians and an AI-powered chatbot, NeuroBot, designed to answer AD-related queries from patients and caregivers. By moving beyond mere classification, this holistic approach creates a practical, scalable, and supportive tool designed to enhance early Alzheimer's detection and improve the overall standard of patient care[11]. Furthermore, to capture non-visual risk factors and cognitive metrics that are crucial for a comprehensive diagnosis, our framework integrates a secondary pathway that leverages a robust FNN to analyze structured patient clinical records.

Although many deep learning models have achieved high diagnostic accuracy for Alzheimer's disease detection, most studies focus primarily on algorithm development and evaluation using experimental datasets. These models are rarely integrated into user-friendly systems that support clinical workflows, patient interaction, or caregiver guidance. As a result, there remains a gap between high-performance diagnostic algorithms and practical systems that can be deployed in real clinical environments. The proposed framework addresses this gap by integrating the diagnostic models within an interactive ecosystem that includes a graphical user interface and an AI-based assistant to support clinicians, patients, and caregivers.

The main contributions of this study can be listed as follows:

- Development of a dual-modality deep learning framework, which uses EfficientNetV2-S for CT image classification, and a Feedforward Neural Network for structured clinical data analysis.
- The application of explainable AI, which uses Grad-CAM for region highlighting on brain CT slices, thereby improving their interpretability.
- Development of a user-centric deployment platform, which includes a GUI and an AI-based chatbot, referred to as NeuroBot.
- The validation of clinical features through statistical analysis, which includes independent sample t-tests and correlation analysis to ascertain the significance of cognitive and lifestyle factors, such as MMSE and ADLs.

2. Literature Review

The ML and DL applied to the neurology practice have completely changed the Alzheimer Disease (AD) diagnostic process and provided meaningful and data-intuitive answers to the current practices. The evolution of these techniques has been rapid, moving from foundational models to highly sophisticated, specialized architectures. Early research in automated AD detection was primarily centered on classical machine learning algorithms. Models such as Support Vector Machines (SVMs), Random Forests, and Decision Trees were applied to neuroimaging data, but their efficacy was often constrained by a reliance on handcrafted feature extraction[12]. This process, which required extensive domain knowledge to manually define and extract relevant features like hippocampal volume or cortical thickness, was not only labor-intensive but also limited the models' ability to discover novel, complex patterns within the data. Despite the need to take such measures, the subsequent emergence of deep learning, and, particularly, the Convolutional Neural Networks (CNNs) became a game-changer in the matter. The CNNs have revolutionised the study of medical images; they enable end-to-end learning of hierarchical features beneath the raw pixel data. This capability has led to a proliferation of studies demonstrating the accurate prediction and diagnosis of AD using a variety of deep learning models, which consistently outperform their traditional ML counterparts[13, 14].

The current research in the area, however, has predominantly been interested in the application of deep learning to analyze neuroimaging images, and, more specifically, in Magnetic Resonance Imaging (MRI) and, more recently, in Computed Tomography (CT) scans. The authors have experimentally shown that applying advanced image processing and enhancement techniques before model training improves AD detection accuracy[15]. Recognizing that a single data source may not capture the complexity of AD pathology, numerous studies have investigated multimodal deep learning methods. These models integrate data from different neuroimaging modalities (e.g., structural MRI, functional MRI, and PET scans) to create a more comprehensive and robust diagnostic picture[16]. The sheer volume of research has produced a rich body of literature that systematically reviews the diverse array of AD detection techniques, highlighting the consistent and rapid progress in diagnostic accuracy and model sophistication[17]. Furthermore, the application of these powerful models has expanded beyond

AD to include the diagnosis of related neurodegenerative conditions, such as tauopathies, thereby demonstrating their broader versatility and clinical potential[18].

Since the field has become mature, the quest to achieve increasingly high accuracy has given rise to the emergence of new and complicated model architectures. Hybrid systems and ensemble-based systems are emerging to the fore. These advanced approaches combine the predictive strengths of multiple deep learning models to create more robust, reliable, and generalizable identification systems that are less prone to the biases of a single model[19, 20]. While neuroimaging remains the primary data source, some innovative models have been designed to diagnose AD based exclusively on structured patient clinical records, providing a valuable alternative or complementary diagnostic pathway that does not require imaging[21]. This has contributed to the development of sophisticated architectures capable not only of binary classification but also of detecting and differentiating between the various stages of AD, from early-stage Mild Cognitive Impairment (MCI) to advanced dementia, which is crucial for tailoring patient care[22, 23].

For these powerful deep learning models to transition from research laboratories to real-world clinical environments, two practical factors have become paramount: computational efficiency and model interpretability. Clinically viable diagnostic tools have been developed to continue to focus on the transfer learning application. The fine-tuning of large models, which have been trained on general image datasets, on smaller medical imaging datasets is a very effective way to achieve state-of-the-art performance with small datasets[24]. This strategy has been a central theme in the broader investigation of deep learning for enhancing early detection and supporting clinical decision-making[25, 26]. Consequently, a significant portion of modern research is focused on developing models for early detection, as this is the stage at which interventions are most likely to be effective[27]. To further improve performance, some researchers have proposed integrative models that combine the strengths of traditional ML with advanced deep learning architectures[28]. Finally, to address the "black-box" problem, there is a growing emphasis on creating explainable AI (XAI). The development of specialized, attention-based explainable networks and custom models, such as ADD-Net, underscores the field's commitment to creating diagnostic tools that are not only accurate and efficient but also transparent and trustworthy for clinical use[29, 30].

Although many researchers have reported positive results in detecting Alzheimer's disease using deep learning models in their literature, a wide range of methodological variations can be seen in their approaches. Most of the existing literature has focused on MRI-based deep learning models using CNN-based architectures and has achieved good accuracy using ADNI databases, ranging from 85% to 96%. However, a few researchers have also used multimodal data for the accurate prediction of Alzheimer's disease using MRI and PET scans. However, these techniques require expensive hardware. In addition, fewer researchers have focused on CT-based deep learning models for detecting Alzheimer's disease. Moreover, few researchers have focused only on prediction models without using interpretability mechanisms or deployment frameworks. Although various researchers have used various models, such as ADD-Net and attention-based models, to explain their models, these models are not used in deployment frameworks. However, in the proposed model, a CT-based deep learning model was used in conjunction with a structured clinical data model, an explainable AI model using a Grad-CAM mechanism, and a user-centric deployment model using a GUI and an AI chatbot support system[31].

3. Methodology and Experiment

The proposed framework integrates a dual-model deep learning pipeline for AD classification with an interactive user interface and an AI chatbot. The architecture is designed for accuracy, interpretability, and user engagement.

3.1. Dataset and System Architecture

The general procedure of our system is illustrated in Figure 1. The training and evaluation dataset consists of 10,240 brain CT slices, evenly distributed between Alzheimer's and non-demented cases. In addition, a clinical dataset containing 2,149 patient records was used for structured data analysis. Figure 2 presents the distribution of diagnoses in the clinical dataset, with 64.6% classified as Non-Demented and 35.4% as Demented. All the images were resized to an average of 224 × 224 pixels and normalised. Random rotations, horizontal flips, and zooming were used as data augmentation techniques to improve the model's robustness.

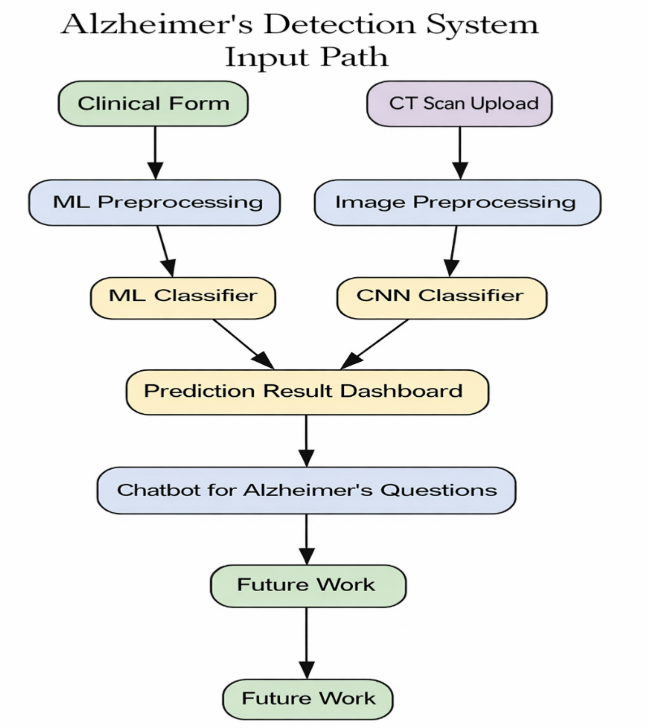


Figure 1. Alzheimer's Detection System Architecture.

A stringent, automated preprocessing pipeline was applied to the clinical data to prepare it for the FNN model. The first split of the features was into numerical (e.g., Age, MMSE) and categorical ones (e.g., Gender, Smoking status). One-hot encoding was then applied to all categorical features in order to put them in numerical form, and all numerical features were rescaled to a normalized range. This ensures that all features of the corresponding role in the prediction of the model are not affected by different scales. All this transformation process was stored as a single object of pipeline in order to make sure that

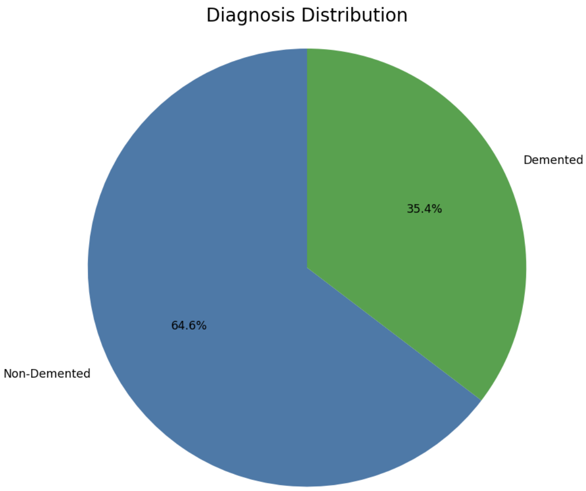


Figure 2. Diagnosis Distribution in the Clinical Dataset.

whenever new information is keyed in the prediction process, it would be processed similarly to the training information.

The CT imaging dataset used in this study consisted of approximately 30,000 brain CT slices derived from the Shrimaan Super Specialty Hospital repository and clinical datasets for Alzheimer's disease research. Of these, 10,240 CT slices were selected for training and evaluation purposes. The images were divided into training, validation, and testing subsets using a 70–10–20 split. In addition, an extended testing pool was used during the final evaluation, resulting in approximately 4,500 CT slices used for performance assessment. Each CT slice in the dataset was linked to a subject-level diagnosis label: Demented or Non-Demented.

To avoid data leakage in the model, the clinical dataset was split at the patient level instead of the slice level. This implies that brain CT slices for a specific patient were only included in either the training or testing dataset. Finally, the clinical dataset was split into 70% for training, 10% for validation, and 20% for testing.

The clinical dataset included 2,149 patient records, which included various physiological measures, cognitive measures such as Mini-Mental State Examination (MMSE), and lifestyle factors. For handling missing values in the clinical dataset, median imputation was applied for numerical features and most frequent imputation for categorical features. This particular process guaranteed the proper transformation of both training and testing data.

3.2. System Deployment Architecture

The framework for deployment is modular, designed to function in real-world environments. It has four components. The first is the user interface, the second is the inference engine, the third is the data management system, and the fourth is the AI chatbot module.

The primary interface with the system is the Graphical User Interface (GUI). In the web-based system, the user can input the CT scan of the brain or the clinical parameters. This is then sent to the model inference server, which then uses the EfficientNetV2-S and the FNN models to perform the prediction. The EfficientNetV2-S is used to perform the prediction with the CT scan, while the FNN is used with the clinical parameters.

There is also an AI chatbot module that interacts with the user in natural language to ask questions and educate the user about Alzheimer's disease. This module is known as the NeuroBot. The backend is designed to support an anonymized clinical dataset and to perform inference requests with the models.

The System deployment architecture of the proposed Alzheimer's detection framework is illustrated in Figure 3. The system integrates a web-based user interface with backend APIs, deep learning inference models (EfficientNetV2-S for brain CT slices and FNN for clinical data), and a chatbot module to support clinical decision-making and patient interaction.

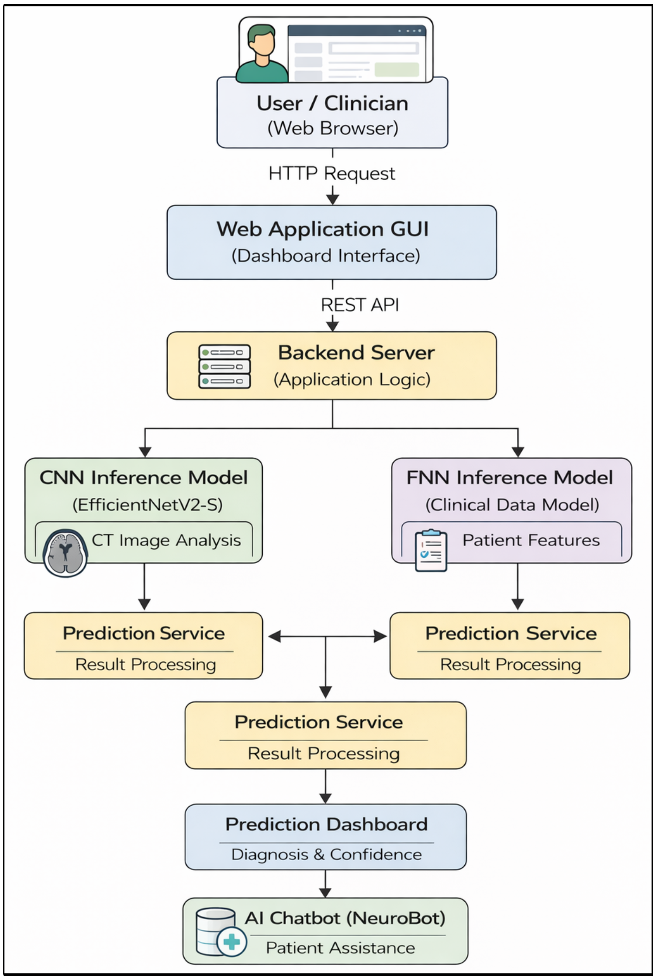


Figure 3. System Deployment Architecture Diagram.

3.3. Diagnostic Model Architectures

To create a comprehensive diagnostic tool, we employ two types of neural nets: Convolutional Neural Networks (CNNs) to process image data and Feedforward Neural Networks (FNNs) to process data based on clinical features.

3.3.1. Architectures for Image-Based Analysis (CNNs)

The fundamental operation in a CNN is the two-dimensional convolution. Convolution is a method of feature extraction, in which a kernel is applied to a given input image or feature map. This is arithmetically measured as:

$$O(i, j) = (I * K)(i, j) = \sum_{m} \sum_{n} I(i - m, j - n) \cdot K(m, n) \quad (1)$$

$O(i, j)$ characterises the output feature map of location (i, j), I is the input, and K is the convolutional kernel. Activation functionality is applied to add non-linearity. Our application is the ReLU, which is:

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

For the image classification task, this study utilizes the EfficientNetV2-S architecture, a powerful and computationally efficient model. This model achieves its efficiency through the use of Fused-MBConv blocks in its early layers.

Fused-MBConv block simplifies the normal inverted residual block by fusing the first 3×3 depthwise convolution of the block with the second 1×1 projection convolution into a standard 3×3 convolution block. This minimizes the production of memory access overhead increases and enhances the training speed on advanced accelerators. The work of a Fused-MBConv block could be explained by the equation:

$$y = \mathrm{BN}\left(\mathrm{Conv}_{3\times3}\left(\mathrm{BN}\left(\mathrm{Conv}_{1\times1}(x)\right)\right)\right) + x \quad (3)$$

In this case, x is a signal sent into the block, and the equation shows the series of a 1×1 expansion convolution, a 3×3 standard convolution (a replacement of the individual depthwise and projection steps) and Batch Normalization (BN). The end result y is gotten by summing the original input x via a residual (skip) connection, which makes this a distinguishing feature that makes it likely to train very deep networks.

3.3.2. Architecture for Clinical Data Analysis (FNN)

A Feedforward Neural Network (FNN) was developed to analyze structured clinical data. This kind of network was trained using a hyperparameter optimization strategy that entailed the optimization of the predictive performance using the Optuna framework.

The FNN is made up of a series of dense (fully connected) layers. The general unit is its dense layer, which is a linear transform of its input vector x, which can be expressed as:

$$y = Wx + b \quad (4)$$

The input is represented by y, the learnable weight matrix is W, and the learnable bias is b.

The result of each of the dense layers is then run through the activation function of the Rectified Linear Unit (ReLU) to introduce non-linearity, as well as permit the model to learn more intricate patterns, and is defined as:

$$f(x) = \max(0, x) \quad (5)$$

Where γ and β are learnt scale and shift parameters, and the constant ϵ is a small number to maintain numerical stability. This is followed by the normalized output going through the (ReLU) activation function.

To avoid overfitting, a ReLU activation will be followed by a Dropout layer. Dropout is random and assigns a part of the input units to 0 at every update time throughout the training period, which assists in augmenting the strength of the network.

The optimized architecture that was discovered by hyperparameterOptimization is an input layer and two hidden layers:

1. The initial hidden layer is a dense layer that contains 57 units, an activation of ReLU, and a Dropout with a rate of 0.38.
2. The second hidden layer is a dense layer with 129 units, and the next layer consists of a ReLU activation and a Dropout layer at a rate of 0.49.

The concluding component is a single output neuron that produces a raw logit, z. This logit is then transformed into a probability using the sigmoid activation function. For improved numerical stability during training, this function is integrated directly into the loss function (BCEWithLogitsLoss). The sigmoid function is defined as:

$$\sigma(z) = \frac{1}{1 + e^{-z}} \quad (6)$$

3.4. Interactive Framework and Training Algorithm

The project's workflow, executed within a Jupyter Notebook, provides two distinct diagnostic pathways: one for image-based analysis and another for clinical feature-based analysis. The classification pipeline for the image-based pathway is detailed in Algorithm 1, while the corresponding process for the clinical data FNN model is outlined in Algorithm 2.

Algorithm 1 AD Classification Pipeline (Image-Based CNN)

```
Input: Raw brain CT image $I_{\text{raw}}$.
Data Loading: The image dataset is loaded from the data/image directory and split into training and validation sets.
Preprocessing & Augmentation:
    $I_{\text{resized}} \Leftarrow \text{RandomResizedCrop}(I_{\text{raw}}, (224, 224))$
    $I_{\text{flipped}} \Leftarrow \text{RandomHorizontalFlip}(I_{\text{resized}})$
    $I_{\text{tensor}} \Leftarrow \text{ToTensor}(I_{\text{flipped}})$
    $I_{\text{norm}} \Leftarrow \text{Normalize}(I_{\text{tensor}})$
Model Loading: Load pre-trained model $M$ (EfficientNetV2-S) and adapt its final layer for binary classification.
Training: The model is fine-tuned on the training set using an Adam optimizer and Binary Cross-Entropy with Logits loss (BCEWithLogitsLoss).
Prediction:
    $Z \Leftarrow M(I_{\text{norm}})$ // Get raw logit output from the CNN
Classification:
    $P \Leftarrow \sigma(Z)$ // Apply sigmoid function to get probability
    If $P > 0.5$, then $C \Leftarrow$ 'Demented'
    Else $C \Leftarrow$ 'Non-Demented'
Output: Return class label $C$.
```

Algorithm 2 Clinical Data Classification Pipeline (FNN)

```
Input: Raw clinical feature set $D_{\text{raw}}$ from alzheimers_disease_data.csv.
Data Splitting: The dataset is split into training and testing sets (80/20 split).
Preprocessing:
    A ColumnTransformer pipeline is fitted on the training set:
        Numerical features are imputed (median) and standardized (StandardScaler).
        Categorical features are imputed (most frequent) and one-hot encoded (OneHotEncoder).
    $D_{\text{train\_processed}} \Leftarrow$ Apply fitted pipeline to training data.
    $D_{\text{test\_processed}} \Leftarrow$ Apply fitted pipeline to test data.
Class Imbalance Handling:
    $D_{\text{train\_resampled}} \Leftarrow$ Apply SMOTE (Synthetic Minority Over-sampling Technique) to $D_{\text{train\_processed}}$ to balance the classes.
Model Loading: Load the pre-trained Feedforward Neural Network (FNN).
Tensor Conversion:
    $T_{\text{input}} \Leftarrow \text{ToTensor}(D_{\text{test\_processed}})$
Prediction:
    $Z \Leftarrow M_{\text{fnn}}(T_{\text{input}})$ // Get raw logit output from the FNN
Classification:
    $P \Leftarrow \sigma(Z)$ // Apply sigmoid function to get probability
If $P > 0.5$, then $C \Leftarrow$ 'Demented'
Else $C \Leftarrow$ 'Non-Demented'
Output: Return class label $C$.
```

4. Results and Discussion

The efficacy of the hybrid framework that was suggested was critically evaluated using the aid of an integrated approach. This was accompanied by a quantitative measure of the classification performance of the deep learning models using a held-out test set of 4,500 brain CT slices, a large-scale measure of the associated clinical data to validate the validity of the features of the models, as well as a qualitative measure of the user interface and the interactivity aspects.

4.1. Performance Metrics

We used quantitative measures of the models using a sample set of standard classification measures that is decided by the elements of the confusion matrix: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN).

- **Accuracy:** The percentage of all the correct predictions.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

- **Precision:** The proportion of positive predictions that were actually correct.

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

- **Recall (Sensitivity):** The proportion of actual positives that were identified correctly.

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

- **F1-Score:** The harmonic mean of Precision and Recall, providing a single score that balances both.

$$F1\text{-}Score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{10}$$

4.2. Model Interpretability and Performance Evaluation

Besides the model's accuracy, interpretability is another critical requirement for clinical decision-support system. In this study, the interpretability of the model was facilitated by the use of a technique called Gradient-weighted Class Activation Mapping (Grad-CAM) in the EfficientNetV2-S CNN model. This technique enables the production of visual maps that highlight areas of the images used in the model that are being used to make a prediction, as illustrated in Figure 4. These maps allow verification that the model focuses on anatomically relevant regions associated with Alzheimer's disease, such as cortical atrophy and ventricular enlargement.

Although Grad-CAM visualizations, statistical validation, and feature-importance analysis provide valuable insights into the model's decision-making process. Future work will investigate quantitative explainability metrics, including localization-based evaluation and clinician-assisted validation of explanation maps, to provide a more rigorous assessment of model interpretability and clinical trustworthiness.

Moreover, the interpretability of the clinical data pathway was facilitated by the use of statistical tests, such as independent sample t-tests and correlation tests, which showed that clinical variables, such as MMSE, Activities of Daily Living (ADL), and functional assessment, significantly vary across the different diagnostic groups.

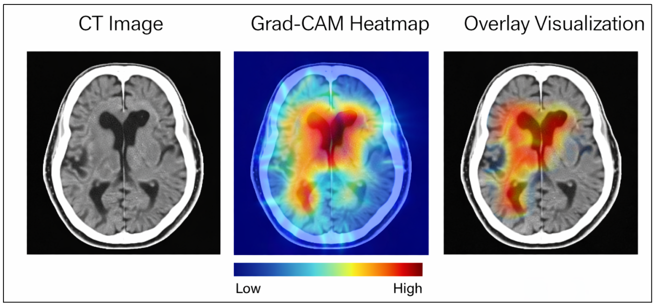


Figure 4. Grad-CAM visualization highlighting important regions in brain CT slices used by the EfficientNetV2-S model for Alzheimer's disease classification. The highlighted areas indicate regions contributing most strongly to the model's prediction.

The EfficientNetV2-S model achieved the best validation performance with an accuracy of 98.19% compared to other architectures, compared with other evaluated architectures, including EfficientNetV2-B0, ResNet50V2, and VGG16, as indicated in Table 1. Figure 5 shows the visualization of the performance of these models in comparison with each other, and it is evident that the EfficientNetV2-S model is more accurate.

Figure 6 displays the training and validation history of the EfficientNetV2-S model and depicts the change in the accuracy and loss after 10 epochs. This trend is convergently stable, and the trend has small overfitting, which is an attribute that indicates good fine-tuning and regularization.

Model	Accuracy (%)
EfficientNetV2-S	98.19%
EfficientNetV2-B0	95.8%
ResNet50V2	94.9%
VGG16	91.2%
DenseNet121	90.7%
MobileNetV2	89.3%

Table 1. Performance Comparison of Deep Learning Models

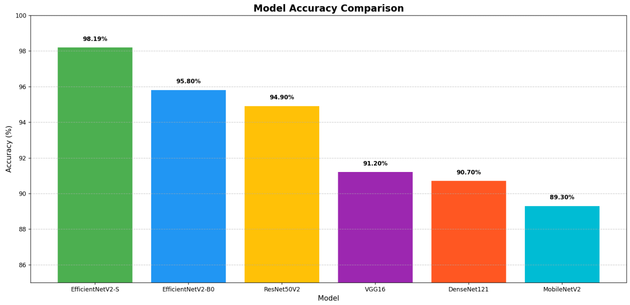


Figure 5. Visualization of Performance Comparison of All Models.

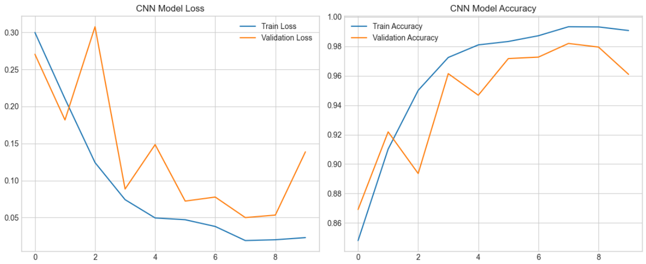


Figure 6. Training and Validation History for the EfficientNetV2-S Model, showing Loss and Accuracy over 10 epochs.

As a method of measuring the reliability of classifications, a confusion matrix of the EfficientNetV2-S model was constructed and presented in Figure 7. The matrix shows the accuracy of the model and recall of both the Demented and Non-Demented classes, that provide us with insight into the discriminatory ability of the model.

The FNN model, after undergoing hyperparameter tuning with Optuna, achieved a final test accuracy of 87.21% on the clinical dataset. This result validates the effectiveness of using structured clinical data for prediction and highlights the benefit of automated hyperparameter optimization.

In the dual-modality framework, there is the incorporation of medical images and clinical data, which makes the diagnosis even better. The CNN examines the images, digging deeper into the changes seen on the CT scan, such as cortical atrophy and enlarged ventricles, which indicate neurodegeneration. On the other hand, the FNN examines the clinical data, which includes cognitive tests such as MMSE, daily living activities, and other relevant demographic factors that may be associated with the condition. Statistical analysis reveals that MMSE, ADL, and other functional measures vary significantly across different diagnostic groups. The dual-modality framework, therefore, examines both visible changes on the brain images and the patient's individual risk factors, making the diagnosis even better and more reliable.

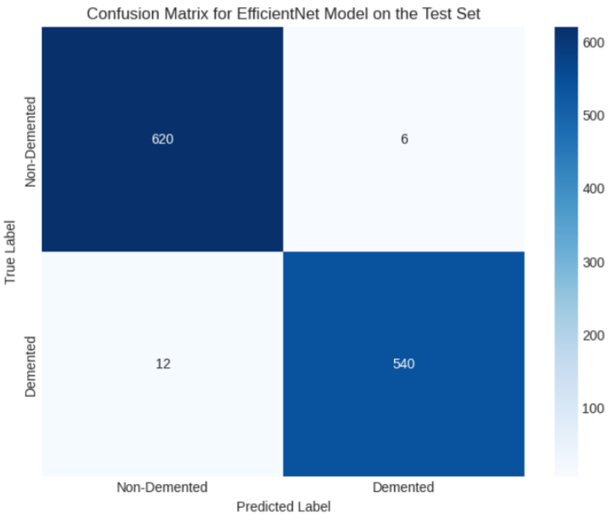


Figure 7. Confusion Matrix for EfficientNetV2-S on the Test Set.

To further evaluate the effectiveness of the proposed framework, its performance was compared with several recent state-of-the-art approaches for Alzheimer's detection reported in the literature. These studies employed various deep learning architectures and imaging modalities, including MRI-based convolutional neural networks and multimodal models. The comparison result presented in Table 2 demonstrates that the proposed framework achieves competitive performance while integrating CT imaging, structured clinical data, and an interactive clinical support system.

It should be noted that direct comparison of classification accuracies across studies should be interpreted with caution because the evaluated datasets, imaging modalities, sample sizes, and experimental protocols differ substantially. Most of the compared studies employed MRI-based datasets, particularly ADNI, whereas the proposed framework was developed using brain CT images and structured clinical data obtained from a hospital-based cohort. MRI generally provides higher soft-tissue contrast than CT imaging, which may influence diagnostic performance, Furthermore, variations in dataset composition, preprocessing procedures, and evaluation strategies can affect reported accuracies. Despite these differences, the proposed framework achieved competitive performance while simultaneously providing explainability through Grad-CAM and a deployment-oriented clinical support platform.

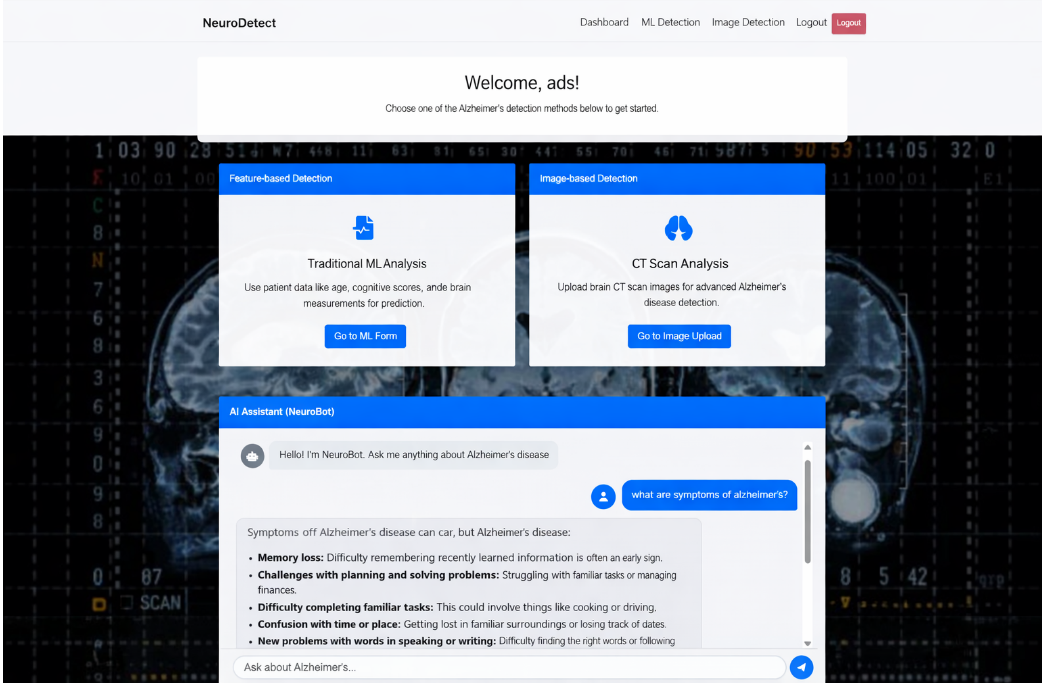
4.3. User Interface and System Interaction

The practical utility of the diagnostic models is realized through an intuitive and user-centric graphical user interface (GUI). The system's main dashboard, shown in Figure 8, provides a clean and accessible entry point for users. It presents two primary options for analysis, feature-based (clinical data) and image-based, alongside the integrated "NeuroBot" AI assistant.

For a feature-based analysis, the user navigates to a comprehensive data input form (Table 3), where they can enter demographic and clinical parameters. After submission, the system presents a detailed results page (Figure 9) that not only displays the diagnosis ('Demented' or 'Non-Demented') but also provides actionable suggestions and precautions tailored to the outcome. This aspect turns the system into a supportive system rather than a basic classifier. For an image-based diagnosis, the

Study	Data Modality	Model / Method	Dataset	Accuracy (%)
Pradhan et al., 2024	MRI	DenseNet + ResNet50	ADNI	95.3
Helaly et al., 2022	MRI	CNN-based Deep Learning	ADNI	93.6
Saleem et al., 2022	MRI	Transfer Learning CNN	ADNI	94.7
Ayus & Gupta, 2024	MRI	Hybrid Ensemble DL	ADNI	96.2
Alwakid et al., 2024	MRI	Image Processing + CNN	ADNI	94.5
Ávila-Jiménez et al., 2024	Clinical Records	Deep Learning Model	Clinical Dataset	85.0
Proposed Method	CT + Clinical Data	EfficientNetV2-S + FNN	Hospital + Clinical Dataset	98.19

Table 2. Comparison of the proposed method with recent state-of-the-art Alzheimer's disease detection approaches.



Welcome, ads!

Choose one of the Alzheimer's detection methods below to get started.

Traditional ML Analysis

Use patient data like age, cognitive scores, ande brain measurements for prediction.

CT Scan Analysis

Upload brain CT scan images for advanced Alzheimer's disease detection.

Hello! I'm NeuroBot. Ask me anything about Alzheimer's disease

what are symptoms of alzheimer's?

Symptoms off Alzheimer's disease can car, but Alzheimer's disease:

- **Memory loss:** Difficulty remembering recently learned information is often an early sign.
- **Challenges with planning and solving problems:** Struggling with familiar tasks or managing finances.
- **Difficulty completing familiar tasks:** This could involve things like cooking or driving.
- **Confusion with time or place:** Getting lost in familiar surroundings or losing track of dates.
- **New problems with words in speaking or writing:** Difficulty finding the right words or following

Figure 8. Main Dashboard Interface with Chatbot Panel.

user interacts with a simple drag-and-drop interface to upload a brain CT scan (Figure 10). The system processes the image and returns a clear results page displaying the diagnosis, a confidence score, and preventive measures or daily activities that may be beneficial (Figure 11). This consistent, informative clinical workflow is formulated to be easily ventured by clinicians, patients, and even caregivers, and this gap between a sophisticated AI model and a practical and usable tool is addressed.

4.4. Statistical Validation and Feature Analysis

To offer a sound clinical basis to our models, the structured clinical dataset underwent a complete statistical analysis. This validation guarantees that the characteristics learnt by the models are associated with accepted clinical markers of Alzheimer's disease.

This was done by the use of an independent samples t-test to identify the clinical characteristics that significantly differed between the Demented and Non-Demented groups. Figure 12 gives the general findings of the t-test in a manner that summarizes the comparative statistics of the two diagnostic groups. In the analysis, it was found that the significant indicators that constitute an overwhelming proportion of the significant indicators were the MMSE score, and that the p-value was 0.0000. These differences in distribution between the Age and MMSE age scores, as represented graphically in Figure 13 and Figure 14 of the boxplot is a very sharp and significant drop in MMSE scores in the demented group compared with Age.

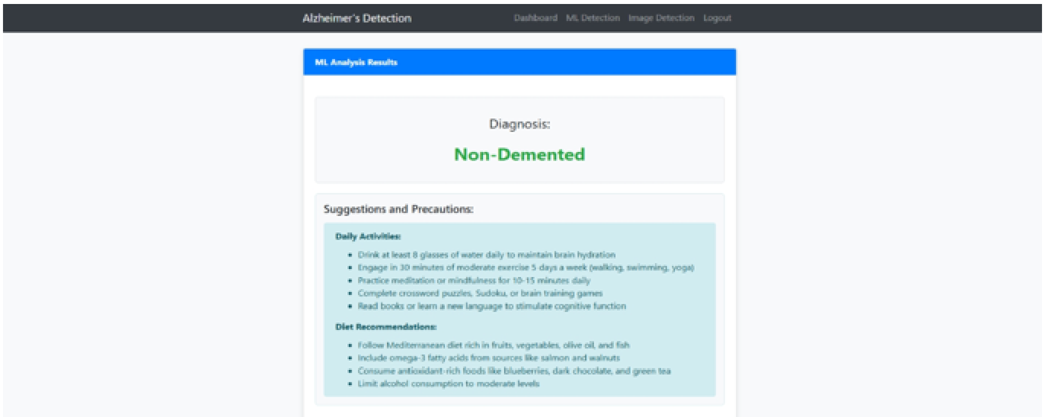
The correlation heatmap in Figure 15 also shows correlations between major clinical characteristics and the final diagnosis, indicating that the MMSE score is most strongly negatively correlated with dementia diagnosis.

Most importantly, all these statistics are directly related to the feature-importance analysis of the FNN model presented in Figure 16. These characteristics, the FunctionalAssessment, ADL (Activities of Daily Living), MemoryComplaints and MMSE, were evaluated as the most effective predictors for the model. Such a high concordance rate assures that the FNN model evidently learns to focus on the clinically significant variables associated with Alzheimer's disease.

In addition to the visual explanations provided by Grad-CAM, quantitative evidence supporting model interpretability was obtained through statistical validation of the clinical variables. Independent sample t-tests demonstrated significant differences between demented and Non-demented groups, with MMSE exhibiting a highly significant association ($p < 0.001$). Correlation analysis further confirmed strong relationships between important clinical features and dementia

Feature	Value	Feature	Value
Age	68	SystolicBP	118
Gender	0	DiastolicBP	78
Ethnicity	0	CholesterolTotal	180
EducationLevel	3	CholesterolLDL	100
BMI	24.5	CholesterolHDL	65
Smoking	0	CholesterolTriglycerides	130
AlcoholConsumption	2	MMSE	29
PhysicalActivity	56	FunctionalAssessment	1.0
DietQuality	9	MemoryComplaints	0
SleepQuality	8	BehavioralProblems	0
FamilyHistoryAlzheimers	0	ADL	1.0
CardiovascularDisease	0	Confusion	0
Diabetes	0	Disorientation	0
Depression	0	Personality Changes	0
HeadInjury	0	Difficulty Completing Tasks	0
Hypertension	0	Forgetfulness	1
Doctor In Charge	221		

Table 3. Sample Input Values for Feature-Based Analysis.



Diagnosis:

Non-Demented

Suggestions and Precautions:

Daily Activities:

- Drink at least 8 glasses of water daily to maintain brain hydration
- Engage in 30 minutes of moderate exercise 5 days a week (walking, swimming, yoga)
- Practice meditation or mindfulness for 10-15 minutes daily
- Complete crossword puzzles, Sudoku, or brain training games
- Read books or learn a new language to stimulate cognitive function

Diet Recommendations:

- Follow Mediterranean diet rich in fruits, vegetables, olive oil, and fish
- Include omega-3 fatty acids from sources like salmon and walnuts
- Consume antioxidant-rich foods like blueberries, dark chocolate, and green tea
- Limit alcohol consumption to moderate levels

Figure 9. Analysis Results Page for Clinical Data Submission.

diagnosis. Furthermore, feature-importance analysis identified MMSE, Functional Assessment, ADL, and Memory Complaints as the most influential used by the FNN model. The agreement between statistical significance and model-derived feature importance provides quantitative evidence that the model focuses on clinically meaningful biomarkers associated with Alzheimer's disease.

4.5. User Engagement Framework

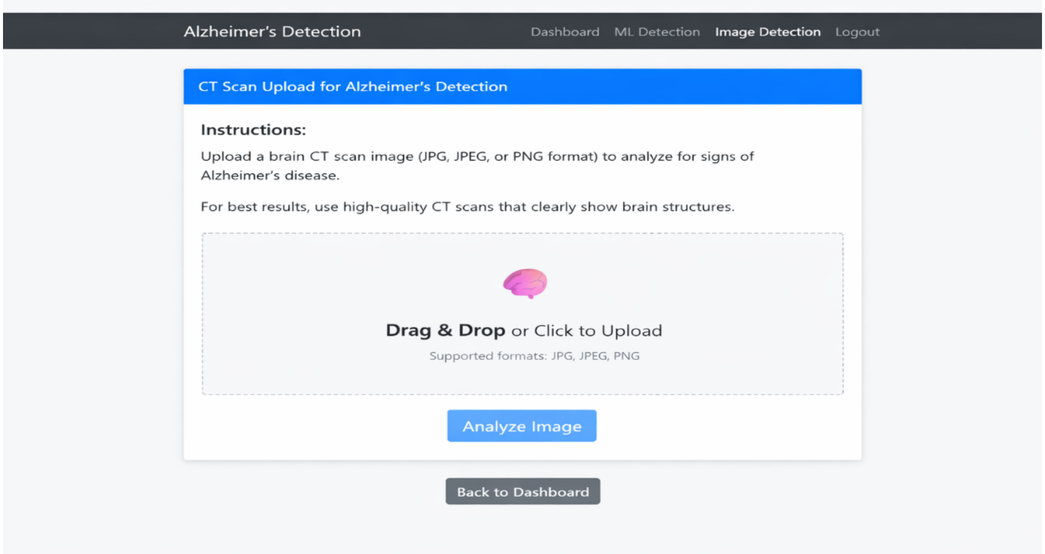
The final component of our evaluation focused on the AI-powered chatbot, NeuroBot, which serves as the primary tool for patient and caregiver support. The chatbot was assessed across five key criteria: Domain Relevance, Medical Accuracy, Politeness and Safety, Responsiveness, and its ability to refuse out-of-scope questions. As illustrated in the radar chart in Figure 17, NeuroBot scored perfectly or near-perfectly on all measures. This lends credibility to its accuracy, safety, and relevancy in delivering pertinent information within the field of Alzheimer Disease. Also, response time analysis revealed that the response to user queries was received in time (85% of queries were attended to within a range of 4.2 to 4.6 seconds), which guaranteed a consistent and interactive user experience.

5. Conclusion and Future Work

This study developed and evaluated a dual-modality framework for the early detection of Alzheimer's disease, where the deep learning pipeline is applied to the CT image processing and the hyperparameter-optimized Feedforward Neural Network (FNN) is applied to the clinical data processing.

The findings demonstrate that using a fine-tuned EfficientNetV2-S architecture for image analysis can achieve high diagnostic accuracy, reaching a peak validation performance of 98.19%. Complementing this, the FNN model, optimized through systematic hyperparameter tuning, achieved a final accuracy of 87.21% on structured clinical data. The statistical analysis of this clinical data further solidified our approach, confirming that the models are learning from features, such as the MMSE score, that are strongly correlated with established indicators of cognitive decline.

One of the major contributions that this research makes is the holistic approach. When putting these diagnostic models into a broader ecosystem, which includes a user-friendly interface and the AI-powered NeuroBot, this piece of work offers a blueprint of an all-inclusive support platform. This usability-based strategy is also essential in translating the gap between the complicated AI technology and the daily clinical practice,



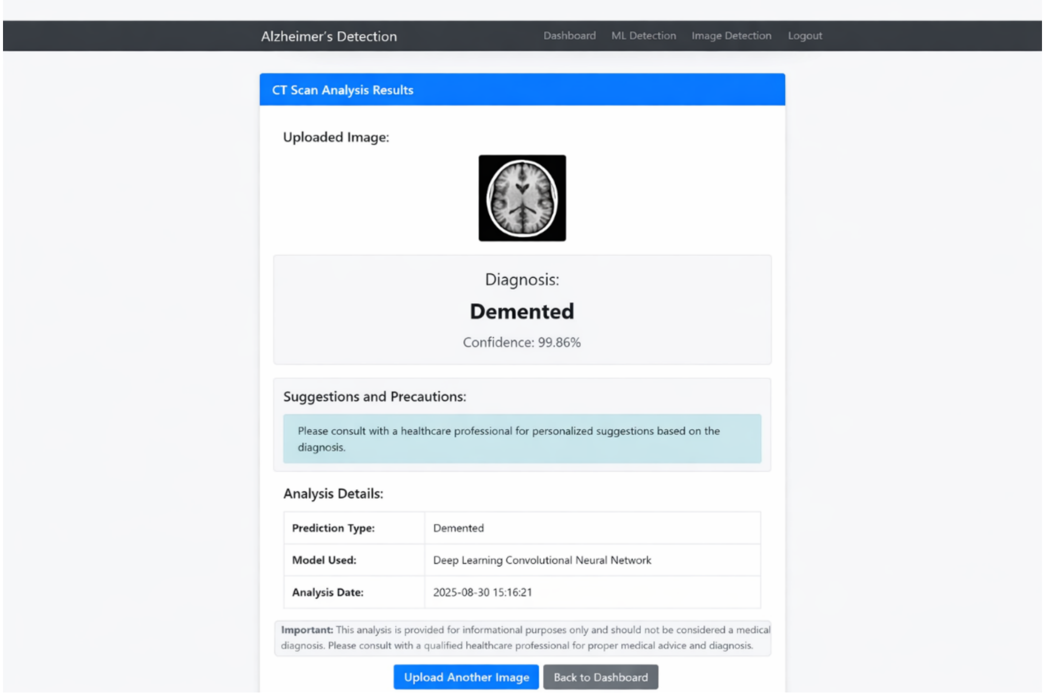
Instructions:

Upload a brain CT scan image (JPG, JPEG, or PNG format) to analyze for signs of Alzheimer's disease.

For best results, use high-quality CT scans that clearly show brain structures.

Supported formats: JPG, JPEG, PNG

Figure 10. Drag-and-Drop Interface for Image-Based Analysis.



Uploaded Image:

Diagnosis:

Demented

Confidence: 99.86%

Suggestions and Precautions:

Please consult with a healthcare professional for personalized suggestions based on the diagnosis.

Analysis Details:

Prediction Type:	Demented
Model Used:	Deep Learning Convolutional Neural Network
Analysis Date:	2025-08-30 15:16:21

Important: This analysis is provided for informational purposes only and should not be considered a medical diagnosis. Please consult with a qualified healthcare professional for proper medical advice and diagnosis.

Figure 11. Analysis Results Page for Image-Based Diagnosis.

and allows patients, caregivers, and clinicians to place the right diagnostics and information right at their fingertips.

The proposed framework offers promising avenues for future development and presents several opportunities for improvement. The capabilities of the diagnostic functions will be expanded to provide a more comprehensive portrait of the patient, with a priority on the integration of multimodal data from MRI scans and genetic markers. We will also make the existing binary classification models extendable so that Alzheimer's can be classified in multiple steps, so that we are in a position to further distinguish between MCI and the onset of further stages of the disease.

The last and most important stage will be the push of the framework into clinical validation by means of large-scale trials. This will be essential for validating the system's performance across diverse populations and is a necessary step for its eventual integration into standard clinical workflows. Concurrently, we will explore the optimization of the models for deployment on edge devices and expand the chatbot's capabilities to include summarizing prediction results and offering multilingual support, thereby increasing the system's accessibility and global impact.

It should be noted that the above evaluation was carried out using controlled experimental data. Although the models were

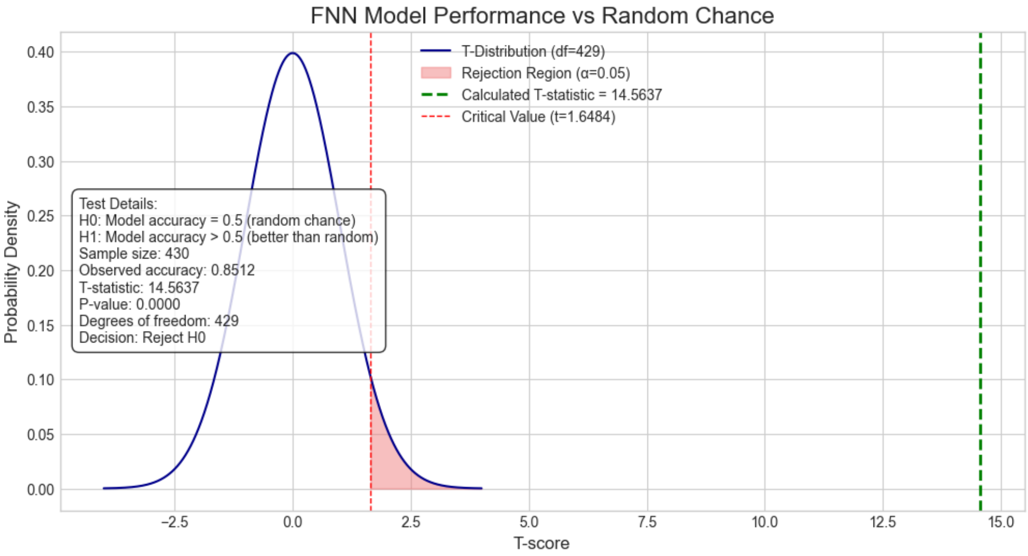


Figure 12. Visualization of T-Test Results Comparing Demented vs. Non-Demented Groups.

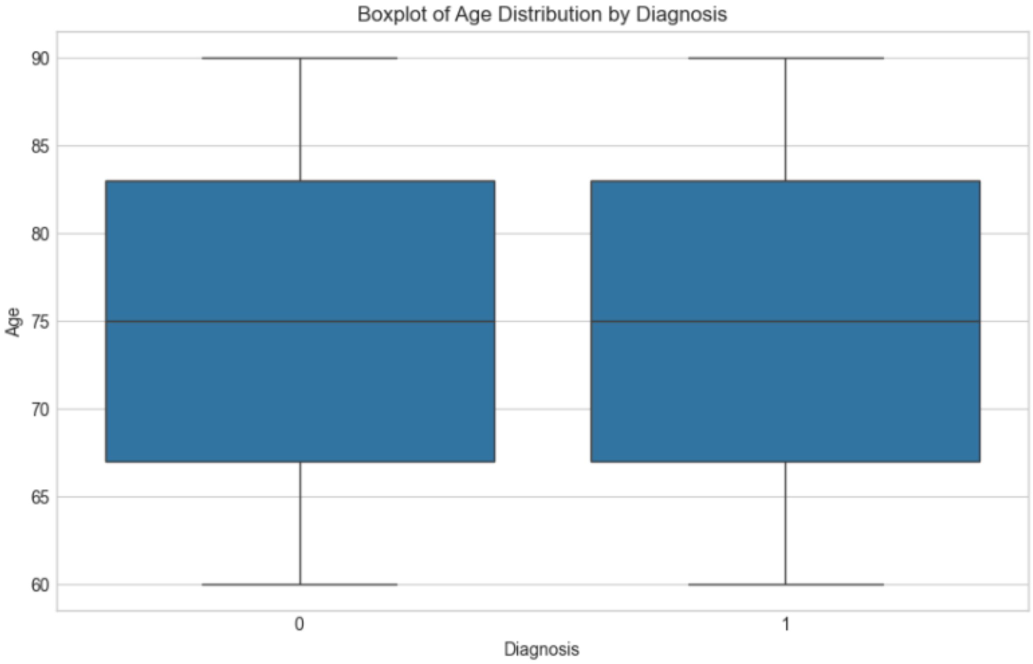


Figure 13. Boxplot of Age Distribution by Diagnosis.

validated using a held-out test set, the models were not validated using any independent clinical data sets. In the future, the focus will be on evaluating the framework using data sets from multiple institutions and clinical settings.

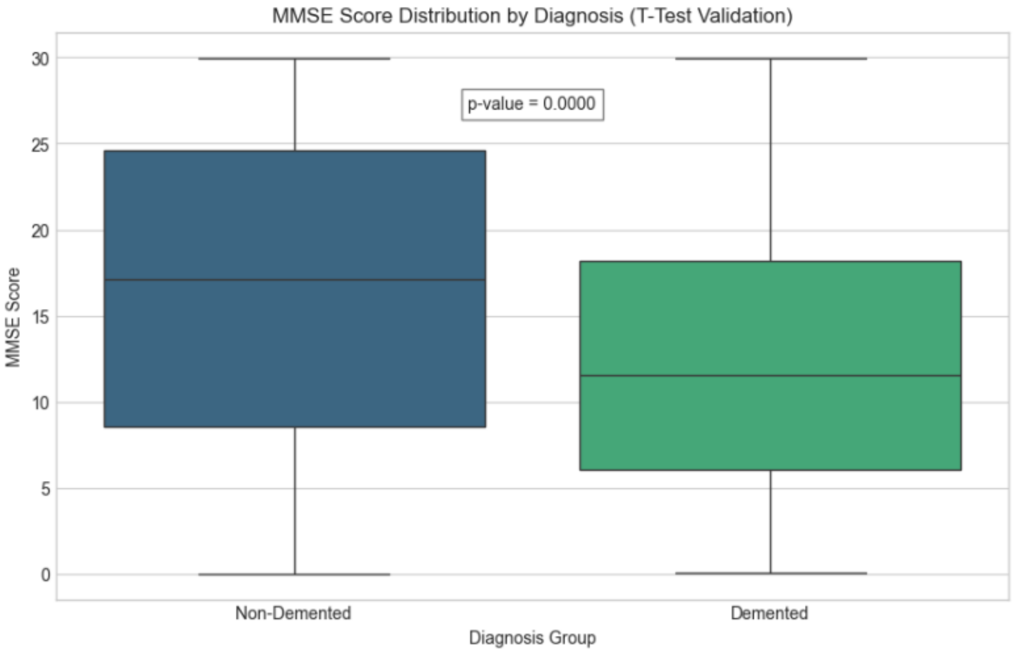
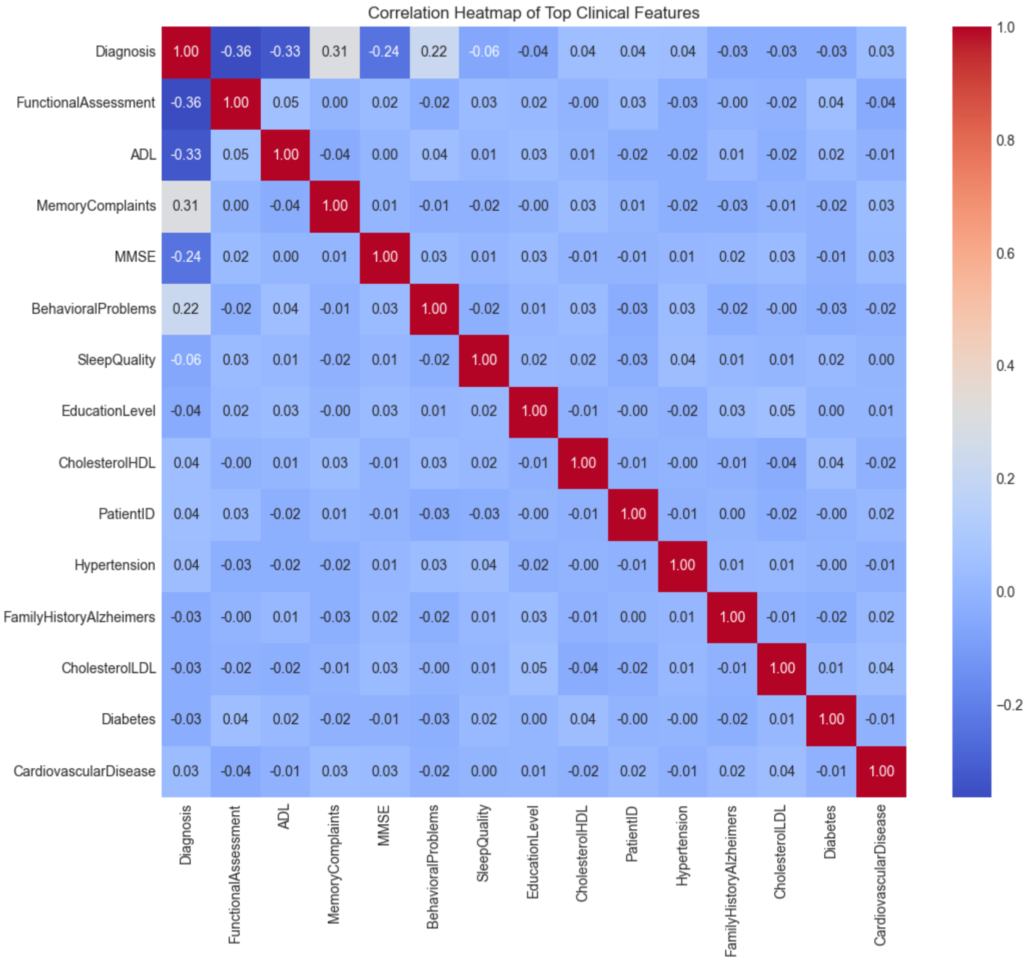


Figure 14. MMSE Score Distribution by Diagnosis (T-Test Validation).



	Diagnosis	FunctionalAssessment	ADL	MemoryComplaints	MMSE	BehavioralProblems	SleepQuality	EducationLevel	CholesterolHDL	PatientID	Hypertension	FamilyHistoryAlzheimers	CholesterolLDL	Diabetes	CardiovascularDisease
Diagnosis	1.00	-0.36	-0.33	0.31	-0.24	0.22	-0.06	-0.04	0.04	0.04	0.04	-0.03	-0.03	-0.03	0.03
FunctionalAssessment	-0.36	1.00	0.05	0.00	0.02	-0.02	0.03	0.02	-0.00	0.03	-0.03	-0.00	-0.02	0.04	-0.04
ADL	-0.33	0.05	1.00	-0.04	0.00	0.04	0.01	0.03	0.01	-0.02	-0.02	0.01	-0.02	0.02	-0.01
MemoryComplaints	0.31	0.00	-0.04	1.00	0.01	-0.01	-0.02	-0.00	0.03	0.01	-0.02	-0.03	-0.01	-0.02	0.03
MMSE	-0.24	0.02	0.00	0.01	1.00	0.03	0.01	0.03	-0.01	-0.01	0.01	0.02	0.03	-0.01	0.03
BehavioralProblems	0.22	-0.02	0.04	-0.01	0.03	1.00	-0.02	0.01	0.03	-0.03	0.03	-0.02	-0.00	-0.03	-0.02
SleepQuality	-0.06	0.03	0.01	-0.02	0.01	-0.02	1.00	0.02	0.02	-0.03	0.04	0.01	0.01	0.02	0.00
EducationLevel	-0.04	0.02	0.03	-0.00	0.03	0.01	0.02	1.00	-0.01	-0.00	-0.02	0.03	0.05	0.00	0.01
CholesterolHDL	0.04	-0.00	0.01	0.03	-0.01	0.03	0.02	-0.01	1.00	-0.01	-0.00	-0.01	-0.04	0.04	-0.02
PatientID	0.04	0.03	-0.02	0.01	-0.01	-0.03	-0.03	-0.00	-0.01	1.00	-0.01	0.00	-0.02	-0.00	0.02
Hypertension	0.04	-0.03	-0.02	-0.02	0.01	0.03	0.04	-0.02	-0.00	-0.01	1.00	0.01	0.01	-0.00	-0.01
FamilyHistoryAlzheimers	-0.03	-0.00	0.01	-0.03	0.02	-0.02	0.01	0.03	-0.01	0.00	0.01	1.00	-0.01	-0.02	0.02
CholesterolLDL	-0.03	-0.02	-0.02	-0.01	0.03	-0.00	0.01	0.05	-0.04	-0.02	0.01	-0.01	1.00	0.01	0.04
Diabetes	-0.03	0.04	0.02	-0.02	-0.01	-0.03	0.02	0.00	0.04	-0.00	-0.00	-0.02	0.01	1.00	-0.01
CardiovascularDisease	0.03	-0.04	-0.01	0.03	0.03	-0.02	0.00	0.01	-0.02	0.02	-0.01	0.02	0.04	-0.01	1.00

Figure 15. Correlation Heatmap of Top Clinical Features

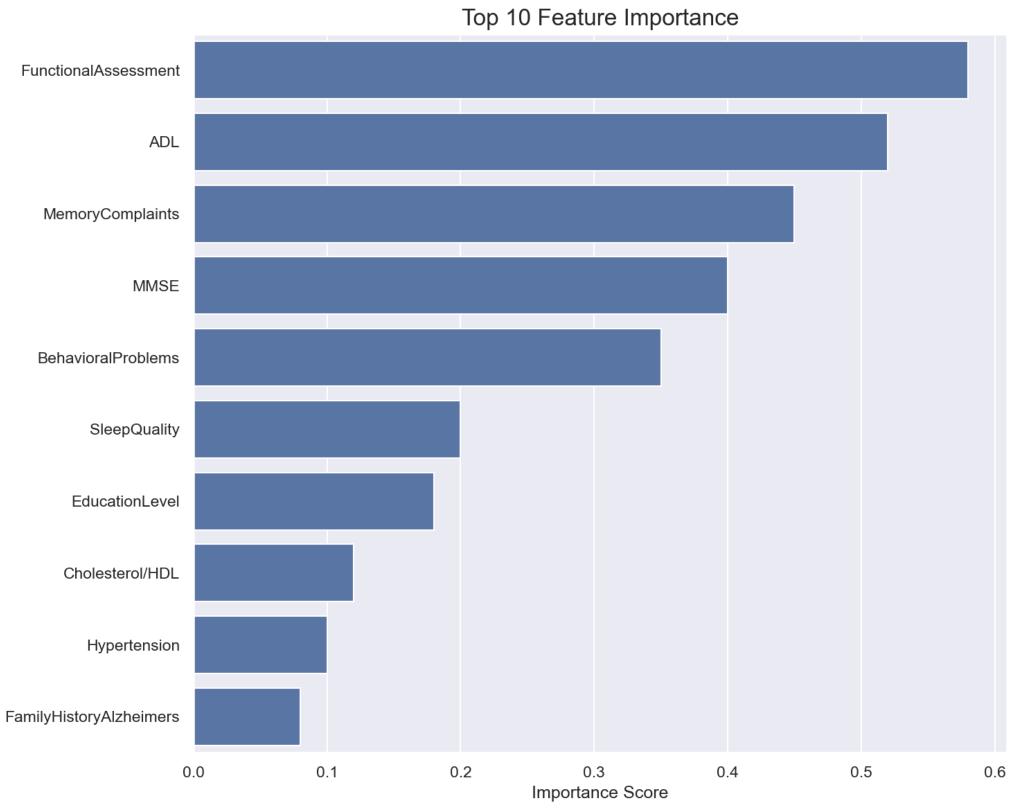


Figure 16. Top 10 Most Important Features.

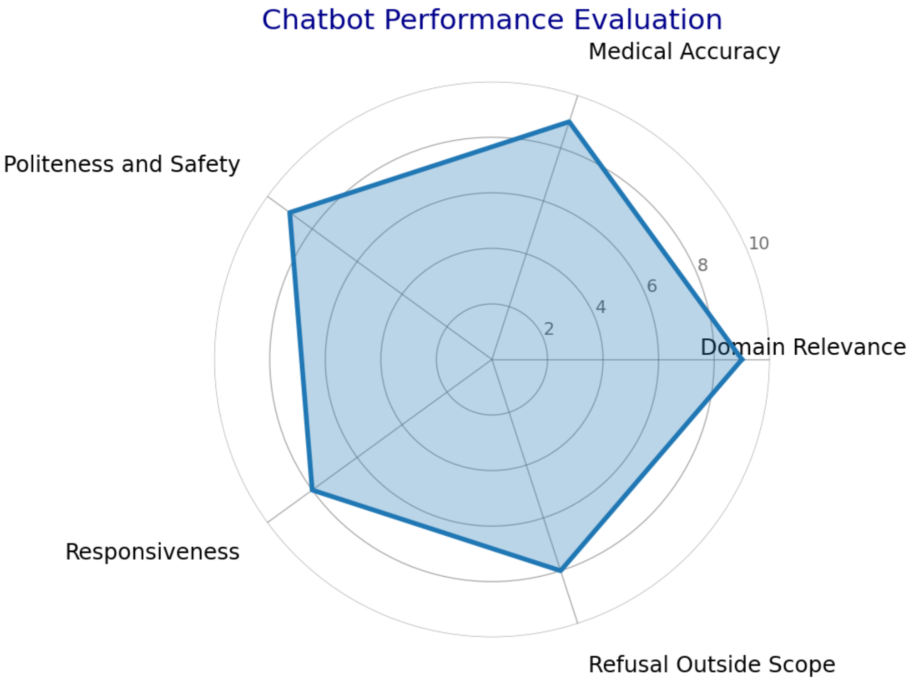


Figure 17. Radar Chart Evaluation of the Chatbot.

Ethical Statement

Ethical approval of this study was obtained from the Institutional Ethics Committee of Lovely Professional University, India (Ref: LPU/IEC-LPU/2025/1/2, dated 15 February 2025). Permission to access clinical and imaging data was granted by Shrimaan Superspeciality Hospital, Jalandhar, India. The study was retrospective and involved analysis of previously collected anonymised clinical and brain CT imaging data. No direct patient contact occurred, and no personally identifiable information was accessed. In accordance with the Institutional Ethics Committee clarification, the requirement for informed consent was waived/not applicable.

This study employed anonymized brain images from the brain CT slices of patients at Shrimann Superspeciality Hospital. No personally identifiable patient information was accessed during this study. The data were used strictly for academic research in accordance with the hospital's data privacy regulations. The proposed framework is intended for use as a clinical decision support tool. However, it is not intended for use as a substitute for clinical expertise. The NeuroBot chatbot provides general information guidance only and does not replace professional advice.

Funding

The scholar was sponsored by Tertiary Education Trust Fund (TETFUND) Nigeria for Higher education Studies only.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability Statements

The data supporting the findings of this study are available from the corresponding author upon reasonable request. However, the data are not publicly available due to privacy or ethical restrictions.

Credit authorship contribution statement

Shehu Mohammed: Conceptualization; Project Administration; Methodology; Data Curation; Software Development; Investigation; Writing – Original Draft Preparation.
Neha Malhotra: Supervision; Writing – Review & Editing; Formal Analysis.
Anmol Singh Rai: Resources; Validation; Visualization.

References

1. S. Alinsaif, "Dca-enhanced alzheimer's detection with shearlet and deep learning integration," *Computers in Biology and Medicine*, vol. 185, p. 109538, 2025, doi: https://doi.org/10.1016/j.compbiomed.2024.109538.
2. M. Srikanth and P. Bellapukonda, "The early detection of alzheimer's illness using machine learning and deep learning algorithms," *Journal of Pharmaceutical Negative Results*, vol. 13, no. 9, pp. 4852–4859, 2022, doi: https://doi.org/10.47750/pnr.2022.13.S09.603. [Online]. Available: https://www.pnrjournal.com/index.php/home/article/view/4470
3. E. M. Mohammed, A. M. Fakhrudeen, and O. Y. Alani, "Detection of alzheimer's disease using deep learning models: A systematic literature review," *Informatics in Medicine Unlocked*, vol. 50, p. 101551, 2024, doi: https://doi.org/10.1016/j.imu.2024.101551.
4. A. D. Arya *et al.*, "A systematic review on machine learning and deep learning techniques in the effective diagnosis of alzheimer's disease," *Brain Informatics*, vol. 10, no. 1, p. 17, 2023, doi: https://doi.org/10.1186/s40708-023-00195-7.
5. N. Pradhan, S. Sagar, and A. S. Singh, "Analysis of mri image data for alzheimer disease detection using deep learning techniques," *Multimedia Tools and Applications*, vol. 83, no. 6, pp. 17 729–17 752, 2024, doi: https://doi.org/10.1007/s11042-023-16256-2.
6. P. Kishore, C. U. Kumari, M. Kumar, and T. Pavani, "Detection and analysis of alzheimer's disease using various machine learning algorithms," *Materials today: proceedings*, vol. 45, pp. 1502–1508, 2021, doi: https://doi.org/10.1016/j.matpr.2020.07.645.
7. 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*, vol. 11, no. 1, p. 149, 2023, doi: https://doi.org/10.3390/biomedicines11010149.
8. S. Mirabian, F. Mohammadian, Z. Ganji, H. Zare, and E. Hasanpour Khalesi, "The potential role of machine learning and deep learning in differential diagnosis of alzheimer's disease and ftd using imaging biomarkers: A review," *The Neuroradiology Journal*, vol. 38, no. 5, pp. 571–587, 2025, doi: https://doi.org/10.1177/19714009251313511.
9. J. Chua *et al.*, "Utilizing deep learning to predict alzheimer's disease and mild cognitive impairment with optical coherence tomography," *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, vol. 17, no. 1, p. e70041, 2025, doi: https://doi.org/10.1002/dad2.70041.
10. 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*, vol. 13, no. 7, p. 1216, 2023, doi: https://doi.org/10.3390/diagnostics13071216.
11. K. Lokesh, N. P. Challa, A. S. Satwik, J. C. Kiran, N. K. Rao, and B. Naseeba, "Early alzheimer's disease detection using deep learning," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 9, 2023, doi: https://doi.org/10.4108/eetpht.9.3966.
12. S. Balne and A. Elumalai, "Machine learning and deep learning algorithms used to diagnosis of alzheimer's," *Materials Today: Proceedings*, vol. 47, pp. 5151–5156, 2021, doi: https://doi.org/10.1016/j.matpr.2021.05.499.
13. S. Baskar, M. L. Prasad, N. Sharma, T. Katale, P. C. S. Reddy *et al.*, "An accurate prediction and diagnosis of alzheimer's disease using deep learning," in *2023 IEEE North Karnataka Subsection Flagship International Conference (NKCon)*. IEEE, 2023, pp. 1–7, doi: https://doi.org/10.1109/nkcon59507.2023.10396132.
14. A. Muydinov, "Advances in deep learning techniques for alzheimer's disease detection using mri images review," in *Proceedings of the 8th International Conference on Future Networks & Distributed Systems*, 2024, pp. 266–269, doi: https://doi.org/10.1145/3726122.3726160.

15. G. N. Alwakid, S. Tahir, M. Humayun, and W. Gouda, “Improving alzheimer’s detection with deep learning and image processing techniques,” *IEEE Access*, vol. 12, pp. 153 445–153 456, 2024, doi: https://doi.org/10.1109/ACCESS.2024.3481238.
16. G. S. Chhabra, A. Guru, B. J. Rajput, L. Dewangan, and S. K. Swarnkar, “Multimodal neuroimaging for early alzheimer’s detection: a deep learning approach,” in *2023 14th international conference on computing communication and networking technologies (ICCCNT)*. IEEE, 2023, pp. 1–5, doi: https://doi.org/10.1109/ICCCNT56998.2023.10307780.
17. S. Afzal *et al.*, “Alzheimer disease detection techniques and methods: A review.” *International journal of interactive multimedia and artificial intelligence*, vol. 6, no. 7, pp. 26–38, 2021, doi: https://doi.org/10.9781/ijimai.2021.04.005.
18. S. Koga, A. Ikeda, and D. W. Dickson, “Deep learning-based model for diagnosing alzheimer’s disease and tauopathies,” *Neuropathology and Applied Neurobiology*, vol. 48, no. 1, p. e12759, 2022, doi: https://doi.org/10.1111/nan.12759.
19. I. Ayus and D. Gupta, “A novel hybrid ensemble based alzheimer’s identification system using deep learning technique,” *Biomedical Signal Processing and Control*, vol. 92, p. 106079, 2024, doi: https://doi.org/10.1016/j.bspc.2024.106079.
20. S. Suganyadevi, A. S. Rajasekaran, N. Satheesh, R. Suganthi, R. Naveenkumar *et al.*, “Alzheimer’s disease diagnosis using deep learning approach,” in *2023 Second International Conference on Electronics and Renewable Systems (ICEARS)*. IEEE, 2023, pp. 1205–1209, doi: https://doi.org/10.1109/ICEARS56392.2023.10085017.
21. J. L. Ávila-Jiménez, V. Cantón-Habas, M. del Pilar Carrera-González, M. Rich-Ruiz, and S. Ventura, “A deep learning model for alzheimer’s disease diagnosis based on patient clinical records,” *Computers in Biology and Medicine*, vol. 169, p. 107814, 2024, doi: https://doi.org/10.1016/j.compbiomed.2023.107814.
22. S. Savaş, “Detecting the stages of alzheimer’s disease with pre-trained deep learning architectures,” *Arabian Journal for Science and Engineering*, vol. 47, no. 2, pp. 2201–2218, 2022, doi: https://doi.org/10.1007/s13369-021-06131-3.
23. W. Al Shehri, “Alzheimer’s disease diagnosis and classification using deep learning techniques,” *PeerJ Computer Science*, vol. 8, p. e1177, 2022, doi: https://doi.org/10.7717/peerj-cs.1177.
24. T. J. Saleem *et al.*, “Deep learning-based diagnosis of alzheimer’s disease,” *Journal of Personalized Medicine*, vol. 12, no. 5, p. 815, 2022, doi: https://doi.org/10.3390/jpm12050815.
25. G. Hcini, I. Jdey, and H. Dhahri, “Investigating deep learning for early detection and decision-making in alzheimer’s disease: A comprehensive review: G. hcini et al.” *Neural Processing Letters*, vol. 56, no. 3, p. 153, 2024, doi: https://doi.org/10.1007/s11063-024-11600-5.
26. S. Jain and R. Srivastava, “Enhanced eeg-based alzheimer’s disease detection using synchrosqueezing transform and deep transfer learning,” *Neuroscience*, vol. 576, pp. 105–117, 2025, doi: https://doi.org/10.1016/j.neuroscience.2025.04.041.
27. H. A. Helaly, M. Badawy, and A. Y. Haikal, “Deep learning approach for early detection of alzheimer’s disease,” *Cognitive computation*, vol. 14, no. 5, pp. 1711–1727, 2022, doi: https://doi.org/10.1007/s12559-021-09946-2.
28. T. Vanaja, K. Shanmugavadivel, M. Subramanian, and C. Kanimozhiselvi, “Advancing alzheimer’s detection: integrative approaches in mri analysis with traditional and deep learning models,” *Neural Computing and Applications*, vol. 37, no. 14, pp. 8527–8546, 2025, doi: https://doi.org/10.1007/s00521-025-10993-1.
29. M. M. S. Fareed *et al.*, “Add-net: an effective deep learning model for early detection of alzheimer disease in mri scans,” *IEEE Access*, vol. 10, pp. 96 930–96 951, 2022, doi: https://doi.org/10.1109/ACCESS.2022.3204395.
30. E. Kina, “Tleablcnn: Brain and alzheimer’s disease detection using attention based explainable deep learning and smote using imbalanced brain mri,” *IEEE Access*, 2025, doi: https://doi.org/10.1109/ACCESS.2025.3539550.
31. M. Taiyeb Khosroshahi *et al.*, “Explainable artificial intelligence in neuroimaging of alzheimer’s disease,” *Diagnostics*, vol. 15, no. 5, p. 612, 2025, doi: https://doi.org/10.3390/diagnostics15050612.