Early Detection of Alzheimer's Disease Using Deep Learning on MRI Images: A Comparative Study of CNN And ResNet50

Nisha Ranjan[1], Prachi[2], Shweta Vishwakarma[3], Akash Sanghi[4]
[1,2,3,4]*Department of Computer Applications, Invertis University, Bareilly, India*
doi.org/10.64643/IJIRTV12I11-207110-459

***Abstract*—Alzheimer's disease is a progressive neurodegenerative disorder that affects memory, cognitive abilities, and daily functioning. Early detection is crucial for timely intervention and effective disease management. Traditional diagnostic methods are often time-consuming and may fail to identify the disease in its early stages. In recent years, deep learning has emerged as a powerful approach for medical image analysis, particularly in the classification of brain MRI images. This study presents a comparative analysis of deep learning models, specifically a custom Convolutional Neural Network (CNN) and transfer learning using ResNet50 model, for the early detection of Alzheimer's disease. The dataset consists of MRI images categorized into four classes: non-demented, very mild, mild, and moderate Alzheimer's. Data preprocessing techniques such as normalization, resizing, and augmentation were applied to enhance model performance and reduce overfitting. The experimental results demonstrate that the CNN model achieved an accuracy of 75.22%, outperforming the ResNet50 model, which achieved 51.52%.**

Additional evaluation metrics, including precision, recall, and F1-score, further confirm the superior and balanced performance of the CNN model. The results highlight that custom CNN architectures are more effective in handling limited and imbalanced medical image datasets compared to transfer learning approaches. This study provides a comparative insight into model performance under imbalanced data conditions and emphasizes the importance of model selection in medical imaging applications. The findings demonstrate the potential of deep learning techniques in supporting early diagnosis and improving healthcare decision-making systems.

***Index Terms*—Alzheimer's Disease, Deep Learning, MRI Imaging, Convolutional Neural Network (CNN), ResNet50, Early Detection, Medical Image Analysis.**

I. INTRODUCTION

Alzheimer's disease is a progressive neurodegenerative disorder that primarily affects memory, thinking ability, and behavior. It is the most common cause of dementia among older adults and poses a significant challenge to global healthcare systems [1]. As the disease progresses, patients experience severe cognitive decline, loss of independence, and difficulty in performing daily activities. According to recent studies, the number of Alzheimer's patients is rapidly increasing worldwide, making early diagnosis and intervention critically important.[1], [4]. As shown in Fig. 1, Alzheimer's disease leads to significant structural changes and brain shrinkage compared to a healthy brain.

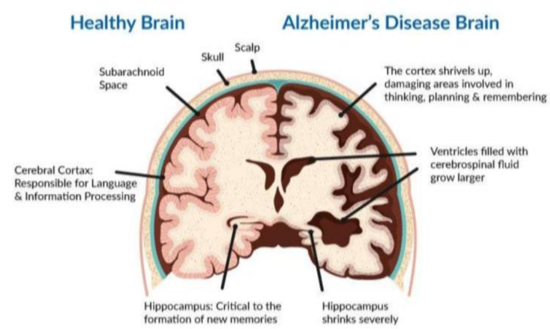

Fig. 1. Comparison between a healthy brain and an Alzheimer's affected brain

This figure highlights the significant structural differences, including cortical shrinkage and enlargement of ventricles, which are key indicators of Alzheimer's disease

Early detection of Alzheimer's disease plays a vital role in slowing down its progression and improving the quality of life for patients. However, traditional diagnostic methods such as clinical assessments and manual analysis of brain scans are often time-consuming, subjective, and less effective in identifying the disease at its early stages. These limitations highlight the need for automated and accurate diagnostic systems [4].

In recent years, advancements in artificial intelligence, particularly deep learning, have shown promising results in the field of medical image analysis [13], [14]. Deep learning models are capable of automatically extracting complex features from medical images, making them highly effective for disease detection. Among these models, Convolutional Neural Networks (CNN) has gained significant attention due to their strong ability to process and analyze image data.[9] Additionally, transfer learning techniques, such as ResNet50, enable the reuse of pre-trained models to improve performance and reduce training time [11], [12].

Magnetic Resonance Imaging (MRI) is widely used for brain imaging as it provides detailed structural information about the brain. Subtle changes in brain structure associated with Alzheimer's disease can be effectively captured through MRI scans. When combined with deep learning techniques, MRI data can be utilized to detect Alzheimer's disease at an early stage with improved accuracy [3], [6]. However, existing studies often rely on large and balanced datasets, and there is limited research comparing the performance of custom CNN models and transfer learning approaches under imbalanced data conditions.

This study aims to perform a comparative analysis of deep learning models, specifically CNN and ResNet50, for the early detection of Alzheimer's disease using MRI images. The objective is to evaluate model performance based on metrics such as accuracy, precision, recall, and F1-score, and to identify the most effective approach for handling imbalanced datasets. By addressing the limitations of existing studies, this research contributes to the development of more reliable and efficient diagnostic systems for Alzheimer's disease [1], [5].

II. LITERATURE REVIEW

Several studies have explored the application of deep learning techniques for the early detection of Alzheimer's disease using MRI images.

Dharshini et al. proposed a CNN-based approach for Alzheimer's detection that effectively extracted structural features from MRI scans and achieved promising accuracy. However, the study was limited by dataset imbalance and small sample size [1].

Abdelbaset and Hariri applied transfer learning using pre-trained models and showed improved classification performance. However, their model required large labeled datasets and high computational resources [2].

Samanta et al. highlighted the importance of preprocessing techniques such as normalization and filtering, which significantly improved classification accuracy [3].

Umer et al. presented a comprehensive survey on deep learning in medical imaging and concluded that deep learning models outperform traditional techniques, although challenges like overfitting and data scarcity remain [4].

Al-Zaabi et al. investigated both CNN and transfer learning models and found that transfer learning performance depends heavily on dataset size and quality [5].

Liu et al. proposed a multimodal deep learning approach combining multiple imaging features, achieving high accuracy but with increased computational complexity [6].

Wang and Cao developed an ensemble-based deep learning framework that improved performance compared to single models, but required more training time and resources [7].

The Alzheimer MRI dataset from Kaggle is commonly used for training deep learning models for Alzheimer's disease detection [8].

Wen et al. demonstrated that CNN architectures are highly effective in medical image classification due to their ability to automatically extract hierarchical features [9].

LeCun et al. explained the fundamental concepts of deep learning and its impact on image-based applications, including healthcare and medical diagnostics [10].

K. He. et al. introduced the ResNet architecture, which uses residual connections to improve training of deep neural networks and is widely used in image classification tasks [11].

Russakovsky et al. described the ImageNet challenge, which significantly contributed to the advancement of deep learning models by providing large-scale annotated datasets [12].

Overall, previous studies indicate that deep learning techniques, particularly CNN-based models, are effective for Alzheimer's disease detection. However, most existing approaches rely on large and balanced datasets, and there is limited research on model performance under imbalanced data conditions, which is addressed in this study

III. METHODOLOGY

This section describes the overall approach used for the early detection of Alzheimer's disease using deep learning techniques on MRI images. The methodology consists of dataset collection, preprocessing, model development, training, and Evaluation.

A. Dataset Description

The dataset used in this study consists of brain MRI images obtained from the Kaggle Alzheimer's MRI dataset, containing approximately 5000–6400 images categorized into four classes: Non-Demented, Very Mild Demented, Mild Demented, and Moderate Demented. Each image represents structural information of the brain, which is useful for identifying patterns associated with Alzheimer's disease. The dataset is imbalanced, with fewer samples in the Moderate Demented class, which may affect model performance and classification accuracy. To evaluate the model effectively, the dataset was divided into 80% training and 20% testing sets [8].

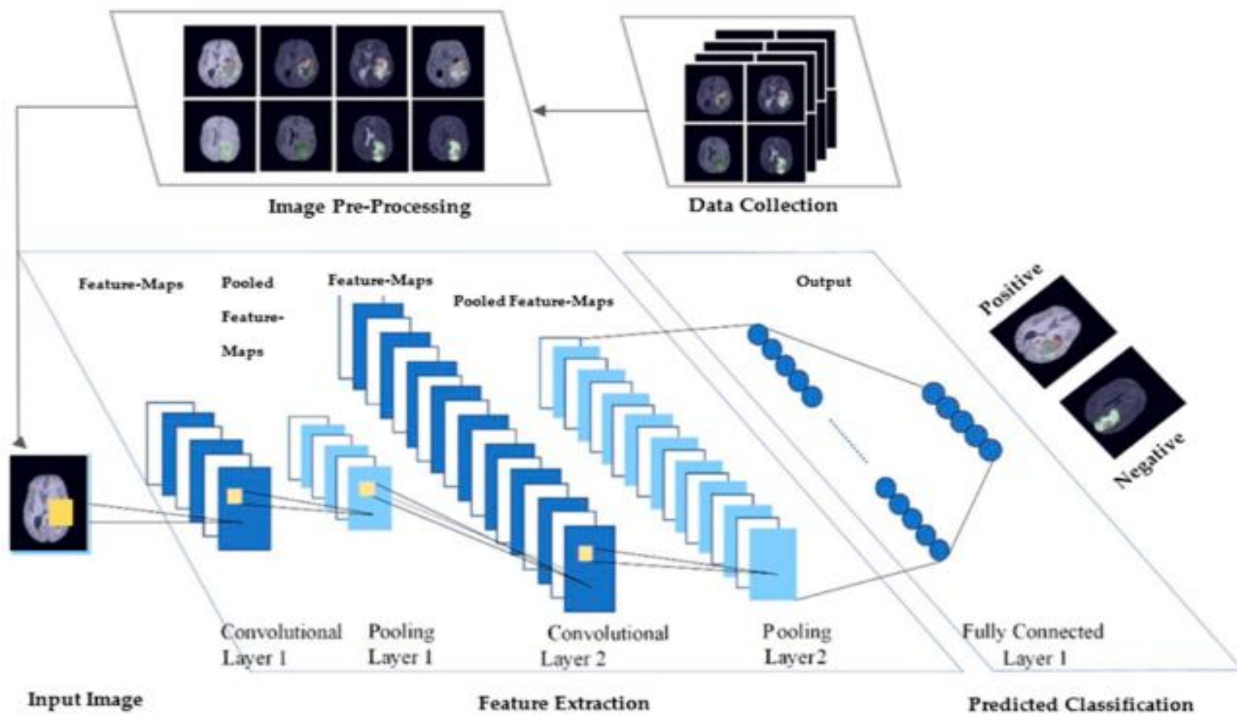

Fig. 2. Workflow of the proposed methodology for Alzheimer's disease detection

B. Data Preprocessing

Data preprocessing is an essential step to prepare the dataset for deep learning models. The following preprocessing techniques were applied:

- *Resizing:* All images were resized to 128×128 pixels to maintain uniform input size.
- *Normalization:* Pixel values were scaled to the range of 0 to 1 to improve model performance.
- *Data Augmentation:* Techniques such as rotation, flipping, and zooming were applied to increase dataset size and reduce overfitting.
- *Train-Test Split:* The dataset was divided into training and testing sets for model evaluation.

These preprocessing steps improve model generalization, enhance feature extraction, and help in reducing overfitting [3].

C. Model Development

Two deep learning models were used in this study:

(i) Convolutional Neural Network (CNN): A custom CNN model was designed for image classification.

The architecture consists of:

- Convolutional layers (3×3 filters) for feature extraction
- Max-pooling layers for dimensionality reduction.
- Batch normalization for stable training.
- Dropout layers(rate=0.5) to prevent overfitting.
- Fully connected dense layers with Softmax activation for multi-class classification.

The CNN model consists of multiple convolutional layers with 3×3 filters followed by ReLU activation functions to introduce non-linearity. Max-pooling layers are applied to reduce spatial dimensions and extract dominant features. Batch normalization is used to stabilize training, while dropout layers (0.5) help prevent overfitting. The final fully connected layer uses Softmax activation for multi-class classification.

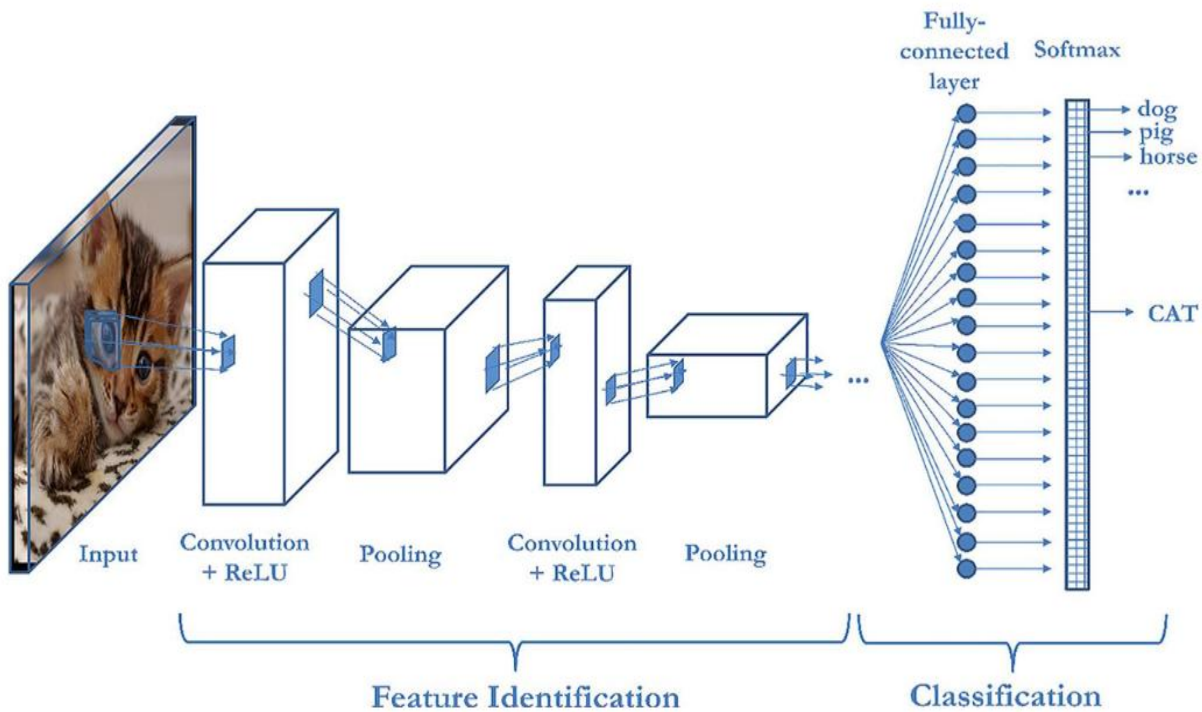

Fig. 3. Architecture of Convolutional Neural Network (CNN)

Fig. 3 illustrates the architecture of the CNN model, showing the process of feature extraction through convolution and pooling layers, followed by classification using fully connected layers [9].

(ii) Transfer Learning using ResNet50: A pre-trained ResNet50 model was used for transfer learning. The top layers of the model were removed and replaced with custom dense layers suitable for multi-class classification. The base model was frozen to retain learned features from ImageNet, while the newly added layers were trained on the MRI dataset. Global Average Pooling was applied before the fully connected layers to reduce overfitting and improve generalization. The final fully connected layer consists of 4 neurons with Softmax activation, corresponding to the four classes of Alzheimer's disease [11].

D. Model Training

Both models were trained using the following parameters:

- Optimizer: Adam
- Learning Rate: 0.001
- Loss Function: Categorical Cross-Entropy
- Batch Size: 32
- Epochs: 100
- Early Stopping was applied to prevent overfitting.

E. Evaluation Metrics

The performance of the models was evaluated using standard metrics:

- Accuracy: Measures overall correctness.
- Precision: Measures correctness of positive predictions
- Recall: Measures ability to detect actual positives
- F1-Score: Harmonic mean of precision and recall
- AUC (Area Under Curve): Measures classification performance [4].

Precision = TP / (TP + FP)

Recall = TP / (TP + FN)

F. Proposed Approach

The proposed approach focuses on evaluating and comparing the performance of deep learning models, specifically Convolutional Neural Network (CNN) and ResNet50, for the early detection of Alzheimer's disease using MRI images. The study aims to analyze how effectively these models can classify different stages of Alzheimer's disease based on structural changes in brain MRI scans.

Special emphasis is placed on understanding model performance under imbalanced dataset conditions, where certain classes have fewer samples. This helps in identifying the most reliable and robust model for real-world medical applications. Furthermore, the approach includes preprocessing techniques, model training, and performance evaluation using key metrics such as accuracy, precision, recall, F1-score, and AUC. The overall goal is to determine which model provides better classification performance and can support early diagnosis and decision-making in healthcare systems.

IV. RESULTS AND DISCUSSION

This section presents a comprehensive performance evaluation of deep learning models for the early detection of Alzheimer's disease using MRI images. The models were assessed using key metrics such as accuracy, precision, recall, F1-score, and AUC to ensure reliable and balanced performance analysis.

A. Results of CNN Model

The Convolutional Neural Network (CNN) model demonstrated robust and reliable performance in classifying MRI images. The model achieved an overall accuracy of 75.22%, indicating its effectiveness in detecting Alzheimer's disease. In addition to accuracy, the CNN model achieved a precision of approximately 0.76, recall of 0.74, and F1-score of 0.75, reflecting a balanced and reliable performance. These metrics indicate that the model is capable of correctly identifying positive cases while minimizing false predictions. The model performed particularly well in detecting Mild and Moderate Alzheimer's stages, which are important for early diagnosis. However, a slightly lower recall was observed for the non-demented class, suggesting that some normal cases were misclassified. This may be due to the similarity in brain structure at early stages of the disease. The AUC score further confirms that the CNN model has good classification capability across different classes. Overall, the CNN model proved to be effective in extracting meaningful features from MRI images and handling the imbalanced dataset. This demonstrates that the CNN model is well-suited for handling imbalanced medical image datasets.

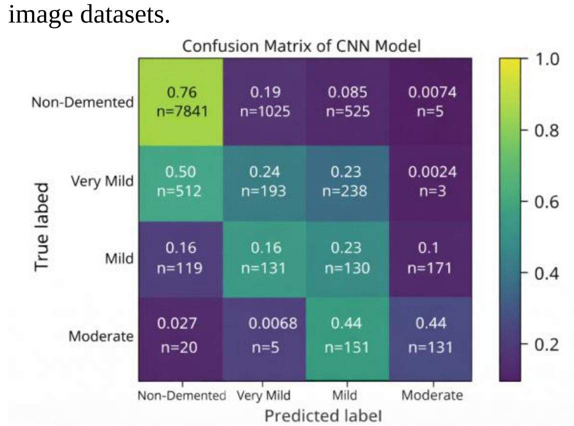

Fig. 4. Confusion matrix of CNN Model for Alzheimer's Disease Classification

The confusion matrix illustrates the classification performance of the CNN model across four classes: Non-Demented, Very Mild, Mild, and Moderate Alzheimer's disease. The diagonal values represent correctly classified samples, indicating that the model performs well, particularly in Non-Demented and Mild categories. However, some misclassification is observed between adjacent classes such as Very Mild and Mild, as well as Mild and Moderate, due to similarities in brain MRI patterns. Overall, the confusion matrix confirms the effectiveness and reliability of the CNN model in multi-class classification [1], [9].

B. Results Of ResNet50 Model

The transfer learning model using ResNet50 showed comparatively lower performance. The model achieved an accuracy of 51.52%, which is significantly lower than the CNN model. Although transfer learning models are generally effective, in this study, the ResNet50 model struggled due to the limited size and imbalance of the dataset. The model showed inconsistent precision and recall values across different classes, especially for Mild and Moderate Alzheimer's stages. The lower F1-score indicates that the model was not able to maintain a proper balance between precision and recall. This suggests that transfer learning models require larger and more balanced datasets for optimal performance [2], [11]. This indicates that transfer learning models may not always be effective when applied to small and imbalanced medical datasets

C. Comparative Analysis

A comparison between the CNN and ResNet50 models based on evaluation metrics such as accuracy, precision, recall, F1-score, and AUC. The CNN model outperformed ResNet50 across all metrics, demonstrating better classification performance and reliability.is presented in Table I.

TABLE I. COMPARATIVE RESULTS OF CNN AND ResNet50

MODEL	CNN	ResNet50
ACCURACY (%)	75.22%	51.52%
PRECISION	0.76	0.52
RECALL	0.74	0.50
F1-SCORE	0.75	0.51
AUC	0.78	0.55
PERFORMANCE	Good	Moderate

This demonstrates that CNN is more effective and reliable for handling imbalanced medical image datasets.

A comparison between the CNN and ResNet50 models clearly shows that the CNN model outperformed ResNet50 in terms of accuracy and overall performance. As illustrated in Fig. 5, the CNN model achieved significantly higher accuracy. This improved performance can be attributed to the ability of custom CNN architectures to learn domain-specific features directly from MRI images. In contrast, ResNet50, being trained on general image datasets, is less effective for specialized medical imaging tasks [5], [9]. Overall, the results demonstrate that CNN is a more reliable and efficient model for Alzheimer's disease detection using MRI images.

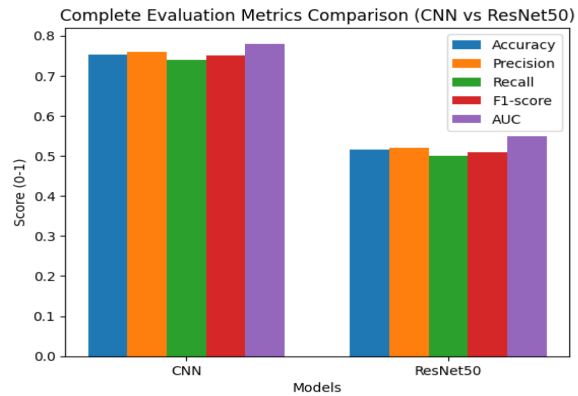

Fig. 5. Complete Evaluation Metrics Comparison (CNN vs ResNet50)

The experimental results highlight that deep learning techniques are effective for Alzheimer's disease detection using MRI images. Among the two models, CNN provided more reliable and consistent results.

One of the key factors affecting performance is the imbalance in the dataset, where certain classes have significantly fewer samples. This leads to biased predictions and reduced accuracy for underrepresented classes.

Additionally, transfer learning models like ResNet50, although powerful, may not always perform well when the dataset is small or significantly different from the pre-training dataset (ImageNet).

Therefore, the study suggests that custom CNN models are more suitable for this type of medical imaging problem. Future improvements can be achieved by using balanced datasets, applying advanced preprocessing techniques, and combining multiple models using hybrid or ensemble approaches [4], [9]. These findings emphasize the importance of selecting appropriate deep learning models for medical imaging applications, particularly when dealing with limited and imbalanced dataset.

V. CONCLUSION AND FUTURE WORK

This study presents a comparative analysis of deep learning models for the early detection of Alzheimer's disease using MRI images. Two models, a custom Convolutional Neural Network (CNN) and transfer learning using ResNet50, were evaluated based on performance metrics such as accuracy, precision, recall, and F1-score [1],[5]. The experimental results clearly demonstrate that the CNN model outperforms the ResNet50 model, achieving an accuracy of 75.22% compared to 51.52%. The CNN model showed better capability in extracting domain-specific features from MRI images and provided more balanced and reliable classification across different stages of Alzheimer's disease. In contrast, the ResNet50 model exhibited lower and inconsistent performance, mainly due to the limited size and imbalance of the dataset [1], [9].

Despite promising results, certain limitations were identified, including dataset imbalance and limited data availability, which may affect model generalization [4]. These challenges highlight the need for improved data quality and more robust modeling techniques. In future work, the performance of Alzheimer's disease detection systems can be enhanced by using larger and more balanced datasets. Advanced approaches such as hybrid models, ensemble learning, and improved preprocessing techniques can be explored to achieve higher accuracy [6], [7]. Additionally, integrating clinical data along with MRI images may further improve model performance and real-world applicability [10]. Overall, this study highlights the practical applicability of deep learning techniques in improving early diagnosis and supporting intelligent healthcare decision-making systems

REFERENCES

[1] S. Dharshini *et al*., "Early detection of Alzheimer's disease using deep learning techniques," *International Journal of Advanced Research in Computer Science (IJARCS)*, 2022.

[2] R. Abdelbaset and W. Hariri, "Transfer learning approach for Alzheimer's disease classification using MRI images," IEEE Access, 2023.

[3] A. Samanta et al., "Improving Alzheimer's detection using image preprocessing techniques," Journal of Medical Imaging and Health Informatics (JMIHI), 2023.

[4] M. Umer et al., "Deep learning in medical imaging: A survey," Computers in Biology and Medicine, 2023.

[5] M. Al-Zaabi et al., "Alzheimer's disease detection using CNN and transfer learning models," Procedia Computer Science, 2020.

[6] S. Liu et al., "Multimodal deep learning for Alzheimer's disease classification," Neurocomputing, 2023.

[7] H. Wang and Y. Cao, "Deep learning framework for Alzheimer's diagnosis using MRI data," IEEE Transactions on Medical Imaging, 2021.

[8] Kaggle, "Alzheimer MRI Dataset." [Online]. Available: https://www.kaggle.com/datasets

[9] J. Wen et al., "Convolutional neural networks for Alzheimer's disease detection," IEEE Reviews in Biomedical Engineering, 2020.

[10] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[12] O. Russakovsky et al., "ImageNet large scale visual recognition challenge," International Journal of Computer Vision, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.

[13] A. Sanghi, R. Agarwal, S. Pandey, A. Kumar, and K. Johari, "A comprehensive machine learning framework for diabetes prediction: Incorporating unique performance metrics and K-fold cross-validation," in Proc. 2025 Seventh International Conference on Computational Intelligence and Communication Technologies (CCICT), Sonepat, India, 2025, pp. 564–569, doi: 10.1109/CCICT65753.2025.00091.

[14] D. K. Singh and A. Sanghi, "Advancing heart attack prediction through AI: A review of ML, DL, and hybrid models," in Proc. 2025 14th International Conference on System Modeling & Advancement in Research Trends (SMART), Moradabad, Uttar Pradesh, India, 2025, pp. 499–505, doi: 10.1109/SMART66937.2025.11389447.