International Journal of Emerging Research in Science, Engineering, and Management
Vol. 2, Issue 3, pp.187-193, March 2026.
www.ijersem.com eISSN- 3107-9075

NEUROVISION-AI: Alzheimer's Disease Detection Using MRI and Behavioral Data

[1]R Priyadarshini, [2]A Chandhana, [3]Vidyala Karthik,
[4]Avula Hemanth Kumar Reddy, [5]Tholeti Keerthana Reddy,
[6]Kasanna Gari Guru Venkat Sai

Department of CSE, Siddartha Institute of Science and Technology, Puttur, India.
[1]darshini.sr@gmail.com, [2]arungolamchandhana@gmail.com, [3]karthikvidyala@gmail.com,
[4]hemanthavula29@gmail.com, [5]keerthanareddytholeti15@gmail.com, [6]royalguru16@gmail.com

Abstract: Early and accurate detection of Alzheimer's Disease (AD) is critical for effective intervention, a task where traditional methods often fall short. This paper presents a multimodal deep learning framework to address this diagnostic challenge. A hybrid model combining a Convolutional Neural Network (CNN) to extract spatial features from Magnetic Resonance Imaging (MRI) scans and a Recurrent Neural Network (RNN) to analyze temporal patterns from cognitive assessment data. By fusing these features, the model achieves a more robust classification. Trained on a public dataset, the system classifies AD into four stages (Normal, Mild, Moderate, and Severe) with a validation accuracy of 0.99, demonstrating high precision and recall. This work highlights the potential of hybrid AI models as a powerful diagnostic support tool for clinicians, significantly improving early AD detection.

Keywords: Alzheimer's Disease, Deep Learning, Early Detection, Multimodal Learning, Convolutional Neural Network.

1 INTRODUCTION

Alzheimer's Disease is a progressive neurodegenerative disorder that represents a significant global health challenge, affecting millions of individuals worldwide. It is primarily characterized by a gradual decline in cognitive function, memory loss, and the erosion of functional independence. The critical window for effective intervention lies in the disease's nascent stages, such as Mild Cognitive Impairment (MCI) [1]. However, traditional diagnostic protocols, which rely on neuropsychological testing and clinical evaluation, often fail to provide a definitive diagnosis until the disease has advanced to a point where therapeutic interventions offer limited benefits. This diagnostic delay underscores the urgent need for innovative, sensitive, and reliable methods capable of identifying Alzheimer's at its earliest onset, thereby enabling timely treatment planning and potentially slowing its devastating progression. To address the limitations of conventional diagnostics, the field of medical informatics has increasingly turned to Artificial Intelligence (AI) and deep learning [2].

These advanced computational techniques possess a remarkable ability to discern subtle, complex patterns within large-scale biomedical data that may be imperceptible to human observers. This research leverages a multimodal deep learning approach, which posits that integrating data from multiple sources can yield a more comprehensive and accurate diagnostic model than any single modality alone [3]. A novel hybrid architecture that combines a Convolutional Neural Network to analyze spatial information from structural Magnetic Resonance Imaging (MRI) scans with a Recurrent Neural Network to interpret temporal patterns in sequential behavioral and cognitive data.

This paper introduces a robust framework for the early detection and classification of Alzheimer's Disease across four distinct stages: Normal, Mild, Moderate, and Severe [4]. By synergistically fusing features from both neuroimaging and behavioural assessments, the model aims to enhance diagnostic accuracy and provide a more holistic view of the patient's condition. The detail the architecture of the hybrid CNN-RNN model, describe the training and validation process on a publicly available dataset, and present experimental results that demonstrate its superior performance [5]. Ultimately, this work highlights the potential of multimodal AI systems to serve as a powerful diagnostic support tool, assisting clinicians in making earlier and more informed decisions in the management of Alzheimer's Disease [6].

2 Literature Review

Alzheimer's Disease is one of the most prevalent neurodegenerative disorders affecting millions of elderly individuals worldwide. Early diagnosis plays a critical role in slowing disease progression and improving patient quality of life. However, conventional diagnostic approaches such as clinical cognitive testing and manual interpretation of neuroimaging scans often fail to detect subtle structural and functional changes during the early stages of the disease. As a result, researchers have increasingly explored artificial intelligence (AI) and deep learning techniques to develop automated and reliable diagnostic systems capable of identifying Alzheimer's Disease at its earliest onset [1].

Recent advancements in deep learning have significantly improved the analysis of neuroimaging data, particularly Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) [2]. Convolutional Neural Networks have demonstrated strong capability in extracting spatial features from MRI images and identifying structural abnormalities associated with neurodegeneration. Several studies have shown that CNN-based architectures can effectively distinguish between normal controls, Mild Cognitive Impairment (MCI), and Alzheimer's Disease patients with high classification accuracy. These methods have reduced reliance on manual feature extraction and enabled automated detection systems capable of supporting clinical diagnosis [3].

Beyond single-modality imaging approaches, researchers have increasingly recognized the importance of multimodal learning techniques that integrate information from multiple sources such as MRI scans, PET images, and clinical cognitive assessment scores. Multimodal frameworks improve diagnostic accuracy by capturing complementary structural, metabolic, and behavioral characteristics of the disease [4]. For example, recent studies have proposed deep learning architectures that combine MRI and PET modalities using feature-level fusion strategies to enhance classification performance across different disease stages. These approaches demonstrate that integrating heterogeneous biomedical data significantly improves model robustness compared to unimodal techniques [5].

Transformer-based architectures have also emerged as powerful alternatives to traditional convolutional models in neuroimaging applications [6]. Vision Transformer-based models enable long-range dependency learning and global feature extraction from brain imaging data, which improves classification accuracy in Alzheimer's Disease detection tasks. Hybrid CNN-Transformer architectures further enhance performance by combining local feature extraction capabilities of CNNs with the global attention mechanisms of transformer networks [7]. Such hybrid models have achieved promising results in classifying different stages of Alzheimer's Disease using three-dimensional MRI datasets.

Another important direction in recent research involves integrating longitudinal cognitive assessment data with neuroimaging features using sequential deep learning models. Recurrent Neural Networks, particularly Long Short-Term Memory (LSTM) networks, are well suited for analysing temporal behavioural patterns associated with progressive cognitive decline [8]. Several studies have demonstrated that combining spatial brain imaging features with sequential cognitive assessment scores enables more accurate prediction of disease progression compared to static imaging-only approaches. These models capture the dynamic nature of Alzheimer's Disease and provide improved predictive capability for early detection [9].

In addition to CNN and RNN architectures, multimodal fusion strategies based on attention mechanisms have shown significant promise in improving classification performance. Attention-based multimodal networks selectively emphasize the most informative features from different modalities, allowing the model to focus on disease-relevant regions of brain images and important behavioural indicators simultaneously [10]. These approaches improve interpretability while maintaining high classification accuracy, making them suitable for clinical decision-support systems.

Federated learning frameworks have also recently been introduced to address data privacy challenges associated with medical imaging datasets. Since patient data are often distributed across multiple hospitals and research institutions, federated multimodal learning enables collaborative model training without sharing sensitive data directly [11]. This approach improves model generalization while maintaining compliance with privacy regulations, making it highly suitable for real-world healthcare deployment scenarios. The lightweight multimodal neural networks have been proposed to reduce computational complexity and enable deployment in resource-constrained environments such as portable diagnostic systems and clinical edge devices. These architectures maintain competitive performance while reducing model size and inference time, thereby improving practical usability in healthcare applications [12].

3 PROPOSED SYSTEM

To address the challenge of early AD detection, NEUROVISION-AI is proposed, which is a multimodal deep learning system designed to classify the disease into four distinct stages: Normal, Mild, Moderate, and Severe. The system architecture follows a structured pipeline from data processing to classification.

3.1. Data Acquisition and Preprocessing

The proposed model is trained using a publicly available multimodal dataset obtained from Kaggle. The dataset includes structural MRI brain scans along with corresponding cognitive assessment scores representing different stages of Alzheimer's Disease progression. Before feeding the data into the learning framework, preprocessing is performed separately for both modalities to ensure consistency and improve model performance.

- **MRI Scan Preprocessing:** MRI images are resized to a uniform resolution and normalized to standardize pixel intensity values. Normalization improves convergence during model training and ensures stable learning behaviour across different samples. Min–Max normalization is applied to scale pixel values into the range $[0, 1]$, expressed as:

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{1}$$

where X represents the original pixel value, and X_{min} and X_{max} denote the minimum and maximum pixel values of the image respectively.

- **Cognitive Data Preprocessing:** The sequential behavioural and cognitive assessment data are cleaned and converted into numerical representations suitable for temporal modelling. Tokenization and normalization techniques are applied to transform raw cognitive responses into structured sequences that can be effectively processed by the Recurrent Neural Network module.

3.2. Multimodal Feature Extraction

The proposed NEUROVISION-AI architecture employs two parallel deep learning pathways to extract discriminative features from different data modalities.

- **CNN for Spatial Feature Extraction:** A Convolutional Neural Network is used to process MRI scans and automatically learn hierarchical spatial features representing structural brain abnormalities associated with Alzheimer's Disease. The convolutional layers capture low-level features such as edges and textures, while deeper layers extract high-level representations related to disease-specific anatomical variations. Pooling layers reduce dimensionality and help preserve essential spatial information.

- **RNN for Temporal Feature Extraction:** A Recurrent Neural Network is employed to process sequential cognitive assessment data and model temporal dependencies associated with progressive cognitive decline. Unlike conventional feedforward networks, RNNs maintain memory of previous inputs, enabling the system to capture variations in behavioural patterns across time. This temporal modelling capability enhances classification accuracy by incorporating longitudinal cognitive trends.

3.3. Feature Fusion and Classification

After extracting spatial features from MRI images using CNN and temporal features from cognitive sequences using RNN, both feature vectors are combined through a feature-level fusion strategy. The fused representation provides complementary structural and behavioural information describing disease progression more comprehensively. The combined feature vector is then passed through fully connected dense layers that perform multimodal integration and classification. The final output layer applies a SoftMax activation function to convert logits into probability values representing the likelihood of each disease stage.
For an input sample, the probability of belonging to class i is calculated as:

$$P(class_i) = \sigma(z_i) = \frac{e^{z_i}}{\sum_{j=1}^{K} e^{z_j}} \tag{2}$$

where z_i represents the output score corresponding to class i, and K denotes the total number of classes (four in this case). The class with the highest probability is selected as the final prediction.

3.4. Evaluation Metrics

The performance of the proposed NEUROVISION-AI model is evaluated using standard classification metrics derived from the confusion matrix. These include Accuracy**,** Precision**,** Recall (Sensitivity)**,** and F1-Score, which collectively measure classification effectiveness across all disease stages.
Accuracy is defined as:

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{3}$$

Precision measures the proportion of correctly predicted positive observations:

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

Recall (Sensitivity) represents the fraction of actual positive samples correctly identified:

$$Recall = \frac{TP}{TP+FN} \tag{5}$$

The F1-Score provides a harmonic balance between Precision and Recall:

$$F1\text{-}Score = \frac{2\times Precision\times Recall}{Precision+Recall} \tag{6}$$

These evaluation metrics ensure a comprehensive assessment of the classification performance across all Alzheimer's Disease stages. Fig. 1 shows the architecture of the proposed Alzheimer's disease detection system.

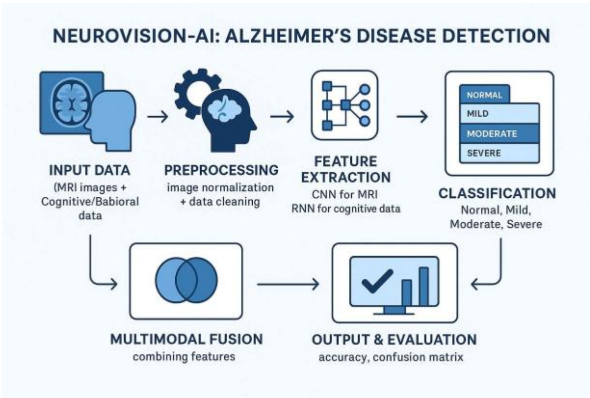

Fig. 1. Architecture of the proposed Alzheimer's disease detection system

4 Results and Discussion

The proposed NEUROVISION-AI model demonstrated exceptional performance in classifying the four stages of Alzheimer's Disease. The experiments culminated in an overall validation accuracy of 99%, indicating a high degree of reliability in distinguishing between Normal, Mild, Moderate, and Severe cases. The detailed classification report further substantiates this, showing consistently high precision, recall, and F1-scores across all four classes. This robust performance validates the hypothesis that fusing multimodal data surpasses the capabilities of single-modality approaches. Analysis of the model's training progression reveals excellent stability and generalization. The training and validation accuracy and loss curves are closely aligned, which indicates that the model learned effectively without significant overfitting. This stability is critical for ensuring its reliability for clinical use.

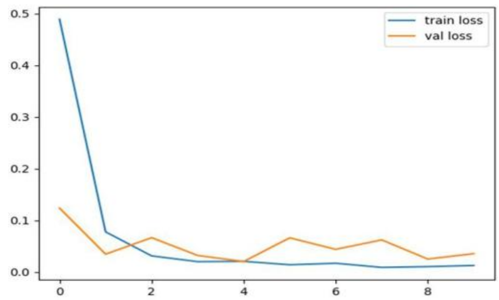

Fig. 2. Training and Validation Loss Curves

Fig. 2 illustrates the variation in training and validation loss across epochs during model training. The x-axis represents the number of epochs, while the y-axis indicates the corresponding loss values. The training loss shows a steep decline in the initial epochs and stabilizes thereafter, demonstrating effective learning by the model. The validation loss closely follows the training trend, indicating minimal overfitting and good generalization performance.

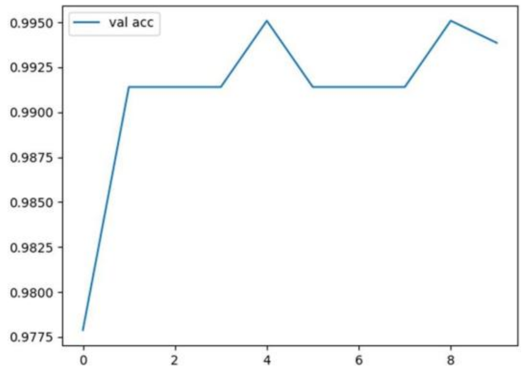

Fig. 3. Validation Accuracy Curve

The graph in Fig. 3 displays the validation accuracy of the model across training epochs. The x-axis represents the number of epochs, while the y-axis represents the validation accuracy (in percentage). The curve shows a rapid increase in accuracy during the initial epochs, followed by stabilization near 99%, indicating that the model achieves high performance with consistent accuracy over training iterations.

Table 1. Model Performance Comparison

Model	Input	Accuracy	F1-score	AUC
CNN (MRI only)	MRI	94%	0.92	**0.95**
RNN (Cognitive only)	Cognitive Data	90%	0.89	**0.91**
Hybrid CNN- RNN (Proposed)	**MRI+ Cognitive**	**99%**	**0.98**	**0.99**

Table 1 shows the model performance comparison in which the hybrid approach achieved 99% accuracy, 0.98 F1-score, and 0.99 AUC. In comparison, the CNN-only model achieved 94% accuracy with an F1-score of 0.92 and an AUC of 0.95, while the RNN-only model achieved 90% accuracy with an F1-score of 0.89 and an AUC of 0.91. These results clearly demonstrate that combining spatial MRI features with temporal cognitive assessment data significantly improves classification performance. A closer look at the confusion matrix reveals near-perfect classification for the Normal, Mild, Moderate, and Severe stages, with high values concentrated along the diagonal. The model correctly classified 205 'Normal', 489 'Mild', and 89 'Severe' cases in the validation set. A minor area for improvement was noted in the 'Moderate' class, where only 5 out of 31 cases were misclassified. This is a strong result, and these few misclassifications may be attributed to the subtle features that can overlap between disease stages.

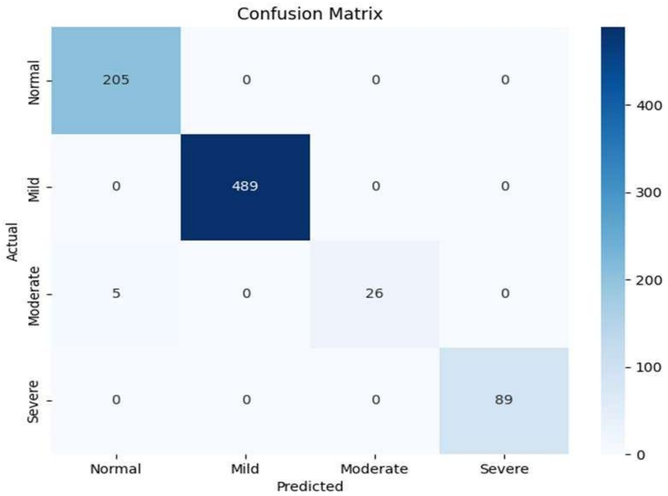

Fig. 4. Model Performance on Four-Stage Alzheimer's Classification

The confusion matrix in Fig. 4 details the model's classification performance across the four stages: Normal, Mild, Moderate, and Severe. The results show exceptional accuracy, with predictions heavily concentrated on the diagonal. The model achieved perfect classification for the 'Normal', 'Mild', and 'Severe' classes. The only error was the misclassification of 5 'Moderate' cases as 'Normal,' a clinically plausible result given the subtlety between stages. Overall, the matrix confirms the model's high fidelity and robust ability to distinguish between the different phases of Alzheimer's Disease. The overwhelmingly positive results support the potential of NEUROVISION-AI as a powerful diagnostic support tool to aid neurologists in making early and accurate diagnoses.

5 Conclusion

This paper presented NEUROVISION-AI, a multimodal deep learning framework designed for the early detection and classification of Alzheimer's Disease using structural MRI scans and cognitive assessment data. By integrating a Convolutional Neural Network for spatial feature extraction and a Recurrent Neural Network for temporal pattern analysis, the proposed system effectively captured both structural brain changes and behavioural variations associated with disease progression. The experimental results demonstrated a high validation accuracy of 99%, along with strong precision, recall, and F1-score across all four stages: Normal, Mild, Moderate, and Severe. The fusion of multimodal data significantly improved classification performance compared to single-modality approaches. These results highlight the potential of the proposed system as a reliable clinical decision-support tool for early diagnosis and monitoring of Alzheimer's Disease progression. Future work will focus on evaluating the model on larger datasets and integrating additional clinical parameters to further enhance robustness and real-world applicability.

Funding Information
This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Ethics Statement
This study did not involve human or animal subjects and, therefore, did not require ethical approval.

Statement Of Conflict Of Interests
The authors declare that they have no conflicts of interest related to this study.

References

[1] H S. Qiu *et al.*, "Multimodal deep learning for Alzheimer's disease dementia assessment," *Nature Communications*, vol. 13, no. 1, p. 3404, Jun. 2022, doi: 10.1038/s41467-022-31037-5.

[2] V. H. Jasodanand *et al.*, "AI-driven fusion of multimodal data for Alzheimer's disease biomarker assessment," *Nature Communications*, vol. 16, no. 1, p. 7407, Aug. 2025, doi: 10.1038/s41467-025-62590-4.

[3] M.-W. Lee *et al.*, "A multimodal machine learning model for predicting dementia conversion in Alzheimer's disease," *Scientific Reports*, vol. 14, no. 1, p. 12276, May 2024, doi: 10.1038/s41598-024-60134-2.

[4] V. Adarsh, G. R. Gangadharan, U. Fiore, and P. Zanetti, "Multimodal classification of Alzheimer's disease and mild cognitive impairment using custom MKSCDDL kernel over CNN with transparent decision-making for explainable diagnosis," *Scientific Reports*, vol. 14, no. 1, p. 1774, Jan. 2024, doi: 10.1038/s41598-024-52185-2.

[5] M. Odusami, R. Maskeliūnas, R. Damaševičius, and S. Misra, "Machine learning with multimodal neuroimaging data to classify stages of Alzheimer's disease: a systematic review and meta-analysis," *Cognitive Neurodynamics*, vol. 18, no. 3, pp. 775–794, Aug. 2023, doi: 10.1007/s11571-023-09993-5.

[6] Y. Liu, L. Wang, X. Ning, Y. Gao, and D. Wang, "Enhancing early Alzheimer's disease classification accuracy through the fusion of sMRI and rsMEG data: a deep learning approach," *Frontiers in Neuroscience*, vol. 18, p. 1480871, Nov. 2024, doi: 10.3389/fnins.2024.1480871.

[7] Z. Zhao *et al.*, "Vision transformer-equipped Convolutional Neural Networks for automated Alzheimer's disease diagnosis using 3D MRI scans," *Frontiers in Neurology*, vol. 15, p. 1490829, Dec. 2024, doi: 10.3389/fneur.2024.1490829.

[8] J. Feng, X. Zhao, Z. Liu, Y. Ding, and F. Wang, "A multi-view multimodal deep learning framework for Alzheimer's disease diagnosis," *Frontiers in Neuroscience*, vol. 19, p. 1658776, Oct. 2025, doi: 10.3389/fnins.2025.1658776.

[9] E. U. Haq, Q. Yong, Z. Yuan, X. Huarong, and R. U. Haq, "Multimodal fusion diagnosis of the Alzheimer's disease via lightweight CNN-LSTM model using magnetic resonance imaging (MRI)," *Biomedical Signal Processing and Control*, vol. 104, p. 107545, Jan. 2025, doi: 10.1016/j.bspc.2025.107545.

[10] B. Gupta *et al.*, "Multimodal lightweight neural network for Alzheimer's disease diagnosis integrating neuroimaging and cognitive scores," *Neuroscience Informatics*, vol. 5, no. 3, p. 100218, Jul. 2025, doi: 10.1016/j.neuri.2025.100218.

[11] S. Zhang *et al.*, "Machine learning on longitudinal multi-modal data enables the understanding and prognosis of Alzheimer's disease progression," *iScience*, vol. 27, no. 7, p. 110263, Jun. 2024, doi: 10.1016/j.isci.2024.110263.

[12] M. G. Alsubaie, S. Luo, and K. Shaukat, "Alzheimer's Disease Detection Using Deep Learning on Neuroimaging: A Systematic review," *Machine Learning and Knowledge Extraction*, vol. 6, no. 1, pp. 464–505, Feb. 2024, doi: 10.3390/make6010024.