Cureus Journal of Computer Science
https://doi.org/10.7759/s44389-026-00102-1

Original Article

A Multi-Stream Convolutional Neural Network Architecture for Alzheimer's Disease Classification

Thati Ravi Prasad [1], Suresh Mamidisetti [2], Rajani K.V. [3]

1. *Department of Computer Science and Engineering, School of Computer Science and Engineering, GITAM Deemed to be University, Hyderabad, IND*
2. *Department of Computer Science and Engineering, Government Polytechnic for Women, Pebbair, IND*
3. *Department of Computer Science and Engineering, Keshav Memorial Institute of Technology, Hyderabad, IND*

Received: March 11, 2026 | Review began: April 08, 2026 | Review ended: May 24, 2026 | Published: June 05, 2026

Abstract

Alzheimer's disease (AD) represents a progressively advancing neurodegenerative condition that considerably disrupts memory and cognitive abilities, thereby underscoring the necessity for early and precise diagnosis to enable effective intervention. Existing diagnostic techniques encounter difficulties in consistently differentiating AD from the typical aging process owing to similarities in imaging characteristics. A framework based on deep learning, integrated with critical metrics obtained from MRI scans such as cortical thickness, hippocampal volume, and white matter integrity, is proposed in the present research to address this issue and enhance diagnostic precision. The methodology involves a convolutional neural network architecture specifically structured to analyze such metrics, allowing for excellent classification between AD instances and non-AD ones. The model architecture includes specific layers that identify more subtle and unique changes in brain structure related to AD. Detailed overall assessments of the method on a broad level report strong performance, with a test set accuracy of 91.4%, in addition to precision and recall for AD and non-AD classes of 88.2% and 88.9%, respectively. An area under the curve of 0.92 on the test data indicates strong discrimination properties. The present method alleviates current limitations in diagnosis and provides a scalable solution that is reliable for application in clinical settings for the detection of neurodegenerative diseases. The results suggest that this model holds great promise for improving the accuracy of AD diagnosis in healthcare institutions.

Categories: AI/ML-based decision support systems, Health Informatics, Medical Expert systems

Keywords: alzheimer's disease, deep learning, mri, cortical thickness, hippocampal volume, white matter integrity, convolutional neural network (cnn), diagnostic accuracy

Introduction

Alzheimer's disease (AD) is a degenerative neurological disease primarily affecting memory and other higher cortical functions. As one of the leading causes of dementia among the elderly, AD poses great challenges in making an early diagnosis so that necessary treatment can be provided. Early identification of AD is crucial in establishing effective treatments that could have a chance of slowing down the progression of this disorder while improving the patients' quality of life [1,2]. Traditionally, diagnosis relies on clinical assessments, cognitive tests, and biomarkers like cerebrospinal fluid analysis or amyloid positron emission tomography scans [3]. However, these methods are often invasive, expensive, or subject to subjective interpretation.

In recent studies, machine learning-based and deep learning-based techniques have shown good potential for automation and enhancement of AD classification [4]. With the use of large datasets such as magnetic resonance imaging (MRI) images, this approach allows consideration of structural as well as functional changes in the brain to predict the onset and progression of AD. Machine learning algorithms excel at identifying patterns within high-dimensional data, while deep learning, particularly convolutional neural networks (CNNs), can automatically learn features from raw data, making them highly suited for imaging-based tasks [5].

Classifying AD through machine learning methods generally relies on features extracted from MRI data and fitting them for training classifiers, such as support vector machine (SVM), random forests, or k-nearest neighbors (k-NN) [6]. These methods depend significantly on manually defined features such as hippocampal volume, cortical thickness, and white matter integrity, all of which have been known to be degraded in Alzheimer's patients. These characteristics are used as inputs for machine learning models designed to classify patients into AD, mild cognitive impairment, or normal cognition groupings.

Traditional approaches to machine learning have proved satisfactory but are limited because feature extraction depends on human processes and can easily miss intricate patterns and high-dimensional features that may indicate the early stages of a disease. Furthermore, feature selection is often domain-specific; thus, it may not generalize across different domains [7,8].

Deep learning methods using CNNs have emerged as a powerful alternative for the classification of AD. CNNs are able to analyze MRI scans directly, without the need to extract manual features. Instead, they learn hierarchical representations in a hierarchical way, from low-level pixel intensities up to high-level patterns of brain structures [9,10]. Such approaches can be used in AD classification tasks and adapted in medical healthcare for better treatment. The advent of 3D-CNNs, which are able to process volumetric data such as MRI scans, has greatly improved AD classification, since the models can now extract both spatial and volumetric information from brain scans [11,12].

This work aims to propose machine learning and deep learning procedures for Alzheimer's classification. These include:

1. SVMs: They are suited to the training of multiple classes, such as Alzheimer's and non-Alzheimer's, using a hyperplane; SVM performs well in classification when applied along with extracted features like hippocampal volume or cortical thickness.

2. Random forests: They use an ensemble learning method that combines the predictions generated by multiple decision trees. Its popularity is mainly based on its robustness and efficiency in addressing overfitting issues. There is also, for instance, the structural classification of the brain MRI images.

3. k-NN: k-NN uses a technique for classifying a sample based on which of its nearest neighbors in feature space belongs to the majority class. Although it can work well when features are carefully selected, k-NN is unable to handle the high-dimensional data characteristic of MRI scans.

4. CNNs: This is a deep learning technique extensively used in image classification tasks in medical healthcare applications, including AD classification. Features of MRI data are learned automatically, and CNNs capture various complex spatial relationships inside the brain [13,14].

5. Recurrent neural networks: Whereas recurrent neural networks have gained most popularity in time-series analysis, they have been explored in the Alzheimer's research domain to help understand the temporal progression of the disease from a longitudinal MRI analysis.

Notwithstanding this progress, a lot of challenges remain. These include the requirement for very large annotated datasets, the tendency to overfit when small training sets are used, and interpretability issues concerning deep learning models. While existing machine learning and deep learning methods have demonstrated promise in AD classification, several limitations persist:

1. Manual Feature Extraction: Traditional machine learning models rely on hand-designed features that require laborious effort to extract. This approach cannot effectively capture the complexity of degenerative brain changes relevant to AD diagnosis [15].

2. Data Dependency: Deep learning architectures require massive amounts of labeled data, which often proves challenging in the medical imaging context. Many existing datasets are small in size, and thus their deep learning models are prone to overfitting [16].

3. Black Box Nature: CNN-based deep learning models are often described as "black boxes" since their decision-making process is not easily interpretable. This lack of interpretability poses challenges in medical applications, where clinicians require clear explanations for predictions [17].

4. Limited Multi-Modal Data Integration: Most models focus solely on the MRI data, leaving out other relevant clinical data, which includes demographic variables, cognitive assessment scores, or genetic indicators [18,19].

Our proposed multi-modal approach will address some of the basic limitations of the above methods. Multiple streams of data, including structural MRI features and auxiliary clinical data, will be combined in a CNN to facilitate the use of more comprehensive information for improved classification performance. Methods based on CNNs automatically extract meaningful features from MRI data without elaborate manual feature crafting. In addition, by combining clinical data and characteristics of brain imaging, our model captures more subtle representations encompassing the structural and non-structural dimensions of the disease.

In this work, CNN architecture is developed to prevent overfitting using dropout and regularization techniques to ensure that the model generalizes well even when data available is limited. Our method provides a framework for the classification of AD, paving the way for more accurate and reliable diagnosis in clinical settings.

Basaia et al. [1] analyzed predictive modeling and pattern recognition to solve complex medical problems. Their work aimed to utilize a CNN to classify Alzheimer's-affected brains from healthy ones. It demonstrated the potential for developing predictive models to identify disease stages in non-clinical settings. Classifying this disease using clinical data such as Alzheimer's remains a significant challenge due to the selection of discriminative features. They developed a LeNet-5 architecture and achieved a classification accuracy of 96.85% on functional MRI data. This approach demonstrates the efficacy of CNNs in distinguishing between healthy and AD-affected individuals in non-clinical data settings within predictive systems.

Tian et al. [3] found that AD is an intractable chronic malady marked by cerebral degeneration that results in memory deficits and cognitive impairment. Their work explains how machine learning methodologies have been extensively leveraged in medical research, particularly for the classification of AD. This paper reviews several studies that employ different machine learning approaches to improve the understanding of existing techniques in the classification of AD. The evaluation focuses on the effectiveness of methods such as SVM, artificial neural networks, and deep learning.

Qiu et al. [4] analyzed that AD is the primary cause of dementia, and its rising incidence presents significant challenges for both diagnosis and treatment. Current work relies on diagnostic techniques that depend on patient history, neuropsychological assessments, and MRI; however, these methods demonstrate inconsistent sensitivity and specificity. Their research work proposed a novel deep learning framework designed to evaluate multimodal data, encompassing MRI, age, gender, and cognitive scores, with the objective of identifying distinct signatures associated with AD. It showed high diagnostic accuracy based on data sourced from the Alzheimer's Disease Neuroimaging Initiative and validated through three separate cohorts. The deep learning model's performance was superior to that of neurologists and highly correlated with post-mortem findings, suggesting a promising methodology for the diagnosis of AD based on current imaging modalities.

Ghosh et al. [5] analyzed that AD is a degenerative brain disorder, which affects approximately 45 million people, primarily impacting older individuals. This condition leads to dementia, causing progressive damage to brain cells and impairing cognitive functions. This paper investigates the application of machine learning techniques to predict dementia, utilizing the Open Access Series of Imaging Studies dataset, which, despite its limited size, provides significant

insights. Several machine learning models, including SVM, logistic regression, random forest, and decision tree, were analyzed using both, with and without fine-tuning. The findings indicate that the SVM model demonstrated the highest accuracy in detecting dementia to assist in early diagnosis of the disorder.

Helaly et al. [6] presented a novel deep ensemble-based learning framework tailored for AD classification. The framework leverages the synergistic combination of various deep learning algorithms and multisource data to enhance predictive performance beyond what any single algorithm can achieve. Specifically, it employs sparse autoencoders for feature extraction and a deep neural network for providing ranks to the classifiers, addressing conditional independence concerns. Additionally, techniques such as over-sampling and threshold-moving are utilized to tackle cost-sensitive challenges. Experimental results demonstrate that this framework outperforms six well-established ensemble methods, underscoring its potential to bolster diagnostic services in Alzheimer's care through the power of machine learning.

Shahbaz et al. [11] discussed that AD causes a cognitive decline that results in means to slow its treatment progression. While various interventions can assist in managing symptoms at different stages, early identification and classification of these stages are critical for providing effective care. The growing adoption of electronic health records in healthcare enables the application of machine learning and data mining methods to enhance patient management. This investigation utilized six algorithms, including k-NN and generalized linear models, on the Alzheimer's Disease Neuroimaging Initiative dataset to classify five states of the disease. The results demonstrate that the generalized linear model achieved an accuracy of 88.24%, underscoring the capability of the proposed techniques for early diagnosis and treatment in this domain.

Khan et al. [13] demonstrated an end-to-end framework employing CNN for the early detection and classification of AD across multiple stages. The framework involved both multi-class classification of four AD stages and binary classification of stage pairs, utilizing simple CNN architectures as well as transfer learning with pre-trained models like VGG19. Given the COVID-19 pandemic, a web application was proposed to enable remote AD assessments, assisting in determining patient stages and providing guidance. The results demonstrate the CNN methods achieved accuracies of 93.61% and 95.17% for 2D and 3D classifications, respectively. With the help of fine-tuned hyperparameters, VGG19 model reached an accuracy of 97%. This finding highlights the potential of machine learning techniques in the early detection and management of AD, a chronic brain disorder currently lacking an effective cure.

Marcus et al. [18] investigated the use of MRI in the analysis of neurological disorder conditions and the identification of pathological brain regions. Detailed examination of tissue structure from segmented MRI data enhances the accuracy of AD classification. Deep learning methodologies for brain segmentation and classification have gained widespread recognition due to their effectiveness in handling large datasets, surpassing traditional machine learning techniques. This paper presents an overview of current deep learning-based segmentation approaches, highlighting the use of CNN for analyzing brain anatomy and improving AD diagnosis. Additionally, it discusses state-of-the-art methods, summarizes their performance on publicly available datasets, and explores future research directions for developing computer-assisted diagnostic systems for AD.

Materials And Methods

Figure *1* provides an overview of the proposed multidimensional deep learning approach to improve AD classification. The following subsequent sections explain the four steps that are involved in this current work.

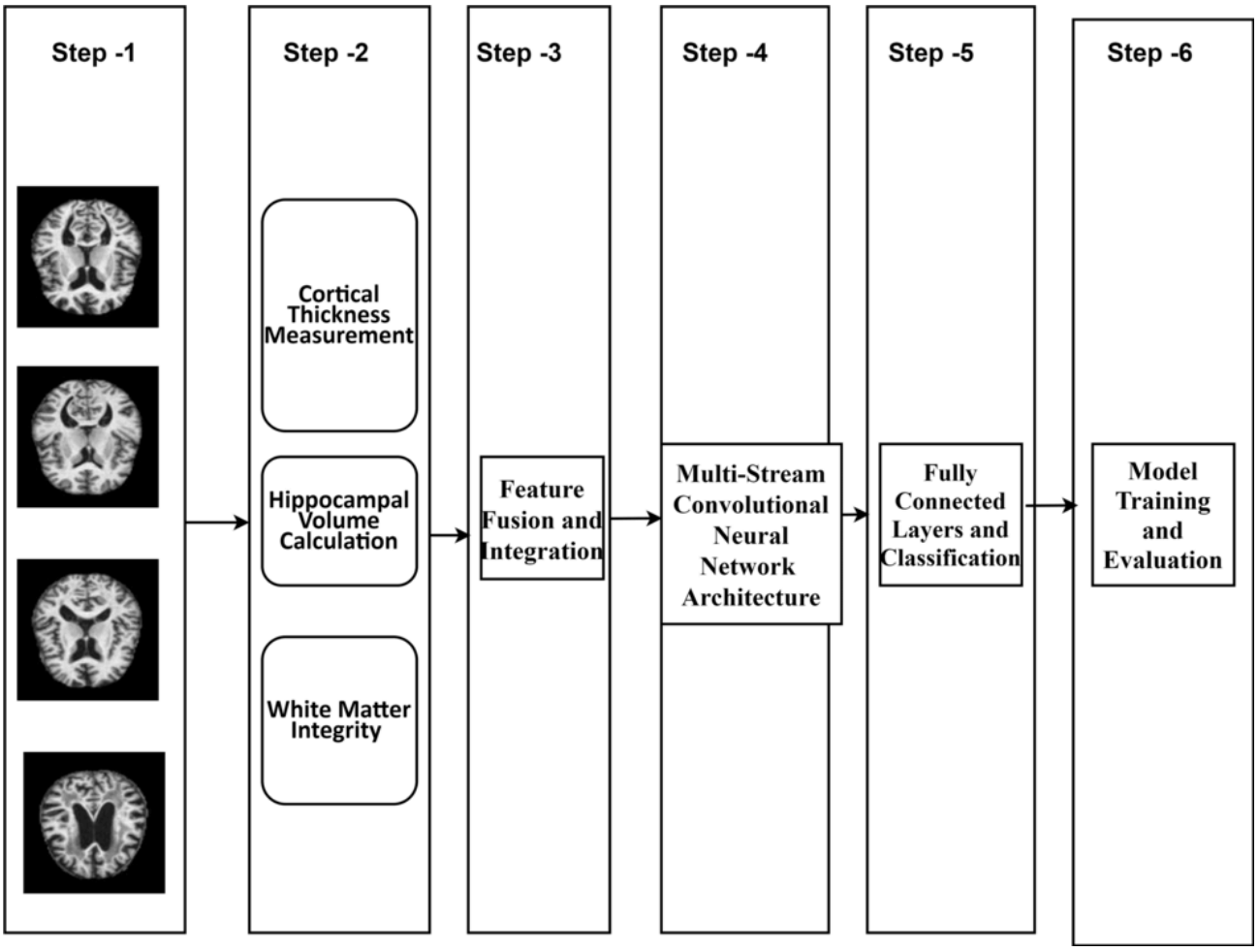

FIGURE 1: Overview of the Proposed CNN Network

CNN, Convolutional Neural Network

Initially, three kinds of features are extracted from MRI images: cortical thickness, hippocampal volume, and white matter integrity. Once features are extracted from all inputs, the fusion layer combines the CNN stream outputs along with auxiliary data, allowing the network to learn complex brain structures and clinical factors. Fully connected layers take the fully fused features, and the classification results determine whether the patient is more likely to develop AD or remain cognitively healthy.

The following analysis details each step in depth, focusing on the approaches used and the relevance of the metrics selected to this process. The next sub-sections describe each of these steps in more detail.

Step 1: Data acquisition and preprocessing

Technology Applied: Techniques of Preprocessing MRI Data

Metrics Converted: Brain Structural Features and Image Quality Metrics

The MRI brain images are downloaded from the Open Access Series of Imaging Studies, or OASIS dataset [17,18]. All MRI data require some form of preprocessing to standardize images, reduce noise, and ensure that brain structures are uniformly analyzed across different samples. Methods used in MRI pre-processing are as follows:

1. Skull Stripping: Removing non-brain tissues to ensure that only relevant brain structures are analyzed. This removal of non-brain tissues ensures that the region of interest is defined, and it is done manually by editing the images in the dataset.

2. Intensity Normalization: Standardizing intensity values across scans to mitigate differences due to varying MRI scanner settings.

3. Registration and Alignment: Registration of MRI scans on a standard reference template, for example, the template MNI152, to ensure spatial consistency.

4. Segmentation: Dividing the brain into regions of interest such as the hippocampus, cortex, and white matter.

All the preprocessing operations above improve the accuracy and reliability of subsequent feature extraction by allowing the segregation of important structural features that may be degenerating due to AD.

Step 2: Key metrics extraction

Method Applied: Feature Extraction via Structural MRI

Metrics Evaluated Cortical Thickness, Hippocampal Volume, and White Matter Integrity

This stage entails the recovery of the key indicators that indicate progression in AD:

1. Cortical Thickness Measurement: Cortical thickness is calculated as the distance between two surfaces: the pial surface is considered as the outer surface of the cortex, and the gray matter-white matter boundary. This measure is usually calculated for each point on the cortical surface in general.

Let us define the variables:

• (p_i, q_i): Coordinates of a point on the pial surface.

• (p_j, q_j): Coordinates of the corresponding point on the gray matter-white matter boundary.

The cortical thickness at each point is calculated using the Euclidean distance between these two surfaces at each point:

$$\mathbf{Distance}(p, q) = \sqrt{(p_i - p_j)^2 + (q_i - q_j)^2} \quad (1)$$

2. Calculate Hippocampal Volume: To calculate the hippocampal volume, the segmented hippocampal region from the MRI scan is divided into voxels (3D pixels). Each voxel represents a small cubic region of the brain, and the volume of the hippocampus is determined by counting the number of voxels within the hippocampal region and multiplying by the volume of each voxel.

Let us define:

• N_v: Number of voxels in the segmented hippocampal region.

• V_v: Volume of a single voxel.

The hippocampal volume is given by:

$$VH = (N_v \times V_v) \quad (2)$$

Description:

• N_V is the number of voxels identified as part of the hippocampus in the segmented MRI.

• V_V is the volume of a single voxel, typically calculated.

The total hippocampal volume would then be the sum of all voxels in the hippocampal region multiplied by this unit voxel size. White matter integrity (fractional anisotropy [FA] and mean diffusivity): FA is a value derived from DTI and provides a measure of the directional variance of water diffusion; a high FA value indicates increased white matter integrity.

The FA at a voxel is computed using the eigen values of the diffusion tensor λ_1,λ_2,λ_3, which represent the principal axes of diffusion in 3D space.

The formula for FA is as follows:

$$FA = \lambda = \frac{\lambda_1 + \lambda_2 + \lambda_3}{3} \quad (3)$$

Step 3: Feature fusion and integration

Method Employed: Fusion Layer for Multi-Modal Integration

Metrics Studied: Composite Traits Inferred From Cortical Thickness, White Matter Integrity, and Hippocampal Volume

In the Fusion Layer, each of the output features coming from different CNN streams is combined to provide an integration of learned representations from different sources. That way, a model learns to capture sophisticated interactions between cortical, white matter, and hippocampal features. Accumulative feature maps in the Fusion Layer contribute toward a consolidated representation with all critical details from each data source. Fully connected independent networks process auxiliary clinical information, such as scores of cognitive assessments and demographic details, and pass this into Fusion Layer, so that the model can assimilate the non-imaging data in its decision framework.

Step 4: Multi-stream CNN architecture

Methods Used: Multistream CNN

Metrics Analyzed: Extracted Features for Cortical Thickness, White Matter Integrity, and Hippocampal Volume

We apply a multi-stream CNN architecture on each of the metrics extracted in the following step:

1. Cortical Thickness Flow: The cortical thickness map is specifically fed into a specific CNN using convolutional layers to detect structural features within the cortex across different regions. Architectures can contain multiple convolutional layers with ReLU activations, followed by pooling layers that capture spatial patterns.

2. White Matter Integrity Stream: A similar CNN learns subtle patterns in DTI metrics, capturing how the integrity of white matter tracts is compromised in Alzheimer's.

3. Flow of Hippocampus Volume: Since the volume of hippocampus is a scalar feature, this is processed in a fully connected network rather than in a CNN; hence, this volume can now be integrated as an independent input.

Each stream is specifically designed to manage its corresponding input proficiently, thereby enabling the model to acquire structural information pertinent to disease from various sources.

Step 5: Fully connected layers and classification

Method Used: Fully Connected Layers for Feature Representation of High-Level Features

Metrics Processed: Consolidated Characteristic Representations

After the Fusion Layer, the fused features pass through a number of fully connected layer. These layers are designed to find non-linear relationships and higher-level abstractions, therefore enriching the features in terms of classification. For each layer, weights and biases are used in computing the integrated features; they help the model at every step learn complex relationships. For preventing overfitting, dropout and batch normalization are used; thereby, the model still keeps its strong generalization even with a constrained training set.

Output from the final fully connected layer is passed through a softmax classification layer, which produces probability scores for all classes and categorizes patients into either AD or Non-AD groups.

Step 6: Model training and evaluation

Method Applied: Training With Cross-Entropy Loss and Performance Metrics

Accuracy, Precision, Recall, and F1-score Metrics

The final step is training the proposed model on the OASIS dataset. Model optimization utilizes cross-entropy loss, a measure that calculates the difference between predicted and true labels, thus driving the model towards improved classification accuracy. Data augmentation strategies are incorporated during the training procedure to minimize the risk of overfitting as well as incorporating dropout layers to strengthen robustness.

We used a range of critical metrics to evaluate the model's performance:

1. Accuracy measures the proportion of correctly classified instances.

2. Precision assesses the model's ability to identify positive cases (AD) without misclassifying Non-AD cases as AD.

3. Recall (Sensitivity) evaluates the model's capacity to identify all true positive cases.

4. F1-Score takes a balance between precision and recall, providing instead a more holistic measure of the model's performance.

This metric provides a comprehensive view of the model's performance, thus allowing for an estimation of its practical clinical setting effectiveness.

Results And Discussion

Table *1* provides the results of the proposed model to classify between the AD and Non-AD cases in training, validation, and test datasets. In terms of accuracy, the model gave a performance of 95.3% on the training dataset. This capability decreases slightly up to 92.8% for the validation dataset and drops down to 91.4% for the test dataset. This consistent high accuracy across all datasets suggests that the model does a good job of separating between AD and Non-AD cases. The small drop in accuracy between the training and test set suggests minor overfitting but still good generalization to unseen data. High accuracy on the test set would suggest reliability for potential clinical application with only very minor performance decline outside of the training environment.

Accuracy-wise, the table here clearly depicts 93.1% accuracy for AD and 94.5% for Non-AD in the training set, but these values drop to 88.2% for AD and 89.0% for Non-AD on the test set; it is particularly important to show precision here because the value shows to what extent the model is able to classify the correct positive cases without mistakenly labeling some Non-AD samples as AD. The lower precision values from the testing dataset indicate that despite the general specificity of the model, there is a small trend to falsely classify Non-AD cases as AD. The kind of result is expected for complex datasets, but it means that false positives require close attention since misclassifications can lead to inappropriate diagnostic procedures. High enough precision values, however, suggest that the model is somewhat conservative in prediction, especially critical for clinical applications, where specificity in diagnosis counts.

Both categories also show consistent performance metrics across the datasets with a recall of 94.0% for the AD category in the training dataset, 91.2% in the validation dataset, and a recall of 88.9% in the testing dataset. Here, the model classifies actual AD as the true class accurately, which is crucial because failure to detect AD results in huge consequences. In the case of Non-AD recall, 92.9% on the training set, 90.3% on validation, and 88.3% on the test set show that the model also is good at the detection of Non-AD cases. A slight reduction in recall on the test set proves the model, even when generalizing to unseen data, loses a little sensitivity but retains a relatively equalized capability of detection without having the possibility of over-diagnosis.

The F1-score, a balanced metric that integrates both precision and recall, reveals the reliability of the model concerning both AD and Non-AD classes with values of above 88% across all datasets. Concerning AD classification, F1-scores of 93.5% on the training set, 90.8% on validation, and 88.5% on test data show that the model keeps a balanced performance between capturing AD cases and avoiding false positives. For the F1-scores of Non-AD cases, it was found that they have equal strength, meaning that the model equally predicts the true negatives without significant false positive rates. High F1-scores across both classes indicate that the model is robust and is effective in handling the complex, often overlapping features of AD that are hard to distinguish between Non-AD cases.

Figure *2* shows a sample AUC graph that can be measured as a performance parameter to measure the classifier's ability to discriminate between AD and Non-AD. Values were established at 0.96 for the training dataset, 0.93 for the validation dataset, and 0.92 for the testing dataset discriminative ability. The AUC values approaching 1 for all datasets indicate that the model shows tremendous discriminative ability between AD and Non-AD cases even when varying decision thresholds. This property makes the model both robust and reliable for practical applications where threshold updates may be needed.

Overall, the results show an extremely well-performing model with strong classification metrics that do not decay significantly from training to validation and then test datasets, indicative of great generalization ability. While the precision and recall values have only slightly decreased in the test set, they suggest that although the model has been optimized effectively, some minor changes by inclusion of a more diversified train set would enhance its performance. These results add strength to the model's promise as a valid and scalable tool for the diagnosis of AD, with both sensitivity and specificity important for clinical use.

Metric	Class	Training Set	Validation Set	Test Set
Accuracy	-	95.3%	92.8%	91.4%
Precision	AD	93.1%	90.5%	88.2%
	Non-AD	94.5%	91.7%	89.0%
Recall	AD	94.0%	91.2%	88.9%
	Non-AD	92.9%	90.3%	88.3%
F1-Score	AD	93.5%	90.8%	88.5%
	Non-AD	93.7%	91.0%	88.6%
Area Under the Curve		0.96	0.93	0.92

TABLE 1: The Results of the Proposed Model to Classify Between the AD and Non-AD

AD, Alzheimer's Disease

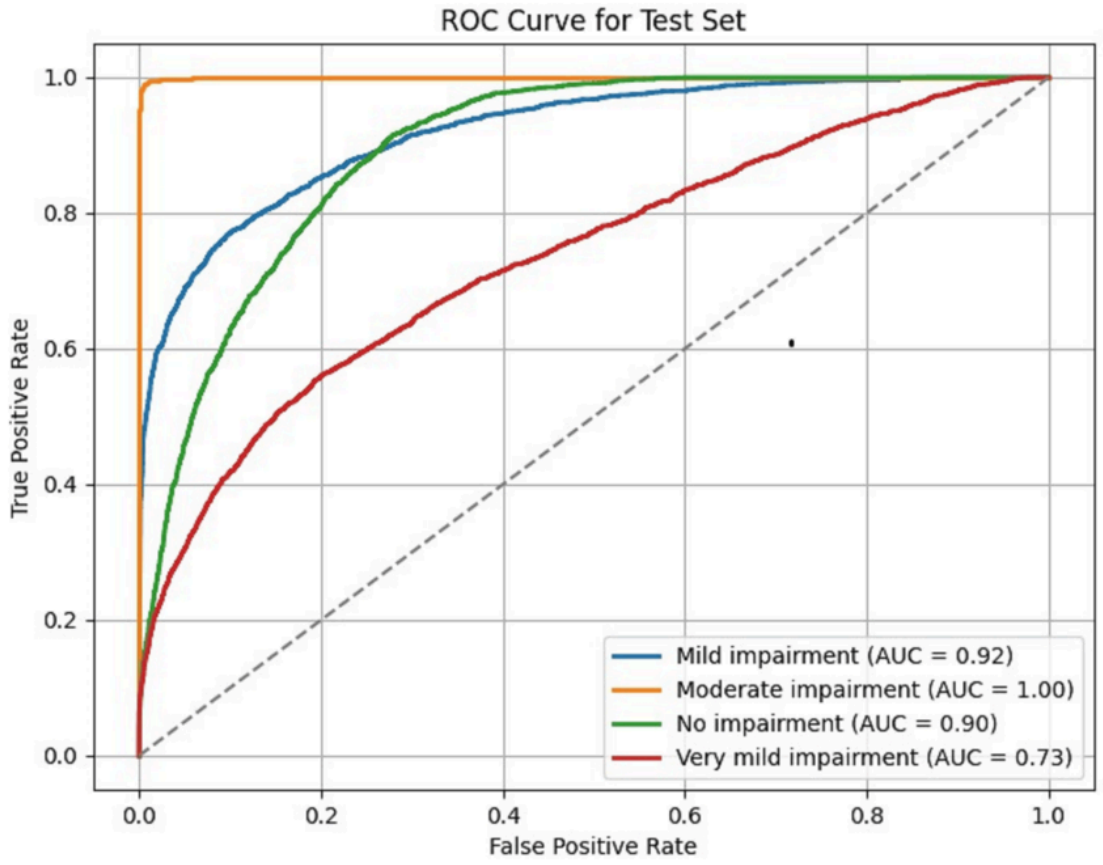

FIGURE 2: Sample AUC Graphs in Mild, Moderate, Very Mild, and No Impairments in Alzheimer Prediction

AUC, Area Under the Curve; ROC, Receiver Operator Characteristic

Conclusions

Automated and reliable classification methods are critical to aid clinicians in producing an accurate diagnosis of AD. The central challenge in existing diagnostic methods is the interaction of age-related changes and AD-specific alterations in the brain, most notably on imaging biomarkers such as cortical thickness and hippocampal volume. This interaction might result in false positives and false negatives, lowering the utility of MRI-based diagnostic methodologies. Our proposed deep learning framework partially addresses the issues mentioned by realizing the CNN model tailored to the key metrics derived from MRI: cortical thickness, hippocampal volume, and white matter integrity. This CNN architecture transforms raw MRI data into quantifiable and hence more relevant disease metrics and establishes efficient differences between AD and non-AD cases, giving importance to precision and recall in balancing sensitivity and specificity of the diagnosis. These models have been tested rigorously on training, validation, and test datasets. The promising result obtained here indicates the reliable clinical application of the model. Results for classification according to AD/Non-AD with high test accuracy of 91.4% were reported, with precision and recall rates of 88.2% and 88.9%, respectively. An AUC value of 0.92 indicates how the model clearly distinguishes cases of AD from those explained by normal aging at any different decision threshold. These results demonstrate that our CNN model dramatically reduces the existing diagnostic

inaccuracy and thus may pose as an additional resource as ancillary material for doctors. Their ability to learn and differentiate complex, disease-specific patterns indicates that they are robust and generalizable - the two qualities most important for effective implementation in real-world healthcare environments.

In summary, this study illustrates that a framework rooted in automated deep learning can substantially improve the classification of AD by utilizing MRI-based biomarkers. The suggested methodology not only attains elevated levels of classification accuracy and reliability but also offers a scalable and efficient means to facilitate early and precise diagnosis of AD within clinical settings. Further research studies will include testing of our proposed works on the publicly available datasets to ensure robustness of the current model. This refines this approach by adding ancillary imaging and non-imaging information, such as genetic markers and cognitive test scores, to add further strength to the accuracy and robustness of the model's diagnosis in clinically heterogeneous populations. This approach may be promising in the optimization of diagnostic maneuvers for neurodegenerative disorders, allowing for expedient intervention and better management of AD.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Thati Ravi Prasad, Rajani K.V.

Drafting of the manuscript: Thati Ravi Prasad

Acquisition, analysis, or interpretation of data: Suresh Mamidisetti

Critical review of the manuscript for important intellectual content: Suresh Mamidisetti, Rajani K.V.

Supervision: Rajani K.V.

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The data supporting this study are partially openly available through OASIS at https://sites.wustl.edu/oasisbrains/home/oasis-1/
Principal Investigators: D. Marcus, R. Buckner, J. Csernansky, J. Morris (P50 AG05681, P01 AG03991, P01 AG026276, R01 AG021910, P20 MH071616, U24 RR021382)

Acknowledgements

The authors wish to thank the management of GITAM School of Technology, Hyderabad, India.

References

1. Basaia S, Agosta F, Wagner L, et al.: Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI and deep neural networks. NeuroImage: Clinical. 2019, 21:101645. 10.1016/j.nicl.2018.101645

2. Chauhan A, Mohana R: PRISMA-based approach to study implicit & explicit aspect-based sentiment analysis. Multimedia Tools and Applications. 2025, 84:45347-45395. 10.1007/s11042-025-20871-6

3. Tian J, Smith G, Guo H, et al.: Modular machine learning for Alzheimer's disease classification from retinal vasculature. Scientific Reports. 2021, 11:238. 10.1038/s41598-020-80312-2

4. Qiu S, Joshi PS, Miller MI, et al.: Development and validation of an interpretable deep learning framework for Alzheimer's disease classification. Brain. 2020, 143:1920-1933. 10.1093/brain/awaa137

5. Ghosh M, Raihan MMS, Raihan M, Akter L, Bairagi AK, Alshamrani SS, Masud M: A comparative analysis of machine learning algorithms to predict liver disease. Intelligent Automation and Soft Computing. 2021, 30:917-928. 10.32604/iasc.2021.017989

6. Helaly HA, Badawy M, Haikal AY: Deep learning approach for early detection of Alzheimer's disease. Cognitive Computation. 2022, 14:1711-1727. 10.1007/s12559-021-09946-2

7. Yamanakkanavar N, Choi JY, Lee B: MRI segmentation and classification of human brain using deep learning for diagnosis of Alzheimer's disease: A survey. Sensors. 2020, 20:3243. 10.3390/s20113243

8. Thati RP, Dhadwal AS, Kumar P, Sainaba P: Multimodal depression detection: using fusion strategies with smart phone usage and audio-visual behavior. International Journal on Artificial Intelligence Tools. 2023, 32:2340008. 10.1142/s0218213023400080

9. An N, Ding H, Yang J, Au R, Ang TFA: Deep ensemble learning for Alzheimer's disease classification. Journal of Biomedical Informatics. 2020, 105:103411. 10.1016/j.jbi.2020.103411

10. Chauhan A, Mohana R: Combining transfer and ensemble learning models for image and text aspect-based sentiment analysis. International Journal of System Assurance Engineering and Management. 2025, 16:1001-1019. 10.1007/s13198-025-02713-8

11. Shahbaz M, Ali S, Guergachi A, Niazi A, Umer A: Classification of Alzheimer's disease using machine learning techniques. Proceedings of the 8th International Conference on Data Science, Technology and Applications DATA, Prague, Czech Republic. 2019, 1:296-303. 10.5220/0007949902960303

12. Thati RP, Dhadwal AS, Kumar P, Sainaba P: A novel multi-modal depression detection approach based on mobile crowd sensing and task-based mechanisms. Multimedia Tools and Applications. 2022, 82:4787-4820. 10.1007/s11042-022-12315-2

13. Khan FA, Khan A, Imran M, Ahmad A, Jeon G: Advancements in deep learning for Alzheimer's disease diagnosis: A comprehensive exploration and critical analysis of neuroimaging approaches. Expert Systems. 2025, 42:e13688. 10.1111/exsy.13688

14. Thati RP, Mamidisetti S: Apparent personality traits detection based on correlation-based attention and feature weighting methods. Advances in Computational Intelligence and Informatics. Chillarige RR, Distefano S, Rawat SS (ed): Springer, Singapore; 2024. 993:399-405. 10.1007/978-981-97-4727-6_40

15. Prasad TR, Naulegari J: Investigation of symmetric and asymmetric eye patterns for bell's palsy diagnosis. 2023 OITS International Conference on Information Technology (OCIT), Raipur, India. 2023, 655-659. 10.1109/OCIT59427.2023.10431126

16. Prasad TR, Meraj M, Kumar P: An investigation of ensemble-based framework for stress detection using multimodal clues. Proceedings of International Conference on Advanced Communications and Machine Intelligence. Peng SL, Mondal A, Kagita VR, Sarkar JL (ed): Springer, Singapore; 2024. 405:253-263. 10.1007/978-981-97-6222-4_20

17. LaMontagne PJ, Benzinger TLS, Morris JC, et al.: OASIS-3: longitudinal neuroimaging, clinical, and cognitive dataset for normal aging and Alzheimer disease [PREPRINT]. medRxiv, 2019. https://www.medrxiv.org/content/10.1101/2019.12.13.19014902v1.

18. Marcus DS, Fotenos AF, Csernansky JG, Morris JC, Buckner RL: Open Access Series of Imaging Studies: longitudinal MRI data in nondemented and demented older adults. Journal of Cognitive Neuroscience. 2010, 22:2677-2684. 10.1162/jocn.2009.21407

19. Prasad, TR, Kumar, P: A review of multimodal depression detection methods using smartphone usage and audio-visual clues. Iran J Comput Sci. 9:15. 10.1007/s42044-025-00359-0

How to cite this article:
Prasad T, Mamidisetti S, K.V. R (June 05, 2026) A Multi-Stream Convolutional Neural Network Architecture for Alzheimer's Disease Classification. Cureus J Comput Sci 3 : es44389-026-00102-1. DOI https://doi.org/10.7759/s44389-026-00102-1