A Hybrid CNN Framework for Robust multiclass Alzheimer's Disease Classification Using ADNI sMRI

Md. Shamim*[1] 22203149@iubat.edu
[1] *4 Embankment Drive Road, Sector-10, Uttara Model Town, Dhaka-1230, Bangladesh*
D.M Arif Afsar[1] afsar.cse@iubat.edu
Shahriar Chowdhury Imran[1] 22203098@iubat.edu
MD. TANVIR AHMED SIDDIQUE[1] 22203097@iubat.edu

Editors: Under Review for MIDL 2026

Abstract

Alzheimer's disease (AD) is a progressive neurodegenerative disorder that severely affects memory, cognition, and daily functioning, making early detection crucial for effective intervention and disease management. This study proposes an interpretable and bias-aware deep learning framework for early Alzheimer's disease classification using structural magnetic resonance imaging (sMRI) data. A hybrid convolutional neural network (CNN) was developed to extract deep spatial features from pre-processed T1-weighted MRI scans, which included global statistical features including the mean and standard deviation of pixel intensities. The dataset was obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) and underwent an extensive preprocessing pipeline, including N4 bias field correction, skull stripping, strict registration to MNI152 space, intensity normalization, and 2D slice extraction. To address class imbalance and fairness concerns, weighted sampling technique was used during training. The proposed model achieved high classification performance with a test accuracy of about 98%, demonstrating strong generalization ability. Furthermore, Grad-Cam++ was applied to generate class-specific attention maps, which enhance interpretability by highlighting anatomically relevant brain regions associated with Alzheimer's pathology. The integration of interpretability and bias-aware techniques reduces the black-box limitations of deep learning models and promotes credible AI in medical diagnosis. Overall, this study contributes to an accurate, interpretable, and ethically responsible framework for early-stage Alzheimer's disease detection. The code is available at `https://github.com/MohammadShamim-29/Alzheimar-Disease-Classification-with-ADNI-Data/tree/main`

Keywords: Alzheimer's Disease, Structural Magnetic Resonance Imaging, Hybrid Convolutional Neural Network, Explainable Artificial Intelligence, Grad-CAM++, Bias-Aware Learning, Neuroimaging Analysis.

1. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder which affects memory, thinking and behavior severely. Structural MRI (sMRI) can be used for early detection which is essential for timely intervention and disease management (Jack et al., 2008). Deep learning models like convolutional neural networks (CNNs) have shown promising performance in AD classification (Jo et al., 2019). However, two critical challenges limit their

* Contributed equally

clinical applicability: lack of interpretability and potential bias in predictions. Deep learning models often operate as black boxes, making it difficult for clinicians to trust their decisions (Arrieta et al., 2020). Additionally, imbalanced datasets may lead to biased performance across different demographic groups, raising concerns about fairness in medical diagnosis (Mehrabi et al., 2021). To address these challenges, we propose a bias-aware and explainable deep learning framework for AD classification. The proposed method integrates a hybrid CNN architecture that combines spatial features with statistical descriptors to enhance representation learning. A bias-aware training strategy is employed to mitigate class imbalance and evaluate fairness across subgroups. Furthermore, Grad-CAM++ is used to provide visual explanations, highlighting regions relevant to model predictions (Chattopadhyay et al., 2018). The main contributions of this work are:

- A hybrid feature fusion framework for improved classification
- Bias-aware training and fairness evaluation
- Explainability through Grad-CAM++ for enhanced model transparency.

2. Methods

Dataset: This study uses structural MRI (sMRI) data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) (Weiner et al., 2010), consisting of 735 T1-weighted scans from 370 male and 365 female subjects. The data is labeled into three classes: Cognitive Normal (CN), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD).

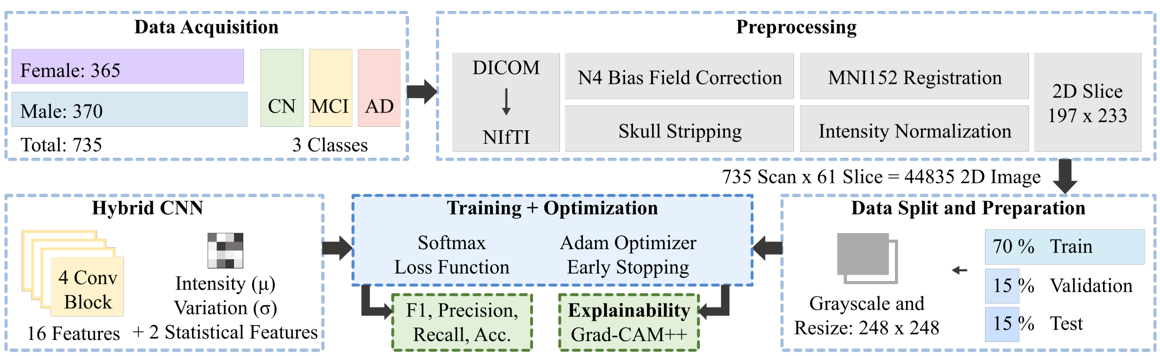

Figure 1: Workflow Diagram

Preprocessing: Each 3D volume is preprocessed and converted into 2D axial slices, resulting in 44,835 images. All MRI scans undergo N4 bias field correction, skull stripping, spatial registration to MNI152 space, and intensity normalization. Informative axial slices are extracted from each volume and resized for model input. The dataset is split into training, validation, and test sets in a 70:15:15 ratio. Use of subject wise split ensures reliability and prevents data leakage.

Hybrid CNN Architecture: A hybrid convolutional neural network (CNN) is proposed to classify MRI slices. The model consists of four convolutional blocks followed by fully connected layers to extract hierarchical spatial features. In addition, two global statistical features—mean and standard deviation of pixel intensities—are computed for each image and concatenated with the learned deep features to form a fused representation for final classification.

Bias-aware Training: This study uses weighted random sampling to address class imbalance. Furthermore, model performance is evaluated across demographic subgroups to assess fairness and reduce potential bias.

Explainability: Grad-CAM++ is used to generate class-specific activation maps, highlighting brain regions that contribute most to model predictions and improving interpretability (Chattopadhyay et al., 2018).

3. Results and Discussion

The proposed model achieved .98 accuracy with .98 precision, .98 recall and .98 F1-score which indicates strong classification performance across all classes. The model maintains high precision, recall, and F1-score across AD (.99, .96, .98), MCI (.97, .98, .98) and CN (.98, .98, .98). Notably, the performance on the MCI class demonstrates the model's ability to capture subtle differences between adjacent classes. The confusion matrix shown in Figure 2 illustrates the number of corrected and incorrected prediction for each class across CN, MCI and AD.

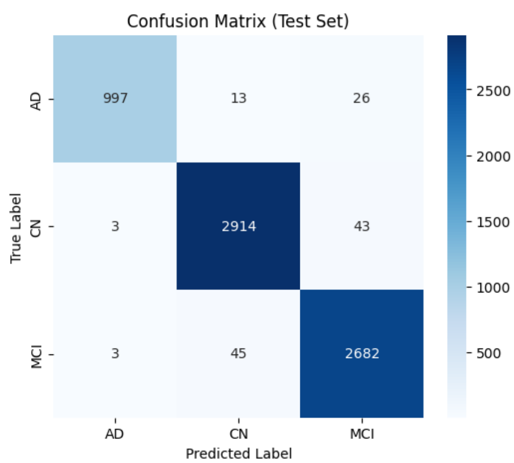

Figure 2: Confusion Matrix

Previous studies on ADNI-based 3-class Alzheimer's disease classification report moderate performance using CNN architectures. For example, Slimi et al. (2025) used CNN models and achieve only .7567 accuracy (Slimi et al., 2025). Ocen et al. (2025) used EfficientNet-B0 and ResNet-50 and achieved accuracy of .968 and .975 respectively (Ocen et al., 2025). On the other hand, the proposed model achieves an accuracy of .98, significantly outperforming existing approaches. This improvement can be attributed to enhanced feature extraction and improved representation of subtle structural variation in brain MRI. A comparative analysis with existing methods is presented in Table 1 to evaluate the effectiveness of the proposed model.

Table 1: Comparison table

Model	**Accuracy**	**Precision**	**Recall**	**F1-score**
Baseline CNN	.7567	.7522	.752	.7554
EfficientNet-B0	.968	.96	.96	.96
ResNet-50	.975	.97	.97	.97
Proposed Model	.98	.98	.98	.98

4. Conclusion

This paper presents a deep learning-based approach for 3-class classification of Alzheimer's disease using MRI data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. The proposed model effectively differentiates between AD, MCI, and CN by overcoming the challenge of intra-class subtle differences. The model achieves .98 accuracy, which is better than several existing CNNs. Specifically, its improved performance in distinguishing between MCI and CN, demonstrating its robust feature representation capability. The use of subject-based data segmentation further ensures reliable evaluation. The use of Grad-CAM++ increases the interpretability of the model by identifying important brain regions that contribute to classification decisions. However, the model has been validated on a single dataset, and future work will focus on testing its generalization to external datasets and incorporating multimodal data.
Overall, the proposed method provides an effective and accurate solution for automatic classification of Alzheimer's disease.

Acknowledgments

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu). As such, the investigators within ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in the analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf.

References

Alejandro Barredo Arrieta, Natalia Díaz-Rodríguez, et al. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*, 58:82–115, 2020.

Aditya Chattopadhyay, Anirban Sarkar, Prantik Howlader, and Vineeth N Balasubramanian. Grad-cam++: Improved visual explanations for deep convolutional networks. *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 839–847, 2018.

Clifford R Jack, Matthew A Bernstein, Nick C Fox, Paul Thompson, Gene Alexander, David Harvey, et al. The alzheimer's disease neuroimaging initiative (adni): Mri methods. *Journal of Magnetic Resonance Imaging*, 27(4):685–691, 2008.

Taehoon Jo, Kwangsik Nho, and Andrew J Saykin. Deep learning in alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data. *Frontiers in Aging Neuroscience*, 11:220, 2019.

Ninareh Mehrabi, Fred Morstatter, Nripsuta Saxena, Kristina Lerman, and Aram Galstyan. A survey on bias and fairness in machine learning. *ACM Computing Surveys*, 54(6):1–35, 2021.

Samuel Ocen, Lawrence Muchemi, and Michaelina Almaz Yohannis. Optimized cnn ensemble with class-balanced mri data augmentation for accurate multi-class dementia diagnosis. *Advances in Alzheimer's Disease*, 14(3):53–76, 2025. doi: 10.4236/aad.2025.143004.

Houmem Slimi, Imen Cherif, Sabeur Abid, and Mounir Sayadi. Biologically inspired hybrid model for alzheimer's disease classification using structural mri in the adni dataset. *Frontiers in Artificial Intelligence*, 8:1590599, 2025. doi: 10.3389/frai.2025.1590599.

Michael W Weiner, Paul S Aisen, Clifford R Jack, William J Jagust, John Q Trojanowski, Leslie Shaw, et al. The alzheimer's disease neuroimaging initiative: progress report and future plans. *Alzheimer's & Dementia*, 6(3):202–211, 2010.