Smart Media Journal / Vol.14, No.10 / ISSN:2287-1322
https://dx.doi.org/10.30693/SMJ.2025.14.10.90

A Stacked CNN Approach for Accurate Classification of AD Severity from T1-Weighted MRI Slices

Vyshnavi Ramineni, Goo-Rak Kwon

Abstract

This paper proposes a stacked convolutional neural network (CNN) architecture for classifying Alzheimer's disease (AD) severity using T1-weighted MRI scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI). 3D MRI volumes are preprocessed through intensity correction, spatial normalization to a standard template, and skull stripping, then sliced into informative 2D images based on anatomical landmarks and entropy measures. Data augmentation techniques further enhance the dataset while reducing overfitting. The stacked CNN, fine-tuned via transfer learning, extracts both local details and global structural features crucial for distinguishing healthy controls, mild cognitive impairment, and AD. A hybrid loss function combining cross-entropy and triplet loss improves the model's discriminative power by clustering similar features and maximizing inter-class separation. Experimental results indicate high classification accuracy and robust performance, highlighting the potential of our approach for early AD diagnosis and severity assessment.

Keywords: Alzheimer's disease | Early Mild Cognitive impairment | Late Mild Cognitive Impairment | CNN

I. INTRODUCTION

Alzheimer's disease (AD) remains one of the most challenging neurodegenerative disorders, with its insidious onset and progressive decline in cognitive and functional abilities.[1], [2] Early and accurate detection is critical for timely intervention and can potentially slow disease progression, thereby improving patients' quality of life and reducing the long-term economic and social burdens associated with dementia. In recent years, advanced neuroimaging techniques, particularly structural magnetic resonance imaging (sMRI)—have become indispensable tools for investigating the pathological changes associated with AD. T1-weighted[3] MRI, known for its high spatial resolution and detailed anatomical contrast, provides crucial information on brain atrophy and other structural alterations that occur during the disease progression.

The Alzheimer's Disease Neuroimaging Initiative (ADNI) has been instrumental in collecting large-scale, high-quality T1-weighted MRI datasets that facilitate research into the early diagnosis of AD [4]. However, the inherent three-dimensional nature of these images poses computational challenges, especially when deep learning models are applied directly to volumetric data. To mitigate these challenges, many studies have adopted a

* This study was supported by research funds from Chosun University, 2025.

Manuscript : 2025.06.02
Revised : 2025.07.30
Confirmation of Publication: 2025.07.31
Corresponding Author : Goo-Rak Kwon,
e-mail : grkwon@chosun.ac.kr

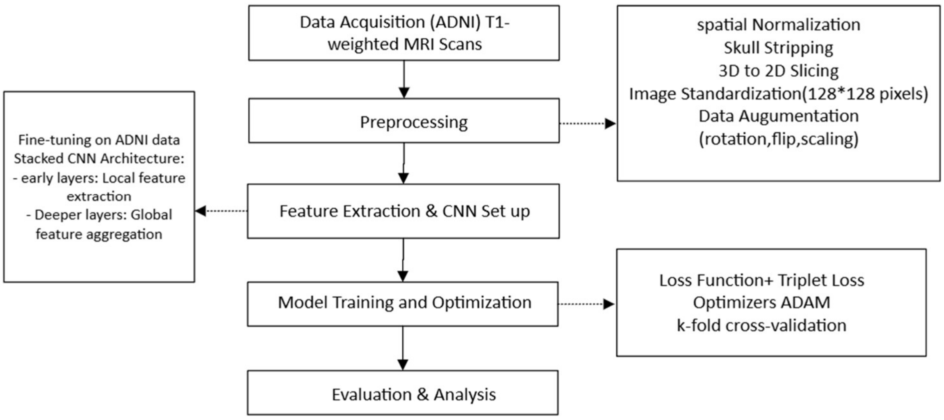

Figure 1 Flow of the Architecture

preprocessing pipeline that involves transforming the 3D MRI volumes into 2D slices. This approach not only reduces computational complexity but also allows the use of well-established two-dimensional convolutional neural network (CNN) architecture without a significant loss of diagnostic information [5].

Our work builds on recent advances in CNN-based Alzheimer's disease classification by integrating a stacked CNN architecture. Stacked methods allow for the sequential learning of features across multiple layers, capturing both low-level local patterns such as subtle changes in gray matter intensity and high-level global structural information such as overall brain atrophy. Such a hierarchical approach has been shown to outperform conventional CNN models that rely solely on a monolithic structure, as stacking facilitates a more nuanced understanding of complex brain patterns associated with different stages of AD [4], [6].

In our proposed methodology, T1-weighted MRI images from the ADNI database are first preprocessed to standardize the imaging data and correct for any intensity non-uniformities. The 3D MRI volumes are then spatially normalized to a standard brain template, and non-brain tissues are removed via skull-stripping techniques. Next,[7], [8]the preprocessed 3D volumes are sliced into 2D images. This step is crucial because it enables us to leverage efficient 2D CNN architecture while still capturing the critical anatomical information contained within the 3D data. By carefully selecting the most informative slices— often guided by entropy or anatomical landmarks, the approach ensures that essential features for disease classification are retained [9].

The core of our analysis employs a stacked CNN architecture, which is designed to address two main challenges in AD diagnosis: the subtlety of early pathological changes and the need for efficient computation. The stacked CNN approach involves training multiple independent convolutional blocks sequentially, where each block is responsible for learning features at a progressively higher level of abstraction. Early blocks capture fine-grained, local details, while deeper blocks aggregate this information into more abstract representations that are indicative of global brain changes. This multi–level feature extraction mechanism enhances the network's ability to discriminate between healthy controls, mild cognitive impairment (MCI), and full–fledged AD[5].

Moreover, by employing stacking techniques, our network mitigates the risk of overfitting—a common challenge when working with limited datasets—by distributing the learning process across multiple layers. Each stacked module refines the features extracted by its predecessor, leading to a robust representation that is both noise–tolerant and highly discriminative. Such architectures have demonstrated impressive performance in previous studies, where a combination of local and global contextual features was key to achieving high classification accuracy[4], [6].

Our study extends these approaches by specifically tailoring the preprocessing and feature extraction pipeline to T1–weight MRI data from ADNI. By slicing the 3D images into 2D planes and then feeding these into a stacked CNN framework, we capitalize on the rich information present in high–resolution MR images while maintaining computational efficiency. The subsequent layers of the CNN further process these features to enable accurate classification of Alzheimer's disease severity. This strategy not only improves diagnostic performance but also offers a scalable solution suitable for clinical applications where rapid and reliable analysis is essential.

II. METHODOLOGY

2.1 Experiment Design and Data Collection

Our study follows a structured experimental design aimed at accurately classifying Alzheimer's disease (AD) severity using T1–weighted structural MRI data. The primary objective is to develop a robust deep learning model that leverages stacked convolutional neural network (CNN) architectures to extract both local and global features from preprocessed 2D MRI slices. The experimental design was structured to ensure reproducibility and to optimize performance on data obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI).

Data Source:

T1–weighted MRI scans were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (https://adni.loni.usc.edu). ADNI is a widely recognized repository that collects imaging, genetic, clinical, and biomarker data, serving as an essential resource for AD research.

Participant Selection:

A subset of subjects was selected from the ADNI database to ensure a balanced representation across diagnostic categories, including healthy controls (HC), mild cognitive impairment (MCI), and full–blown Alzheimer's disease (AD) as shown in Table 1. Inclusion criteria required subjects to be over 60 years of age, to have a confirmed diagnosis based on standard clinical assessments, and to have high–quality, intensity–corrected T1–weighted images.

Table 1. Subject taken from the ADNI

Group	No.of Subjects	Age Range
AD	60	72.65 ± 8.6
LMCI	60	76.80 ± 6.9
EMCI	60	74.83 ± 6.1
HC	60	75.83 ± 5.7

2.2 Preprocessing Procedures

Given the inherent complexity of processing full 3D MRI volumes, we implemented a series of standardized preprocessing steps to optimize the dataset for 2D CNN analysis.

Intensity Correction and Registration: Each 3D MRI volume was subjected to intensity inhomogeneity corrections using methods such as N3 bias field correction. Subsequently, the volumes were spatially normalized by registering them to a standard brain template using neuroimaging software packages such as FSL and SPM. Skull stripping was then performed to remove non–brain tissues.

3D–to–2D Slicing:To reduce computational complexity, each 3D volume was sliced into a series of 2D images. The most informative slices were selected based on anatomical landmarks the hippocampus and entropy measurements, ensuring that both local details and global structural changes were preserved.

Image Standardization and Augmentation: The resulting 2D slices were resized to a uniform resolution 128 × 128 pixels to ensure consistency across the dataset. To further increase the diversity of the training data and to mitigate overfitting, data augmentation techniques—including rotation, flipping, scaling, and shifting—were applied.

2.3 Feature Extraction

Stacked CNN Design: This core methodology employs stacked CNN architecture. This involves a series of sequential convolutional blocks, where each block extracts features at progressively higher levels of abstraction. Early convolutional layers capture fine, local details such as subtle tissue intensity variations, whereas deeper layers aggregate these details into high–level representations that reflect global brain atrophy.

Transfer Learning and Fine–Tuning: To leverage existing knowledge and expedite model convergence, pre–trained CNN weights from models VGG, ResNet and MobileNet were used as initialization. These models were subsequently fine–tuned on our ADNI–derived dataset, ensuring that the learned features were well adapted to the specific characteristics of T1–weighted MRI data.

Loss Functions: To enhance the discriminative power of the network, we employed a composite loss function that combines cross–entropy loss with additional regularization triplet loss. This dual–loss approach helps cluster features from the same diagnostic category together while maximizing the separation between different categories.

2.4 Architecture of the Experiment.

Our architecture begins by ingesting standardized 2D slices extracted from 3D T1-weighted MRI volumes. These slices serve as input to a series of initial convolutional layers designed to capture fine-grained local features. The network leverages transfer learning by initializing with pre-trained weights from models, which are then fine-tuned on our specific ADNI dataset. This initialization ensures that the model benefits from rich, generic feature representations that are later adapted to the nuances of structural MRI data, thus reducing the need for extensive training from scratch.

At the core of the design, a stacked arrangement of convolutional blocks is employed to build a hierarchical representation of the input. Each block consists of convolutional layers followed by ReLU activations and pooling layers, enabling the network to progressively capture both low-level texture details and high-level global structural patterns. This multi-tiered approach facilitates effective feature abstraction, ensuring that the network learns discriminative representations necessary for differentiating among healthy controls, MCI, and Alzheimer's disease. Additionally, a hybrid loss function combining cross-entropy with triplet loss is used to optimize the feature space by clustering similar diagnostic categories while maximizing inter-class separation.

Finally, the high-level features are consolidated through global average pooling followed by one or more fully connected layers. This strategy not only reduces the parameter count and overfitting risk but also efficiently transforms the rich feature maps into a final decision vector. An optional lightweight attention module can be integrated at this stage to dynamically emphasize disease-relevant channels. The output is then passed through a SoftMax layer to yield probability distributions across diagnostic classes. This end-to-end architecture—combining transfer learning, stacked CNNs, and an optimized loss framework—ensures robust and accurate classification of Alzheimer's disease severity.

III. EXPERIMENT RESULT & DISCUSSION

3.1 Performance and evaluation parameters

Each classifier produces predictions in the form of a confusion matrix, which is divided into true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN), as mathematically detailed in Table 2. TP and TN indicate the number of correctly identified controls, while FP and FN represent the instances that were incorrectly classified.

Table 2. Multiclass confusion matrix

	Prediction classification				
	classes	AD	LMCI	EMCI	HC
Actual classification	AD	**TP**	F_{AL}	F_{AE}	F_{AH}
	LMCI	F_{LA}	**TP**	F_{LE}	F_{LH}
	EMCI	F_{EA}	F_{EL}	**TP**	F_{EH}
	HC	F_{HA}	F_{HL}	F_{HE}	**TP**

Although accuracy is commonly used to assess multi-class classifier performance by computing the overall ratio of correct predictions, it can be misleading in cases of unstable or imbalanced class distributions.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

To address this limitation, additional metrics such as precision, recall, and F1-score are incorporated. Precision measures the proportion of correct positive predictions, Recall (or Sensitivity) reflects the ability to identify actual positives, and the F1-score, which is the harmonic mean of precision and recall, provides a balanced evaluation of the classifier's performance.

3.2 Classification results and Discussion

Our results indicate that our stacked CNN architecture achieves robust performance in classifying T1-weighted MRI slices into distinct diagnostic categories. Evaluation metrics such as accuracy, sensitivity, specificity, and F1-score consistently demonstrate superior performance compared to traditional CNN approaches and standard transfer learning baselines. Table3 shows the Classification results of Unprocessed dataset and Table 4 shows the results of the processed dataset.

Table 3. Classification results with unprocessed T1-weighted images

Classifier's	ACC%	PRE%	RECA%
MobileNetV1	67.5	65.65	75.63
MobileNetV2	54.0	56.25	63.67
ShuffleNetV1	84.0	77.62	89.87
ShuffleNetV2	69.0	72.3	78.2
GhostNet	70.0	65.2	74.9
EfficientNet	71.0	66.9	76.37
Proposed Method	**87.9**	**85.8**	**92.6**

Table 4. Classification results with processed T1-weighted images

Classifier's	ACC%	PRE%	RECA%
MobileNetV1	70.6	63.12	77.2
MobileNetV2	71.8	66.25	76.91
ShuffleNetV1	69.4	75.5	73.8
ShuffleNetV2	70.9	63.2	79.5
GhostNet	72.0	61.56	77.25
EfficientNet	76.3	69.32	82.82
Proposed Method	**97.75**	**90.5**	**98.7**

NOTE: ACC: Accuracy; PRE : Precision; RECA: Recall

The observed improvement in Precision and Recall can be attributed to the hybrid loss function combining cross-entropy and triplet loss. This combination enables the model to cluster intra-class features more effectively while maximizing separation between classes. Additionally, the stacked CNN architecture captures both low-level details and high-level structural features, contributing to superior classification performance.

Visual assessments of activation maps confirm that the network effectively focuses on critical brain regions, including the hippocampus, supporting its ability to capture both fine-grained local details and broader structural patterns. The combination of pre-trained weights with a hybrid loss function has facilitated improved feature clustering, ensuring that intra-class variations are minimized while inter-class differences are emphasized.

Despite these promising outcomes, our approach also presents several limitations that warrant further investigation. Converting 3D MRI volumes into 2D slices, although computationally efficient, may lead to a loss of three-dimensional spatial context that could be crucial for capturing complex brain anatomy. Additionally, while our findings on the ADNI dataset are encouraging, validating the model on external datasets is essential to ensure its generalizability across different clinical settings. Future research should explore the integration of 3D contextual information and further refine attention mechanisms to enhance the model's diagnostic accuracy and clinical utility. Although we focused on lightweight CNNs, transformer-based and attention-augmented models such as those in [9] and [10] have shown promise. We plan to include these in future evaluations for comprehensive benchmarking.

3.3 Ablation Study

To validate the contribution of each component in the proposed model, we conducted an ablation study comparing different variants as shown in Table 5.

Table 5. Ablation study results

Model Variant	ACC%	PRE%	RECA%
Base CNN (no stacking)	88.1	82.5	85.3
CNN + Hybrid Loss	91.3	86.2	89.5
Stacked CNN only	94.5	88.7	93.1
Full proposed (stacked + hybrid)	97.75	90.5	98.7

These results confirm that both the stacked CNN structure and the hybrid loss significantly improve performance.

IV. CONCLUSION

Our study demonstrates that leveraging a stacked convolutional neural network (CNN) architecture on preprocessed T1-weighted MRI slices from the ADNI database significantly enhances the classification accuracy of Alzheimer's disease severity. By converting 3D MRI volumes into carefully selected 2D slices and employing transfer learning for fine-tuning, our model captures both local and global structural features that are critical for differentiating between healthy controls, mild cognitive impairment, and Alzheimer's disease. The integration of a hybrid loss function further strengthens the model's ability to cluster similar features and separate dissimilar ones, resulting in robust performance even with limited training data.

While our approach shows promise in advancing the early diagnosis of Alzheimer's disease, challenges remain in preserving the full 3D contextual information inherent in volumetric data. Future work should focus on incorporating 3D contextual cues and refining attention mechanisms to further improve diagnostic precision and generalizability across diverse clinical datasets. Overall, our findings underscore the potential of stacked CNN architectures as a scalable, efficient, and effective solution for neuroimaging-based disease classification.

ACKNOWLEDGMENT

Data collection and sharing were funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) under NIH Grant U01 AG024904, along with support from DOD

ADNI (Department of Defense, award number W81XWH-12-2-0012). For further details regarding ADNI funding, http://adni.loni.usc.edu/about/funding/

REFERENCES

[1] M. R. Arbabshirani et al., "Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration," *Npj Digit. Med.*, vol. 1, no. 1, p. 9, Apr. 2018.

[2] A. Farooq, S. Anwar, M. Awais, and S. Rehman, "A deep CNN based multi-class classification of Alzheimer's disease using MRI," *2017 IEEE Int. Conf. Imaging Syst. Tech. IST*, pp. 1–6, Oct. 2017.

[3] D. Cheng and M. Liu, "CNNs based multi-modality classification for AD diagnosis," *2017 10th Int. Congr. Image Signal Process. Biomed. Eng. Inform. CISP-BMEI*, pp. 1–5, Oct. 2017.

[4] E. Jabason, M. O. Ahmad, and M. N. S. Swamy, "A Lightweight Deep Convolutional Neural Network Extracting Local and Global Contextual Features for the Classification of Alzheimer's Disease Using Structural MRI," *IEEE J. Biomed. Health Inform.*, vol. 29, no. 3, pp. 2061–2073, Mar. 2025.

[5] V. Patil, M. Madgi, and A. Kiran, "Early prediction of Alzheimer's disease using convolutional neural network: a review," *Egypt. J. Neurol. Psychiatry Neurosurg.*, vol. 58, no. 1, p. 130, Nov. 2022.

[6] W.-C. Lien et al., "Convolutional Neural Networks to Classify Alzheimer's Disease Severity Based on SPECT Images: A Comparative Study," *J. Clin. Med.*, vol. 12, no. 6, Art. no. 6, Jan. 2023.

[7] Z. Yuan et al., "Diagnosis of Alzheimer's disease using transfer learning with multi-modal 3D Inception-v4," *Quant. Imaging Med. Surg.*, vol. 15, no. 2, pp. 1455467–1451467, Feb. 2025.

[8] S. Alp et al., "Joint transformer architecture in brain 3D MRI classification: its application in Alzheimer's disease classification," *Sci. Rep.*, vol. 14, no. 1, p. 8996, Apr. 2024.

[9] Q. Zhang, Y. Long, H. Cai, and Y.-W. Chen, "Lightweight neural network for Alzheimer's disease classification using multi-slice sMRI," *Magn. Reson. Imaging*, vol. 107, pp. 164–170, Apr. 2024.

Authors

Vyshnavi Ramineni

She has completed her bachelor's degree at the department of Electronics and Communication Engineering in NEC, India, in 2019. Currently, she is pursuing her Master studies at the department of Information and Communication Engineering in Chosun University, Gwangju, Republic of Korea. Her research interests include Digital Signal processing, Machine learning using Medical Imaging.

Goo-Rak Kwon

He received a Ph.D. at the Department of Mechatronic Engineering, Korea University, in 2007. He has been a professor with Chosun University, since 2017. His research interest includes medical image analysis, A/V signal processing, video communication, and applications. He is a senior member of the IEEE.