Journal of Innovations in Computer Science and Engineering
Shahid Beheshti Unversity

June 2024, Special Issue on AI 4 All- 1

Aβ42/40 ratio prediction using MRI images features for Alzheimer's Early Detection

Atefe Aghaei✉, Code Orcid: 0000-0001-5257-9147
Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran, a_aghaei@sbu.ac.ir
Mohsen Ebrahimi Moghaddam, Code Orcid: 0000-0002-7391-508X
Faculty of Computer Science and Engineering, Shahid Beheshti University, Tehran, Iran, m_moghadam@sbu.ac.ir

Abstract— Alzheimer's disease (AD) is a progressive neurodegenerative disorder characterized by cognitive decline and the accumulation of amyloid-beta plaques. Early detection is crucial for timely intervention, and the Aβ42/Aβ40 ratio is a key biomarker for identifying amyloid deposition. In this study, we propose a method to predict the Aβ42/Aβ40 ratio using the extracted features from MRI images using 3D Convolutional Neural Network (3D CNN). Moreover, Random Forest Regression is employed to obtain the relationship between MRI features and the Aβ42/Aβ40 ratio. Our results demonstrate a strong correlation (r = 0.72) between the predicted and actual Aβ42/Aβ40 ratios, effectively predicting amyloid accumulation. This result also makes the proposed feature extraction model more reliable. By leveraging MRI and molecular biomarkers such as the Aβ42/Aβ40 ratio, the proposed method provides valuable insights into disease progression and early diagnosis.

Keywords—3DCNN, Alzheimer's Disease, Aβ42/Aβ40 ratio, MRI, Random Forest Regression

I. Introduction

Alzheimer's disease (AD) is a neurodegenerative disorder characterized by the accumulation of amyloid-beta plaques and tau tangles in the brain, leading to progressive cognitive decline [1–3]. Early detection of AD is critical for timely intervention and treatment, as amyloid-beta (Aβ) deposition typically precedes the onset of clinical symptoms. Among various biomarkers, the Aβ42/Aβ40 ratio has emerged as a reliable indicator of amyloid plaque deposition. This ratio is particularly valuable in the preclinical stages of AD, where amyloid accumulation occurs without apparent cognitive deficits [4–6]. Magnetic Resonance Imaging (MRI), on the other hand, captures structural changes in the brain, such as hippocampal atrophy and cortical thinning, which are hallmark features of AD progression [7–9]. MRI-based features of gray matter volume and cortical thickness have been widely used to assess neurodegeneration in AD, with particular emphasis on the medial temporal lobe, where the hippocampus resides, as a critical region affected in the early stages of the disease. Combining MRI features with the Aβ42/Aβ40 ratio in a multimodal framework has been shown to improve the predictive power for Alzheimer's diagnosis. Combining these biomarkers Aβ42/Aβ40 ratio and MRI-derived structural changes offers a powerful approach for early diagnosis and prediction of AD progression [10,11].

The Aβ42/Aβ40 ratio is widely recognized as a sensitive biomarker for predicting amyloid plaque deposition, an essential pathology in Alzheimer's disease (AD). Multiple studies have demonstrated its superiority over Aβ42 alone, making it a cornerstone in AD diagnosis. For example, the study by [12] showed that plasma Aβ42/Aβ40 ratios are significantly associated with amyloid positivity on PET imaging, making the ratio a reliable marker for detecting amyloid burden in preclinical AD and mild cognitive impairment (MCI) stages. Similarly, [13] found that using the Aβ42/Aβ40 ratio in cerebrospinal fluid (CSF) increased the accuracy of predicting amyloid pathology. Their results indicated a stronger correlation between amyloid burden and hippocampal atrophy when using the ratio rather than Aβ42 alone.

Moreover, structural changes in the brain, such as hippocampal atrophy, cortical thinning, and ventricular enlargement, are well-established MRI features used in diagnosis Alzheimer's disease. Based on this, many papers have used features obtained from 3D and 2D T1-weighted MRI images with deep learning methods [14–16]. To provide a **non-invasive** alternative to costly and invasive methods like **PET scans** and **cerebrospinal fluid (CSF) tests**, which are traditionally used to detect Aβ buildup in the brain, [17] predict **amyloid-beta (Aβ) positivity** (Aβ+) in Alzheimer's disease (AD) patients using **3D convolutional neural networks (CNNs)** and **transfer learning** based on **3D T1-weighted brain MRI** scans. Other studies, such as [18] focuses on predicting amyloid PET positivity using **multimodal MRI** (T1, T2, and FLAIR scans) combined with a CNN architecture. They demonstrate that **MRI alone** can be a viable alternative to amyloid PET for detecting amyloid buildup. This study highlights the potential of MRI and CNNs in non-invasive AD diagnosis. Another study [19] investigated the relationship between **white matter abnormalities** and **cortical amyloid deposition** in older adults without dementia, considering racial differences. Using **MRI** (to measure white matter hyperintensities and microstructural integrity) and **PET imaging** (to measure amyloid levels),

Submit Date: 2024-10-10
Accept Date: 2025-04-21
✉ Corresponding author

the researchers found a modest positive association between white matter hyperintensity volume and elevated amyloid. The association was stronger in **Black participants** than in White participants, though the race interaction was not statistically significant. No significant association was found between white matter microstructure and amyloid.

While **amyloid PET** imaging is a direct method to assess amyloid burden, it is expensive and invasive, limiting accessibility. Based on the literature, since **Aβ42/Aβ40 ratio** is valuable as a **non-invasive biomarker**, providing earlier detection of amyloid buildup before it reaches the levels seen on PET scans, in this study, we aim to predict **Aβ42/Aβ40 ratio 1)** to provide a more accurate reflection of

TABLE 1: DEMOGRAPHIC INFORMATION OF USAGE SUBJECTS

Group	# Subjects	Age (years)	% Gender (M/F)
CN	232	76.5	46/54
MCI	178	76.45	54/46
AD	136	75.2	53/47

the utility of the extracted features from our proposed method in clinical practice and interpretability, and 2) to show the relationship between these biomarkers to use in Alzheimer's early detection. It also has broader applicability in clinical settings due to its more accessible and less costly measurement. We propose using a 3D convolutional neural network (3D-CNN) to extract features from MRI scans of cognitively normal (CN), MCI, and AD subjects. These features are then used in a regression-based model, Random Forest Regression, to predict the Aβ42/Aβ40 ratio for each subject. By linking the MRI features to the amyloid-beta ratio, this approach seeks to offer an integrated diagnostic tool that leverages both structural imaging and molecular biomarkers for early Alzheimer's detection

The results of this study indicate a strong correlation of 0.72 between the predicted Aβ42/Aβ40 ratio from MRI-based features and the actual measured ratio. This high correlation shows that the machine learning model, using 3D CNNs, can accurately estimate amyloid-beta levels non-invasively, making it a promising tool for predicting amyloid burden in Alzheimer's disease. These findings demonstrate the potential of MRI-based predictions in early Alzheimer's detection, reducing reliance on invasive procedures like PET scans or CSF analysis.

II. Materials and Methods

III. Data Description

In this study, T1-weighted MRI data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) were used to develop and test the proposed method. ADNI is a longitudinal study that tracks Alzheimer's progression using clinical, imaging, genetic, and biochemical data. The MRI data, collected across four ADNI phases, were acquired with a 1.5 Tesla scanner using a T1-weighted MPRAGE sequence. The dataset includes subjects classified as cognitively normal (CN), mild cognitive impairment (MCI), and Alzheimer's disease (AD), and underwent preprocessing such as B1 correction and brain extraction using FreeSurfer. Information about the utilized data is presented in Table 1.

IV. Data Pre-processing

To optimize computational efficiency and enhance accuracy, a subset of 32 MRI image slices containing the entire brain are selected from the middle. This selection is made due to the lack of meaningful information and the absence of brain tissue in other slices. The chosen images are subsequently resized to dimensions of $64 \times 64 \times 32$. Also, data augmentation is subsequently employed to augment the dataset. Flipping and rotation have been selected as augmentation techniques for the brain image, as each pixel contains valuable information. Thus, the images undergo both vertical and horizontal flipping. To rotate the

Table 2: The details of the proposed 3DCNN model

Layer	Output Shape	Param #
Conv3D	62, 62, 30, 32	896
LeakyReLU	62, 62, 30, 32	0
MaxPooling3D	31,31, 15, 32	0
BatchNormalization	31,31, 15, 32	128
Dropout	31,31, 15, 32	0
Conv3D	29, 29, 13, 32	27680
LeakyReLU	29, 29, 13, 32	0
MaxPooling3D	14, 14, 6, 32	0
BatchNormalization	14, 14, 6, 32	128
Dropout	14, 14, 6, 32	0
Conv3D	12, 12, 4, 64	55360
LeakyReLU	12, 12, 4, 64	0
MaxPooling3D	6, 6, 2, 64	0
BatchNormalization	6, 6, 2, 64	256
Dropout	6, 6, 2, 64	0
Flatten	4608	0
Dense	10	46090
LeakyReLU	10	0
BatchNormalization	10	40
Dropout	10	0
Dense	3	33

images, a vector of rotation angles (-10, -5, 5, and 10) is defined. Each time, one of these angles is randomly chosen. At the end, all of the data are normalized to (0,1).

The proposed 3D CNN is trained using both original and created images. Test sets are used to assess the model's ability to generalize.

V. Feature Extraction Model

In this paper, a 3D CNN model is used for feature extraction. The proposed 3D CNN model consists of three blocks, in each block, 3D convolutions apply to extract spatial features from volumetric 3D medical scans. Each Conv3D layer is followed by LeakyReLU for non-linearity, MaxPooling3D to reduce spatial dimensions and control overfitting, and Batch-Normalization for stabilizing training. Dropout layers further prevent overfitting. After extracting features, the model flattens the data and passes it through two Dense layers for classification, with the final layer outputting three units, corresponding to three class labels.

VI. Aβ42/Aβ40 ratio prediction

Aβ42/Aβ40 ratio prediction using Random Forest Regression involves training the model on **MRI-derived features** to predict the ratio of amyloid-beta peptides. **Random Forest** is an ensemble learning method that combines multiple decision trees, where each tree is trained on random subsets of the data. By averaging the predictions of these trees, the model captures complex relationships

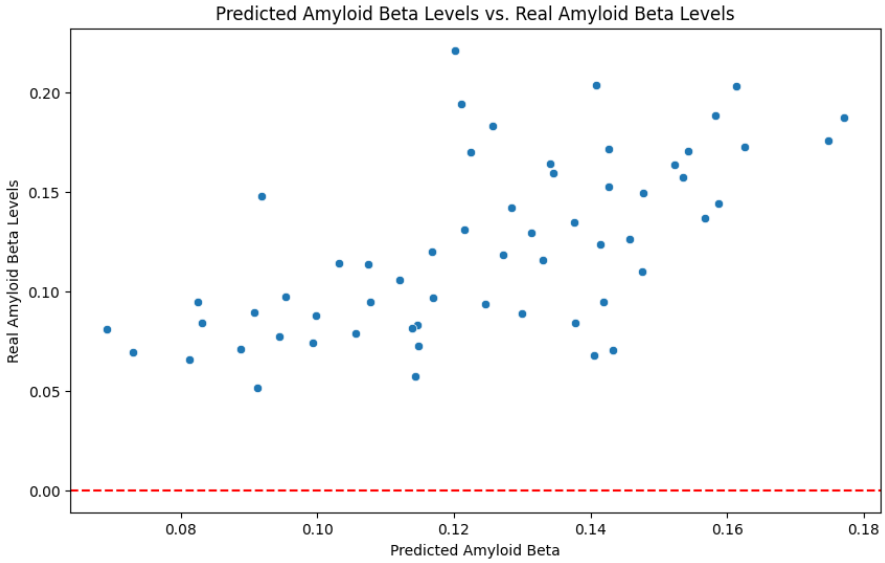

Figure 1: The scatter plot of Real and Predicted Aβ42/Aβ40 ratio

between MRI features and the amyloid ratio. This method benefits from **non-linear data modeling** and reduces overfitting, making it a powerful approach for estimating amyloid burden non-invasively.

Since we do not have **Aβ42/Aβ40 ratio for all of the subjects, part of the datast is used in this section. We use 405 images and do not separate CN, MCI, and AD subjects.**

VII. results

VIII. 3D Feature Extraction

Based on Table1, the dataset includes 546 subjects and about 2000 images, however, after data augmentation data are increased to 4074 images. The 3D CNN model begins with a Conv3D layer that applies 32 filters with a kernel size of 3×3×3 to the input, followed by LeakyReLU activation for non-linearity. This output is passed through MaxPooling3D to reduce the spatial dimensions, and BatchNormalization to stabilize the learning process. A Dropout layer is added to prevent overfitting. The next Conv3D layer, with 32 filters, follows the same pattern of activation, pooling, and regularization. Another Conv3D layer with 64 filters further extracts spatial features, before applying the same operations.

Once the convolutions are completed, the output is flattened into a 1D vector. This flattened vector is fed into a Dense layer with ten units, again followed by LeakyReLU, Batch-Normalization, and Dropout to maintain regularization. The final Dense layer consists of 3 units, corresponding to the number of classes for the output. The model is optimized using the Adam optimizer, and Cross Entropy loss is applied for classification tasks.

The model is trained using the Adam optimizer (learning rate: 0.001) with Cross Entropy loss, and it converges after 100 epochs. This architecture is implemented using the Keras library in Python, running on a Tesla T4 GTX GPU.

We applied normalization techniques that scaled the MRI intensity values to a standardized range of 0 to 1. This step is essential to enhance the model's performance and ensure that it learns from data that is uniformly represented. To facilitate the training of our 3DCNN model, all images were resized to dimensions $64 \times 64 \times 32$,which ensures that the input data maintains consistency in shape. We split our data including 546 subjects into Train/Validation and test data. Twenty percent of the whole data are selected as test data and twenty percent of remaining data are randomly selected as validation. The results obtained from our proposed method for feature extraction on test data are demonstrated in Table 3

Table 3: The results obtained from our proposed method for feature extraction on test data

Class	Precision	Recall	F1-Score	Average Accuracy
CN	0.92	0.81	0.86	0.83
MCI	0.82	0.75	0.78	
AD	0.70	0.89	0.78	

IX. Aβ42/Aβ40 ratio prediction

We calculate Pearson Correlation Coefficient between actual **Aβ42/Aβ40 ratio and the Aβ42/Aβ40 ration obtained from MRI images features using Random Forest model.** The correlation between the Aβ42/Aβ40 ratio and our model's predicted values yielded a significant value of 0.72, demonstrating a strong positive relationship between the extracted MRI features and this key Alzheimer's biomarker. The Aβ42/Aβ40 ratio is critical in early Alzheimer's disease detection, as a lower ratio often signifies amyloid plaque accumulation in the brain. This strong correlation suggests that our model is capturing relevant brain features that closely predict this biomarker, which is indicative of early pathological changes associated with Alzheimer's disease.
Additionally, the scatter plot of predicted versus real values is in Figure 1. Figure 1 shows a close alignment, further supporting the model's ability to predict amyloid-beta ratios with high precision. These results demonstrate that the proposed model can be a valuable tool for early detection by leveraging MRI data to assess disease progression through clinically relevant biomarkers such as the Aβ42/Aβ40 ratio.

X. Conclusion and Future Works

This study introduces a method that combines MRI-based features and machine learning to predict the Aβ42/Aβ40 ratio, a key biomarker for Alzheimer's disease. Using 3D CNN for feature extraction and Random Forest Regression for prediction, we achieved a high correlation (r = 0.72) between the predicted and real Aβ42/Aβ40 ratios. This method provides a non-invasive alternative to costly and invasive procedures like PET scans, offering promise for early Alzheimer's detection. Future research will focus on expanding the dataset and incorporating other imaging modalities, such as fMRI and DTI, to enhance model accuracy. Additionally, we plan to explore multi-class classification for predicting different stages of Alzheimer's progression and integrate more biomarkers like tau proteins for a more comprehensive analysis.

1.1.1.1.1 Acknowledgment

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

1.1.1.1.2 References

1. Jack Jr, C. R. *et al.* Update on hypothetical model of Alzheimer□s disease biomarkers. *Lancet Neurol.* **12,** 207 (2013).
2. Selkoe, D. J. & Hardy, J. The amyloid hypothesis of Alzheimer's disease at 25 years. *EMBO Mol. Med.* **8,** (2016).
3. Zhang, Y., Chen, H., Li, R., Sterling, K. & Song, W. Amyloid β-based therapy for Alzheimer's disease: challenges, successes and future. *Signal Transduction and Targeted Therapy* **8,** (2023).
4. Hampel, H. *et al.* The amyloid-β pathway in Alzheimer's disease: a plain language summary. *Neurodegenerative Disease Management* **13,** (2023).
5. Hampel, H. *et al.* The Amyloid-β Pathway in Alzheimer's Disease. *Molecular Psychiatry* **26,** (2021).
6. Iliyasu, M. O., Musa, S. A., Oladele, S. B. & Iliya, A. I. Amyloid-beta aggregation implicates multiple pathways in Alzheimer's disease: Understanding the mechanisms. *Frontiers in Neuroscience* **17,** (2023).
7. Georgakas, J. E., Howe, M. D., Thompson, L. I., Riera, N. M. & Riddle, M. C. Biomarkers of Alzheimer's disease: Past, present and future clinical use. *Biomarkers in Neuropsychiatry* **8,** (2023).
8. Kälin, A. M. *et al.* Subcortical shape changes, hippocampal atrophy and cortical thinning in future Alzheimer's disease patients. *Front. Aging Neurosci.* **9,** (2017).
9. Rao, Y. L. *et al.* Hippocampus and its involvement in Alzheimer's disease: a review. *3 Biotech* **12,** (2022).
10. Tosun, D. *et al.* Detection of β-amyloid positivity in Alzheimer's Disease Neuroimaging Initiative participants with demographics, cognition, MRI and plasma biomarkers. *Brain Commun.* **3,** (2021).
11. Jack, C. R. *et al.* Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to Alzheimer's disease. *Brain* **133,** (2010).
12. Doecke, J. D. *et al.* Total Aβ42/Aβ40 ratio in plasma predicts amyloid-PET status, independent of clinical AD diagnosis. *Neurology* **94,** (2020).
13. Baiardi, S. *et al.* Antemortem CSF Aβ42/Aβ40 ratio predicts Alzheimer's disease pathology better than Aβ42 in rapidly progressive dementias. *Ann. Clin. Transl. Neurol.* **6,** (2019).
14. Aghaei, A. & Moghaddam, M. E. Smart ROI Detection for Alzheimer's Disease prediction using explainable AI. *arXiv Prepr. arXiv2303.10401* (2023).
15. Aghaei, A., Ebrahimi Moghaddam, M. & Initiative, A. D. N. Brain age gap estimation using attention-based ResNet method for Alzheimer□s disease detection. *Brain Informatics* **11,** 16 (2024).

16. Aghaei, A., Ebrahimi Moghaddam, M. & Malek, H. Interpretable ensemble deep learning model for early detection of Alzheimer's disease using local interpretable model-agnostic explanations. *Int. J. Imaging Syst. Technol.* **32,** 1889–1902 (2022).
17. Chattopadhyay, T., Ozarkar, S. S., Buwa, K., Thomopoulos, S. I. & Thompson, P. M. Predicting Brain Amyloid Positivity from T1 weighted brain MRI and MRI-derived Gray Matter, White Matter and CSF maps using Transfer Learning on 3D CNNs*. *bioRxiv* (2023).
18. Sungkyu Bang Hanyang University; M.S. Byun, M. D. P. D. Y. P. D. Y. L. M. D. P. J. M. L. P. Deep Learning-based Prediction of Amyloid PET Positivity from Multimodal Magnetic Resonance Imaging of Old-aged Individuals. in *2018 Conference on Machine Intelligence in Medical Imaging* (2018).
19. Walker, K. A. *et al.* Brain white matter structure and amyloid deposition in black and white older adults: The aric-pet study. *J. Am. Heart Assoc.* **10,** (2021).