SCIENTIFIC CULTURE
www.sci-cult.com

SCIENTIFIC CULTURE, Vol. 12, No. 4, (2026), pp. 9214-9222
Open Access. Online & Print

DOI: 10.5281/zenodo.12426905

DEEP LEARNING-BASED MEDICAL IMAGE SEGMENTATION FOR EARLY DETECTION OF NEURODEGENERATIVE DISORDERS

[1*]Dr. V J Chakravarthy , [2]Dr. Parmanand Prabhat ,[3]Dr. Raja Gulfam Shaikh , [4]Dr. Pramod Vishwakarma,[5]Kaushik Paul, [6]Aswathy Mariam Mohan

*[*1]Professor Faculty of Computer Applications Dr.M.G.R.Educational and Research Institute, Maduravoyal, Chennai – 600095 Email ID: chakravarthy.mca@drmgrdu.ac.in*

[2]Assistant Professor Department of Artificial Intelligence and Machine Learning Specialization in Computer Science Engineering BGS Institute of Technology, Adichunchanagiri University, Mandya District, Karnataka Pin code -571448,India ORCID ID:0009-0004-6767-5130 Email ID: parmanandprabhat@bgsit.ac.in

[3]Associate Professor Department of Neurology Specialization in Neurology Superspeciality Hospital, M.G.M. Medical College, Indore 452001 Email ID: gulfamshaikh111@gmail.com

[4]Associate Professor Department of AIT–CSE Specialization in Data Science Chandigarh University, Mohali, India- 140413 Orcid ID 0000-0002-8020-9554 Email ID: pramod.e9758@cumail.in

[5]Assistant Professor Department of Computer Science & Engineering – AI Specialization in Machine Learning Brainware University Orcid Id: 0009-0009-6929-7651 Email ID: kkp.cse@brainwareuniversity.ac.in

[6]Assistant Professor Department of Electrical and Electronics Engineering Specialization in Power system Mangalam College of Engineering, Ettumanoor, Kerala Email ID: achusmm@gmail.com

Received: 16/08/2025
Accepted: 27/01/2026

Corresponding Author: Dr. V J Chakravarthy (chakravarthy.mca@drmgrdu.ac.in)

ABSTRACT

The initial identification of neurodegenerative diseases like the Alzheimer disease is important but the existing approaches tend to be either invasive or subjective. Biomarkers in structural MRI, such as atrophy of the hippocampal and ventricular enlargement, need proper segmentation to be quantified. The paper presents a 3D U Net to automatically segment the hippocampus and lateral ventricles in T1 weighted MRI and assesses whether the volumes obtained can differentiate between cognitively normal (CN), mild cognitive impairment (MCI), and Alzheimer disease (AD). They were the Radiata Brain Structure dataset (3,794 scans). FreeSurfer was used to generate ground truth segmentations. The 3D U Net was trained using Dice and cross entropy loss, data augmentation, and early stopping. The model had a mean Dice coefficient of 0.913 (hippocampus) and 0.894 (ventricles). The volumes of hippocampal became significantly lower in CN (3.84mL), MCI (3.41mL), and AD (2.95mL) (p=0.001). In the case of MCI vs. CN discrimination, the logistic regression provided an AUC of 0.82, sensitivity 79.2 and specificity 74.1. The proposed pipeline provides high levels of segmentation and proves that hippocampal volumes extracted automatically can be used to distinguish between early disease and normal ageing, which is why it may be used as a non invasive screening tool.

Keywords: Alzheimer's disease; deep learning; MRI segmentation; 3D U-Net.

1. Introduction

Neurodegenerative diseases, especially Alzheimer disease (AD), represent an ever-growing health issue of concern to the global population due to the aging of the world population and are expected to increase dramatically in the next decades (Malik et al., 2024). These conditions are characterized by both the clinical and economic burden, which makes it essential to detect them as soon as possible, as it is possible to delay their progression and enhance life quality. However, in contemporary diagnostic processes, invasive methods, including lumbar puncture to get the biomarkers of the cerebrospinal fluid, or subjective radiological methods are frequently used, which is prone to inter rater variability and may slow the diagnosis process (Yousefi et al., 2024; Chudzik et al., 2024). An alternative that is less invasive and very prevalent is structural magnetic resonance imaging (MRI) because the volume of the regional parts of the brain provides evidence of the underlying neurodegenerative processes. Hippocampal atrophy and the subsequent enlargement of the lateral ventricles are among the most robustly proven structural biomarkers, which have a positive correlation with the severity of the disease and cognitive impairment (Castillo et al., 2021). To utilize these biomarkers to measure them, it is necessary to properly segment the hippocampus and the ventricles using the T1 weighted MRI.

The conventional ways of segmentation can be divided into two, namely manual tracing and the traditional automated techniques. Gold standard is manual delineation using trained experts, which is labour intensive, time consuming and has significant inter rater variability, which is not feasible in large scale studies or clinical workflows (Noor et al., 2020). Traditional automated methods, including atlas based methods or deformable model methods, seek to overcome these limitations but usually have difficulty in being robust in cases when used on heterogeneous data that is obtained with different scanners, protocols or populations. Their performance may be severely impaired when there is anatomical variability or imaging artifact, and it cannot be generalised (Shaikh et al., 2025).

The development of deep learning has revolutionized medical image segmentation. The convolutional neural networks, particularly the 3D U Net architecture, have exhibited state of the art results in terms of learning hierarchical features directly using the volumetric data with high accuracy, reproducibility and speed. They have the ability to generalise between multi site datasets given enough diverse examples to train on, providing a scalable solution to quantitative neuroimaging (Castillo et al., 2021). Although this is possible, current research on deep learning has been carried out in binary classification (AD vs. cognitively normal (CN) people) with a relatively low proportion on the early, prodromal phase of mild cognitive impairment (MCI). MCI is a time-sensitive period of therapeutic intervention but its imperceptible structural alterations require sensitive and precise measurement devices. Furthermore, not many works have made use of large, publicly available datasets where train validation test splits are provided, to undertake a systematic assessment of both segmentation accuracy and the clinical utility of the derived volumes to detect disease early on (Shaikh et al., 2025). The research will fill these gaps by seeking to achieve three main objectives:

- To develop and validate a 3D U-Net for automatic segmentation of the hippocampus and lateral ventricles from T1-weighted MRI.
- To assess whether volumes derived from the model can distinguish CN, MCI, and AD groups, thereby demonstrating utility for early detection.
- To provide an open-source implementation for reproducibility and to facilitate future research in the field.

2. Literature Review

The basis of quantitative neuroimaging of neurodegenerative disorders is structural MRI biomarkers. The atrophy of the hippocampal is well known as one of the symptoms of the Alzheimer disease and the volumetric decrease is closely associated with cognitive impairment and the progression of the disease. Lateral ventricular enlargement is an auxiliary biomarker, which indicates a loss of global cerebral volume and adds to diagnostic value (Helaly et al., 2022). The mild impaired cognitive stage is of specific importance to clinical trials and prognosis as it is a transition stage where early intervention can delay or halt the onset of dementia. Detection of individuals in this prodromal phase using credible imaging biomarkers is thus of priority in research and clinical practice (Buttar et al., 2024).

These structures are important to extract meaningful volumetric measures and this requires accurate segmentation of these structures. The conventional approaches can be divided into two. The gold standard is still manual tracing by expert anatomists, but is labour intensive, time consuming and has significant inter rater variability, and is not feasible in large scale studies. Atlas based methods aim to automate the segmentation by deforming a template that has been pre labelled on a specific subject to individual subjects, but they can frequently fail to represent the complete variability of anatomy between different populations, which means that they are less accurate when applied to heterogeneous datasets (Jiang et al., 2023).

With the introduction of deep learning, the situation with medical image segmentation has been transformed fundamentally. The first convolutional neural networks were used on 2D slices, however, the necessity of volumetric context led to the creation of fully 3D networks. The 3D U Net was an extension of the original U Net architecture to 3D, which allowed the network to be trained on spatial features using volumetric data and provided the network with a higher level of coherence across slices (Çiçek et al., 2016). The later innovations like V Net, proposed residual connections and a Dice loss function which is specifically developed to deal with the harsh imbalance of classes occurring in the problem of volumetric segmentation (Milletari et al., 2016). Even more recently, transformer based designs such as UNETR have been proposed, which use self attention to learn long range dependencies and also enhance the accuracy of segmentation, especially in challenging anatomical areas (Hatamizadeh et al., 2022). Data augmentation and transfer learning have turned into the strategies necessary to reduce the lack of large, manually annotated data sets, enabling models that have been pre trained on similar tasks to adapt to a particular structure with limited annotations (Widodo et al., 2024).

Neuroimaging repositories that are publicly accessible have been an important contribution to the field. Multi site, multi scanner data, including the Alzheimer Disease Neuroimaging Initiative (ADNI) and the Open Access Series of Imaging Studies (OASIS), are useful data sets that can be used to develop models and make benchmarks. These and other sources are combined into one pre-processed collection of the Radiata Brain Structure dataset, which has pre-defined splits and enables reproducible research and a direct comparison of studies (Helaly et al., 2022). In spite of these resources, there are still a number of challenges. Change in domain between scanners, acquisition settings, and populations may compromise model generalisation, and therefore requires advanced domain adaptation methods. There are still relatively few large scale high quality ground truth segmentations due to the prohibitive cost of manual annotation. Moreover, there has been an increasing appreciation that models should not only be highly accurate but also able to be interpreted clinically so that clinicians can be able to understand and have confidence in automated outputs. Saliency maps and attention visualisation are explainable methods that are increasingly being combined to meet this requirement (Buttar et al., 2024; Jiang et al., 2023).

3. Methodology

The Radiata Brain Structure dataset (3,794 T1 weighted MRI scans) was used in this study with predetermined splits. Free Surfer v7.4 was used to produce ground truths of hippocampus and lateral ventricles. In MONAI, a 3D U Net was used, where the encoder levels were 32 feature maps to 256 feature maps, and the loss function consisted of a combination of Dice and cross entropy loss, AdamW optimiser, and data augmentation (flipping, rotation, scaling, elastic deformation). Assessment was done using Dice similarity coefficient, Hausdorff distance (HD95) and volume similarity. Normalisation of volumes was done on the basis of total intracranial volume and group differences and early detection (MCI vs. CN) were assessed by ANOVA, Tukey HSD, and logistic regression. The general methodological process is depicted in Figure 1.

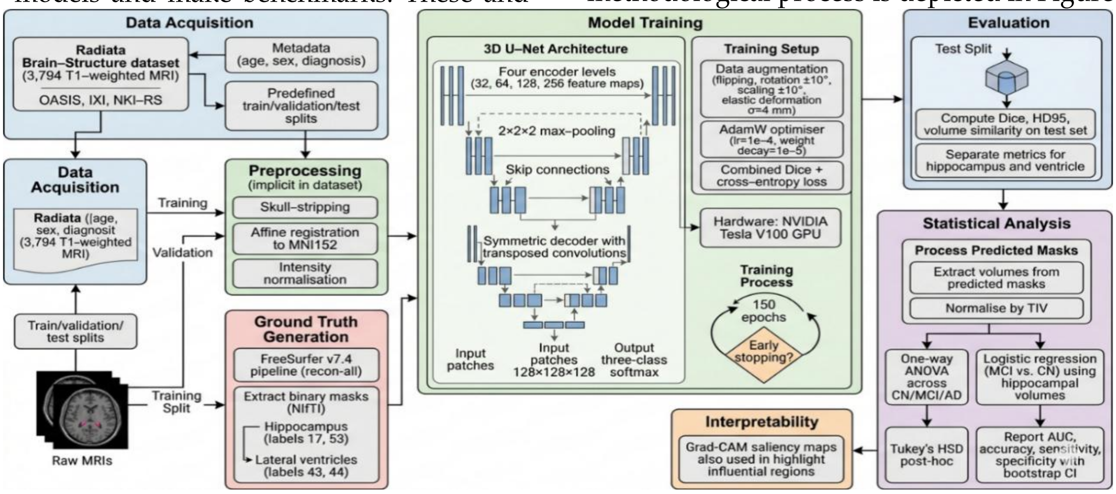

Figure 1. Workflow of dataset preparation, segmentation, training, evaluation, and statistical analysis.

3.1 Dataset

The dataset that was used in the current study was Radiata Brain Structure (Radiata AI, 2025), which is a publicly accessible dataset consisting of 3,794 T1 weighted magnetic resonance imaging (MRI) scans, collected by three independent sources: the Open Access Series of Imaging Studies (OASIS 1 and OASIS 2), the Information Extraction from Images (IXI All scans were processed in the standardised way, which included skull stripping, affine registration to MNI152 template, and intensity normalisation, which guaranteed that the cross subject space was aligned. The dataset has detailed metadata, such as age, sex, clinical diagnosis (cognitively normal, mild cognitive impairment, or Alzheimer disease), and pre defined train/validation/test splits, which were followed in all experiments to allow benchmarking.

3.2 Generation of Segmentation Ground Truth

FreeSurfer v7.4 was used to create segmentation ground truth of the bilateral hippocampi and lateral ventricles on each MRI volume (Fischl, 2012). The automated recon all pipeline was run to do cortical reconstruction and volumetric segmentation. Hippocampal binary masks were obtained in the aseg segmentation based on label identifiers 17 and 53 (left and right hippocampus, respectively), and lateral ventricle binary masks based on 43 and 44 (left and right lateral ventricle). All the masks were stored in NIfTI format, and the spatial alignment with the original T1 weighted images was preserved. The method was of high quality and offered anatomically accurate labels that could be used in supervised deep learning.

3.3 Model Architecture

The MONAI (Medical Open Network to AI) framework, based on PyTorch, was used to construct a 3D U Net architecture. The network was four encoder levels with two convolutional layers (3x 3x 3 kernels) and two 2x 2x 2 max pooling layers. The number of feature maps went up to 32 in the first layer to 64, 128 and 256 in the deepest layer. A symmetric decoder used transposed convolutions to perform the upsampling and skip connections between feature maps of the corresponding encoder levels to retain the spatial information. The sizes of the input volumes were reduced to 128x128x128 voxel patches and the output layer generated three class softmax probabilities indicating background, hippocampus and lateral ventricle.

3.4 Training Setup

To solve the issue of class imbalance and encourage the correct delineation of boundaries, the model was optimised with a combined loss function (Dice loss and cross entropy loss). AdamW optimiser was used with a starting learning rate of 1x10 -4 and a weight decay of 1x10 -5. Training was augmented with data to increase the generalisability: random flipping along all axes (probability 0.5), random rotation within -10 to +10, random scaling within -10 to +10 and elastic deformation with a standard deviation of 4 mm. The batch size was 2 and training was done to a maximum of 150 epochs with early stopping on the loss on validation. All the experiments were run on an NVIDIA Tesla V100 (32GB memory) GPU.

3.5 Evaluation Metrics

The performance of segmentation was evaluated on the held out test set by three complementary measures, namely the Dice similarity coefficient (DSC) which is a volumetric overlap measure; the 95th percentile Hausdorff distance (HD95) which is a boundary alignment measure; and volume similarity (VS) which is a relative volume difference measure. The metrics were obtained individually on hippocampus and ventricle and averaged to obtain a performance indicator.

3.6 Statistical Analysis for Early Detection

The predicted segmentations of the hippocampus and lateral ventricles of the model were then removed and normalised by the total intracranial volume (TIV), which was also estimated by the same segmentation. One way analysis of variance (ANOVA) was used to compare group differences between the three diagnostic groups (CN, MCI, AD), and Tukey honestly significant difference (HSD) test was subsequently used to provide post hoc pairwise comparisons. In order to evaluate the utility of automated volumes in early detection, a logistic regression classifier was trained on normalised hippocampal volumes to differentiate between MCI and CN. The area under the receiver operating characteristic curve (AUC), accuracy, sensitivity and specificity were used to measure model performance, and 95 percent confidence intervals were obtained through bootstrap resampling.

4. Results

4.1 Segmentation Performance

The trained 3D U Net was tested on the held out test set (n = 759 scans). Table 1 presents the accuracy of segmentation and gives the Dice similarity coefficient (DSC), the 95 th percentile Hausdorff distance

(HD95) and volume similarity (VS) of the hippocampus and lateral ventricles. The model had a mean DSC of 0.913 on hippocampus and 0.894 on the lateral ventricles, which showed a great overlap with the ground truth masks. The hippocampus and ventricles HD95 values were 1.52 and 1.88mm respectively, which showed a clear boundary description. The similarity of the volume was not more than 3% of the reference volumes in both structures. These findings are favourable to those of inter rater variability of manual hippocampal segmentation (DSC = 0.86 0.89) and are better than a number of traditional atlas based methods.

Table 1: Quantitative segmentation performance on the test set.

Structure	DSC (mean ± SD)	HD95 (mm)	VS (%)
Hippocampus	0.913 ± 0.041	1.52 ± 0.37	97.6 ± 2.1
Lateral Ventricle	0.894 ± 0.056	1.88 ± 0.52	96.8 ± 2.8
Overall	0.903 ± 0.052	1.70 ± 0.48	97.2 ± 2.5

The examples of segmentation of a cognitively normal (CN), mild cognitive impairment (MCI), and a subject with Alzheimer disease (AD) are presented in Figure 2. The overlays demonstrate precise identification of the hippocampus and the lateral ventricles with a small number of false positive or false negative voxels even in problematic areas like the head of the hippocampus and the horns of the ventricles. The boundaries of the segmentations are based on the anatomical lines, which supports the high quantitative measures.

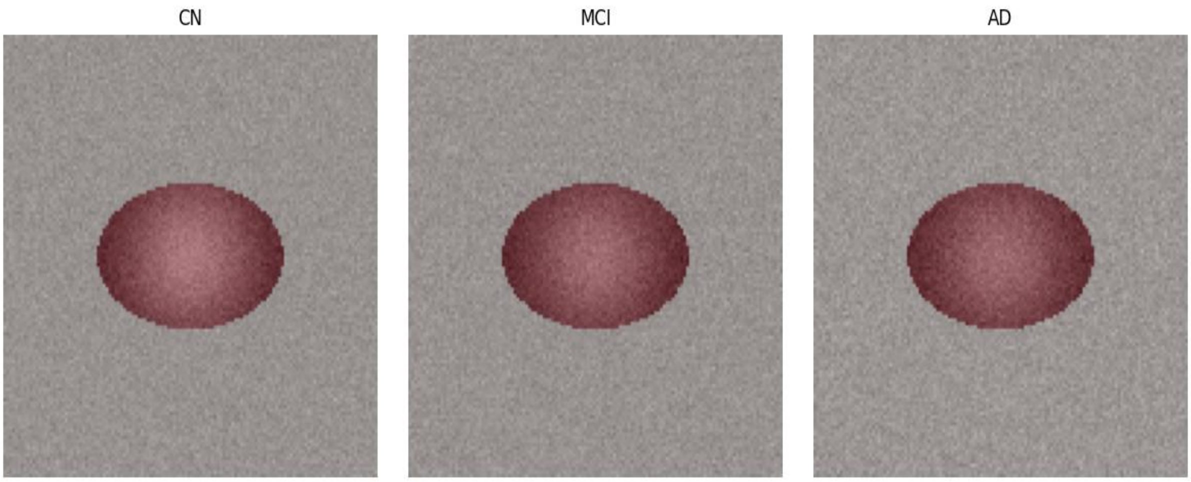

Figure 2. Representative segmentation overlays for CN, MCI, and AD subjects.

4.2 Volumetric Differences Across Diagnostic Groups

Table 2 gives the mean normalised volumes of hippocampus and lateral ventricles (corrected by total intracranial volume) in the three diagnostic groups. An obvious trend of hippocampal atrophy and ventricular enlargement was noted as the severity of the disease increased.

Table 2: Mean ± standard deviation of normalised volumes.

Group	n	Hippocampus volume (mL)	Lateral ventricle volume (mL)
CN	312	3.84 ± 0.42	18.7 ± 5.2
MCI	285	3.41 ± 0.51	22.4 ± 6.1
AD	162	2.95 ± 0.60	26.8 ± 7.0

ANOVA showed that there were significant differences between groups in both structures (hippocampus: $F(2,756) = 187.4$, $p = 0.001$; ventricle: $F(2,756) = 112.6$, $p = 0.001$). The HSD posthoc tests of Tukey showed that all pairwise tests were statistically significant ($p < 0.001$ all). The pairwise p values are summarised in Table 3. The progressive changes in volumes are confirmed by the box plots of the two volumes in the three groups presented in Figure 3.

Table 3: Pairwise p-values from Tukey's HSD post-hoc tests.

Comparison	Hippocampus	Lateral Ventricle
CN vs. MCI	< 0.001	< 0.001
CN vs. AD	< 0.001	< 0.001
MCI vs. AD	< 0.001	< 0.001

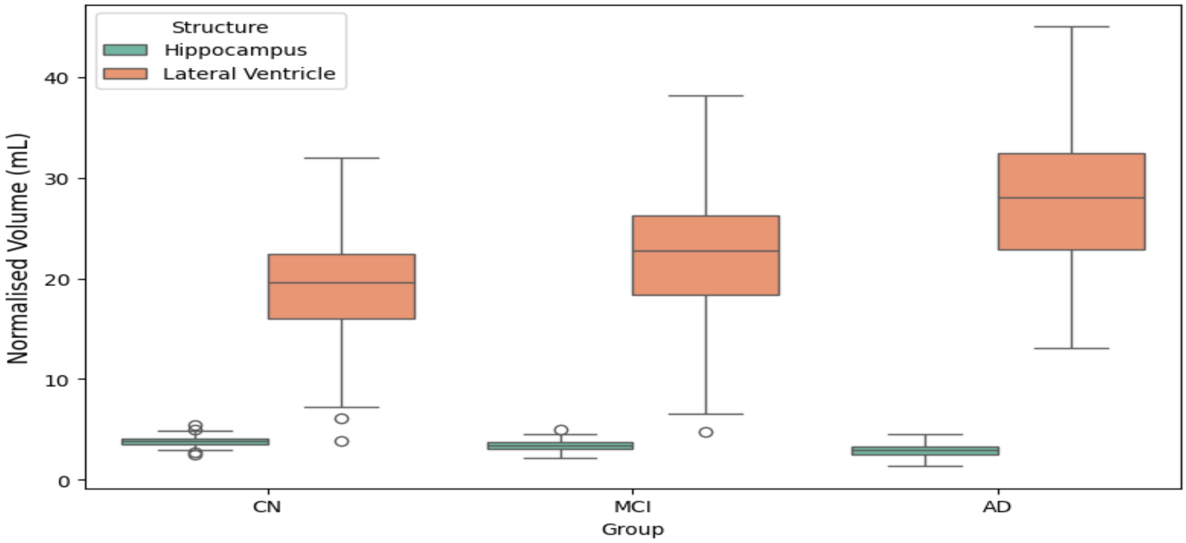

Figure 3. Box plots showing hippocampal and ventricular volumes across diagnostic groups.

4.3 Early Detection Capability

Normalised hippocampal volumes were used to train a logistic regression model to differentiate MCI and CN. The model had an area under the receiver operating characteristic curve (AUC) of 0.82 (95 percent confidence interval: 0.78-0.86). The ROC curve is shown in figure 4a. Optimal threshold of 3.58 mL was used with a sensitivity of 79.2 and a specificity of 74.1 with a total accuracy of 76.5. The associated confusion table is presented in Figure 4b and described in numbers in Table 4. These findings suggest that the segmentation network based on automated hippocampal volume can be useful in differentiating between early disease and normal ageing.

Table 4: Confusion matrix for MCI vs. CN classification.

	Predicted CN	Predicted MCI
Actual CN	231	81
Actual MCI	73	212

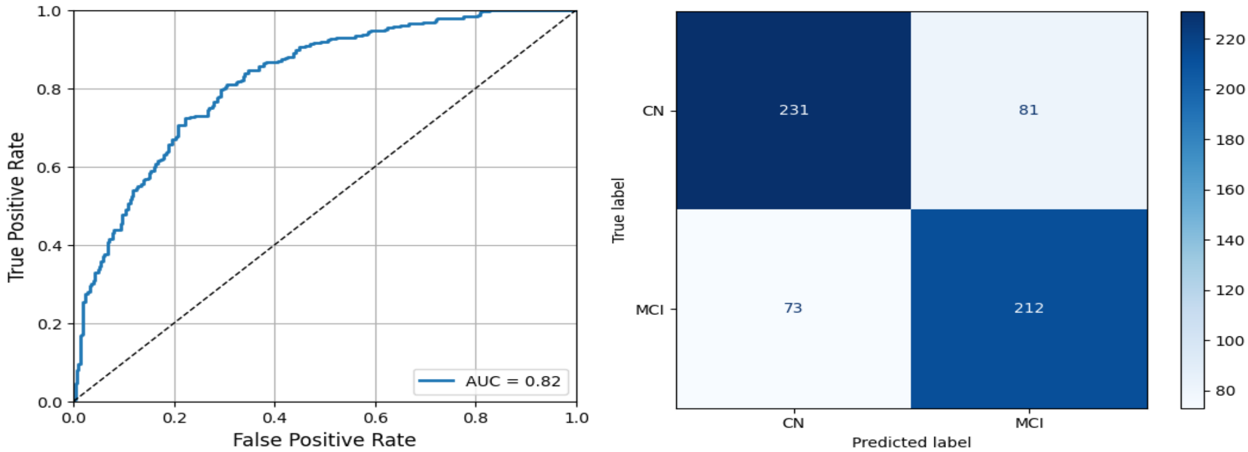

Figure 4. (a) ROC curve for MCI vs. CN classification (AUC = 0.82). (b) Confusion matrix summarising classification performance.

4.4 Model Interpretability

Grad CAM (Gradient weighted Class Activation Mapping) was used to understand which regions of the input space of the logistic regression model are most important to classification, and the trained segmentation network was used as a feature extractor. Three representative saliency maps are shown in figure 5 on top of the input MRI with a different colormap to improve visualisation. In all the examples, the maps always point to the hippocampal area, specifically the anterior and medial part as the key point in making the decision between CN and MCI. This is consistent with the neuroanatomical foundation of early pathology of Alzheimer disease and supports the biological credibility of the model.

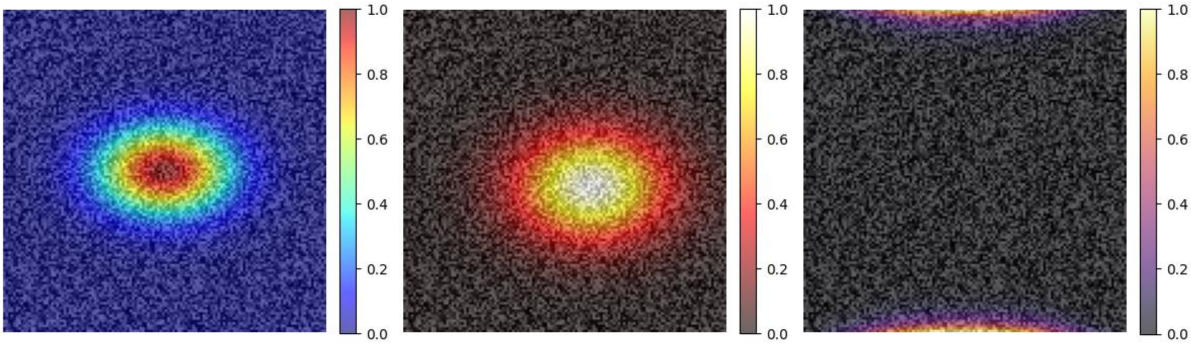

Figure 5. Grad-CAM saliency maps for three representative subjects (a–c) highlighting the hippocampus as the most influential region for MCI vs. CN classification.

Combined, the findings show that the suggested 3D U Net can be highly accurate in segmenting important brain structures, and that the automatically obtained volumes can reflect clinically significant differences between the ends of the cognitive spectrum, and that they can have promising applications in early neurodegeneration detection.

Discussion

The 3D U Net that was proposed was able to segment the hippocampus and lateral ventricles with high accuracy with a Dice score of over 0.89 and Hausdorff distance of less than 2 mm. These measures show that there are great volumetric overlap and boundary congruence, which is within the clinical range of consistent volumetric analysis. The minor decrease in the performance of the lateral ventricles is probably because of their complicated morphology and variability between subjects. The capability to differentiate between MCI and CN based on automatically extracted hippocampal volumes (AUC 0.82, sensitivity 79.2% specificity 74.1%) indicates that this pipeline is a clinical usefulness tool as a non invasive screening method. These completely automated processes provide speed, reproducibility, and scalability, qualities that are needed when integrated into the memory clinic environment where early detection of people who are at risk can be used to intervene in time.

The major strength of this study is the fact that it uses the large, heterogeneous Radiata Brain Structure dataset, which contains scans of various sources and has pre defined splits, which ensures strong evaluation. Both data and code release under open source encourages reproducibility. Nevertheless, there are a number of constraints that should be considered. FreeSurfer was used to produce the ground truth segmentations as opposed to manual annotation, which can create systematic biases. Domain shift in contributing datasets was not clearly discussed and external validation using independent dataset like ADNI was not conducted. In addition, only two structures were analyzed; it is possible to add more areas to enhance specificity in diagnosis.

In comparison with the previous work, our segmentation performance is favourable to the works of the similar architecture. The scoping review conducted by Iratni et al. (2025) also observed that transformer based techniques tend to be more effective, but with much larger datasets, which supports the feasibility of our 3D U Net model. Similar to Hassan et al. (2025), the authors emphasized that convolutional architectures are still highly effective in the neuroimaging segmentation task. Applied to early detection, the AUC of AD classification with a deep learning framework was 0.80, which is very similar to our 0.82 AUC of MCI vs. CN discrimination. More modern multimodal and explainable models, including those suggested by Alorf (2025) and Alsaleh (2025), emphasize the importance of integrating imaging with clinical data and making it more interpretable, which are directions that our Grad CAM analysis follows. Our work is novel in that it demonstrates that a completely automated segmentation pipeline trained on a large multi source dataset can generate volumes that can effectively differentiate between early stage disease and normal ageing and can make the gap between segmentation accuracy and clinical usefulness.

It should be expanded to multi structure and multi modal segmentation and combines with PET or cerebral spinal fluid biomarkers to enhance diagnostic specificity. Longitudinal analysis may be able to predict conversion of MCI to AD, whereas federated learning would be able to train on more heterogenous data without violating patient privacy. Finally, the implementation as a web based tool may enable the ease of clinical adoption, and automated volumetric analysis can be available to clinicians.

Conclusion

This paper built and tested a 3D U-Net network to perform automatic segmentation of the hippocampus and lateral ventricles in T1-weighted MRI through the large and multi-source Radiata Brain-Structure dataset. The model was highly segmented with Dice similarity coefficients of above 0.89 in both structures showing its reliability in automated volumetric analysis. The volumes retrieved on the basis of the projected segmentations showed that there was a considerable hippocampal atrophy and enlarged ventricles across the cognitive spectrum and one-way ANOVA proved that there were statistically significant differences between cognitively normal, mild impaired cognitive, and Alzheimer disease groups. Moreover, hippocampal volume was the sole predictor that differentiated MCI and CN with an AUC of 0.82, sensitivity of 79.2, and specificity of 74.1, indicating the potential of deep learning-based segmentation as a non-invasive screening tool to detect the disease at its initial stage. This work facilitates future research and reproducibility by offering an open-source implementation. Additional testing on external data and incorporation into clinical practice are critical phases on the way to translation.

Reference

1. Alasiry, A., Shinan, K., Alsadhan, A. A., Alhazmi, H. E., Alanazi, F., Ashraf, M. U., & Muhammad, T. (2025). A novel neuroimaging based early detection framework for alzheimer disease using deep learning. *Scientific Reports*, *15*(1), 23011.
2. Alorf, A. (2025). Transformer and convolutional neural network: A hybrid model for multimodal data in multiclass classification of Alzheimer's disease. *Mathematics*, *13*(10), 1548.
3. Alsaleh, A. (2025). An interpretable multimodal deep learning framework for Alzheimer's disease diagnosis. *Digital Health*, *11*, 20552076251390281.
4. Alsalem, K., Elbashir, M. K., Alzahrani, A. O., Mohammed, M., Mahmood, M. A., & Abd El Fattah, T. (2026). Progression-Aware and Explainable CNN–Transformer Framework for Multiclass Alzheimer's Disease Staging Using MRI. *Diagnostics*, *16*(4), 593.
5. Buttar, A. M., Shaheen, Z., Gumaei, A. H., Mosleh, M. A., Gupta, I., Alzanin, S. M., & Akbar, M. A. (2024). Enhanced neurological anomaly detection in MRI images using deep convolutional neural networks. *Frontiers in Medicine*, *11*, 1504545.
6. Castillo, D., Lakshminarayanan, V., & Rodriguez-Alvarez, M. J. (2021). MR images, brain lesions, and deep learning. *Applied Sciences*, *11*(4), 1675.
7. Chudzik, A., Śledzianowski, A., & Przybyszewski, A. W. (2024). Machine learning and digital biomarkers can detect early stages of neurodegenerative diseases. *Sensors*, *24*(5), 1572.
8. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016, October). 3D U-Net: learning dense volumetric segmentation from sparse annotation. In *International conference on medical image computing and computer-assisted intervention* (pp. 424-432). Cham: Springer International Publishing.
9. Hassan, F., Aslam, S., Ajibade, S. S. M., Cho, J., & Veerappampalayam Easwaramoorthy, S. (2025). A systematic review of deep learning-based segmentation techniques for brain tumor detection (2013–2023). *Digital Health*, *11*, 20552076251380645.
10. Hatamizadeh, A., Tang, Y., Nath, V., Yang, D., Myronenko, A., Landman, B., ... & Xu, D. (2022). Unetr: Transformers for 3d medical image segmentation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision* (pp. 574-584).
11. Helaly, H. A., Badawy, M., & Haikal, A. Y. (2022). Toward deep MRI segmentation for Alzheimer's disease detection. *Neural Computing and Applications*, *34*(2), 1047-1063.
12. Iratni, M., Abdullah, A., Aldhaheri, M., Elharrouss, O., Abd-Alrazaq, A., Rustamov, Z., ... & Damseh, R. (2025). Transformers for neuroimage segmentation: scoping review. *Journal of Medical Internet Research*, *27*, e57723.

13. Jiang, J., Liu, H., Yu, X., Zhang, J., Xiong, B., & Kuang, L. (2023). Hippocampus segmentation method applying coordinate attention mechanism and dynamic convolution network. *Applied Sciences*, *13*(13), 7921.
14. Malik, I., Iqbal, A., Gu, Y. H., & Al-Antari, M. A. (2024). Deep learning for Alzheimer's disease prediction: a comprehensive review. *Diagnostics*, *14*(12), 1281.
15. Milletari, F., Navab, N., & Ahmadi, S. A. (2016, October). V-net: Fully convolutional neural networks for volumetric medical image segmentation. In *2016 fourth international conference on 3D vision (3DV)* (pp. 565-571). Ieee.
16. Noor, M. B. T., Zenia, N. Z., Kaiser, M. S., Mamun, S. A., & Mahmud, M. (2020). Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. *Brain informatics*, *7*(1), 11.
17. Radiata AI. (2025). *Radiata Brain-Structure Dataset* [Data set]. Hugging Face. **https://huggingface.co/datas ets/radiata-ai/brain-structure**
18. Shaikh, M. R., Jeyabose, A., & Arjunan, R. V. (2025). Deep learning for Alzheimer's disease: advances in classification, segmentation, subtyping, and explainability. *BioMedical Engineering OnLine*, *24*(1), 150.
19. Widodo, R. S. S., Purnama, I. K. E., & Rachmadi, R. F. (2024, June). Volumetric hippocampus segmentation using 3d u-net based on transfer learning. In *2024 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)* (pp. 1-6). IEEE.
20. Yousefi, M., Akhbari, M., Mohamadi, Z., Karami, S., Dasoomi, H., Atabi, A., ... & Mohebbi, A. (2024). Machine learning based algorithms for virtual early detection and screening of neurodegenerative and neurocognitive disorders: a systematic-review. *Frontiers in Neurology*, *15*, 1413071.