NEUROSCIENCE RESEARCH NOTES

OPEN ACCESS | RESEARCH NOTES ISSN: 2576-828X

Comparison of deep learning convolutional neural networks method with conventional volume-based morphometry measurement of hippocampal volume in Alzheimer's disease

Nur Shahidatul Nabila Ibrahim [1], Subapriya Suppiah [1,2,3,4]*, Buhari Ibrahim [1], Nur Hafizah Mohad Azmi [1], Vengkatha Priya Seriramulu [1], Mazlyfarina Mohamad [5], Marsyita Hanafi [6], Hakimah Mohammad Sallehuddin [4,7], Rizah Mazzuin Razali [8] and Noor Harzana Harrun [9]

[1] Department of Radiology, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, Selangor, Malaysia.
[2] Nuclear Imaging Unit, Hospital Sultan Abdul Aziz Shah, Universiti Putra Malaysia, Selangor, Malaysia.
[3] Centre for Diagnostic Nuclear Imaging, Universiti Putra Malaysia, Selangor, Malaysia.
[4] Malaysia Research Institute on Ageing (MyAgeing), Universiti Putra Malaysia, Selangor, Malaysia.
[5] Centre for Diagnostic and Applied Health Sciences, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia.
[6] Faculty of Engineering, Universiti Putra Malaysia, Selangor, Malaysia.
[7] Department of Geriatric, Hospital Sultan Abdul Aziz Shah, Universiti Putra Malaysia, Selangor, Malaysia.
[8] Department of Medicine, Hospital Kuala Lumpur, Ministry of Health of Malaysia, Kuala Lumpur, Malaysia.
[9] Pandamaran Health Clinic, Ministry of Health of Malaysia, Selangor, Malaysia.
* Correspondence: subapriya@upm.edu.my; Tel.: +603-97692512

Received: 22 April 2023; **Accepted:** 12 October 2023; **Published:** 23 December 2023
Edited by: Indranath Chatterjee (Tongmyong University, South Korea)
Reviewed by: Sukumal Chongthammakun (Mahidol University, Thailand);
Siti Aishah Abdul Aziz (Universiti Sains Malaysia, Malaysia)
https://doi.org/10.31117/neuroscirn.v6i4.248

Abstract: Dementia is a spectrum of diseases characterised by a progressive and irreversible decline in cognitive function. Appropriate tools and references are essential for evaluating individuals' cognitive levels, especially hippocampal volume, as it is the commonly used biomarker in detecting Alzheimer's disease (AD). It is important to note that while there is no cure for dementia, early intervention and support can greatly improve the lives of those affected. Our ongoing AD research is being conducted to develop new treatments and improve our understanding of the disease by using voxel-based morphometry (VBM) to compare sensitivity and specificity with the HippoDeep toolbox. We validated AD's hippocampal volume compared to age-matched healthy controls (HC) based on the HippoDeep Model by comparing it with VBM as the reference standard. Significant differences between hippocampal volume in AD and HC have been detected using VBM and HippoDeep analysis.

Keywords: Hippocampal volume; Alzheimer's disease; Deep learning; Convolutional neural network

1.0 INTRODUCTION

Alzheimer's Disease (AD) is the most common neurodegenerative disorder affecting older adults, leading to significant morbidity. A combination of clinical assessment, neuropsychological tests, and diagnostic imaging to help increase diagnostic accuracy (Piersson et al., 2021). One of the hallmarks of AD is hippocampal atrophy, which is detected using magnetic resonance imaging (MRI). In both MRI-based and histology-based investigations, the scientific literature indicates that the average volume of the hippocampus falls within this range of 1.73 to 5.68 cm^3 (Honeycutt & Smith, 1995). Nevertheless, the numerous measurement methods can give rise to variable results.

Historically, the hippocampal volume measurement was made using 2D slices on MRI images acquired on a high-resolution T1-weighted MRI sequence. Eventually, this progressed into developing the hand-drawn volume of interest (VOI), which segmented the brain and selected the hippocampus for volumetric assessment. Several image processing software have also been designed to evaluate hippocampal volume; for example, the Statistical Parametric Mapping version 12 (SPM12) tool can be used to process raw structural data to measure the volume of the hippocampus. SPM can be used to assess voxel-based morphometry (VBM) of the hippocampus in a semi-automated manner. VBM is an example of an image analysis technique that has gained popularity since its inception (Wright et al., 1995; Ashburner & Friston, 2000) primarily because it is relatively simple to implement and has produced biologically plausible results (Whitwell, 2009).

Nevertheless, this technique has inherent limitations as it is time-consuming, repetitive, and necessitates specialised training. According to Kurth et al. (2015), the acquisition methods and modelling of VBM analysis were presented, which included three fundamental pre-processing steps: tissue classification, spatial normalisation, and spatial flattening, followed by the statistical analysis itself. The measurements are potentially inconsistent and depend on the person performing the measurements. Thus, artificial intelligence (AI) may provide a better solution to enable more consistent, accurate, and reproducible results. Using deep learning model, the convolutional neural network (CNN) algorithm permits an automated pipeline to measure more data within a shorter period. CNN involves using novel computational models, which are based on 3D-deep learning frameworks, and can potentially incorporate biomarker-based serial labelled data for improved diagnostics of AD (Ibrahim et al., 2021a).

HippoDeep, a new CNN-based algorithm, has emerged as a rapid and reliable hippocampal segmentation method (Thyreau et al., 2018). HippoDeep utilises hippocampal "appearance" rather than a singular atlas-based method. HippoDeep is a recently developed algorithm for automated hippocampal segmentation that has not yet been tested on AD patients. HippoDeep does not warp individual images into an atlas; instead, it uses a hippocampal appearance model learned from existing FreeSurfer v5.3 labelled online data sets and synthetic data (Thyreau et al., 2018). Moreover, HippoDeep has a better spatial agreement with manual segmentations than FreeSurfer version 6.0, "sum of subfields" segmentation in healthy ageing populations (Zavaliangos-Petropulu et al., 2020).

Among all lateralised brain regions, the hippocampus performs a unique and crucial function as a precursor to the broader asymmetrical development of the human brain. Our study aimed to compare the diagnostic performance of HippoDeep compared to the conventional VBM method for measuring hippocampal volume and predicting AD. We hypothesise that the HippoDeep Model had a good correlation and comparable diagnostic accuracy with the VBM method.

2.0 MATERIALS AND METHODS

2.1 Subjects' recruitment and setting

This cross-sectional case-control study was conducted from February 2021 to March 2023. The AD subjects were recruited from the memory clinics at the Geriatric Clinic in Hospital Kuala Lumpur, Klinik Kesihatan Pandamaran in Klang, and Hospital Sultan Abdul Aziz Shah, Universiti Putra Malaysia. These AD subjects were diagnosed based on the clinicians' examination and neuropsychological testing using various cognitive assessment instruments, such as the Mini-Mental State Examination (MMSE) and the Montréal Cognitive Assessment (MoCA) questionnaires, overall being guided by the DSM-5 criteria for major neurodegenerative disorders (First et al., 2018).

The cognitively healthy older adults (HC) subjects were recruited through advertisements on local bulletin boards and flyers distributed to the community. All the AD subjects were recruited for this study based on the clinician's diagnosis. The clinicians used DSM-5 criteria, which included information gathered from interviews with the subjects and their caregivers. They also used neuropsychological tests as a guide, such as MMSE and

MoCA. MMSE range for mild AD: 21–26, moderate AD: 10–14, severe AD: less than 10. Meanwhile, for MoCA scores, mild AD is 18-25, moderate AD is 10-17, and severe AD is less than 10. The Clinical Dementia Ratings (CDR) scores were utilised to evaluate the cognitive performance of the subjects' accustomed functioning in everyday tasks, indicating their cognitive decline. CDR 0 indicates no dementia, CDR 0.5 indicates mild cognitive impairment (MCI), whereas CDRs 1, 2, and 3 indicate mild, moderate, and severe dementia.

2.2 Sample size determination

The calculated sample size was based on the data presented by Ibrahim et al. (2021b). Whereby the prevalence of AD in this region was 25%. We utilised a web-based sample size calculator downloaded from http://wnarifin.github.io., giving a minimum sample size of 30 subjects per group.

2.3 Ethical clearance

This study was approved by the Medical Research Ethics Committee (MREC) of the National Medical Registration Registry (NMRR) Malaysia (NMRR-19-2719-49105) and the Ethics Committee for Research Involving Human Subjects of Universiti Putra Malaysia (JKEUPM-2019-328).

2.4 Neuropsychological testing

All the recruited subjects underwent our trained researchers' MMSE, MoCA and CDR assessments. The values were interpreted as MoCA: >26 as cognitively healthy; 16.3-25 as moderate cognitive impairment; <16.2 as having severe cognitive impairment, MMSE: 24-30 as cognitively normal; 18-23: moderate cognitive impairment and 0-17: severe cognitive impairment: As for the CDR that evaluated the effect on the activities of daily living, the assessment was based on a questionnaire with, No Dementia (CDR = 0), Questionable cognitive impairment or MCI (CDR = 0.5), mild dementia (CDR = 1), moderate dementia (CDR = 2), and Severe dementia (CDR = 3).

2.5 Protocol for structural MRI imaging

High resolution, T1-weighted MRI scans were performed using a 3.0-T MAGNETOM Prisma scanner (Siemens Healtineers, Germany) with software version Syngo MR D13D, having repetition time (TR): 2300 ms; echo time (TE): 2.27 ms; Inversion time: 900 ms; Slice thickness: 1 mm; Number of slices: 160; Flip Angle: 8.

2.6 Pre-processing using SPM Matlab

All the T1-weighted images were pre-processed using the SPM Toolbox: (https://www.fil.ion.ucl.ac.uk/spm/download/restricted/eldorado/spm12.zip) implemented in MATLAB (R2021a).

2.7 Processing VBM

The steps were adopted from the VBM protocol, beginning with segmentation (spatial pre-processing): using the native space option, the researcher can generate a tissue class image aligned with the original. Next, DARTEL was executed to generate templates from selected images to be distorted. The next step involved normalisation, where the tissue class image was transformed into a standard space, followed by smoothing to improve the signal-to-noise ratio. These steps were crucial for the preparation of the data for further analysis. Images with the same dimensions, orientation, and voxel size were selected for statistical analysis using the factorial design specification. The hippocampal volume was calculated for each subject using a design specification file from the computer's file system. The Automated Anatomical Labeling (AAL) toolbox was superimposed with the Montreal Neurological Institute (MNI) template to extract the volume to cluster level.

2.8 Processing for HippoDeep using FSL

HippoDeep relies on a hippocampal appearance model learned from existing FreeSurfer v5.3 labelled online data sets and synthetic data rather than warping individual images to an atlas (Thyreau et al., 2018). The HippoDeep CNN was trained using two different kinds of synthetic data. The first artificial data set consisted of a manual segmentation of a synthetic high-resolution image of the hippocampus created from an average of 35 consecutive MRI scans of a healthy participant. The goal of segmenting the hippocampus on a high-resolution image (0.6 mm isotropic resolution) was to give the CNN more detailed boundary information that may need clarification on a lower-resolution image. The second type of synthetic data used to train the HippoDeep CNN is FreeSurfer v5.3 training data that has been artificially and geometrically warped. While some distortions were outside the range of clinically acceptable values, it is still realistic enough for a human rater to distinguish them. This distorted data was used to give the CNN practical training recommendations. Details on how the synthetic data were generated may be found in the study by Thyreau et al. (2018).

HippoDeep hippocampus segmentation was installed using a Repo for Ubuntu/Linux and Mac OS compatible: https://github.com/bthyreau/hippodeep_pytorch.

Meanwhile, for a Windows-based system, the repo used was https://github.com/bfoe/hippodeep_pytorch. We utilised an executable Windows pre-compiled app: https://github.com/bfoe/hippodeep_pytorch/releases/download/v0.3/HippoDeep_Windows_v0.3.zip. It is a plug-and-play pre-compiled app for use on the Windows platform HippoDeep_Windows_v0.3.zip. Once we installed the HippoDeep toolbox, we opened the high-resolution T1-weighted images and followed the instructions in the pop-up box. The processed images and reports were saved in the original T1-weighted file.

2.9 Statistical analysis

The statistical tests were conducted using Statistical Package for the Social Sciences (SPSS software Version 23.0, Chicago, USA) and MedCalc, a statistical software package designed for the biomedical sciences. It has an integrated spreadsheet for data input and can import files in several formats. The level of significance was set at p-value < 0.05. Spearman correlation was used to determine the association between the VBM and HippoDeep measurements of bilateral hippocampi volumes of AD and HC subjects. ANCOVA test was used to analyse the association between the hippocampi volume and age. The ROC curve was used in this study to compare the accuracy of hippocampus volume measurements between VBM and HippoDeep methods.

3.0 RESULTS

The final data was lower than our calculated sample size because we selected the best-suited age-matched subjects for this evaluation. **Table 1** shows age distribution and neuropsychological test scores based on AD and HC groups' mean values and standard deviation. The age range for our AD and HC have a mean value of 74.27 years and 71.2 years, respectively. This indicates that our two groups are age-matched and that any differences in the cognitive performance or brain imaging findings among the AD subjects with the HC subjects are not attributable to the normal ageing process. Furthermore, the significance of the low MMSE and MoCA scores among the AD subjects indicates that the AD subjects have significantly impaired cognitive function compared to our HC subjects.

In **Table 2**, Spearman rank correlation was computed to assess the association of hippocampus volume between the VBM method and the HippoDeep Toolbox method. There was a moderately positive correlation between VBM and the HippoDeep, r = 0.569, p < 0.001 for the measurement of the left hippocampus, while there was also a moderately positive correlation between VBM and HippoDeep, r = 0.453, p < 0.001 for the measurement of the right hippocampal volume.

Table 1: Distribution data for healthy controls (HC) and Alzheimer's disease patients (AD)

	HC	AD	p-value
n	15 (F= 10, M=5)	15 (F= 9, M=6)	-
Age	71.2 (7.68)	74.27 (10.05)	0.223
MoCA	27.33 (3.13)	14.13 (6.82)	<0.05*
MMSE	27.53 (2.61)	14.07 (7.79)	<0.05*
CDR	0 (0.00)	2.07 (0.88)	<0.05*

MMSE = Mini-Mental State Examination; MoCA = Montréal Cognitive Assessment; CDR = Clinical Dementia Ratings; Value = Mean (standard deviation); Significant value p < 0.05*.

Table 2. Bilateral hippocampal volume difference in Voxel-based morphometry (VBM) and HippoDeep.

Variable	Mean (SD)		p-value
	VBM	HippoDeep	
Left Hippocampus Volume	1.39 (0.43)	2.55 (0.60)	<0.001
Right Hippocampus Volume	1.47 (0.38)	2.68 (0.66)	<0.001

Value = Mean (standard deviation); Significant value p < 0.05*.

A One-way ANCOVA was conducted to determine whether there was any significant difference between the hippocampal volume of the AD and HC based on VBM and HippoDeep methods using age as a covariate. There was a significant difference between the left hippocampal volume between the AD and HC groups using VBM analysis, F (1, 29) = 0.942, p < 0.001. There was also a significant difference in the left hippocampal of AD and HC using the HippoDeep toolbox, F (1, 29) = 1.08, p < 0.001. One-way ANCOVA also detected a statistically significant difference between right hippocampal volumes in HC and AD groups using the VBM analysis and HippoDeep method. There was a significant difference between the right hippocampus volume of AD compared to HC using VBM analysis, F (1, 29) = 0.795, p < 0.001. While using the HippoDeep toolbox, a significant difference was detected between the right hippocampal volume of the AD compared to HC groups, F (1, 29) = 1.403, p < 0.001.

Therefore, the HippoDeep model allows for an automated segmentation of the hippocampus bilaterally, and can aid in the measurement of the right hippocampal volume (red volume of interest) and left hippocampal volume (green volume of interest), while enabling comparison between AD and cognitively healthy subjects (**Figure 1**).

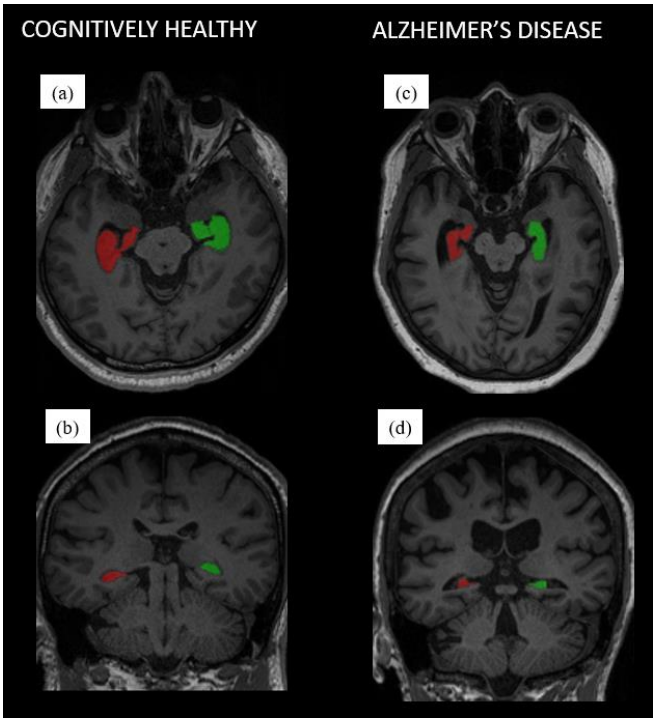

Figure 1. Automated MRI brain image segmentation using HippoDeep, showing the right hippocampus (red volume of interest) and left hippocampus (green volume of interest). Cognitively healthy brain segmentation in (**a**) axial view and (**b**) coronal view. Alzheimer's disease brain segmentation in (**c**) axial and (**d**) coronal views.

Figure 2 and **Figure 3** show the scatter plot pairs of numerical data of hippocampal volume distribution according to age for the VBM and HippoDeep-based methods, comparing the AD and HC groups' distribution to demonstrate their relationship. The linear line shows the pattern of decrease in hippocampus volume with increasing age, as this variable is inversely proportional to age. Furthermore, the right and left hippocampal volume values, measured using both methods, showed that these values were lower (showing more accelerated atrophy) in the AD group compared to the age-matched HC group.

4.0 DISCUSSION

Individuals with dementia are assessed for cognitive dysfunction using the Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA). Clinicians and researchers require norms for cognitive tests to assess the cognitive levels of subjects referred for evaluation (Riello et al., 2021). MMSE and MoCA were significantly different in AD and HC based on the study by Wei et al. (2022), which shows that these tests done by measuring data pooled from a large community-based study allow for the correct dichotomisation of the cognitively healthy and dementia patient.

Fundamentally, the most salient structural change ubiquitously present in AD patients is the atrophy of the hippocampus in the medial temporal lobe. AD subjects' hippocampal volume is significantly smaller than HC, even after being corrected for increasing age. A study by Hu et al. (2023) recommended that MRI-estimated hippocampal volume is a superior structural biomarker compared to estimating the total temporal lobe volume or entorhinal cortex volume. The European Federation of the Neurological Societies (EFNS) (Hort et al., 2010), the EMA (Hill et al., 2014), the National Institute on Aging and the Alzheimer's Association (NIAAA) (Albert et al., 2011), and the International Working Group (IWG) (Dubois et al., 2007) also recommended using the hippocampal volume as a supplementary biomarker facilitating the clinical diagnosis of AD. Thus, we explored HippoDeep, a new CNN-based hippocampal segmentation method, for our study because it does not rely exclusively on a singular atlas-based approach. This algorithm can be used to measure the hippocampal volume in a simplified and automated manner.

This interest stemmed from recognising that the hippocampal structures play a crucial role in memory formation and retrieval. Understanding how these substructures are affected in AD could provide valuable insights into this disease's underlying mechanisms of memory impairment (Sarica et al., 2018).

The hippocampus can be segmented automatically due to the features extracted using AI techniques (Hurtz et al., 2019), specifically the HippoDeep pipeline. Automating segmentation can eliminate errors caused by non-standardised hippocampal volume drawing. Among the available methods are AI clustering, region expansion, and thresholding algorithms. In this

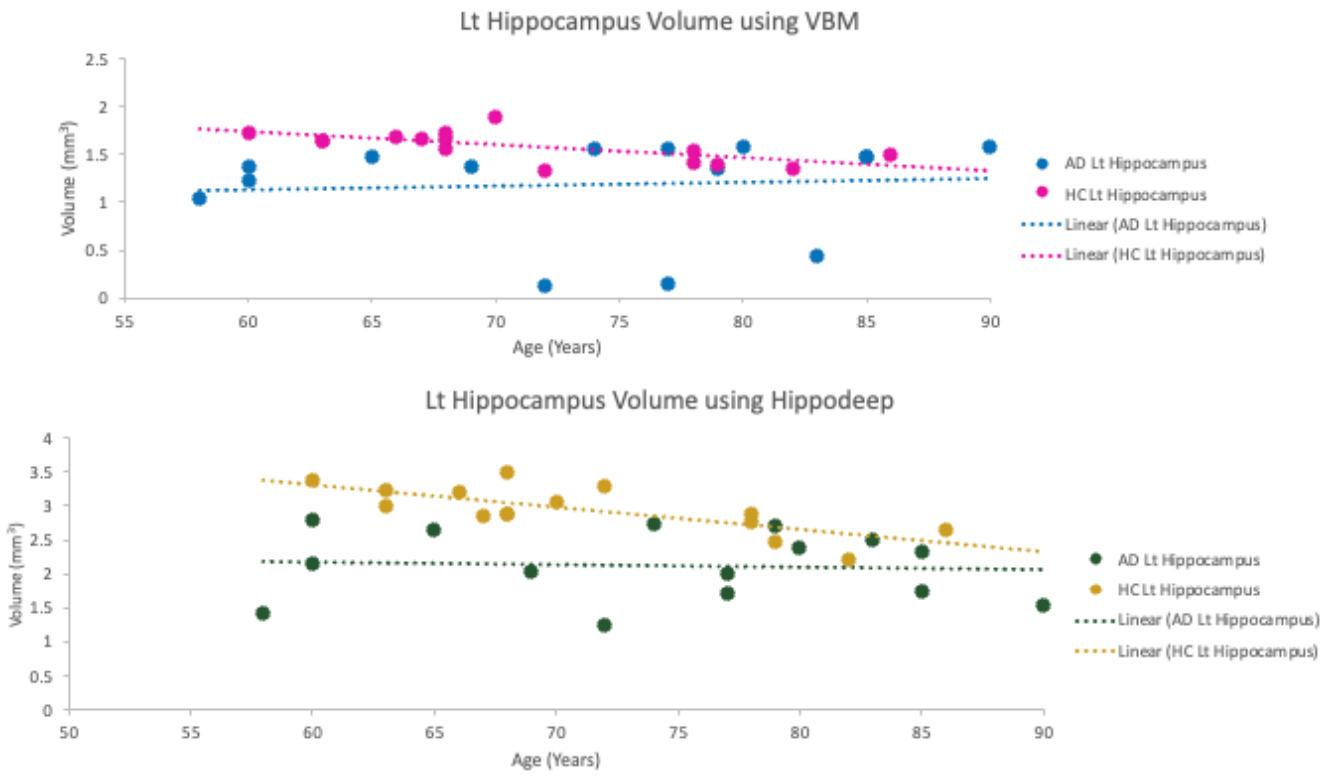

Figure 2. Left hippocampal volume distribution according to age using VBM and HippoDeep analysis for AD compared to HC groups.

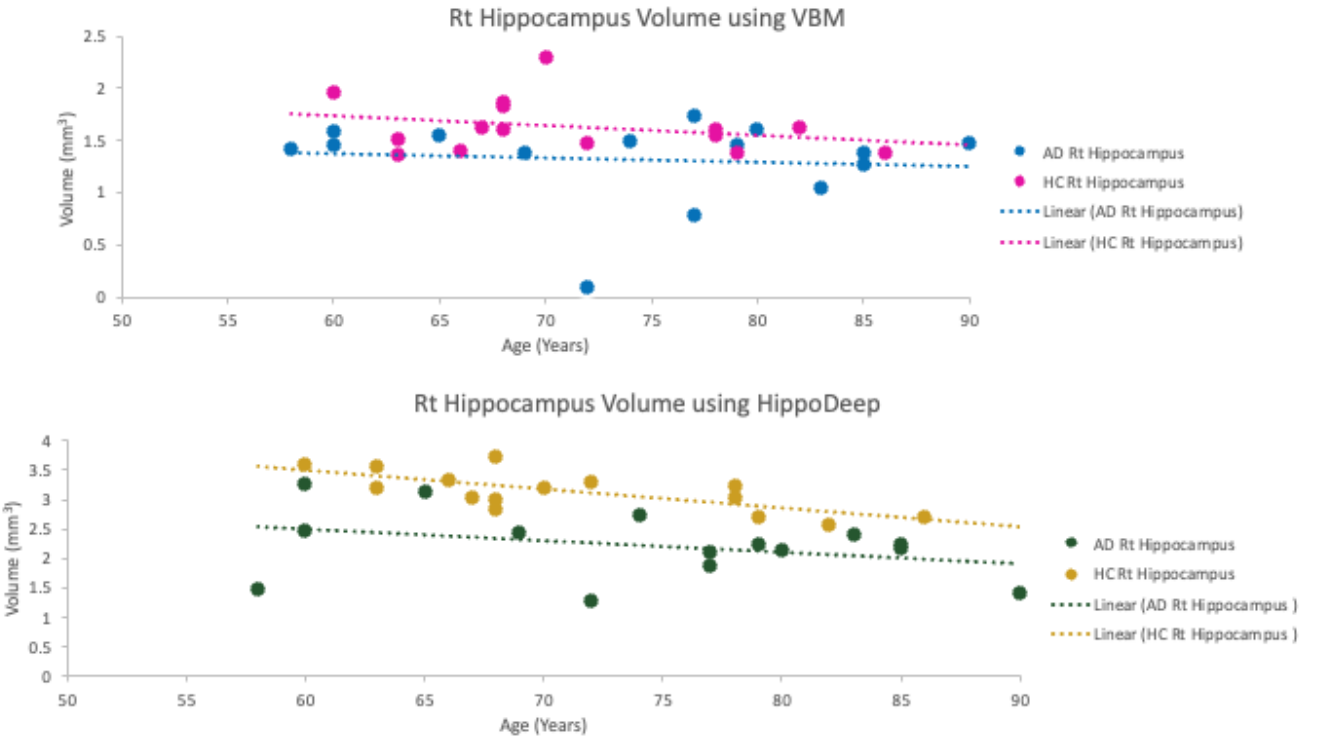

Figure 3. Right hippocampal volume distribution according to age using VBM and HippoDeep analysis for AD compared to HC groups

instance, we employed the clustering technique, which combines pixels that are related in some way or other. Thus, region-growing algorithms expand regions from a seed point until they reach an outer margin that defines the hippocampal volume.

In contrast, the VBM method makes the assumption of spatially normalising high-resolution images from all subjects in the study into the same stereotactic space and is a relatively straightforward procedure (Schell et al., 2023). Albeit, the VBM technique may cause over-corrections based on the Caucasian atlas utilised. After segmenting the GMV from the spatially normalised images, the segments were then smoothed.

The voxel-by-voxel parametric statistical comparisons of the smoothed grey matter images from the two groups were performed. Subsequently, multiple comparison corrections were made using the theory of Gaussian random fields.

HippoDeep could predict and classify AD better than HC, as evidenced by the ROC curve of left hippocampal volume with a sensitivity of 93.33%, specificity of 80.00% and area under the curve (AUC) = 0.927 (**Figure 4**). Besides, the right hippocampal volume had a sensitivity of 80.00%, specificity of 100.00% and AUC = 0.911. In contrast, the VBM method gave left hippocampal volume sensitivity of 100.00%, specificity of 53.33% and AUC = 0.800 for the left hippocampus, while for the right hippocampal volume, the sensitivity was 73.33%, specificity was 66.67 and AUC = 0.747. A study by DeMarshall et al. (2016) proved that it is possible to detect disease specificity of the selected biomarkers of AD using ROC curve assessment. A study by Suzuki et al. (2023) noted that AUC was more than 0.8, indicating a good sensitivity and specificity in general for classifying AD. HippoDeep can better predict and classify AD compared to HC by measuring the left hippocampus volume with an accuracy of 92.7%, compared to the VBM method, which has 80.0% accuracy for the same side of hippocampal volume.

Interestingly, our study also detected that the left hippocampus demonstrated a more significant decrease in volume than the right, which is the magnitude of non-directional hippocampal asymmetry that the decreasing cognitive state may accelerate. Moreover, *in vivo* MRI volumetry research consistently showed that the right hippocampus is larger than the left in large population studies (Pedraza et al., 2004), which may explain our observed phenomenon. Furthermore, we hypothesise that a more marked loss of left hippocampal volume is consistent with the verbal and language impairments observed in our AD subjects. This finding is comparable to that of Ezzati et al. (2016), who observed that the left hippocampal volume is more atrophying in subjects with verbal cognitive decline. According to Ezzati et al. (2016), right hippocampal volume loss was prevalent among individuals with spatial memory impairment. In our study, Alzheimer's disease patients did not exhibit a significant decline in spatial memory. Thus, our research did not detect a significant decrease in the right hippocampal volume.

The limitation of our study is the relatively small sample size. Thus, we recommend more extensive studies to evaluate this CNN technique to diagnose AD. Furthermore, larger representative samples may be able to suggest cutoff values for determining the difference between mild, moderate, and severe dementia. This may identify patients in their stages of the disease and help to give better prognostic markers to aid in the management.

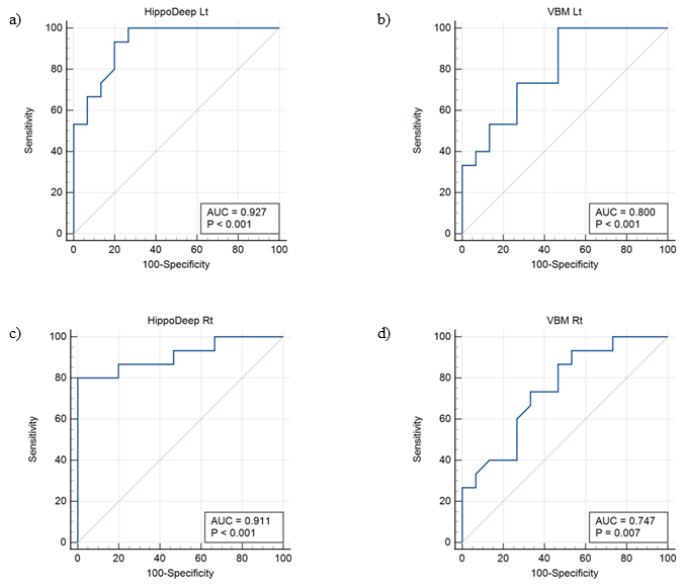

Figure 4. The ROC curve of the HippoDeep accuracy compared to the VBM model for classifying Alzheimer's disease based on the measurements of bilateral hippocampal volumes. (**a**) HippoDeep Left Hippocampus Volume, (**b**) VBM Left Hippocampus Volume, (**c**) HippoDeep Right Hippocampus Volume and (**d**) VBM Right Hippocampus Volume.

5.0 CONCLUSION

HippoDeep has a higher sensitivity and specificity than the VBM analysis method for detecting specifically left hippocampal atrophy accelerated in Alzheimer's disease

with verbal cognitive decline, enabling it to accurately classify Alzheimer's patients while consuming less time and requiring minimal training.

5.1 Limitations and future recommendation

Our sample size is insufficient to differentiate the stages of Alzheimer's disease. A larger sample size would allow us to distinguish between the various stages of Alzheimer's disease accurately. In addition, a larger sample size would yield more comprehensive data for a more accurate analysis of the progression of the disease.

Our recommendation for future research is to use VBM analysis and HippoDeep to diagnose different stages of AD (mild, moderate, and severe) to obtain accurate hippocampal volume among participants. Using VBM analysis and HippoDeep, researchers can obtain accurate hippocampal volume measurements in AD participants at various stages. This would allow for a more accurate diagnosis of the severity of AD, facilitating the development of targeted treatment strategies. In addition, incorporating a diverse range of participants across various demographic factors would increase the generalizability of the findings and enhance our understanding of the progression of AD in different populations.

Acknowledgements: We thank the Ministry of Higher Education Malaysia for providing the research grant under the Fundamental Research Grant Scheme (FRGS) with the reference code number FRGS/1/2019/SKK03/UPM/02/4 and project code 04-01-19-2119FR and project number 5540244, that was awarded to Associate Professor Dr Subapriya Suppiah and helped to support this research.

Author Contributions: SS conceptualised the study design. NSNI carried out the literature search, data extraction, and quality assessment. BI, NHMA and VPS carried out the data collection. NSNI wrote the manuscript's first draft. SS, MM, MH, and HMS edited the manuscript, verified the data, and provided critical feedback to help shape the research. RMR and NHH are clinicians who refer their patients as research subjects in this study and provide critical feedback on the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

Albert, M. S., DeKosky, S. T., Dickson, D., Dubois, B., Feldman, H. H., Fox, N. C., Gamst, A., Holtzman, D. M., Jagust, W. J., Petersen, R. C., Snyder, P. J., Carrillo, M. C., Thies, B., & Phelps, C. H. (2011). The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging–Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia, 7*(3), 270–279. https://doi.org/10.1016/j.jalz.2011.03.008

Ashburner, J., & Friston, K. J. (2000). Voxel-Based Morphometry–The Methods. *NeuroImage, 11*(6), 805–821. https://doi.org/10.1006/nimg.2000.0582

DeMarshall, C. A., Nägele, E. P., Sarkar, A., Acharya, N. K., Godsey, G., Goldwaser, E. L., Kosciuk, M., Thayasivam, U., Han, M., Belinka, B., & Nagele, R. G. (2016). Detection of Alzheimer's disease at mild cognitive impairment and disease progression using autoantibodies as blood-based biomarkers. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 3*(1), 51–62. https://doi.org/10.1016/j.dadm.2016.03.002

Dubois, B., Feldman, H. H., Jacova, C., DeKosky, S. T., Barberger-Gateau, P., Cummings, J., Delacourte, A., Galasko, D., Gauthier, S., Jicha, G., Meguro, K., O'Brien, J., Pasquier, F., Robert, P., Rossor, M., Salloway, S., Stern, Y., Visser, P. J., & Scheltens, P. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS–ADRDA criteria. *Lancet Neurology, 6*(8), 734–746. https://doi.org/10.1016/s1474-4422(07)70178-3

Ezzati, A., Katz, M. J., Zammit, A. R., Lipton, M. L., Zimmerman, M. E., Sliwinski, M. J., & Lipton, R. B. (2016). Differential association of left and right hippocampal volumes with verbal episodic and spatial memory in older adults. *Neuropsychologia, 93*, 380–385. https://doi.org/10.1016/j.neuropsychologia.2016.08.016

First, M., Williams, J., Karg, R., Spritzer, R. (2018). *Structured clinical interview for DSM-5-Research version (SCID-5-RV).* American Psychiatric Association, Arlington.

Hill, D. L. G., Schwarz, A. J., Isaac, M., Pani, L., Vamvakas, S., Hemmings, R., Carrillo, M. C., Yu, P., Sun, J., Beckett, L., Boccardi, M., Brewer, J., Brumfield, M., Cantillon, M., Cole, P. E., Fox, N., Frisoni, G. B., Jack, C., Kelleher, T., Luo, F., Novak, G., Maguire, P., Meibach, R., Patterson, P., Bain, L., Sampaio, C., Raunig, D., Soares, H., Suhy, J., Wang, H., Wolz, R., & Stephenson, D. (2014). Coalition against major diseases/European medicines agency biomarker qualification of hippocampal volume for enrichment of clinical trials in predementia stages of Alzheimer's disease. *Alzheimer's & Dementia, 10*(4), 421– 429, e3. https://doi.org/10.1016/j.jalz.2013.07.003

Honeycutt, N. A., Smith, C. D. (1995). Hippocampal volume measurements using magnetic resonance imaging in normal young adults. *Journal of Neuroimaging, 5*(2), 95–100. https://doi.org/10.1111/jon19955295

Hort, J., O'brien, J. T., Gainotti, G., Pirttila, T., Popescu, B. O., Rektorova, I., Sorbi, S., & Scheltens, P. (2010). EFNS guidelines for the diagnosis and management of Alzheimer's disease. *European Journal of Neurology, 17*(10), 1236–1248. https://doi.org/10.1111/j.1468-1331.2010.03040.x

Hu, X., Meier, M., & Pruessner, J. (2023). Challenges and opportunities of diagnostic markers of Alzheimer's disease based on structural magnetic resonance imaging. *Brain and Behavior, 13*(3), e2925, https://doi.org/10.1002/brb3.2925

Hurtz, S., Chow, N., Watson, A. E., Somme, J. H., Goukasian, N., Hwang, K., Morra, J., Elashoff, D., Gao, S., Petersen, R. C., Aisen, P. S., Thompson, P. M., & Apostolova, L. G. (2019). Automated and manual hippocampal segmentation techniques: Comparison of results, reproducibility and clinical applicability. *NeuroImage: Clinical, 21*, 101574. https://doi.org/10.1016/j.nicl.2018.10.012

Ibrahim, B., Suppiah, S., Ibrahim, N., Mohamad, M., Hassan, H. A., Nasser, N. S., & Saripan, M. I. (2021a). Diagnostic power of resting-state fMRI for detection of network connectivity in Alzheimer's disease and mild cognitive impairment: A systematic review. *Human Brain Mapping, 42*(9), 2941–2968. https://doi.org/10.1002/hbm.25369

Ibrahim, B., Suppiah, S., Piersson, A. D., Razali, R. M., Mohamad, M., Hassan, H. A., & Ibrahim, N. (2021b). Cardiovascular risk factors of Alzheimer's disease and other neurocognitive disorders in Malaysia. *Medical Journal of Malaysia, 76*(3), 291–297. https://www.e-mjm.org/2021/v76n3/alzheimers-disease.pdf

Kurth, F., Luders, E. & Gaser, C. (2015). Voxel-based morphometry. *Brain Mapping, 1*, 345-349. https://doi.org/10.1016/B978-0-12-397025-1.00304-3

Pedraza, O., Bowers, D., & Gilmore, R. L. (2004). Asymmetry of the hippocampus and amygdala in MRI volumetric measurements of normal adults. *Journal of the International Neuropsychological Society, 10*(5), 664–678. https://doi.org/10.1017/s1355617704105080

Piersson, A. D., Ibrahim, B., Suppiah, S., Mohamad, M., Hassan, H. A., Omar, N., Ibrahim, M. S., Yusoff, A. N., Ibrahim, N., Saripan, M. I., & Razali, R. M. (2021). Multiparametric MRI for the improved diagnostic accuracy of Alzheimer's disease and mild cognitive impairment: Research protocol of a case-control study design. *PLOS ONE, 16*(9), e0252883. https://doi.org/10.1371/journal.pone.0252883

Riello, M., Rusconi, E., & Treccani, B. (2021). The role of brief global cognitive tests and neuropsychological expertise in the detection and differential diagnosis of dementia. *Frontiers in Aging Neuroscience, 13*, 648310. https://doi.org/10.3389/fnagi.2021.648310

Sarica, A., Vasta, R., Novellino, F., Vaccaro, M. G., Cerasa, A., & Quattrone, A. (2018). MRI asymmetry index of hippocampal subfields increases through the continuum from the mild cognitive impairment to the Alzheimer's disease. *Frontiers in Neuroscience, 12*, 576. https://doi.org/10.3389/fnins.2018.00576

Schell, M., Foltyn-Dumitru, M., Bendszus, M., & Vollmuth, P. (2023). Automated hippocampal segmentation algorithms evaluated in stroke patients. *Scientific Reports, 13*, 11712. https://doi.org/10.1038/s41598-023-38833-z

Suzuki, Y., Tsubaki, T., Nakaya, K., Kondo, G., Takeuchi, Y., Aita, Y., Murayama, Y., Shikama, A., Masuda, Y., Suzuki, H., Kawakami, Y., Shimano, H., Arai, T., Hada, Y., & Yahagi, N. (2023). New balance capability index as a screening tool for mild cognitive impairment. *BMC Geriatrics, 23*, 74. https://doi.org/10.1186/s12877-023-03777-6

Thyreau, B., Sato, K., Fukuda, H., & Taki, Y. (2018). Segmentation of the hippocampus by transferring algorithmic knowledge for large cohort processing. *Medical Image Analysis, 43*, 214–228. https://doi.org/10.1016/j.media.2017.11.004

Wei, Y. C., Chen, C. K., Lin, C., Chen, P.Y., Hsu, P.C., Lin, C.P., Shyu, Y. C. & Huang, W. Y. (2022). Normative data of mini-mental state examination, Montreal cognitive assessment, and Alzheimer's disease assessment scale-cognitive subscale of community-dwelling older adults in Taiwan. *Dementia and Geriatric Cognitive Disorders, 51*(4), 365–376. https://doi.org/10.1159/000525615

Whitwell, J. L. (2009). Voxel-Based Morphometry: an automated technique for assessing structural changes in the brain. *The Journal of Neuroscience, 29*(31), 9661–9664. https://doi.org/10.1523/jneurosci.2160-09.2009

Wright, I. C., McGuire, P., Poline, J., Travère, J., Murray, R. M., Frith, C. D., Frackowiak, R. S. J., & Friston, K. J. (1995). A Voxel-Based method for the statistical analysis of gray and white matter density applied to schizophrenia. *NeuroImage, 2*(4), 244–252. https://doi.org/10.1006/nimg.1995.1032

Zavaliangos-Petropulu, A., Tubi, M. A., Haddad, E., Zhu, A. H., Braskie, M. N., Jahanshad, N., Thompson, P. M., & Liew, S. L. (2020). Testing a convolutional neural network-based hippocampal segmentation method in a stroke population. *Human Brain Mapping, 43*(1), 234–243. https://doi.org/10.1002/hbm.25210