


Automatic Detection of Cognitive Impairment in Patients With White Matter Hyperintensity Using Deep Learning and Radiomics

American Journal of Alzheimer's Disease & Other Dementias®
Volume 40: 1–10
© The Author(s) 2025
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/15333175251325091
journals.sagepub.com/home/aja


Junbang Feng, MM^{1,*}, Xingyan Le, MM^{1,*}, Li Li, MM^{2,*}, Lin Tang, MM³, Yuwei Xia, MM⁴, Feng Shi, MD⁴, Yi Guo, MM¹, Yueqin Zhou, MM¹, and Chuanming Li, MD² 

Abstract

White matter hyperintensity (WMH) is associated with cognitive impairment. In this study, 79 patients with WMH from hospital 1 were randomly divided into a training set (62 patients) and an internal validation set (17 patients). In addition, 29 WMH patients from hospital 2 were used as an external validation set. Cognitive status was determined based on neuropsychological assessment results. A deep learning convolutional neural network of VB-Nets was used to automatically identify and segment whole-brain subregions and WMH. The PyRadiomics package in Python was used to automatically extract radiomic features from the WMH and bilateral hippocampi. Delong tests revealed that the random forest model based on combined features had the best performance for the detection of cognitive impairment in WMH patients, with an AUC of 0.900 in the external validation set. Our results provide clinical doctors with a reliable tool for the early diagnosis of cognitive impairment in WMH patients.

Keywords

white matter hyperintensity, cognitive impairment, MRI, deep learning, radiomics

¹Medical Imaging Department, Chongqing Emergency Medical Center, Chongqing University Central Hospital, School of Medicine, Chongqing University, Chongqing, China

²Pathology Department, Chongqing University Central Hospital, Chongqing Emergency Medical Center, Chongqing, China

³Department of Radiology, Chongqing University Cancer Hospital, School of Medicine, Chongqing University, Chongqing, China

⁴Department of Research and Development, Shanghai United Imaging Intelligence, Co., Ltd. Shanghai, China

*These authors have contributed equally to this work.

Corresponding Authors:

Chuanming Li, MD, Medical Imaging Department, Chongqing Emergency Medical Center, Chongqing University Central Hospital, School of Medicine, Chongqing University. No. 1, Jiankang Road, Yuzhong District, Chongqing 400014, China.

Email: licm@cqu.edu.cn

Yueqin Zhou, MM, Medical Imaging Department, Chongqing Emergency Medical Center, Chongqing University Central Hospital, School of Medicine, Chongqing University. No. 1, Jiankang Road, Yuzhong District, Chongqing 400014, China.

Email: 308902549@qq.com

Yi Guo, MM, Medical Imaging Department, Chongqing Emergency Medical Center, Chongqing University Central Hospital, School of Medicine, Chongqing University. No. 1, Jiankang Road, Yuzhong District, Chongqing 400014, China.

Email: yiguo_0909@sina.com

Data Availability Statement included at the end of the article



Creative Commons Non Commercial No Derivs CC BY-NC-ND: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits non-commercial use, reproduction and distribution of the work as published without adaptation or alteration, without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (<https://us.sagepub.com/en-us/nam/open-access-at-sage>).

Significance Statements

- Our VB-Net could segment the WMH and whole-brain subregions automatically and accurately.
- RF models based on radiomic features of WMH and bilateral hippocampi performed excellently in detecting cognitive impairments.
- Age, education level and Hachinski score were risk factors for cognitive impairment in patients with WMH.

Introduction

White matter hyperintensities (WMH) are common in the elderly population, and the WMH burden typically increases with age.¹ According to previous reports, more than 80% of people aged 60 to 70 years suffered from WMH, and the prevalence of WMH among people aged 80 to 90 years could reach 100%.² Many studies had shown that WMH was a risk factor for cognitive impairment.³ WMH could increase the risk of all-cause dementia by 14% and accelerate the transformation of cognitive impairment into dementia.⁴ Interventions in the early stages of cognitive impairment could help delay the onset of irreversible dementia. However, at present, the diagnosis of cognitive impairment in WMH patients still relies on clinical symptoms and neuropsychological tests, which are highly subjective, time-consuming, and easily influenced by patient cooperation.⁵ Therefore, investigating new technologies for the rapid and accurate identification of cognitive impairment in WMH patients is highly important to help patients receive targeted treatment in time, improve their prognosis and improve quality of life.

Magnetic resonance imaging (MRI) can be used to evaluate the brain structure and function of patients with WMH in a noninvasive, radiation-free, and high-resolution manner.⁶ Radiomics, developed by Dutch professor Philippe Lambin in 2012,⁷ enables the quantitative analysis of physiological and pathological changes in lesions by collecting many invisible image features through high throughput from raw MRI data. Previous studies had shown that hippocampus-based radiomics could be used for imaging-based diagnosis of mild cognitive impairment (MCI) and Alzheimer's disease (AD).⁸ Hanseeuw BJ et al reported reduced hippocampal volume (HV), cortical metabolism and thickness in patients with MCI.⁹ The internal texture features of WMH had also been found to be closely related to the occurrence of cognitive impairment.¹⁰ However, the diagnosis of cognitive impairment in patients with WMH by combining analyses of WMH and the hippocampus in MR images had not been reported. At present, the segmentation of brain subregions and WMH relies mostly on manual or semiautomatic delineation, which is very time-consuming and subjective. In this study, patients with WMH with Fazekas scores of 2 or 3 were included and divided into a group with cognitive impairment and a group without cognitive impairment based on their cognitive status. A convolutional network was applied to automatically perform

whole-brain subregion and WMH segmentation. Radiomic features of WMH and bilateral hippocampi were extracted and trained through machine learning methods to develop rapid and accurate models to diagnose cognitive impairment in patients with WMH.

Methods

Participants

A total of 122 patients with WMH admitted to the Second Affiliated Hospital of Chongqing Medical University (Hospital 1) between September 2019 and June 2023 were retrospectively included. WMH was evaluated using the modified Fazekas scale on T₂ fluid attenuated inversion recovery (T2-FLAIR) sequences (0 = non-lesion, 1 = punctate lesion, 2 = lamellar fusion, and 3 = extensive fusion). The Fazekas score for each patient was determined by 2 radiologists with more than 10 years of experience who were blinded to the patients' clinical information and cognitive status. The inclusion criteria were as follows: (1) aged 45-80 years and (2) Fazekas grade 2 or 3. The exclusion criteria were as follows: (1) acute intracranial macrovascular diseases such as ischaemic stroke and cerebral haemorrhage; (2) metabolic encephalopathy, ischaemic hypoxic encephalopathy, or other nonvascular white matter lesions; (3) other coexisting intracranial lesions such as tumours, dementia, craniocerebral trauma, or other diseases; (4) incomplete clinical data; (5) incomplete imaging data; and (6) severe imaging artefacts. All patients completed a formal neuropsychological assessment including the following tests: the Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), Clinical Dementia Rating (CDR), Geriatric Depression Scale (GDS), Activities of Daily Living (ADL) and Hachinski Ischaemia Index scale (HIS). According to the cognitive status based on neuropsychological test results, all patients were subdivided into 2 groups: WMH, with cognitive impairment, and WMH, without cognitive impairment. Finally, 42 patients with cognitive impairment and 37 patients without cognitive impairment from hospital 1 were included in this study. They were randomly divided into a training set (62 patients) and an internal validation set (17 patients) at a ratio of 8:2. In addition, we collected MRI data from 29 patients, including 14 patients with cognitive impairment and 15 patients without cognitive impairment, from Chongqing University Central Hospital (Hospital 2) between February 2020 and July 2023 as an external validation set. The inclusion and exclusion criteria were the same as those described above. The flow chart of subject recruitment was shown in [Figure 1](#).

Clinical and Laboratory Characteristics

Clinical and laboratory characteristics, including sex, age, smoking status, alcohol abuse status, diabetes status, hypertension status, coronary heart disease status, and body mass

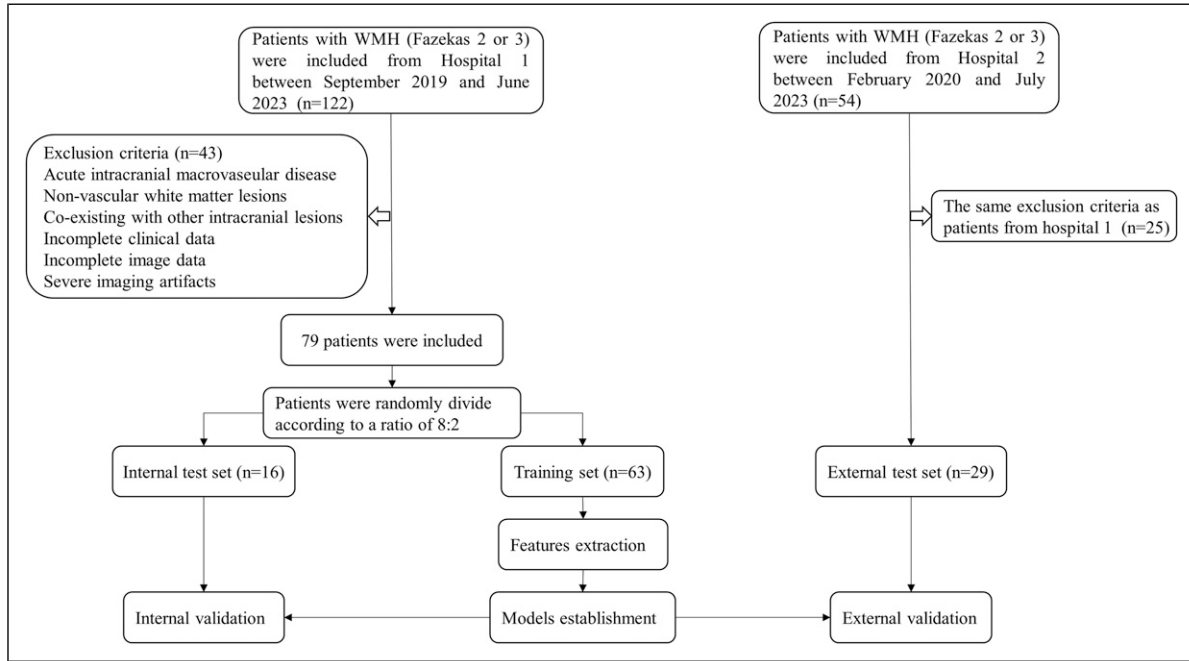


Figure 1. Flow chart of subject recruitment.

index (BMI), were directly obtained via standardized forms. The laboratory characteristics included fasting blood glucose (GLU), glycosylated haemoglobin (HbA1c), total cholesterol (TC), low density lipoprotein (LDL), high density lipoprotein (HDL), free fatty acid (FFA), triglyceride (TG), thrombin time (TT), prothrombin time (PT), activated partial thrombin time (APTT), fibrinogen (FIB) and D-dimer.

MRI Scanning

MR images of all the subjects treated at Hospital 1 were obtained using a 3.0T MRI scanner (Achieva 3.0T scanner, Phillips, Netherlands). The scanning sequence included T₁-weighted imaging (T₁WI), T2-FLAIR, and diffusion-weighted imaging (DWI). The parameters of the T₁WI sequence were as follows: repetition time (TR) = 7.9 ms, echo time (TE) = 39 ms, field of view (FOV) = 256 × 256, layer thickness = 2 mm, and layer spacing = 1 mm. The parameters of the T2-FLAIR sequence were TR = 4800 ms, TE = 279 ms, FOV = 256 × 256, layer thickness = 1.6 mm, and layer spacing = 0.8 mm. The parameters of the DWI sequence were as follows: TR = 6274 msec, TE = 68 msec, FOV = 220 × 220, matrix = 88 × 88, thickness = 2.5 mm, interslice gap = 30 mm. MRI data from Hospital 2 were obtained using a 1.5T MRI scanner (uMR560 1.5T, United Imaging Limited, China). The parameters of the T₁WI sequence were as follows: TR = 10.4 ms, TE = 4.4 ms, FOV = 256 × 240, layer thickness = 1.0 mm, and layer spacing = 0 mm. The parameters of the T2-FLAIR sequence were TR = 6500 ms, TE = 464.5 ms, FOV = 256 × 240, layer thickness = 1 mm, and layer spacing = 0 mm. The parameters of the DWI sequence were TR = 3000 msec,

TE = 97.6 msec, FOV = 230 × 230, matrix = 128 × 100, thickness = 6 mm, and interslice gap = 30 mm.

Brain Subregion Segmentation and Volume Extraction

The preprocessing steps included skull stripping, bias correction, and image resampling to 1 mm isotropic resolution. A 3D VB-Net was trained for brain subregion segmentation,¹¹ which takes each sample T1 image as input and outputs a corresponding brain map label. By adjusting the network parameters based on the difference between the actual brain partition and the output brain partition, continuous training was performed until the network converged, ensuring that the output label image closely matched the corresponding partition image of the sample. Throughout the training process, we adopted a cascade strategy of layer-by-layer segmentation to more accurately capture the complexity and difficulty of the brain segmentation problem. Additionally, the segmentation performance of the network was enhanced by providing additional information to the lower network, progressively achieving the fine division of brain regions, midbrain regions and brain structures. This model was trained on 1800 T1WI images from the publicly available Consortium for Reliability and Reproducibility (CoRR) dataset and the Chinese Brain Molecular and Functional Mapping (CBMFM) project,^{12,13} with an average overlap rate of 0.92 between the dice and the fundamental truth. The entire brain was automatically divided into 109 subregions according to the DK atlas (Supplemental material 1),¹⁴ including 20 subregions in the frontal lobe, 22 subregions in the temporal lobe, 12 subregions in the parietal lobe, 12 subcortical nuclei, white matter structures and

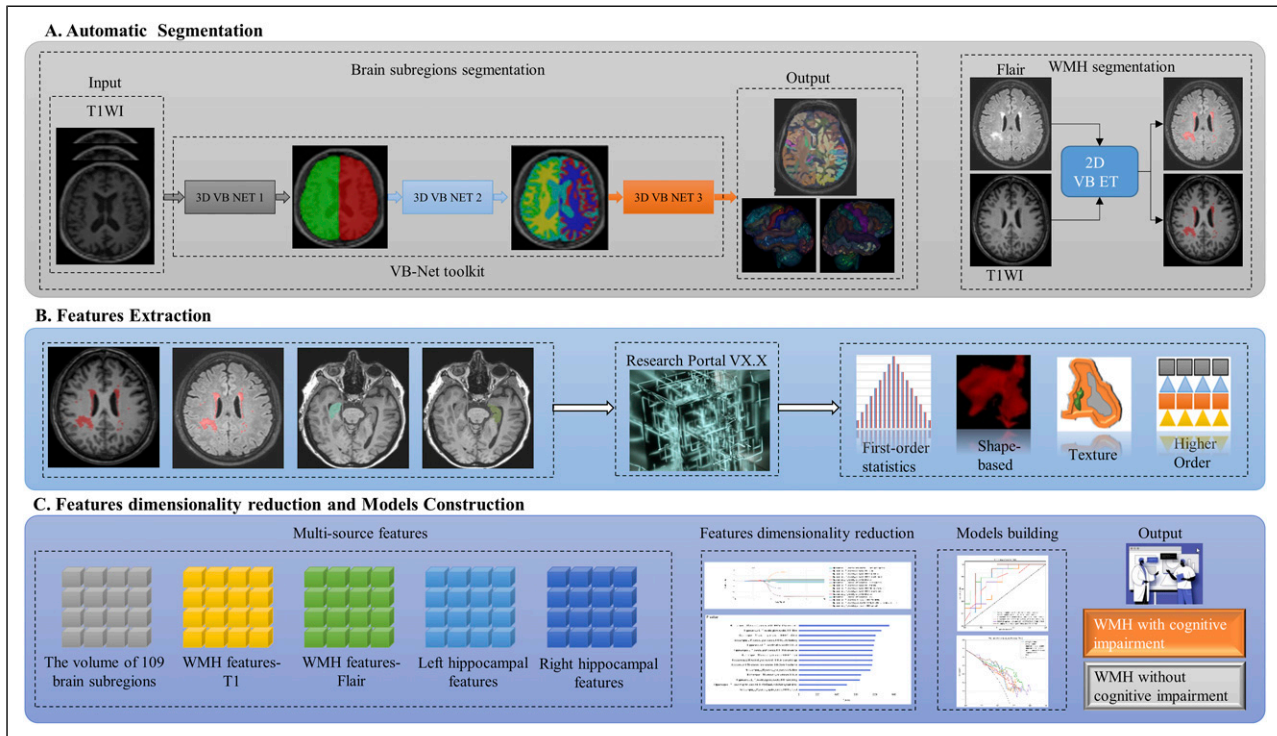


Figure 2. Flow chart of WMH and brain subregion segmentation (A), feature extraction from WMH, bilateral hippocampi (B), and feature dimension reduction and model construction (C).

other structures. The segmentation process took less than half a minute for each patient. Then, the volume of 109 brain regions was extracted automatically.

WMH Segmentation

The WMH on FLAIR images was automatically segmented using a 2D VB-Net, which incorporated the advantages of an efficient encoder-decoder framework for feature embedding, residual connections for information flow, and bottleneck layers for model compression.¹⁵ This 2D VB-Net was trained on a large MR dataset of 1045 WMH subjects.¹⁶ To handle differences in pixel spacing between scanners, the image and labelled areas were resampled evenly in the x and y directions with a spacing of 0.5. The pixel spacing in the z direction remained unchanged. The input two-dimensional block for training was 256×256 , which was randomly sampled from the image body. The kaim initialization strategy was adopted for weight initialization, and the optimizer was adaptive moment estimation, with the first moment coefficient = 0.9, the second moment coefficient = 0.999, the learning rate = 0.001, and the mini-batch size = 48. The number of bottleneck structures of the high levels was reduced. The vertical depth was the level, with the original input and output depth as level 1; after downsampling 4 times as level 5, a lower level indicated a greater spatial resolution but fewer feature maps and vice versa. Because the number of segmentation

categories is relatively small, reducing the bottleneck structure would not impact the segmentation accuracy but would help reduce the model parameters and hence improve the robustness. The output block was adjusted, and 2 convolutional layers were added to the original basis to generate the segmentation probability map. The WMH segmentation model was directly applied to the FLAIR images to identify the sites of WMH ROIs. T1WI images were registered to FLAIR using the medical image registration software Advanced Normalization Tools (ATNs) with affine registration.

WMH and Hippocampal Feature Extraction

Using the PyRadiomics package in Python (version 2.1.2, <https://pyradiomics.readthedocs.io/>), 824 radiomics features were automatically extracted from the WMH regions in the FLAIR/T1WI sequences and the bilateral hippocampi in the T1WI sequence of each patient. These features included first-order statistics and texture features derived from the original images and wavelet transformations of the original images. Eight types of wavelet features were obtained and labelled LLL, LLH, LHL, LHH, HLL, HLH, HHL, and HHH. Additionally, 14 shape features were extracted from the original images. Finally, z score normalization was used to standardize all the radiomic features, and the reproducibility of the features was assessed using a pipeline that met the recommendations of the Image Biomarker Standardization Initiative.

Table 1. Patients' Characteristics in the Training Set, Testing Set and External Testing Set.

| Variables | Training and internal testing set | | External testing set | |
|------------------------|-----------------------------------|----------------------------------|-------------------------------|----------------------------------|
| | WMH-with cognitive impairment | WMH-without cognitive impairment | WMH-with cognitive impairment | WMH-without cognitive impairment |
| Sex | | | | |
| Female | 34 | 25 | 5 | 6 |
| Male | 8 | 12 | 10 | 8 |
| Age(range) | 74 (49-84) | 67 (47-80)* | 69.6 (53-79) | 69.2 (54-79) |
| Education | | * | | |
| Illiteracy | 4 | 0 | 0 | 1 |
| Primary | 7 | 5 | 5 | 3 |
| Middle | 25 | 21 | 8 | 8 |
| University | 6 | 11 | 2 | 2 |
| Smoking | | | | |
| 0 | 38 | 30 | 10 | 10 |
| 1 | 4 | 7 | 5 | 4 |
| Alcohol abuse | | | | |
| 0 | 38 | 32 | 14 | 10 |
| 1 | 4 | 5 | 1 | 4 |
| Diabetes | | | | |
| 0 | 29 | 29 | 11 | 10 |
| 1 | 13 | 8 | 4 | 4 |
| Hypertension | | | | |
| 0 | 7 | 10 | 4 | 5 |
| 1 | 35 | 27 | 11 | 9 |
| Coronary heart disease | | * | | |
| 0 | 26 | 37 | 10 | 9 |
| 1 | 16 | 0 | 5 | 5 |
| Hachinski | 3(0-14) | 2(0-5)* | 4.6(1-8) | 3.8(0-8) |
| MMSE | 26(18-30) | 27(0-30) | 26(18-30) | 24(18-30) |
| MoCA | 21(6-27) | 26(13-30) | 16(10-21) | 25(25-28) |
| GDS | 0(0-10) | 1(0-10) | 1(0-3) | 1(0-6) |
| ADL | 20(19-28) | 20(19-21) | 22(21-28) | 23(23-28) |
| GLU | 6.26(4.6-10.6) | 6.43(4.4-14.31)* | 7.78(5.2-11.55) | 9.33(5.02-15.36) |
| HbA1c | 6.1(4.5-8.1) | 6.29(5.3-9.1)* | 6.74(5.2-8.7) | 7.7(6.6-9.7) |
| TC | 4.585(1.75-6.47) | 4.55(2.27-7.01)* | 4.92(2.96-7.87) | 4.11(2.43-5.77) |
| LDL-C | 2.4(0.75-4.42) | 2.4(1.09-4.2)* | 2.49(1.44-4.18) | 2.14(0.92-3.44) |
| HDL-C | 1.23(0.63-5.6) | 1.13(0.6-2)* | 1.36(0.84-1.96) | 1.15(0.76-1.53) |
| FFA | 0.44(0.13-1.14) | 0.45(0.06-1.04)* | 0.89(1.71-1.07) | 1.25(0.49-2.67) |
| TG | 1.53(0.64-3.39) | 1.55(0.36-3.11)* | 2.61(0.62-13.04) | 1.47(0.69-2.86) |
| PT | 12.85(11.4-14) | 12.9(11.5-22.4)* | 12.89(11.6-13.9) | 11.29(0.94-15.4) |
| APTT | 35.62(28.6-42.6) | 35.5(30.8-51.2)* | 33.95(28.7-52.1) | 35.37(29.4-57.2) |
| TT | 16.55(1.04-20.1) | 16.8(15-22.4) | 17.07(15.5-18.4) | 16.55(13.5-19.2) |
| FIB | 3.295(1.85-5.57) | 3.4(2.07-4.87)* | 3.21(2.34-4.15) | 3.08(2.37-4.38) |
| D-dimer | 168.9(40-1936.4) | 163.9(40-397.1)* | 58(18-217) | 68(15-517) |

Note. *P* value was calculated from two-sample *t* test for continuous variables and from Chi-squared test for discrete variables. *means *P* values less than 0.05. MMSE, mini-mental state examination; MoCA, montreal cognitive assessment; GDS, geriatric depression scale; ADL, activities of daily living; GLU, Glucose; TC, total cholesterol; LDL-C, low density lipoprotein cholesterol; HDL-C, high density lipoprotein cholesterol ; FFA, free fatty acids; TG, triglyceride; PT, prothrombin time; APTT, activated partial thromboplastin time; TT, thrombin time; FIB, fibrinogen.

Table 2. Features Retained After Dimensionality Reduction by LASSO.

| ROI | Features |
|--------------------------------|--|
| The volume of brain subregions | Parietal_Inf_R_Volume, Temporal_Sup_Banks_R_Volume |
| WMH_Flair | WMH_Flair_original_glcmlnverseVariance, WMH_Flair_wavelet_firstorder_wavelet-HLL-Kurtosis, WMH_Flair_wavelet_glcmlwavelet-HLH-Contrast, WMH_Flair_wavelet_glcmlwavelet-HHL-DifferenceVariance, WMH_Flair_wavelet_glcmlwavelet-HHL-Id |
| WMH_T1 | WMH_T1_wavelet_glcmlwavelet-HHL-Contrast, WMH_T1_wavelet_glcmlwavelet-HHL-InverseVariance, WMH_T1_wavelet_gldm_wavelet-LLH-DependenceNonUniformityNormalized |
| Right_hippocampus | Hippocampus_R_T1_wavelet_firstorder_wavelet-LHH-Range, Hippocampus_R_T1_wavelet_glcmlwavelet-LLH-Imc2, Hippocampus_R_T1_wavelet_glcmlwavelet-HHL-Correlation, Hippocampus_R_T1_wavelet_glcmlwavelet-HHH-Autocorrelation, Hippocampus_R_T1_wavelet_glcmlwavelet-HHH-Idmn, Hippocampus_R_T1_wavelet_ngtdm_wavelet-HHH-Complexity |
| Left_hippocampus | Hippocampus_L_T1_wavelet_glcmlwavelet-LHL-Idmn, Hippocampus_L_T1_wavelet_glszm_wavelet-LHH-SizeZoneNonUniformityNormalized |

Feature Dimensionality Reduction, Model Building, and Evaluation

The volumes of 109 brain subregions, radiomics features of WMH-T1WI, WMH-FLAIR and each hippocampus were dimensionality reduced and modelled, and all the features were then fused to establish a combined model. The dimensionality reduction process included Max-Relevance and Min-Redundancy (MRMR) and the least absolute shrinkage and selection operator (LASSO) methods. Finally, the random forest (RF) classifier was trained on the training set, adjusted on the internal validation set and evaluated on the independent set.¹⁷ The flow chart of our study is shown in [Figure 2](#).

Statistical Analysis

Statistical analyses were performed using the R language (version 4.0.4) and Python software using the scikit-learn library. The Mann–Whitney U test or Student's t test was used to test the normal distribution of continuous variables. The chi-square test was used to compare categorical variables. A two-tailed *P* value <.05 indicated statistical significance. The performance of the classifier model on the test subset was assessed by the average accuracy, sensitivity/recall, specificity, precision and 95% confidence intervals (CIs) based on the case probability cut-off of 0.5, as well as the F score indicator and the area under the curve (AUC). The Delong test was used to compare the performance of each model.

Results

A total of 79 patients from Hospital 1 and 29 patients from Hospital 2, whose demographic and clinical characteristics were summarized in [Table 1](#) were included in this study. In the training set, sex, smoking status, alcohol abuse status, diabetes status, hypertension status, or free fatty acid status did not significantly differ between WMH patients with cognitive

impairment and those without cognitive impairment (*P* > 0.05). Age, education level, coronary heart disease, HIS, fasting blood glucose, glycated haemoglobin, total cholesterol, low-density lipoprotein, high-density lipoprotein, free fatty acid, triglyceride, homocysteine, prothrombin time, activated partial thrombin time, thrombin time, fibrinogen and D-dimer significantly differed between the 2 groups (*P* < 0.05). Multivariate Cox regression analysis revealed that age, education level and Hachinski status were risk factors for cognitive impairment (*P* < 0.1).

Our 3D VB-Net realized automatic segmentation and volume measurement of 109 subregions of the whole brain. Our 2D VB-Net algorithm demonstrated good performance (Dice = 0.789, lesion F1 = 0.764) in WMH segmentation. After dimensionality reduction, 2 brain subregion volume features, 3 WMH-T₁WI radiomics features, 5 WMH-FLAIR radiomics features, 3 left hippocampal radiomics features and 6 right hippocampal radiomics features were retained ([Table 2](#)). The Delong test revealed that the RF model based on combined features (radiomic features of WMH and bilateral hippocampi and volumes of the right parietal and temporal regions) had the best performance. The areas under the curve (AUCs) of the training set, internal validation set and external validation set were 0.976 (95% CI: 0.940-1.000), 0.937 (95% CI: 0.789-1.000), and 0.900 (95% CI: 0.755-0.900), respectively. ([Figure 3](#)). The calibration curves showed good consistency and stability between the model predictions and real results ([Figure 4](#)). The clinical decision curve revealed a significant positive effect of the RF models, with a wide range of patient benefit probabilities and practical value ([Figure 5](#)).

Discussion

Many studies had shown that WMH could cause cognitive impairment and accelerate the transition from mild cognitive impairment to dementia.^{18,19} Once a patient is diagnosed with

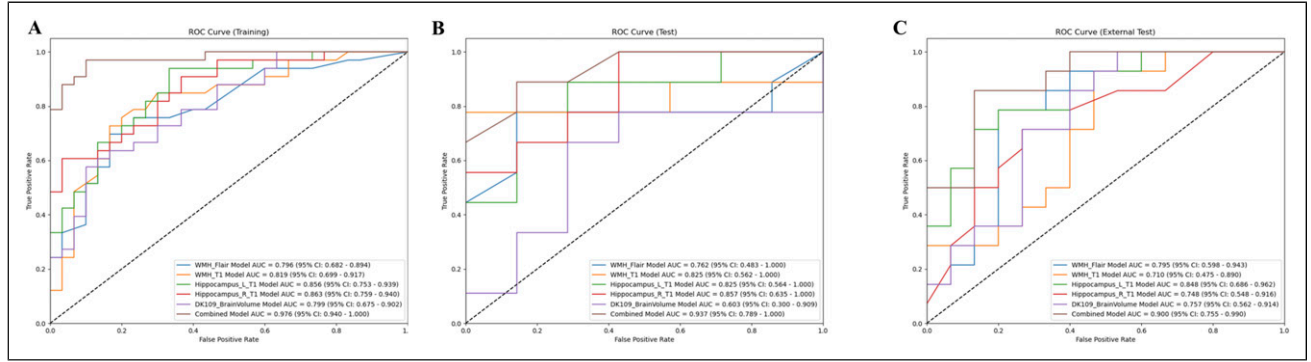


Figure 3. Receiver operating characteristic (ROC) curves for diagnostic models in the training set (A), testing set (B) and external testing set (C).

irreversible dementia,²⁰ families and societies face a heavy burden in both developed and developing countries.^{21,22} Therefore, early and accurate identification of cognitive impairment in WMH patients has important clinical significance. At present, cognitive impairment in patients with WMH is mainly diagnosed through clinical symptoms and neuropsychological assessments, which are highly subjective and time-consuming, and their clinical applications are greatly limited. In this study, we used deep learning to automatically identify and segment brain subregions and WMH and established models using radiomics and machine learning to achieve automatic diagnosis. Our results indicated that the RF model based on radiomic features of WMH and bilateral hippocampi had excellent accuracy, with AUC values of 0.976, 0.937, and 0.900 in the training set, internal validation set, and external validation set, respectively. Our method, for the first time automatically detected cognitive impairment in patients with WMH without the need for human intervention and could be completed within 3 minutes. This method was objective and accurate and had important clinical application significance. Physicians only needed to import the MR images of patients to quickly and automatically obtain accurate diagnostic results. This strategy greatly improved the efficiency

of clinical practice. Furthermore, external independent validation was utilized to ensure the reliability and universality of the results. In the past, Chu T et al used the volume and texture features of the hippocampus based on T1-weighted MR images to establish a support vector machine model to diagnose MCI, with an AUC of 0.90. Their method needed more than 4 hours to process a patient, making it a time-consuming strategy.²³ Liu M et al proposed a multimodel deep learning framework for automatic hippocampal segmentation and MCI diagnosis using structural MRI data, with an accuracy of 76.2%.²⁴ Leandrou S et al extracted radiomic features from the entorhinal cortex and hippocampus to distinguish NC, MCI and AD. The results revealed that the F1 scores of XGBoost for the discrimination of NC vs AD, MC vs MCI, and MCI vs AD were 0.949, 0.818 and 0.810, respectively.²⁵ Our research was the first to diagnose cognitive impairment in WMH patients using artificial intelligence.

MRI can be used to quantitatively evaluate changes in brain structure and is popular for WMH imaging in clinical practice. Previously, Maillard et al²⁶ suggested that early detection of microstructural damage in the brain needed advanced MRI techniques, such as diffusion tensor imaging (DTI). Our method was based on basic T1WI and FLAIR sequences,

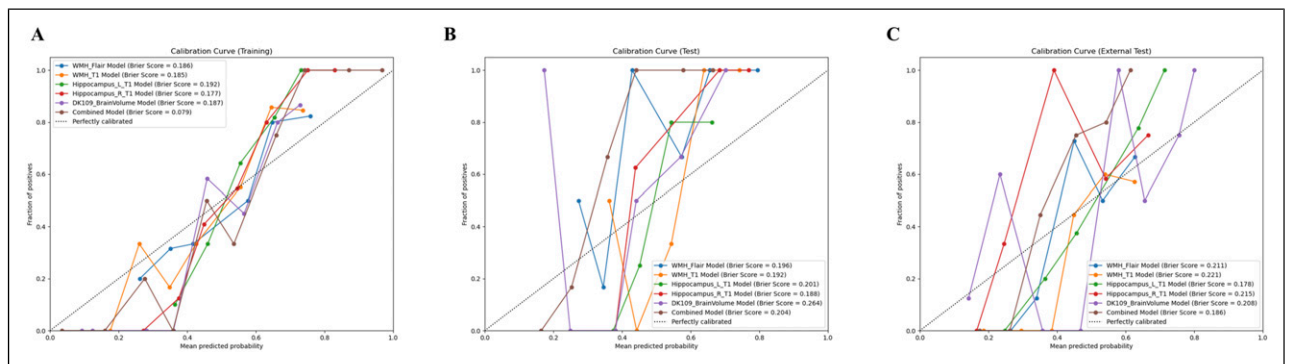


Figure 4. Calibration curves for nomogram goodness of fit in the training set (A), testing set (B) and external testing set (C). The 45° line indicated that the probability predicted by the model matches the actual probability. The closer the distance between the 2 curves was, the higher the accuracy.

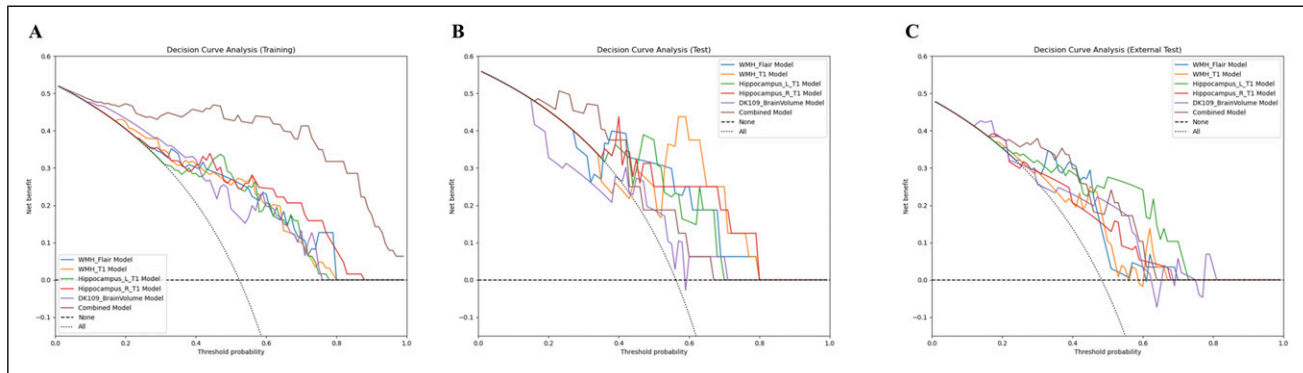


Figure 5. Decision curves of different models in the training set (A), testing set (B) and external testing set (C). The Y-axis represented the net benefit, and the X-axis represented the threshold probability. The RF model had a greater overall net gain in detecting cognitive impairment in patients with WMH.

which were routine scanning sequences used in clinical practice and did not need advanced or expensive additional imaging sequences, thereby reducing the economic burden on patients. Most previous studies on WMH were based on 2D FLAIR sequences.²⁷ In this study, thin-layer 3D FLAIR sequences were used for the first time, which ensured that the results were reliable. Obtaining many 3D images would require at least 30 minutes when using traditional manual drawing methods to segment WMH and hippocampi layer by layer, greatly reducing the efficiency of clinical diagnosis. On the other hand, long-term manual sketching could easily lead to visual fatigue and operational errors. The 3D-VB network we developed was based on U-Net, which could automatically segment WMH and brain subregions. It achieved a Dice consistency of 0.878 with manual segmentation by 2 observers, indicating high accuracy.

In this study, after dimensionality reduction, 19 features were considered to have important diagnostic value and were used for model building, including 2 first-order features and 14 texture features from WMH and the hippocampus²⁸ and 2 volume features from the temporal and parietal lobes. Previous studies had also revealed that radiomic features of first-order features and textural features played important roles in the diagnosis of other diseases, such as dementia and Parkinson's disease. The first-order features describe the distribution of voxel intensity in the image area through basic metrics, thereby reflecting changes in WMH and the internal structure of the hippocampus. The remaining 14 texture features were obtained by calculating the statistical correlation between adjacent voxels, providing a measure of voxel intensity spatial arrangement. The grey level co-occurrence matrix (GLCM) quantifies the incidence of voxels with the same intensity in a fixed direction within a predetermined distance.²⁹ The grey level dependency matrix (GLDM) was used to quantify the number of connected voxels that depended on centrosomes.³⁰ The grayscale size region matrix (GLSZM) was defined as the number of connections that shared the same grayscale intensity. When the microstructure

of the WMH or hippocampus changes, the intensity and continuity of voxels change and are reflected by these texture features. GLSZM and GLCM can reflect mild pathological changes, including increased water content and decreased myelin sheath content, which made white matter fibres rough and blurry. In addition, we also reported that 2 volume features of the temporal and parietal lobes were preserved.³¹ Although changes in the hippocampus and temporal lobe related to cognitive impairment had been well documented in studies, changes in the parietal lobe had not been fully explored. Verfaillie SC et al reported that the parietal cortex was thinner in patients with subjective cognitive decline (SCD) who progressed to non-AD dementia than in patients with stable SCD.³² Delvenne JF et al reported that white matter fibres projecting from the posterior region of the corpus callosum to the parietal lobe region were damaged in MCI and AD patients, as reflected by the relevant diffusion parameters of DTI.³³

This study also identified several clinical risk factors, such as age, education level, and Hachinski scores, for cognitive impairment in patients with WMH. Many previous studies had shown that WMH burden could significantly increase with age and that old age was a key risk factor for cognitive impairment. Educational level might be a protective factor because the level of education inversely correlates with the likelihood of cognitive impairment. Harkness et al³⁴ noted that good educational experience and rich intellectual activities strengthen an individual's ability to resist cognitive impairment. In addition, this study revealed for the first time that the Hachinski score was also a risk factor for cognitive impairment in patients with WMH. The Hachinski scale is a common clinical tool used to identify vascular dementia.³⁵ Our finding reflected that the impairment of cerebral blood vessels were also important factors of cognitive impairment.

This study had several limitations. First, due to the retrospective design and small sample size, selection biases might occurred. Second, all patients in the sample were Asian. Therefore, multiethnic populations from different countries

needed to be examined to further validate the model. Finally, to achieve automatic and rapid diagnosis, this study did not include DTI sequences. Future research should develop automatic analysis methods for DTI images and comprehensively use multimodal MRI data to further improve diagnostic accuracy.

Conclusion

In this study, we established an automated and rapid method to identify cognitive impairment in WMH patients. Deep learning of VB-Nets was used to automatically segment white matter hyperintensities and brain subregions. Machine learning and radiomics were used to accurately diagnose cognitive impairments. The results of our study provide doctors with a reliable tool for the early diagnosis of cognitive impairment in practice.

Acknowledgments

The authors thank the participants and referring technicians for their selfless and valuable assistance in this study.

Author Contributions

All the authors contributed to the study conception and design. Junbang Feng and Xingyan Le: Provision of study materials or patients. Junbang Feng, Xingyan Le, Li Li and Lin Tang: Collection and assembly of data. Yuwei Xia, Feng Shi: Data analysis and interpretation. All authors: Manuscript writing. All authors: Final approval of the manuscript.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the Fundamental Research Funds for the Central Universities of China (Project NO. 2022CDJYGRH-004); the Natural Science Foundation Project of Chongqing (CSTB2024NSQ-MSX1265); the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJQN202400117); and the Chongqing medical scientific research project (Joint project of Chongqing Health Commission and Science and Technology Bureau) (2022QNXM013).

Ethical Statement

Ethical Approval

The study was conducted in accordance with the World Medical Association Declaration of Helsinki and was approved by the Institutional Review Board of Chongqing University Central Hospital (2022-30).

Informed Consent

All participants agreed to participate in this study and sign the informed consent. All participants agreed to the publication of research data.

ORCID iD

Chuanming Li  <https://orcid.org/0000-0002-4006-9411>

Data Availability Statement

The dataset used and analysed in this study is available from the corresponding author upon request.

Supplemental Material

Supplemental material for this article is available online.

References

1. Sassi C. White matter hyperintensities and neurodegenerative dementias. *Aging*. 2019;11(10):2912-2913.
2. Wardlaw JM. Prevalence of cerebral white matter lesions in elderly people: a population based magnetic resonance imaging study: the Rotterdam Scan Study. *J Neurol Neurosurg Psychiatry*. 2001;70(1):2-3.
3. d'Arbeloff T, Elliott ML, Knodt AR, et al. White matter hyperintensities are common in midlife and already associated with cognitive decline. *Brain Commun*. 2019;1(1):fcz041.
4. Hu HY, Ou YN, Shen XN, et al. White matter hyperintensities and risks of cognitive impairment and dementia: a systematic review and meta-analysis of 36 prospective studies. *Neurosci Biobehav Rev*. 2021;120:16-27.
5. Kalaria RN. Neuropathological diagnosis of vascular cognitive impairment and vascular dementia with implications for Alzheimer's disease. *Acta Neuropathol*. 2016;131(5):659-685.
6. Pipe JG. High-value MRI. *J Magn Reson Imaging*. 2019;49(7):e12-e13.
7. Lambin P, Rios-Velazquez E, Leijenaar R, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur J Cancer*. 2012;48(4):441-446.
8. Park YW, Choi D, Park M, et al. Predicting amyloid pathology in mild cognitive impairment using radiomics analysis of magnetic resonance imaging. *J Alzheimers Dis*. 2021;79(2):483-491.
9. Hanseeuw BJ, Jacobs HIL, Schultz AP, et al. Association of pathologic and volumetric biomarker changes with cognitive decline in clinically normal adults. *Neurology*. 2023;101(24):e2533-e2544.
10. Feng Q, Ding Z. MRI radiomics classification and prediction in Alzheimer's disease and mild cognitive impairment: a review. *Curr Alzheimer Res*. 2020;17(3):297-309.
11. Shi F, Hu W, Wu J, et al. Deep learning empowered volume delineation of whole-body organs-at-risk for accelerated radiotherapy. *Nat Commun*. 2022;13(1):6566.

12. Zuo XN, Anderson JS, Bellec P, et al. An open science resource for establishing reliability and reproducibility in functional connectomics. *Sci Data*. 2014;1:140049.
13. Gu D, Shi F, Hua R, et al. An artificial-intelligence-based age-specific template construction framework for brain structural analysis using magnetic resonance images. *Hum Brain Mapp*. 2023;44(3):861-875.
14. Desikan RS, Ségonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*. 2006;31(3):968-980.
15. Zhu W, Huang H, Zhou Y, et al. Automatic segmentation of white matter hyperintensities in routine clinical brain MRI by 2D VB-Net: a large-scale study. *Front Aging Neurosci*. 2022;14:915009.
16. Feng J, Hui D, Zheng Q, et al. Automatic detection of cognitive impairment in patients with white matter hyperintensity and causal analysis of related factors using artificial intelligence of MRI. *Comput Biol Med*. 2024;178:108684.
17. Hu J, Szymczak S. A review on longitudinal data analysis with random forest. *Brief Bioinform*. 2023;24(2):bbad002.
18. Tubi MA, Feingold FW, Kothapalli D, et al. White matter hyperintensities and their relationship to cognition: effects of segmentation algorithm. *Neuroimage*. 2020;206:116327.
19. Li TR, Li BL, Xu XR, et al. Association of white matter hyperintensities with cognitive decline and neurodegeneration. *Front Aging Neurosci*. 2024;16:1412735.
20. Gale SA, Acar D, Daffner KR. Dementia. *Am J Med*. 2018;131(10):1161-1169.
21. Prince M. Dementia in developing countries. A consensus statement from the 10/66 Dementia Research Group. *Int J Geriatr Psychiatry*. 2000;15(1):14-20.
22. Wehling M, Groth H. Challenges of longevity in developed countries: vascular prevention of dementia as an immediate clue to tackle an upcoming medical, social and economic stretch. *Neurodegener Dis*. 2011;8(5):275-282.
23. Chu T, Liu Y, Gui B, et al. Hippocampal subregions volume and texture for the diagnosis of mild cognitive impairment. *Exp Aging Res*. 2024;51:125-136.
24. Liu M, Li F, Yan H, et al. A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. *Neuroimage*. 2020;208:116459.
25. Leandrou S, Lamnisis D, Bougias H, et al. A cross-sectional study of explainable machine learning in Alzheimer's disease: diagnostic classification using MR radiomic features. *Front Aging Neurosci*. 2023;15:1149871.
26. Maillard P, Fletcher E, Harvey D, et al. White matter hyperintensity penumbra. *Stroke*. 2011;42(7):1917-1922.
27. Lee S, Rieu Z, Kim RE, et al. Automatic segmentation of white matter hyperintensities in T2-FLAIR with AQUA: a comparative validation study against conventional methods. *Brain Res Bull*. 2023;205:110825.
28. Guan Y, Li W, Jiang Z, et al. Value of whole-lesion apparent diffusion coefficient (ADC) first-order statistics and texture features in clinical staging of cervical cancers. *Clin Radiol*. 2017;72(11):951-958.
29. Wang Q, Huang W, Zhang X, Li X. GLCM: global-local captioning model for remote sensing image captioning. *IEEE Trans Cybern*. 2023;53(11):6910-6922.
30. Wang C, Ong HH, Chiba S, Rajapakse JC. GLDM: hit molecule generation with constrained graph latent diffusion model. *Brief Bioinform*. 2024;25(3):bbae142.
31. Shirzadi Z, Schultz AP, Properzi M, et al. Greater white matter hyperintensity volume is associated with the number of microhemorrhages in preclinical Alzheimer's disease. *J Prev Alzheimers Dis*. 2024;11(4):869-873.
32. Verfaillie SC, Tijms B, Versteeg A, et al. Thinner temporal and parietal cortex is related to incident clinical progression to dementia in patients with subjective cognitive decline. *Alzheimers Dement*. 2016;5:43-52.
33. Delvenne JF, Scally B, Rose Burke M. Splenium tract projections of the corpus callosum to the parietal cortex classifies Alzheimer's disease and mild cognitive impairment. *Neurosci Lett*. 2023;810:137331.
34. Harkness K, Heckman GA, McKelvie RS. The older patient with heart failure: high risk for frailty and cognitive impairment. *Expert Rev Cardiovasc Ther*. 2012;10(6):779-795.
35. Pantoni L, Inzitari D. Hachinski's ischemic score and the diagnosis of vascular dementia: a review. *Ital J Neurol Sci*. 1993;14(7):539-546.

Appendix

Abbreviations

| | |
|-------------------|--|
| WMH | white matter hyperintensity |
| MRI | Magnetic resonance imaging |
| RF | Random forest |
| AUC | area under the curve |
| ROI | Region of interest |
| T2-FLAIR | T ₂ fluid attenuated inversion recovery |
| T ₂ WI | T ₂ -weighted image |
| MMSE | Mini-Mental State Examination |
| MoCA | Montreal Cognitive assessment |
| CDR | Clinical Dementia Rating |
| GDS | Geriatric depression scale |
| ADL | Activities of daily living |
| HIS | Hachinski Ischaemia Index scale |
| GLCM | Grey-level co-occurrence matrix |
| GLRLM | Grey-level run length matrix |
| GLSZM | Grey level size zone matrix |
| NGTDM | Neighbouring grey tone difference matrix |
| GLDM | Grey-level dependent matrix |