



A deep-learning model for predicting post-stroke cognitive impairment based on brain network damage

Chen Bai^{1,2,3#}, Yilin Leng^{1,2,3#}, Haixing Xiao⁴, Lei Li^{1,2}, Wenju Cui^{1,2,3}, Tan Li⁴, Yuefang Dong^{1,2}, Jian Zheng^{1,2,3}, Xiuying Cai⁴

¹School of Biomedical Engineering (Suzhou), Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China; ²Medical Imaging Department, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, Suzhou, China; ³Shandong Laboratory of Advanced Biomaterials and Medical Devices in Weihai, Weihai, China; ⁴Department of Neurology, The First Affiliated Hospital of Soochow University, Soochow University, Suzhou, China

Contributions: (I) Conception and design: C Bai, Y Leng; (II) Administrative support: J Zheng, T Li; (III) Provision of study materials or patients: H Xiao, W Cui, X Cai; (IV) Collection and assembly of data: L Li, Y Dong; (V) Data analysis and interpretation: C Bai; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Jian Zheng, PhD. Medical Imaging Department, Suzhou Institute of Biomedical Engineering and Technology, Chinese Academy of Sciences, 88 Keling Road, Suzhou 215163, China; School of Biomedical Engineering (Suzhou), Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China; Shandong Laboratory of Advanced Biomaterials and Medical Devices in Weihai, Weihai, China. Email: zhengj@sibet.ac.cn; Tan Li, MM. Department of Neurology, The First Affiliated Hospital of Soochow University, Soochow University, 899 Pinghai Road, Suzhou 215006, China. Email: 504535473@qq.com.

Background: Post-stroke cognitive impairment (PSCI) is a common and severe complication following acute lacunar stroke (ALS). Due to the limitations of current assessment tools and imaging methods, the early diagnosis of PSCI within 3 months of ALS remaining challenging. This study aimed to develop an effective, reliable, and accurate deep-learning method to predict PSCI within 3 months of ALS.

Methods: In total, 100 ALS patients were enrolled in the study, of whom 39 were diagnosed with PSCI and 61 were non-PSCI. First, we quantified three-dimensional (3D) gray-matter damage and white-matter tract disconnection, providing both regional damage (RD) and structural disconnection (SDC) higher-dimensional insights into brain network disruption. Second, we developed a novel deep-learning model based on ResNet18, integrating 3D RD, SDC, and diffusion-weighted imaging (DWI) to provide a comprehensive analysis of brain network integrity and predict PSCI. Finally, we compared the performance of our method with other methods, and visualized brain network damage associated with PSCI.

Results: Our model showed strong predictive performance and had a mean accuracy (ACC) of 0.820 ± 0.024 , an area under the curve (AUC) of 0.795 ± 0.068 , a sensitivity (SEN) of 0.746 ± 0.121 , a specificity (SPE) of 0.869 ± 0.044 , and a F1-score (F1) of 0.760 ± 0.050 in the five-fold cross-validation, outperforming existing models. In the PSCI patients, brain network damage significantly affected the salience, default mode, and somatic motor networks.

Conclusions: This study not only established a model that can accurately predict PSCI, it also identified potential targets for symptom-based treatments, offering new insights into PSCI.

Keywords: Structural disconnection (SDC); regional damage (RD); cognitive impairment; stroke; deep learning

Submitted Sep 21, 2024. Accepted for publication Mar 04, 2025. Published online Apr 21, 2025.

doi: 10.21037/qims-24-2010

View this article at: <https://dx.doi.org/10.21037/qims-24-2010>

Introduction

Post-stroke cognitive impairment (PSCI) is a common complication of acute lacunar stroke (ALS) (1,2), and PSCI leads to unfavorable prognoses such as recurrent stroke, and death (3). Approximately two-thirds of patients of ALS develop cognitive impairment within 3 months of stroke (4). Additionally, recurrent strokes or secondary neurodegenerative changes may result in delayed-onset PSCI, which typically emerges between 3 and 6 months post stroke. Notably, significant improvements in executive function and attention in patients with PSCI occur within 3 to 6 months. Thus, the first 3 months are a key period for early diagnosis (5). The development of effective methods for predicting PSCI within 3 months of ALS is crucial for timely intervention and improved patient outcomes (6).

In clinical settings, the Montreal Cognitive Assessment (MoCA) is the most common and widely recognized method for assessing PSCI. With advancements in magnetic resonance imaging (MRI) technology, diffusion-weighted imaging (DWI) has emerged as the primary tool for detecting small vessel diseases due to its high sensitivity and specificity (7). For example, Moulton *et al.* (8) showed that applying predictive models that reflect the lesion and the surrounding changes to DWI data can predict long-term functional outcomes post stroke with considerable accuracy. However, despite these advantages, DWI alone is insufficient for capturing the broader, network-level changes involved in PSCI (9).

Recent studies (10,11) indicate that PSCI may involve complex disruptions in brain network integrity, particularly involving the structural disconnection (SDC) of white-matter tracts and regional damage (RD) of gray matter (10-14). These factors are increasingly recognized for their broader effects post stroke, which extend beyond localized lesions to widespread connectivity issues across the brain (15,16). For example, Salvalaggio *et al.* (17) showed that SDC is closely associated with impairments in specific cognitive domains, such as attention, memory, and executive function. Similarly, RD has been linked to deficits in applied cognitive functions, SDC and RD offer a more granular view of the brain's structural vulnerability to stroke (18). The predictive value of SDC and RD is further supported by their ability to explain functional connectivity disruptions, often crucial factors in cognitive decline. Given these findings, predictive methods that integrate SDC and RD need to be established to enhance prediction accuracy. However, current methods predominantly focus on either isolated brain damage or

two-dimensional (2D) region-based SDC (11,19,20), often overlooking the integrated influence of three-dimensional (3D) SDC and RD on post-stroke outcomes. Compared to the 2D region-based SDC matrix, 3D SDC and RD images provide richer, higher-dimensional data, making them more suitable for predicting PSCI. Despite this, few studies have attempted to evaluate the predictive value of 3D SDC and RD (17). Thus, integrating 3D SDC, RD, and DWI could provide a more comprehensive understanding of the underlying mechanisms of PSCI.

An accurate prediction model for PSCI could help physicians to decide on a proper management strategy after a stroke. In recent years, deep learning has attracted tremendous attention from radiologists because of its outstanding image recognition capability, and has shown promise in the diagnosis of PSCI (7,8,21). Unlike traditional machine-learning methods, deep learning can learn adaptively and extract high-dimensional features in a fully data-driven way. For example, recent studies (21,22) have shown the potential of deep learning in predicting PSCI using fluorodeoxyglucose positron emission tomography or diffusion tensor imaging. However, as DWI can identify ischemic regions with fast imaging speed and has lower costs than these imaging techniques, it represents a more practical choice for integrating deep learning into clinical workflows. Moreover, explainable deep-learning approaches, such as those that employ class activation mapping (CAM), have been shown to be instrumental in advancing both scientific research and healthcare tasks. Recent studies (23,24) have successfully used CAM-based techniques in deep-learning models for brain disease analysis, showing their utility in identifying discriminative regions and enhancing the interpretability of deep-learning models. These advancements strengthen the application and understanding of deep learning in medical imaging, paving the way for more reliable and transparent predictions in clinical settings.

To this end, we developed a 3D deep-learning model to predict PSCI within 3 months of ALS. The contributions of this study are as follows: (I) it combined 3D DWI, RD, and SDC information to enhance the prediction of PSCI by exploring changes at the global (whole brain), regional (gray-matter connectivity), and structural (white-matter tracts connectivity) levels; and (II) it visualized the damaged brain networks associated with PSCI, thereby offering insights into the underlying mechanisms of how these network disruptions contribute to PSCI through the deep-learning model. We present this article in accordance

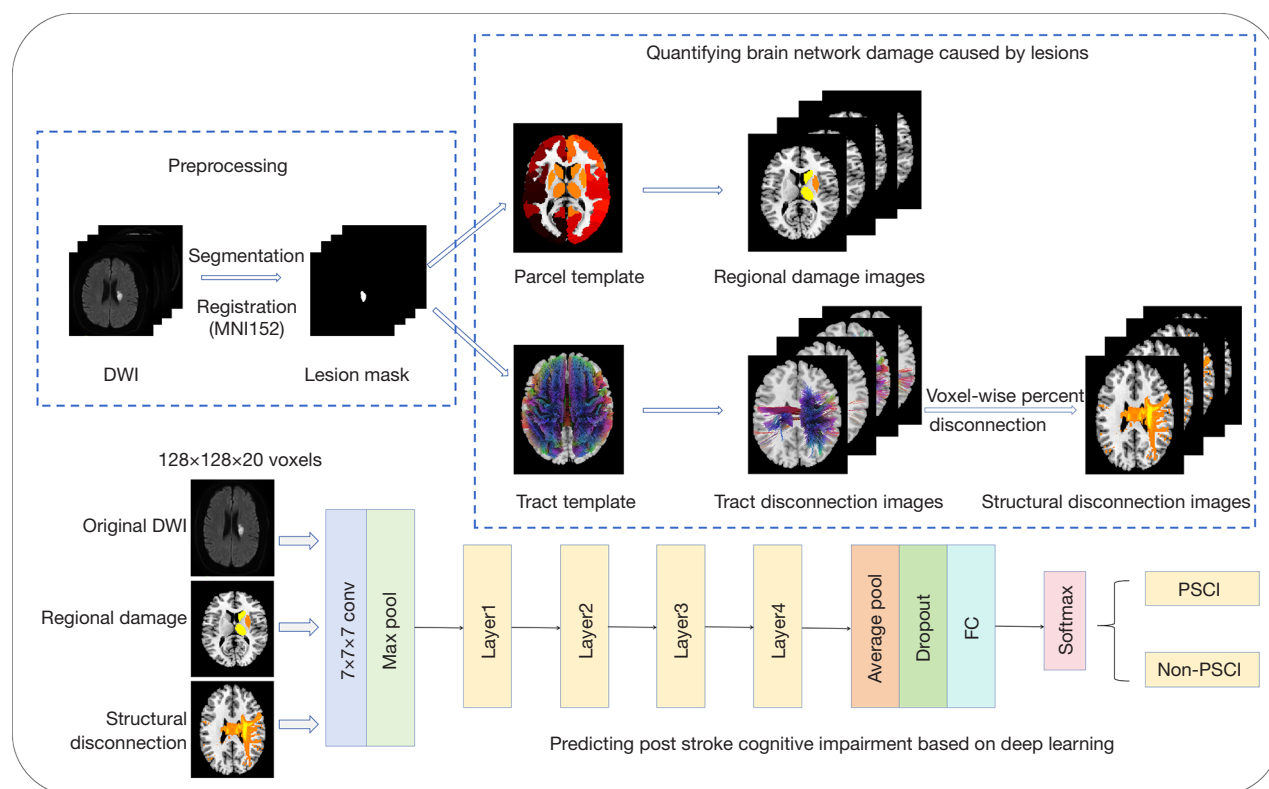


Figure 1 Framework of our proposed method. (I) First, we segmented lesions on DWI from each patient and registered them to the MNI152 template; (II) second, we quantified regional damage and structural disconnection for each patient, respectively; (III) finally, we constructed the deep-learning model to predict PSCI using regional, structural, and global information as input. DWI, diffusion-weighted imaging; FC, fully connected layer; MNI152, Montreal Neurological Institute 152; PSCI, post-stroke cognitive impairment.

with the TRIPOD+AI reporting checklist (available at <https://qims.amegroups.com/article/view/10.21037/qims-24-2010/rc>).

Methods

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Medicine Ethics Committee of The First Affiliated Hospital of Soochow University (No. 2021359), and the requirement for written informed consent was waived due to the retrospective nature of the study. This study developed a 3D deep-learning model to predict PSCI within 3 months of ASL. First, gray-matter parcel damage caused by lesions was extracted to reflect brain RD. Second, white-matter tract disconnection-affected lesions were quantified as SDC. Finally, RD, SDC, and original DWI were combined as inputs of the deep-learning model to

predict PSCI (refer to *Figure 1*).

Patients selection

In total, 199 patients with imaging proven lacunar infarction who were admitted to The First Affiliated Hospital of Soochow University between April 2018 and May 2022 were enrolled in this study. ALS was defined as an acute ischemic lesion with a diameter of ≤ 20 mm on DWI sequence in the perforator artery supply area (3). To be eligible for inclusion in the study, the patients had to meet the following inclusion criterion (3): have had their demographic and clinical data collected, including gender, age, medical history (cerebral infarction, hypertension, and diabetes), cerebral microbleed, deep hyperintensity, overall burden score, perivascular space (grades 2–4), presence of microbleeds, white-matter volume, National Institutes of Health Stroke Scale (NIHSS) score, and MoCA score,

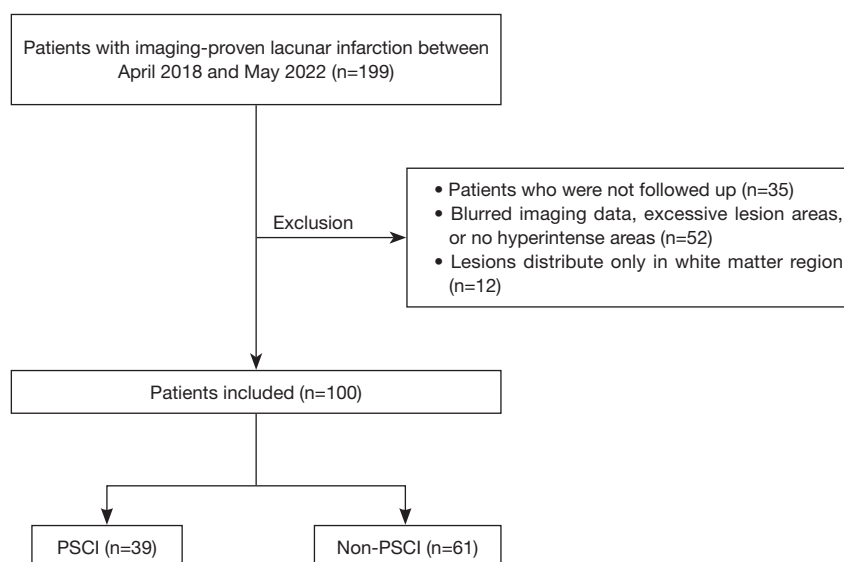


Figure 2 Flowchart showing the patient selection process. PSCI, post-stroke cognitive impairment.

within 48 hours of hospital admission.

Of the 199 patients, 35 were excluded from the study who were not followed up, 52 were excluded due to blurred imaging data, excessive lesion areas, or no hyperintense areas, and 12 were excluded because their lesions were distributed only in white matter (as such lesions alone may not provide sufficient information for generating comprehensive disconnection maps, which require damage to both gray and white matter to effectively reflect brain structural and regional disruptions). Ultimately, the imaging data of 100 patients with complete baseline and follow-up data were collected. Based on the MoCA score during the 3-month follow-up, the 100 patients were categorized as PSCI (39 subjects) and non-PSCI (61 subjects) (Figure 2). The baseline characteristics of the study sample are shown in Table 1. PSCI was defined as the development of cognitive impairment within 3 months of stroke onset, while non-PSCI was defined as no cognitive impairment within 3 months of stroke onset.

Cognitive assessment

The patients were followed up via an outpatient visit after discharge. Experienced physicians used the MoCA to assess the cognitive status of the patients 3 months after the stroke. According to the Chinese revised protocol of the Vascular Impairment of Cognition Classification Consensus Study, the assessment criteria for cognitive impairment

were as follows: for uneducated patients, an MoCA score of ≤ 13 ; for primary school-educated patients, an MoCA score of ≤ 19 ; for junior high school or higher-educated patients, an MoCA score of ≤ 24 .

MRI acquisition and preprocessing

All the magnetic resonance (MR) images were acquired using either a 1.5T (Philips Medical Systems, Best, The Netherlands) or a 3.0T (Philips and Siemens Medical Systems, Erlangen, Germany) scanner. Each MRI scan included DWI for identifying acute infarcts and T1-weighted fluid attenuated inversion recovery (T1W-FLAIR). For the 3T scanner, the DWI images were acquired using the following parameters—b value: 1,000 s/mm²; repetition time (TR): 2,782 ms; echo time (TE): 98 ms; field of view (FOV): 23 cm; and slice thickness: 5 mm. For the 1.5T scanner, the DWI images were acquired using the following parameters—b value: 1,000 s/mm²; TR: 3,616 ms; TE: 110 ms; FOV: 23 cm; and slice thickness: 5 mm. For the 1.5T and 3T scanners, the T1W-FLAIR images were acquired using the following parameters—TR: 2,000 ms; TE: 20 ms; FOV: 23 cm; and slice thickness: 6 mm.

The lesions were manually segmented on the DWI scans using the ITK-SNAP (25) software (<http://www.itksnap.org>) by two experienced physicians. The lesion masks were supervised and corrected by a neurologist. The mask images were then registered using the Statistical Parametric

Table 1 Baseline characteristics of the study sample

Characteristics	Non-PSCI (n=61)	PSCI (n=39)	P
Age (years)	60.87±11.03	66.90±9.74	0.006
Gender			
Female	17 (27.9)	15 (38.5)	–
Male	44 (72.1)	24 (61.5)	–
Cerebral microbleeds			
No	47 (77.0)	26 (66.7)	–
Yes	14 (23.0)	13 (33.3)	–
Sequela of lacunar infarct			
No	30 (49.2)	13 (33.3)	–
Yes	31 (50.8)	26 (66.7)	–
Periventricular hyperintensity	1.15±1.08	1.97±0.99	<0.001*
Deep hyperintensity	0.59±0.84	1.51±1.07	<0.001*
White-matter volume (mL)	4.83±8.31	9.13±14.66	0.064
MoCA	23.02±4.418	17.64±4.743	<0.001*
NIHSS	2.49±2.57	3.92±3.23	0.0160
Overall burden score	1.62±1.07	2.49±1.05	<0.001*
Perivascular space			
0 (grades 0–1)	37 (60.7)	13 (33.3)	–
1 (grades 2–4)	24 (39.3)	26 (66.7)	–

The continuous variables are presented as the mean ± SD, and the categorical variables as the number (percentage). *, variables that are significant between the PSCI and non-PSCI patients (P values less than 0.001). NIHSS, National Institutes of Health Stroke Scale; MoCA, Montreal Cognitive Assessment; PSCI, post-stroke cognitive impairment; SD, standard deviation.

Mapping 12 (<https://www.fil.ion.ucl.ac.uk/spm/>), a clinical toolbox in MATLAB R2023a (<https://www.mathworks.cn/campaigns/products/>), as follows: (I) each T1W-FLAIR image was registered to the Montreal Neurological Institute 152 template (26) with a voxel size of 1 mm × 1 mm × 1 mm; and (II) the corresponding DWI image and its mask image were co-registered to the T1W-FLAIR image with a size of 182×218×182 voxels. Finally, all the DWI images and masks were checked to ensure accurate alignment.

Estimating SDC

To investigate the effect of lesions on brain structural connectivity, we used the Lesion Quantification Toolkit (LQT) (27) to calculate the severity of SDCs caused by the lesions. The LQT integrated large-scale standardized

connectome maps constructed from high-spatial and high-angular resolution diffusion MRI data from 842 healthy participants in the Human Connectome Project-842 (HCP-842). The HCP-842 cohort comprised 372 males and 470 females, aged 22 to 36 years. The input of the LQT was a preprocessed binary mask of each patient. The following procedure was used to estimate the SDC images: (I) embed the lesion into a whole-brain streamline tractography template that comprises 70 white-matter tract streamlines; (II) retain a subset of streamlines that intersect the volume occupied by the lesion, and output the density of disconnected streamlines for each voxel; and (III) estimate the percentage of streamlines in each voxel as the SDC image using tract disconnection imaging (*Figure 1*). Finally, binary SDC maps from all patients were overlaid to illustrate the overall incidence of lesion-related disconnections.

Estimating RD

Similarly, we used the LQT and incorporated previous information on regional boundaries to calculate the voxel-level gray-matter lesion load for each patient. The selected atlas was the Schaefer-Yeo atlas (28,29), which was divided into seven functional cortical networks (see [Tables S1,S2](#)) and consisted of 135 cortical parcels, as well as cerebellar parcels. The Schaefer template was used because it is explicitly defined to maximize regional homogeneity in terms of functional signals and connectivity patterns. The inputs to the LQT were the processed mask images and the Schaefer-Yeo atlas. The percentage of voxels within each parcel of the atlas that intersected with the lesion were estimated to generate the RD images ([Figure 1](#)).

Estimating the region-based SDC matrix

Using the LQT, we incorporated the SDC information between two parcels to establish a region-based SDC matrix. The following procedure was used to calculate the region-based SDC matrix: (I) create an atlas structural connectivity matrix using the HCP-842 streamline tractography atlas and the Schaefer-Yeo atlas; (II) embed the lesion into the HCP-842 streamline tractography atlas, and filter the number of atlas streamlines that bilaterally terminate within both gray-matter parcels; and (III) convert a percentage disconnection severity matrix relative to the atlas structural connectivity matrix as the region-based SDC matrix. Finally, we visualized the region-based SDC matrix to analyze in terms of the expected disconnection severity between gray-matter parcels.

Development of the deep-learning model

ResNet18 (30-32) is one of the most commonly used deep-learning models for image classification, and is known for its ability to retain low-scale features due to its shallow network structure. In this study, we developed the model based on the 3D ResNet18. The network started with a 7×7×7 convolutional layer, followed by four layers of processing. Each of these layers contained two residual blocks to enhance feature learning. The network concluded with global average pooling and a fully connected layer that outputs the final prediction. The DWI, SDC, and RD images were resized to 128×128×20 voxels to ensure uniform size. Further, the DWI, SDC, and RD images were concatenated along the channel dimension to form the

model input. To balance the data among different groups, the study employed a five-fold cross-validation approach, dividing the training and validation sets into 80 and 20 patients, respectively. In the five-fold cross-validation, the dataset was randomly divided into five equal-sized folds. The model was trained and validated five times; each iteration used a different fold as the validation set, while the remaining four folds were used as the training set. This process ensured that every data point was used for validation exactly once. The final evaluation metric was computed as the average performance across all five iterations. Stratified random sampling was used to construct each fold, ensuring that the proportion of the distribution of PSCI patients and non-PSCI patients was consistent across all folds. The five-fold cross-validation allocation details are shown in [Table S3](#). The hyperparameters were selected based on the training data, and the model was evaluated using the validation set (33).

For the training of the deep-learning algorithm, the following hyperparameters were used: 400 epochs, a batch size of 4, an Adaptive Moment Estimation (Adam) (34) optimizer with momentum coefficients ranging from 0.9 to 0.999, a step learning rate scheduler with a step size of 25, and binary cross-entropy with logit loss. Finally, softmax was used to calculate the prediction scores as output.

To mitigate overfitting, we implemented data augmentation and added a dropout layer (dropout rate: 0.5) preceding the final fully connected layer during training. The following augmentation hyperparameters were used: an affine transformation probability of 0.5, a rotate range ($\pi/36$, $\pi/18$, $\pi/18$), a translate range [4, 6, 6], and bilinear interpolation. Model training and validation were conducted using Python 3.9 with Pytorch 1.13 (<https://pytorch.org>) on an NVIDIA GeForce RTX 3090Ti.

Comparative experiments

To confirm the superiority of our model, we conducted a comparative analysis of different input models based on ResNet18, different fusion strategies based on ResNet18, and three prevalent PSCI predictive classification models.

Comparison of different input information

To validate the effectiveness of introducing the SDC and RD in predicting PSCI, we compared the experimental results of four different inputs. The following inputs were used: (I) DWI + ResNet18; (II) SDC images + ResNet18;

(III) RD images + ResNet18; and (IV) our model: DWI + SDC images + RD images + ResNet18 (see *Figure 1*).

Comparison of different fusion strategies based on the ResNet18 model

To explore the optimal fusion strategy based on the ResNet18 model, we compared the early fusion method (our model), late fusion method, and score fusion method as proposed by Seeland and Mäder (35). For late fusion, we employed a fully connected layer fusion strategy, while for score fusion, we adopted the maximum score strategy. Both late fusion and score fusion employed weight sharing to reduce the number of trainable parameters. The training hyperparameters were consistent with those used for our model, except that the dropout rate was set to 0 to prevent the loss of useful information. SDC and RD data are sparse, and many pixel values are zero. The late fusion strategy based on fully connected layers and the score fusion strategy based on softmax still resulted in many zero-value features in the fully connected layer. Setting the dropout rate to 0.5 can result in valuable information being randomly discarded. For our early fusion model, we combined the three input images (SDC, RD, and DWI) along the channel dimension, such that each input contributes information to every channel to provide a more information-rich representation. In this case, the dropout rate was set to 0.5 to mitigate overfitting.

Comparison of three different prediction models

We compared our model with the following three common PSCI prediction models: (I) the clinical data model (36) based on multilayer perceptron (MLP); (II) the radiomics features model (37) based on support vector machine (SVM); and (III) the predictive factors model (38) based on random forest (RF).

Comparison 1: clinical data model based on MLP

Patient age, gender, the presence of microbleeds, periventricular hyperintensities, deep white-matter hyperintensities, the overall burden score, old infarcts, perivascular spaces, and the NIHSS were selected for inclusion in this prediction model. The MLP model consisted of three fully connected layers: the first layer with 64 units; the second layer with 32 units; and an output layer with a single neuron. Rectified linear unit activation functions were applied to the first two layers, while the

output layer used a sigmoid activation function for binary classification. The model was trained using 100 epochs, a batch size of 16, an Adam optimizer, and binary cross-entropy with logit loss as the loss function.

Comparison 2: radiomics feature model based on SVM

In total, 957 radiomics features were extracted from the DWI, including first-order, shape-based, wavelet filtering, logarithmic, and gradient features. We trained and validated a radiomics feature model based on SVM using five-fold cross-validation. For this task, we employed a radial basis function kernel, enabling the SVM to manage non-linear relationships among the input features by mapping them into a higher-dimensional space.

Comparison 3: predictive factor model based on RF

We identified the following two predictive factors based on voxel symptom mapping: the SDC score (SDC_{score}), and the RD score (RD_{score}). The voxel symptom mapping analyzed the relationship between tissue damage and behavior on a voxel-by-voxel basis using NiiStat (<http://www.nitrc.org/projects/niistat>). The predictive factor procedure calculation was conducted as described previously (38); however, the binarization threshold for SDC and RD images were set to 0 when computing the significant clusters associated with PSCI using NiiStat. As our goal was to account for all damage, we incorporated the total voxels. The significant voxels survived the voxel-level family-wise error correction with a corrected Z threshold of 1.6. The SDC_{score} and RD_{score} were used as inputs for the RF model for predicting PSCI. The parameters of RF were set as follows—number of decision trees: 4; minimum number of samples contained in the leaf nodes: 1; and minimum number of samples by which a node could be split: 2.

Statistical analysis

In the statistical analysis, we compared the baseline characteristics between the PSCI and non-PSCI patients using SPSS 24.0 software (<https://www.ibm.com/spss>). A one-way analysis of variance was used for the continuous variables (39). Percentage statistics were used to present the distribution among various categories for the categorical variables. To evaluate the performance of our model, we used several metrics, including accuracy (ACC), balanced accuracy (BACC), the area under the curve (AUC), sensitivity (SEN), specificity (SPE), and the F1-score (F1). Additionally, a pairwise *t*-test based on the AUC was

conducted to compare the performance of our proposed model with that of the other models. A P value <0.05 was considered statistically significant for all the statistical tests.

Model visualization

To further interpret how our deep learning model identifies relevant PSCI information, we employed CAM (40) for the model visualization. CAM was used to generate class-specific heat maps by linking features from the last convolutional layer of the model with its output. These heat maps revealed the attention of the deep-learning model and highlighted the input image regions that affected predictions. To enhance visualization, we applied a jet color scheme to the heat maps.

Results

Analysis of patient characteristics

We analyzed the baseline characteristics of the 100 patients (Table 1). The patients diagnosed with PSCI accounted for 39% of the cohort and exhibited distinct characteristics. Notably, the PSCI cohort had a higher mean age (66.90 ± 9.74 years), a greater proportion of males (61.5%), a higher incidence of old lacunar infarcts (66.7%), and a higher proportion of perivascular space at grades 2–4 (66.7%). Moreover, the patients in the PSCI cohort had lower MoCA scores (17.64 ± 4.743) and higher NIHSS scores (3.92 ± 3.23). Significant differences between the non-PSCI and PSCI patients were observed in terms of periventricular hyperintensity ($P < 0.001$), deep hyperintensity ($P < 0.001$), and the overall burden score ($P < 0.001$). However, no significant difference was observed in the white-matter volume between the two groups ($P = 0.064$).

Results of estimating SDC and RD

The results of analyzing the lesion, SDC, and RD images for all patients, and for PSCI patients minus non-PSCI patients are shown in Figure 3. In our cohort, the frequency and coverage of damage were almost consistent between the left and right hemispheres for all the patients. The lesions involved the midbrain, posterior limb of the internal capsule, and corona radiata region (Figure 3A). Specifically, SDC covered nearly the entire brain (Figure 3C). While RD mainly involved the caudate nucleus, thalamus, and putamen (Figure 3E).

In the comparison of the PSCI and non-PSCI patients, the lesion and RD (Figure 3B,3F) locations were almost consistent with those of all patients, but they mainly involved the left hemisphere. The SDC mainly involved the parietal, temporal lobe, and centrum semioval in the left hemisphere (Figure 3D). The white-matter tracts (Figure 4) mainly involved the left parietopontine tract, superior cerebellar peduncle (SCP), left rubrospinal tract, right rubrospinal tract, and left corticospinal tract (CST_L). Due to the spatial distribution of white-matter tracts, lesions that are spatially distant may affect the same fiber bundle, resulting in damage or disconnection.

To illustrate the variability of SDC across the gray-matter regions, we visualized the region-based SDC matrix constructed from SDC and RD among all PSCI (Figure 5A), non-PSCI (Figure 5B) patients, and PSCI minus non-PSCI patients (Figure 5C). In the brain maps, the nodes and edges separately represent gray-matter regions and white-matter tracts SDC of the brain. Larger nodes and edges signify a higher degree of SDC between brain regions. We observed that the PSCI patients exhibited severe SDC between the: (I) right parietal and frontal lobes; (II) right parietal and left occipital lobes; and (III) right temporal and left parietal lobes. We also explored the functional regions affected by brain network damage, according to the seven functional networks of the brain (Figure 6). Our findings revealed that brain network damage significantly affected the salience, default mode, and somatic motor networks in both the left and right hemispheres, as well as the visual networks in the right hemisphere.

Comparative experiments

Comparison of different input information

Table 2 shows the mean performance of the four different input models in the five-fold cross-validation. Notably, the prediction performance of the DWI images was superior to that of the SDC and RD images. The DWI model had lower results than our model (ACC: 0.770 ± 0.023 , BACC: 0.754 ± 0.054 , AUC: 0.735 ± 0.178 , SEN: 0.671 ± 0.205 , and SPE: 0.836 ± 0.102). Conversely, the SDC model had the lowest performance across most metrics (ACC: 0.730 ± 0.022 , BACC: 0.700 ± 0.033 , AUC: 0.648 ± 0.064 , and SEN: 0.564 ± 0.108). Overall, our model presented the best performance in terms of ACC (0.820 ± 0.024), BACC (0.808 ± 0.040), the AUC (0.795 ± 0.068), SEN (0.746 ± 0.121), and SPE (0.869 ± 0.044), demonstrating its superior ability to classify and predict the target variable, with a balanced

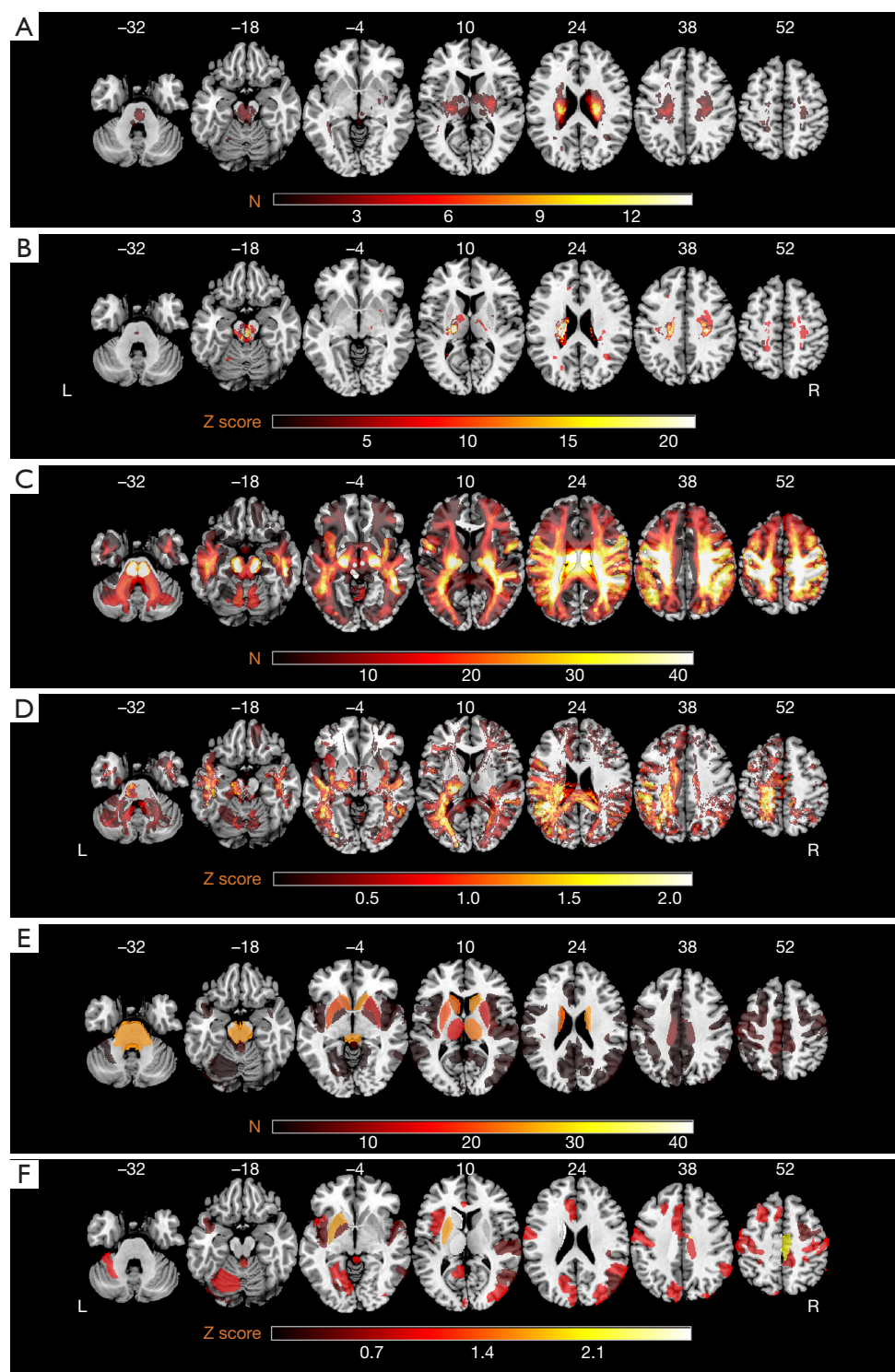


Figure 3 Overlap images of all patients and subtraction images showing PSCI minus non-PSCI patients. (A) Overlap map of lesion masks for all patients; (B) subtraction map of lesion masks; (C) overlap of all binary SDC images; (D) subtraction map of SDC images; (E) overlap of all binary RD images; (F) subtraction map of RD images. The damage frequency of the entire sample is denoted as N, where the Z score represents the standardization value. The axial coordinates refer to the MNI space in mm. L and R separately indicate left and right. MNI, Montreal Neurological Institute; PSCI, post-stroke cognitive impairment; RD, regional damage; SDC, structural disconnection.

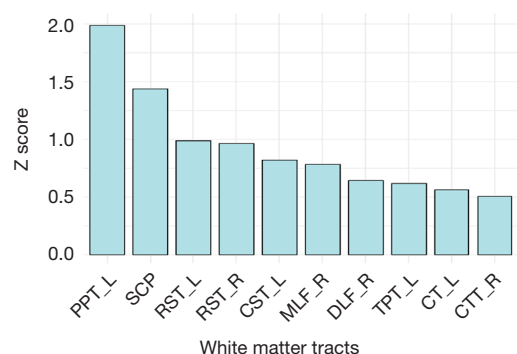


Figure 4 The top 10 white-matter tract disconnections. All PSCI white-matter tract disconnections minus all non-PSCI white-matter tract disconnections. The Z score represents the standardization value. CST_L, left corticospinal tract; CT_L, left corticothalamic pathway; CTT_R, right central tegmental tract; DLF_R, right dorsal longitudinal fasciculus; MLF_R, right medial longitudinal fasciculus; PPT_L, left parietopontine tract; RST_L, left rubrospinal tract; RST_R, right rubrospinal tract; PSCI, post-stroke cognitive impairment; SCP, superior cerebellar peduncle; TPT_L, left temporopontine tract.

trade-off between sensitivity and specificity (for details, see Table S4). A significant difference was observed between our model and the SDC ($P=0.017$) and RD ($P=0.036$) single-input models.

Comparison of different fusion strategies based on the ResNet18 model

Table 3 presents the results of different fusion strategies based on the ResNet18 model. Among these approaches, our (early fusion) model showed superior performance compared to both the late fusion and score fusion models. Specifically, the late fusion model had an ACC of 0.770 ± 0.040 , a BACC of 0.736 ± 0.063 , and an AUC of 0.724 ± 0.078 , while the score fusion model had an ACC of 0.770 ± 0.045 , a BACC of 0.746 ± 0.072 , and an AUC of 0.733 ± 0.066 . Although the differences were not statistically significant compared to the late ($P=0.144$) and score fusion ($P=0.255$) models, our model had higher SEN and F1 scores in practical applications, highlighting its advantage in balancing mistake classifications and supporting its validity as the optimal fusion strategy. The inference time of our

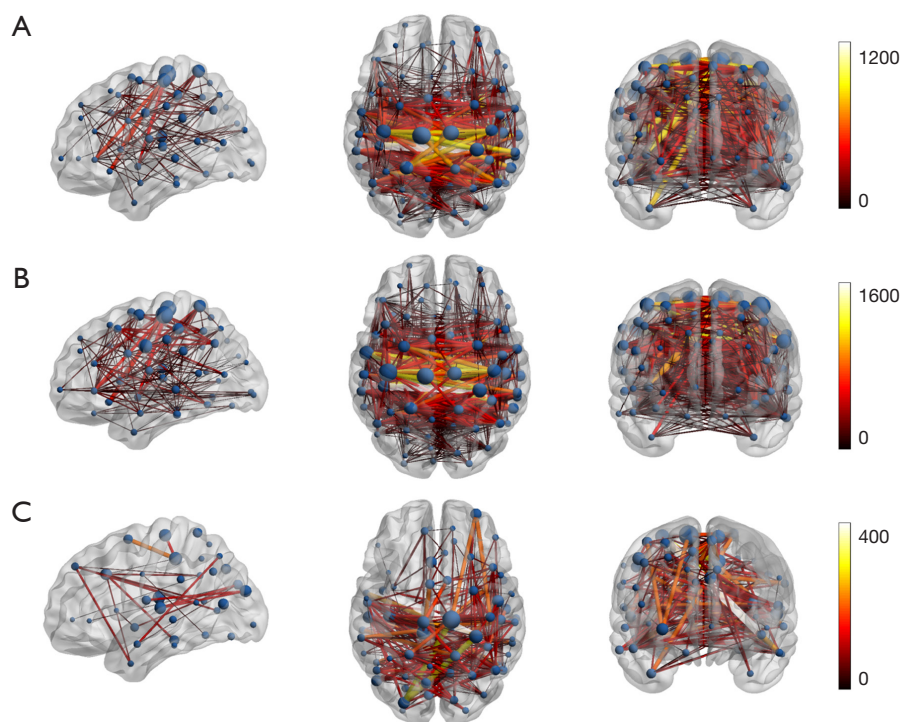


Figure 5 3D region-based SDC brain maps. (A) Region-based SDC brain maps for all PSCI patients; (B) region-based SDC brain maps for all non-PSCI patients; (C) region-based SDC brain maps for the PSCI minus non-PSCI. In the brain maps, the nodes and edges separately represent gray-matter regions and white-matter tracts SDC of the brain. 3D, three-dimensional; SDC, structural disconnection; PSCI, post-stroke cognitive impairment.

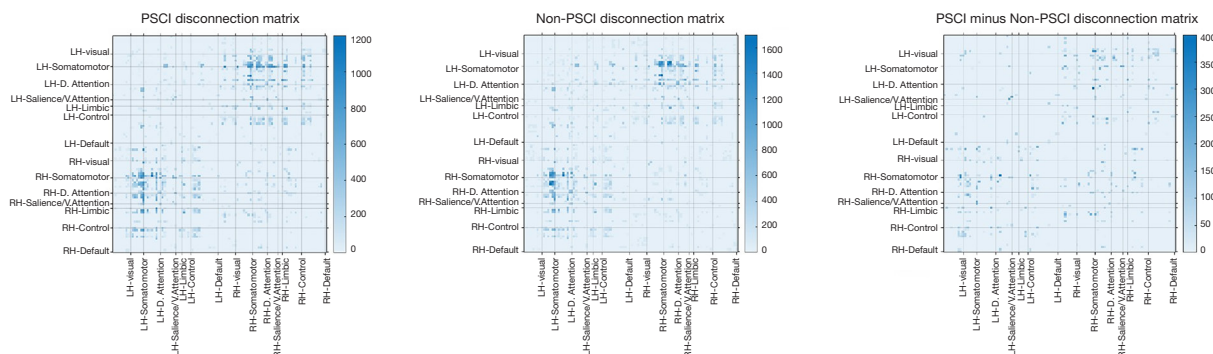


Figure 6 Functional correspondence region-based SDC matrix. The points in the matrix represent the degree of SDC between two brain regions, with brighter points indicating more severe damage. LH, left hemisphere; RH, right hemisphere. SDC, structural disconnection; D, Attention, dorsal attention; Salience/V.Attention, salience/ventral attention; PSCI, post-stroke cognitive impairment.

Table 2 The mean performance of the four models based on the Resnet18 architecture in the validation set (P values for comparing the significant difference between other models and our model)

Model	ACC	AUC	F1	SEN	SPE	BACC	P
DWI	0.770±0.023	0.735±0.178	0.681±0.091	0.671±0.205	0.836±0.102	0.754±0.054	0.355
SDC	0.730±0.022	0.648±0.064	0.615±0.068	0.564±0.108	0.834±0.061	0.700±0.033	0.017
RD	0.741±0.048	0.669±0.079	0.634±0.074	0.589±0.137	0.838±0.124	0.714±0.046	0.036
Our model	0.820±0.024	0.795±0.068	0.760±0.050	0.746±0.121	0.869±0.044	0.808±0.040	–

Data are presented as mean ± SD. ACC, accuracy; AUC, area under the curve; BACC, balanced accuracy; DWI, diffusion-weighted imaging; F1, F1-score; RD, regional damage; SD, standard deviation; SDC, structural disconnection; SEN, sensitivity; SPE, specificity.

Table 3 The mean performance of different fusion strategies based on the ResNet18 model in the validation set

Model	ACC	AUC	F1	SEN	SPE	BACC	P
Late fusion	0.770±0.040	0.724±0.078	0.649±0.095	0.571±0.182	0.900±0.070	0.736±0.063	0.144
Score fusion	0.770±0.045	0.733±0.066	0.662±0.111	0.621±0.229	0.870±0.107	0.746±0.072	0.255
Our model	0.820±0.024	0.795±0.068	0.760±0.050	0.746±0.121	0.869±0.044	0.808±0.040	–

Data are presented as mean ± SD. ACC, accuracy; AUC, area under the curve; BACC, balanced accuracy; F1, F1-score; SD, standard deviation; SEN, sensitivity; SPE, specificity.

model was the shortest compared with the late fusion and score fusion models (*Table 4*), which is crucial in clinical practice.

Comparison of the three different prediction models
Comparison 1 and Comparison 2: clinical data model based on MLP and radiomics feature model based on SVM

As *Figure 7* shows, in terms of its mean performance, the clinical data model based on MLP had an ACC of 0.759±0.046, a SEN of 0.618±0.193, and an AUC of 0.768±0.105. The radiomics feature model based on SVM

performed poorly and had a SEN of 0.386±0.092 and an ACC of 0.589±0.098.

Comparison 3: predictive factor model based on RF

As the results in *Figure 8* show, the clusters of lesions (*Figure 8A*) and SDC images (*Figure 8B*) were mainly observed in the left fornix. The RD images (*Figure 8C*) were primarily concentrated in the left paracentral lobule, left corpus callosum, left supplementary motor area, and left midcingulate. The SDC and RD images contained more surviving voxels related to PSCI than lesions. We found that only one training set in the five-fold cross-validation

produced a significant cluster. This suggested that individual differences affected the results of the symptom-based mapping method.

We trained and validated six predictive factor models based on RF. The clinical data aligned with the Comparison 1 data. *Table 5* shows the performance of the six predictor models in the validation set. When solely employing the SDC_{score} and RD_{score} , all the evaluation metrics exhibited subpar performance, characterized by a SEN of only 0.25. However, integrating the predictive factors with the clinical

data resulted in a significant improvement. Specifically, the fusion model achieved an ACC of 0.762 and a SEN of 0.500. Conversely, our model achieved higher performance without clinical data, relying solely on 3D SDC and RD images.

Model visualization

Figure 9 shows the heat maps generated by our deep-learning model using the CAM technique. The intensity in the heat maps represents the importance of different regions in MR images for predicting PSCI. To better understand the focal regions of the model's attention, we overlaid the SDC and RD images onto the DWI image, aligning them with the model visualization heat map. In the input images, SDC is depicted in red, and RD is shown in blue. The heat map analysis revealed that the regions corresponding to most of the SDC and RD were prominently highlighted in the CAM images, indicating the network's heightened attention to these areas. Our model was able to identify and adaptively learn inter-patient variations in the SDC and RD images.

Discussion

This study developed and validated a deep-learning model based on brain network damage for predicting PSCI within

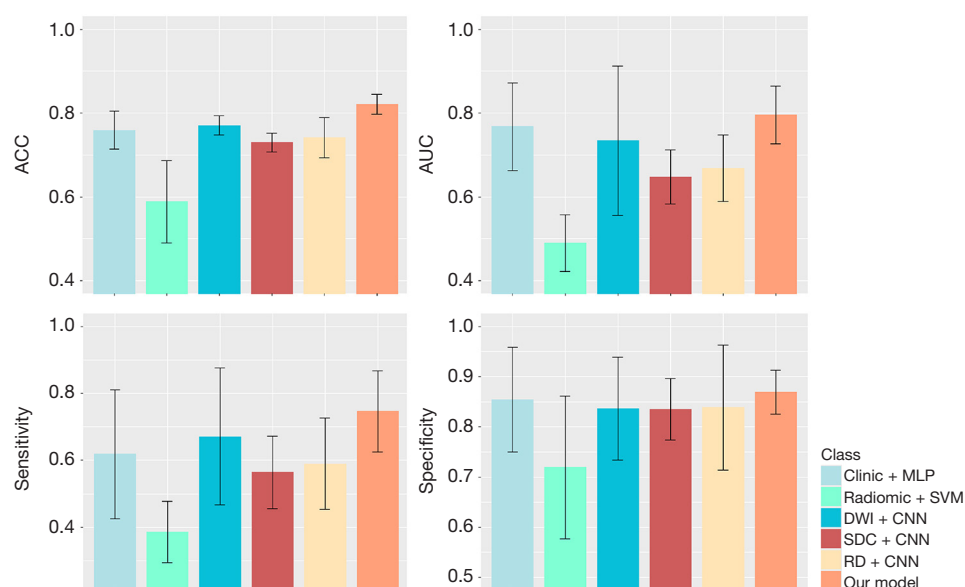


Figure 7 The average performance and standard deviation of the six different models. ACC, accuracy; AUC, area under the curve; CNN, convolutional neural network; DWI, diffusion-weighted imaging; MLP, multilayer perceptron; RD, regional damage; SDC, structural disconnection; SVM, support vector machine.

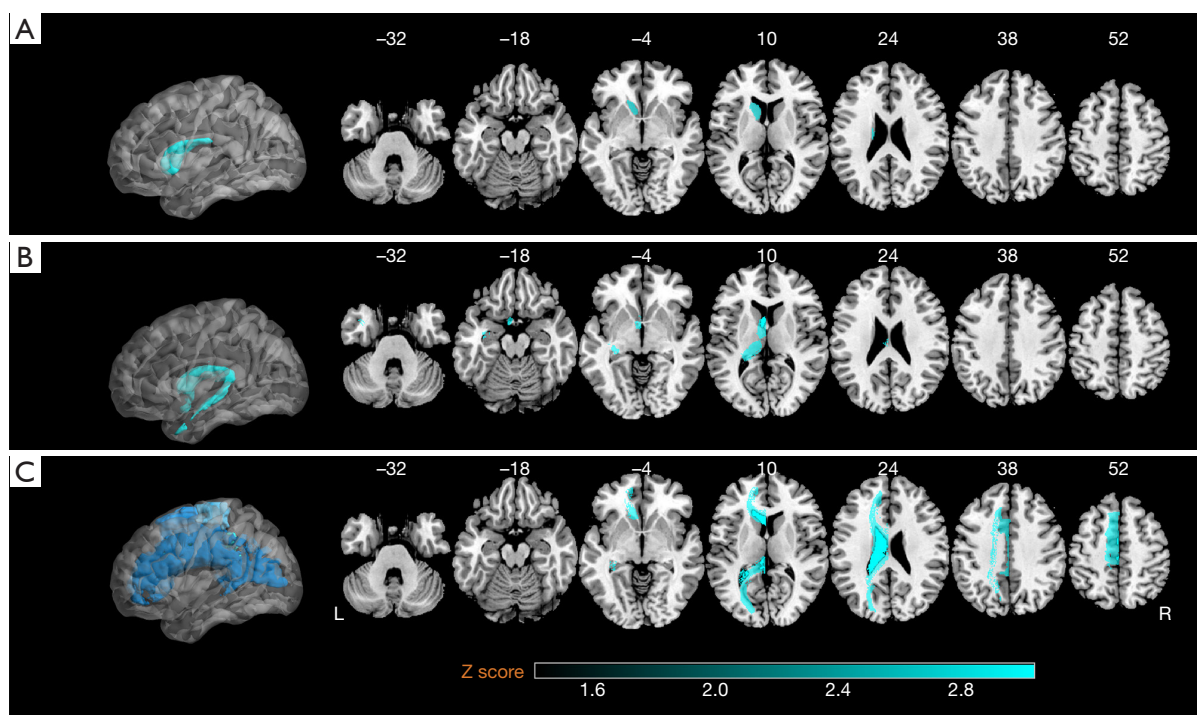


Figure 8 Significant clusters map of training set patients based on voxel symptom mapping (right), 3D visualization of significant clusters map (left). (A) Significant clusters from lesion-symptom mapping; (B) significant clusters from binary SDC symptom mapping; (C) Significant clusters from binary RD symptom mapping. FWE correct Z threshold of 1.6; the axial coordinates refer to the MNI space in mm. L, left; R, right; 3D, three-dimensional; FWE, family-wise error; MNI, Montreal Neurological Institute; RD, regional damage; SDC, structural disconnection.

Table 5 The performance of the six models based on RF in the validation set

Model	Train strategy	ACC	AUC	F1	SEN	SPE	BACC
SDC _{score}	1-fold	0.429	0.394	0.250	0.250	0.538	0.394
RD _{score}	1-fold	0.667	0.587	0.364	0.250	0.923	0.587
SDC _{score} + RD _{score}	1-fold	0.476	0.433	0.267	0.250	0.615	0.433
SDC _{score} + Clinic	1-fold	0.714	0.745	0.700	0.875	0.615	0.745
RD _{score} + Clinic	1-fold	0.619	0.596	0.500	0.500	0.692	0.596
SDC _{score} + RD _{score} + Clinic	1-fold	0.762	0.712	0.615	0.500	0.923	0.712
Our model (mean ± SD)	5-fold	0.820±0.024	0.795±0.068	0.760±0.050	0.746±0.121	0.869±0.044	0.808±0.040

1-fold represents a training set in five-fold cross-validation; five-fold represents five-fold cross-validation. ACC, accuracy; AUC, area under the curve; BACC, balanced accuracy; F1, F1-score; RD_{score}, regional damage scores; RF, random forest; SD, standard deviation; SDC_{score}, structural disconnection scores; SEN, sensitivity; SPE, specificity.

3 months of ALS. The individual models based on SDC, RD, and DWI all demonstrated significant diagnostic efficiency. However, the fusion model incorporating SDC, RD, and DWI showed the highest performance in our study, providing complementary insights by capturing global,

regional, and structural perspectives of brain changes.

To explore the changes associated with PSCI, we first analyzed the baseline characteristics of the patients and conducted a brain network damage assessment. We observed that the PSCI patients had a higher mean age

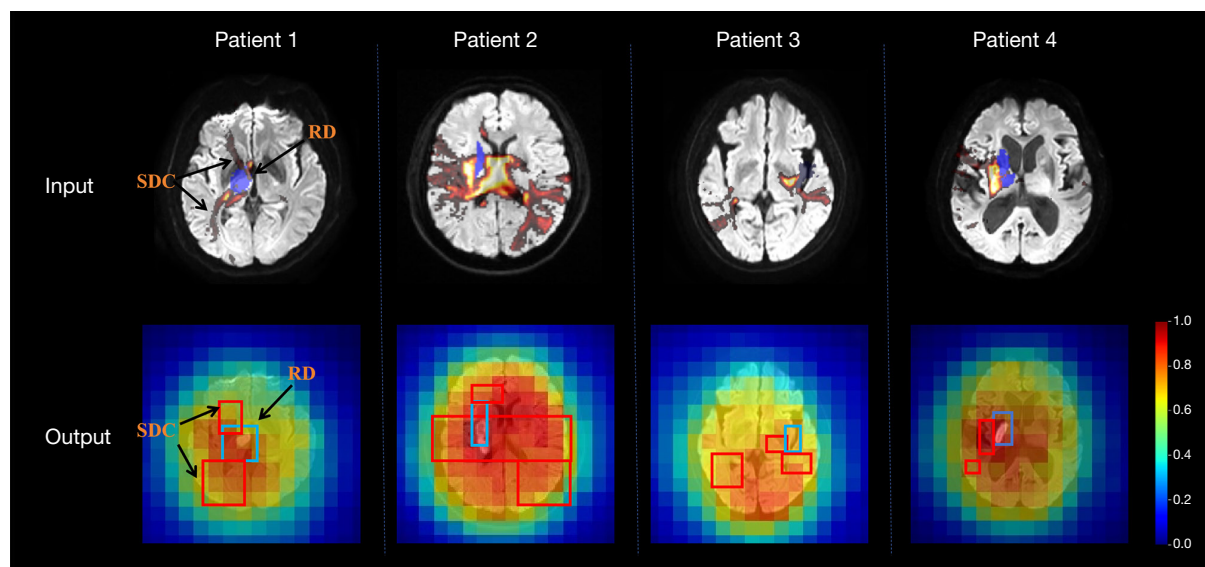


Figure 9 Output of our model visualization using the CAM technique. (I) In the input images, SDC is depicted in red, and RD is shown in blue, overlaying on DWI. (II) In the output images, the red box signifies SDC from the input image, while the blue box indicates RD from the input image. Deeper colors on the heatmap indicate regions with a higher predictive value for PSCI. CAM, class activation mapping; DWI, diffusion-weighted imaging; PSCI, post-stroke cognitive impairment; RD, regional damage; SDC, structural disconnection.

(66.90 ± 9.74 years) than the non-PSCI patients, which is consistent with the findings of a previous study (41). Thus, individuals over 65 years are more likely to develop PSCI. This correlation may be explained by age-related neurodegenerative processes, such as the reduction in cortical neurons and axons, which accelerate the onset and progression of PSCI. In relation to brain network damage, a comparison between the PSCI and non-PSCI patients revealed that RD mainly involved the caudate nucleus, thalamus, and putamen (Figure 3F), while SDC mainly involved the parietopontine tract, SCP, rubrospinal tract, and CST_L (Figure 4). Notably, Maeshima and Osawa (42) provided complementary insights. They suggested that the cognitive dysfunction may result from disrupted cortical pontocerebellar tract fibers and damage to the brainstem reticular regulatory system. Further, our analysis showed that RD and SDC predominantly affected the left hemisphere of the PSCI patients (Figure 3), potentially reflecting limitations in the MoCA's sensitivity to right-hemisphere stroke-induced cognitive impairments (43). Future studies should seek to augment the MoCA using additional tools to enhance diagnostic accuracy and comprehensiveness.

SDC and RD serve as discriminative factors in predicting PSCI. Zhou *et al.* (37) reported strong correlations

between specific radiomics features and PSCI; however, their classification performance remained suboptimal. In our study, the models that used SDC ($\text{SEN} = 0.564 \pm 0.108$) and RD ($\text{SEN} = 0.589 \pm 0.137$) outperformed the radiomics model ($\text{SEN} = 0.386 \pm 0.092$) (Figure 7). The deep learning-based RD models showed similar classification performance to the clinical data, but had smaller error ranges, indicating more stable model training. Additionally, quantitative evaluations showed that SDC and RD were more objective than the clinical data.

The effective use of SDC and RD offers a promising approach for enhancing the diagnosis of PSCI. Traditional symptom mapping methods (38), such as those based on the $\text{SDC}_{\text{score}}$ and RD_{score} , typically focus on commonalities across patients. However, these methods often overlook individual differences that are crucial for accurately predicting PSCI. Consequently, the models derived from traditional symptom mapping demonstrated suboptimal performance (Table 5). Conversely, our model adeptly captured both the distinctions and similarities of 3D SDC and RD across various patients, as evidenced by the visualized model output (Figure 9). Moreover, 3D SDC and RD can effectively extract high-dimensional structural and regional information from the brain, which is essential for understanding the complexities of PSCI.

However, the models that relied solely on SDC and RD underperformed compared to those that used the original DWI data (*Table 2*), emphasizing the limitations of using these features independently. The integration of DWI data significantly improved model performance, underscoring the importance of including global information for accurate predictions. Our (early fusion) method outperformed the late fusion and score fusion methods (*Table 3*), likely due to the unique characteristics of the SDC and RD data. When lesion areas are small, the majority of pixel values in SDC and RD are zero, which increases feature redundancy in late fusion and score fusion methods, thereby reducing model performance. Conversely, early fusion acts as a form of data augmentation, enabling the model to better adapt to the task and enhancing overall performance from the model's insights. Future research should seek to optimize the integration of SDC, RD, and DWI data to further improve the accuracy and diagnostic capabilities of PSCI classification models.

Notably, our model only required DWI images for prediction, eliminating the need for additional data such as clinical information. This characteristic enhances the feasibility of our model for clinical applications, making it a practical choice for widespread use. To mitigate potential overfitting, we employed shallower network architectures and data augmentation techniques. Five-fold cross-validation was used to assess the stability and robustness of the model. Following validation, our model demonstrated superior performance on limited samples. This success highlighted the model's potential as a valuable reference for future training on larger, multi-center datasets, which could enhance its generalizability.

Incorporating regional and structural damage into post-stroke care enables a more sensitive and specific characterization of PSCI based on the severity of brain network damage (44). It advances our understanding of the underlying mechanisms of PSCI, helping clinicians and researchers to better describe the effects of stroke lesions on brain structure and function. Beyond acute stroke care, neuroimaging advancements also inform functional cognitive rehabilitation strategies for chronic injuries. For patients at risk of cognitive impairment, such knowledge empowers healthcare professionals to schedule follow ups, inform patients and caregivers about potential cognitive decline, and develop proactive care plans, bridging the gap between acute care and long-term recovery.

The study had several limitations. First, the exclusion of patients with lesions confined to white matter might

have introduced bias, as significant variations in lesion characteristics and their effects on diagnostic accuracy might have been overlooked. To address this limitation, studies should be conducted with more diverse patient cohorts. Second, while age was not included as a feature in our model, the potential effect of age differences between the PSCI and non-PSCI groups warrants attention. Changes in brain network at different ages could offer a new perspective on PSCI prediction. Thus, further research should explore how age-related brain network variations affect PSCI. Third, the attention weights indicated by CAM images do not always directly correspond to the RD and SDC zones in all patients. This discrepancy (45) arises from the relatively coarse spatial resolution of the final convolutional layer, which might cause the CAM technique to highlight broader regions. Moreover, the relatively small size of our dataset represents another study limitation. Despite using a shallower deep-learning model, the model still contains millions of parameters, which increases the risk of overfitting, especially in our small dataset. The use of larger datasets would improve the model's generalizability and robustness. Finally, the manual segmentation of lesion masks is a time-intensive process. To streamline this aspect of the workflow and enhance efficiency, there is a need to advance the development of automated segmentation networks tailored to the complexities of brain structure.

Conclusions

Our model effectively predicted PSCI 3 months after ALS, surpassing the performance of models based on clinical data, radiomics, and traditional predictive factors. The integration of the SDC and RD images played a critical role in enhancing the model's accuracy in PSCI prediction. This approach not only strengthened the model's predictive capability but also offered deeper insights into the mechanisms underlying PSCI.

Acknowledgments

None.

Footnote

Reporting Checklist: The authors have completed the TRIPOD+AI reporting checklist. Available at <https://qims.amegroups.com/article/view/10.21037/qims-24-2010/rc>

Funding: This work was supported in part by the National Natural Science Foundation of China (grant No. 62371449); the Department of Science and Technology of Shandong Province (grant No. SYS202208); the Science Fund for Distinguished Young Scholars of Jiangsu Province (grant No. JSSA2024YB03); and the Cross Disciplinary Project of Soochow University (project No. P112211023).

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-24-2010/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the Medicine Ethics Committee of The First Affiliated Hospital of Soochow University (No. 2021359), and the requirement for written informed consent was waived due to the retrospective nature of the study.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Pasi M, Cordonnier C. Clinical Relevance of Cerebral Small Vessel Diseases. *Stroke* 2020;51:47-53.
2. Rost NS, Brodtmann A, Pase MP, van Veluw SJ, Biffi A, Duering M, Hinman JD, Dichgans M. Post-Stroke Cognitive Impairment and Dementia. *Circ Res* 2022;130:1252-71.
3. Li T, Ye M, Yang G, Diao S, Zhou Y, Qin Y, Ding D, Zhu M, Fang Q. Regional white matter hyperintensity volume predicts persistent cognitive impairment in acute lacunar infarct patients. *Front Neurol* 2023;14:1265743.
4. Sivakumar L, Kate M, Jeerakathil T, Camicioli R, Buck B, Butcher K. Serial montreal cognitive assessments demonstrate reversible cognitive impairment in patients with acute transient ischemic attack and minor stroke. *Stroke* 2014;45:1709-15.
5. El Husseini N, Katzan IL, Rost NS, Blake ML, Byun E, Pendlebury ST, Aparicio HJ, Marquine MJ, Gottesman RF, Smith EE; American Heart Association Stroke Council; Council on Cardiovascular and Stroke Nursing; Council on Cardiovascular Radiology and Intervention; Council on Hypertension; and Council on Lifestyle and Cardiometabolic Health. Cognitive Impairment After Ischemic and Hemorrhagic Stroke: A Scientific Statement From the American Heart Association/American Stroke Association. *Stroke* 2023;54:e272-91.
6. Ball EL, Shah M, Ross E, Sutherland R, Squires C, Mead GE, Wardlaw JM, Quinn TJ, Religa D, Lundström E, Cheyne J, Shenkin SD. Predictors of post-stroke cognitive impairment using acute structural MRI neuroimaging: A systematic review and meta-analysis. *Int J Stroke* 2023;18:543-54.
7. Mijajlović MD, Pavlović A, Brainin M, Heiss WD, Quinn TJ, Ihle-Hansen HB, et al. Post-stroke dementia - a comprehensive review. *BMC Med* 2017;15:11.
8. Moulton E, Valabregue R, Piotin M, Marnat G, Saleme S, Lapergue B, Lehericy S, Clarençon F, Rosso C. Interpretable deep learning for the prognosis of long-term functional outcome post-stroke using acute diffusion weighted imaging. *J Cereb Blood Flow Metab* 2023;43:198-209.
9. Liu S, Zhang B, Fang R, Rueckert D, Zimmer VA. Dynamic graph neural representation based multi-modal fusion model for cognitive outcome prediction in stroke cases. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*; Springer, Cham; 2023:338-47.
10. Reber J, Hwang K, Bowren M, Bruss J, Mukherjee P, Tranel D, Boes AD. Cognitive impairment after focal brain lesions is better predicted by damage to structural than functional network hubs. *Proc Natl Acad Sci U S A* 2021;118:e2018784118.
11. Yoon KJ, Park CH, Rho MH, Kim M. Disconnection-Based Prediction of Poststroke Dysphagia. *AJNR Am J Neuroradiol* 2023;45:57-65.
12. Kasties V, Karnath HO, Sperber C. Strategies for feature extraction from structural brain imaging in lesion-deficit modelling. *Hum Brain Mapp* 2021;42:5409-22.
13. Olafson ER, Jamison KW, Sweeney EM, Liu H, Wang D, Bruss JE, Boes AD, Kuceyeski A. Functional connectome reorganization relates to post-stroke motor recovery and

- structural and functional disconnection. *Neuroimage* 2021;245:118642.
14. Lim JS, Lee JJ, Woo CW. Post-Stroke Cognitive Impairment: Pathophysiological Insights into Brain Disconnectome from Advanced Neuroimaging Analysis Techniques. *J Stroke* 2021;23:297-311.
 15. Pan C, Chen G, Jing P, Li G, Li Y, Miao J, Sun W, Wang Y, Lan Y, Qiu X, Zhao X, Mei J, Huang S, Lian L, Zhu Z, Zhu S. Incremental Value of Stroke-Induced Structural Disconnection in Predicting Global Cognitive Impairment After Stroke. *Stroke* 2023;54:1257-67.
 16. Bowren M, Bruss J, Manzel K, Edwards D, Liu C, Corbetta M, Tranel D, Boes AD. Post-stroke outcomes predicted from multivariate lesion-behaviour and lesion network mapping. *Brain* 2022;145:1338-53.
 17. Salvalaggio A, De Filippo De Grazia M, Zorzi M, Thiebaut de Schotten M, Corbetta M. Post-stroke deficit prediction from lesion and indirect structural and functional disconnection. *Brain* 2020;143:2173-88.
 18. Kuceyeski A, Navi BB, Kamel H, Raj A, Relkin N, Toglietta J, Iadecola C, O'Dell M. Structural connectome disruption at baseline predicts 6-months post-stroke outcome. *Hum Brain Mapp* 2016;37:2587-601.
 19. Griffis JC, Metcalf NV, Corbetta M, Shulman GL. Structural Disconnections Explain Brain Network Dysfunction after Stroke. *Cell Rep* 2019;28:2527-2540.e9.
 20. Fagerholm ED, Hellyer PJ, Scott G, Leech R, Sharp DJ. Disconnection of network hubs and cognitive impairment after traumatic brain injury. *Brain* 2015;138:1696-709.
 21. Lee R, Choi H, Park KY, Kim JM, Seok JW. Prediction of post-stroke cognitive impairment using brain FDG PET: deep learning-based approach. *Eur J Nucl Med Mol Imaging* 2022;49:1254-62.
 22. Binzer M, Hammernik K, Rueckert D, Zimmer VA. Long-term cognitive outcome prediction in stroke patients using multi-task learning on imaging and tabular data. *Predict Intell Med* 2022:137-48.
 23. Hussain T, Shouno H. Explainable Deep Learning Approach for Multi-Class Brain Magnetic Resonance Imaging Tumor Classification and Localization Using Gradient-Weighted Class Activation Mapping. *Information* 2023;14:642.
 24. Yildiz A, Zan H, Said S. Classification and analysis of epileptic EEG recordings using convolutional neural network and class activation mapping. *Biomedical Signal Processing and Control* 2021;68:102720.
 25. Yushkevich PA, Yang Gao, Gerig G. ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images. *Annu Int Conf IEEE Eng Med Biol Soc* 2016;2016:3342-5.
 26. Rorden C, Bonilha L, Fridriksson J, Bender B, Karnath HO. Age-specific CT and MRI templates for spatial normalization. *Neuroimage* 2012;61:957-65.
 27. Griffis JC, Metcalf NV, Corbetta M, Shulman GL. Lesion Quantification Toolkit: A MATLAB software tool for estimating grey matter damage and white matter disconnections in patients with focal brain lesions. *Neuroimage Clin* 2021;30:102639.
 28. Yeo BT, Krienen FM, Sepulcre J, Sabuncu MR, Lashkari D, Hollinshead M, Roffman JL, Smoller JW, Zöllei L, Polimeni JR, Fischl B, Liu H, Buckner RL. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J Neurophysiol* 2011;106:1125-65.
 29. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo XN, Holmes AJ, Eickhoff SB, Yeo BTT. Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI. *Cereb Cortex* 2018;28:3095-114.
 30. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition*; 2016:770-8.
 31. Talo M, Yildirim O, Baloglu UB, Aydin G, Acharya UR. Convolutional neural networks for multi-class brain disease detection using MRI images. *Comput Med Imaging Graph* 2019;78:101673.
 32. Mehnatkesh H, Jalali S M J, Khosravi A, Nahavandi S. An intelligent driven deep residual learning framework for brain tumor classification using MRI images. *Expert Syst Appl* 2023;213:119087.
 33. Wong TT. Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation. *Pattern Recogn* 2015;48:2839-46.
 34. Kingma DP, Ba J. Adam: A method for stochastic optimization. *arXiv preprint* 2014;arXiv:1412.6980.
 35. Seeland M, Mäder P. Multi-view classification with convolutional neural networks. *PLoS One* 2021;16:e0245230.
 36. Ji W, Wang C, Chen H, Liang Y, Wang S. Predicting post-stroke cognitive impairment using machine learning: A prospective cohort study. *J Stroke Cerebrovasc Dis* 2023;32:107354.
 37. Zhou Y, Wu D, Yan S, Xie Y, Zhang S, Lv W, Qin Y, Liu Y, Liu C, Lu J, Li J, Zhu H, Liu WV, Liu H, Zhang G, Zhu W. Feasibility of a Clinical-Radiomics Model to Predict

- the Outcomes of Acute Ischemic Stroke. *Korean J Radiol* 2022;23:811-20.
38. Pan C, Li G, Jing P, Chen G, Sun W, Miao J, Wang Y, Lan Y, Qiu X, Zhao X, Mei J, Huang S, Lian L, Wang H, Zhu Z, Zhu S. Structural disconnection-based prediction of poststroke depression. *Transl Psychiatry* 2022;12:461.
 39. Larson MG. Analysis of variance. *Circulation* 2008;117:115-21.
 40. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. 2016 IEEE Conference on Computer Vision and Pattern Recognition; Las Vegas, NV; 2016;2921-9.
 41. Ma Y, Yang Y, Wang X, Huang Y, Nan J, Feng J, Yan F, Han L. Prevalence and Risk Factors of Poststroke Cognitive Impairment: A Systematic Review and Meta-Analysis. *Public Health Nurs* 2025;42:1047-59.
 42. Maeshima S, Osawa A. Cognitive impairment caused by subcortical lesion. *J Rehabil Neurosci* 2019;19:3-9.
 43. Chan E, Altendorff S, Healy C, Werring DJ, Cipolotti L. The test accuracy of the Montreal Cognitive Assessment (MoCA) by stroke lateralisation. *J Neurol Sci* 2017;373:100-4.
 44. Holguin JA, Margetis JL, Narayan A, Yoneoka GM, Irimia A. Vascular Cognitive Impairment After Mild Stroke: Connectomic Insights, Neuroimaging, and Knowledge Translation. *Front Neurosci* 2022;16:905979.
 45. Jiang PT, Zhang CB, Hou Q, Cheng MM, Wei Y. LayerCAM: Exploring Hierarchical Class Activation Maps for Localization. *IEEE Trans Image Process* 2021;30:5875-88.

Cite this article as: Bai C, Leng Y, Xiao H, Li L, Cui W, Li T, Dong Y, Zheng J, Cai X. A deep-learning model for predicting post-stroke cognitive impairment based on brain network damage. *Quant Imaging Med Surg* 2025;15(5):3964-3981. doi: 10.21037/qims-24-2010