Appendix 1 Segmenting the brain with CAT12

In the study, the Computational Anatomy Toolbox (CAT12) (http://www.neuro.uni-jena.de/cat/) was adopted for segmenting brain images (17). Specifically, image processing in CAT12 can be separated into a mandatory voxel-based stream and an optional subsequent surface-based processing stream (17). The surface-based processing stream is not applicable in our study. The mandatory voxel-based stream processing stream in CAT12 was adopted for segmenting brain images.

The voxel-based stream consists of two main modules: one for tissue segmentation and another one for spatial registration, resulting in spatially registered gray matter/whiter matter segments (17).

Tissue segmentation (17): The process is initiated by applying a spatially adaptive nonlocal means (SANLM) denoising filter (45), followed by SPM's standard unified segmentation (46). The resulting output serves as a starting point for further optimizations and CAT's tissue segmentation steps: first, the brain is parcellated into the left and right hemispheres, subcortical areas, ventricles, and cerebellum. In addition, local white matter hyperintensities are detected (to be later accounted for during the spatial registration and the optional surface processing). Second, a local intensity transformation is performed to reduce effects of higher gray matter intensities in the motor cortex, basal ganglia, and occipital lobe. Third, an adaptive maximum a posteriori (AMAP) segmentation is applied which does not require any a priori information on the tissue probabilities (47). The AMAP segmentation also includes a partial volume estimation (48).

Spatial registration (17): Geodesic Shooting (49) is used to register the individual tissue segments to standardized templates in the ICBM 2009c Nonlinear Asymmetric space (MNI152NLin2009cAsym; https://www.bic.mni.mcgill.ca/ ServicesAtlases/ ICBM152NLin2009), hereafter referred to as MNI space. While MNI space is also used in many other software packages, enabling cross-study comparisons, users may also choose to use their own templates.

The above were cited from study (17).

Appendix 2 Radiomics features

We computed a series of well-defined radiomics features for each brain region based on (19). The 47 radiomics features extracted were shown in *Table S1*. The final 25 features retained were shown in *Table S2*.

Appendix 3 Construction of radMBN

We construct the brain network radMBN on our own dataset based on study (18). The construction process is as follows:

For each individual T1WI scan, the whole brain gray matter image was segmented using the CAT12 toolbox (http://www.neuro.uni-jena.de/cat/)(18).

A series of well-defined radiomics features (N=25), including intensity and texture features, were computed for automated anatomical labeling (AAL) (N=90) derived regions of interest (ROI) (https://github.com/YongLiulab). Radiomics features were normalized across all ROIs using a min-max approach. Each radiomics feature was normalized for all ROIs using the min-max method. Finally a 90×25 node feature matrix is obtained for each subject (18).

Suppose $y_i^r \in R^F$ represents the F dimension feature of the rth brain region for the ith subject, where $\mathrm{X}_i^r = [y_i^1, \cdots, y_i^{R-1}, y_i^{R+1}, \cdots, y_i^R] \in R^{F \times R-1}$ is all the features except for the rth ROI for the ith subject, with F being the dimension of the feature (18). The group sparse radiomics representation network model can alleviate this problem by jointly estimating the non-zero connections among subjects with $l_{2,1}$-norm regularization based on group sparse representation (GSR) (50,51), as shown in Equation (1):

$$\mathrm{W}^r = \arg\min_{\mathrm{w}^r} \sum_{i=1}^{N} \left(\frac{1}{2}\left\|\mathrm{y}_i^r - \mathrm{X}_i^r \mathrm{w}_i^r\right\|_2^2\right) + \lambda \left\|\mathrm{W}^r\right\|_{2,1} \quad [1]$$

where $\mathrm{W}^r = [w_1^r, \cdots, w_i^r, \cdots w_N^r] \in R^{R-1 \times N}$ denotes the relationship between the rth ROI and the other ROIs for all

https://dx.doi.org/10.21037/qims-24-584

subjects, N is the number of subjects, and λ controls the sparsity of the network (18). The brain network constructed by group sparse representation possesses analogous topological patterns (by enforcing similar non-zero or zero connections across subjects) (18). Specifically, for the ith subject, the brain network is formed as $\mathrm{A}_i = [a_i^1, a_i^2, \cdots, a_i^R] \in R^{R \times R}$, where $a_i^r = [\mathrm{W}_i^{r,1}, \cdots, \mathrm{W}_i^{r,r-1}, 0, \mathrm{W}_i^{r,r+1}, \cdots, \mathrm{W}_i^{r,R}] \in R^R$. Then the max-min normalization is implemented column by column. Since the matrix A_i is usually asymmetric, a further symmetric operation $\mathrm{A}_i = \frac{\mathrm{A}_i + \mathrm{A}_i^{\mathrm{T}}}{2}$ can be performed to achieve a symmetric matrix. The final weighted adjacency matrix of the brain network for each of subjects is obtained (18).

Appendix 4 Confidence interval

In statistics, a confidence interval is an interval estimate of some overall parameter for a probability sample. It exhibits the degree to which the true value of this parameter has a certain probability of falling around the measurement, i.e., it gives the degree of confidence in the measured value of the parameter being measured. The concept of confidence intervals is applicable to a variety of contexts, including scatterplot analysis. In scatterplot analysis, confidence intervals can help to understand the range of variations between data points and the extent to which these variations can be considered reliable (52). When plotting *Figure 8*, we set the confidence interval to the default value of 95%.

Appendix 5 ROC curves for three categorization tasks

We plotted ROC curves for the three classification tasks as shown in *Figure S1*. The figure showed the area under the curve for our proposed method and the comparison methods. A larger area under the curve represents a better diagnostic performance of the model.

References

45. Manjón JV, Coupé P, Martí-Bonmatí L, Collins DL, Robles M. Adaptive non-local means denoising of MR images with spatially varying noise levels. Journal of Magnetic Resonance Imaging 2010;31:192-203.
46. Ashburner J, Friston KJ. Unified segmentation. neuroimage 2005;26:839-51.
47. Rajapakse JC, Giedd JN, Rapoport JL. Statistical approach to segmentation of single-channel cerebral MR images. IEEE transactions on medical imaging 1997;16:176-86.
48. Tohka J, Zijdenbos A, Evans A. Fast and robust parameter estimation for statistical partial volume models in brain MRI. Neuroimage 2004;23:84-97.
49. Ashburner J, Friston KJ. Diffeomorphic registration using geodesic shooting and Gauss–Newton optimisation. NeuroImage 2011;55:954-67.
50. Wee C-Y, Yap P-T, Zhang D, Wang L, Shen D. Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification. Brain Structure Function 2014;219:641-56.
51. Yu R, Zhang H, An L, Chen X, Wei Z, Shen D. Connectivity strength-weighted sparse group representation-based brain network construction for M CI classification. Human brain mapping 2017;38:2370-83.
52. Sim J, Reid N. Statistical inference by confidence intervals: issues of interpretation and utilization. Physical Therapy 1999;79:186-95.

https://dx.doi.org/10.21037/qims-24-584

Table S1 47 radiomics features extracted for each brain region, including 14 intensity features and 33 texture features

Intensity features (14)	Textural features (33)
Energy	Autocorrelation
Entropy	Cluster Prominence (CP)
Kurtosis	Cluster Shade
Maximum	Cluster Tendency
Mean	Contrast
Mean Absolute Deviation (MAD)	Correlation
Median	Difference Entropy
Minimum	Dissimilarity
Range	Energy
Root Mean Square (RMS)	Entropy
Skewness	Homogeneity1
Standard Deviation	Homogeneity2
Uniformity	Informational Measure of Correlation 1 (IMC1)
Variance (Var)	Informational Measure of Correlation 2 (IMC2)
	Inverse Difference Moment Normalized (IDMN)
	Inverse Difference Normalized (IDN)
	Inverse Variance
	Maximum Probability
	Sum Average
	Sum Entropy
	Sum Variance
	Variance
	Short Run Emphasis (SRE)
	Long Run Emphasis (LRE)
	Gray Level Nonuniformity (GLN)
	Run Length Nonuniformity (RLN)
	Run Percentage (RP)
	Low Gray Level Run Emphasis (LGLRE)
	High Gray Level Run Emphasis (HGLRE)
	Short Run Low Gray Level Emphasis (SRLGLE)
	Short Run High Gray Level Emphasis (SRHGLE)
	Long Run Low Gray Level Emphasis (LRLGLE)
	Long Run High Gray Level Emphasis (LRHGLE)

https://dx.doi.org/10.21037/qims-24-584

Table S2 25 radiomics features retained, including 7 intensity features and 18 texture features.

Intensity features (7)	Textural features (18)
Energy	Autocorrelation
Kurtosis	Cluster prominence
Maximum	Cluster Shade
Mean Absolute Deviation (MAD)	Cluster Tendency
Minimum	Contrast
Skewness	Correlation
Entropy	Energy
	Entropy
	Homogeneity1
	Informational Measure of Correlation 1 (IMC1)
	Maximum probability
	Sum entropy
	Short run-length emphasis
	Long run-length emphasis
	Gray-level nonuniformity
	Low gray-level run-length emphasis
	High gray-level run-length emphasis
	High gray-level long run-length emphasis

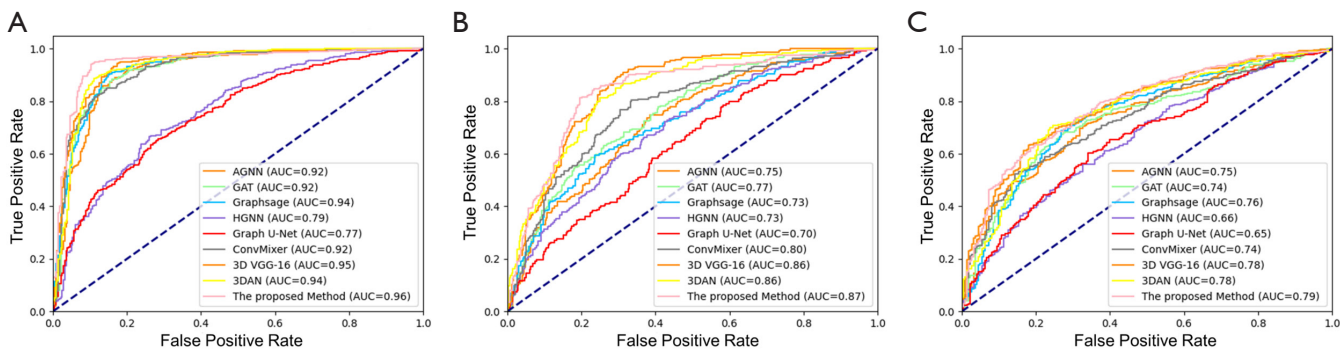

Figure S1 ROC graphs for the three categorization tasks. (A): Classification performance of different methods for AD and NC in the ADNI cohort. (B): Classification performance of different methods for LMCI and EMCI in the ADNI cohort. (C): In the ADNI cohort, pMCI and sMCI were classified based on AD vs. NC classifier. AUC, area under the curve; ROC, receiver operating characteristic; AD, Alzheimer's disease; NC, normal control; pMCI, progressive mild cognitive impairment; sMCI, stable mild cognitive impairment; LMCI, late mild cognitive impairment; EMCI, early mild cognitive impairment; ADNI, Alzheimer's Disease Neuroimaging Initiative database.

https://dx.doi.org/10.21037/qims-24-584