A multispatial information representation model emphasizing key brain regions for Alzheimer's disease diagnosis with structural magnetic resonance imaging

Pengzhi Nan[1], Lin Li[2], Zhiwei Song[1], Yi Wang[1], Chuanzhen Zhu[1], Fang Hu[3], Qiang Zheng[1]

[1]School of Computer and Control Engineering, Yantai University, Yantai, China; [2]Department of Radiology, Yantaishan Hospital Affiliated to Binzhou Medical University, Yantai, China; [3]Key Laboratory of Medical Imaging and Artificial Intelligence of Hunan Province, Xiangnan University, Chenzhou, China

Contributions: (I) Conception and design: P Nan, Z Song; (II) Administrative support: Q Zheng; (III) Provision of study materials or patients: P Nan, Z Song, C Zhu; (IV) Collection and assembly of data: P Nan, Z Song, C Zhu; (V) Data analysis and interpretation: Z Song, C Zhu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscripts: All authors.

Correspondence to: Qiang Zheng, PhD. School of Computer and Control Engineering, Yantai University, No. 30 Qingquan Road, Laishan District, Yantai 264005, China. Email: zhengqiang@ytu.edu.cn.

Background: Structural magnetic resonance imaging (sMRI) can reflect structural abnormalities of the brain. Due to its high tissue contrast and spatial resolution, it is considered as an MRI sequence in diagnostic tasks related to Alzheimer's disease (AD). Thus far, most studies based on sMRI have only focused on pathological changes in disease-related brain regions in Euclidean space, ignoring the association and interaction between brain regions represented in non-Euclidean space. This non-Euclidean spatial information can provide valuable information for brain disease research. However, few studies have combined Euclidean spatial information in images and graph spatial information in brain networks for the early diagnosis of AD. The purpose of this study is to explore how to effectively combine multispatial information for enhancing AD diagnostic performance.

Methods: A multispatial information representation model (MSRNet) was constructed for the diagnosis of AD using sMRI. Specifically, the MSRNet included a Euclidean representation channel integrating a multiscale module and a feature enhancement module, in addition to a graph (non-Euclidean) representation channel integrating a node feature aggregation mechanism. This was accomplished through the adoption of a multilayer graph convolutional neural network and a node connectivity aggregation mechanism with fully connected layers. Each participants' gray-matter volume map and preconstructed radiomics-based morphology brain network (radMBN) were used as MSRNet inputs for the learning of multispatial information. Other than the multispatial information representation in MSRNet, an interactive mechanism was proposed to connect the Euclidean and graph representation channels by five disease-related brain regions which were identified based on a classifier operated on with two feature strategies of voxel intensities and radiomics features. MSRNet focused on disease-related brain regions while integrating multispatial information to effectively enhance disease discrimination.

Results: The MSRNet was validated on four publicly available datasets, achieving accuracies 92.8% and 90.6% for AD in intra-database and inter-database cross-validation, respectively. The accuracy of MSRNet in distinguishing between late mild cognitive impairment (MCI) and early MCI, and between progressive MCI and stable MCI, reached 79.8% and 73.4%, respectively. The experiments demonstrated that the model's decision scores exhibited good detection capability for MCI progression. Furthermore, the potential of decision scores for improving diagnostic performance was exhibited by combining decision scores with other clinical indicators for AD identification.

Conclusions: The MSRNet model could conduct an effective multispatial information representation

in the sMRI-based diagnosis of AD. The proposed interaction mechanism in the MSRNet could help the model focus on AD-related brain regions, thus further improving the diagnostic ability.

Introduction

Alzheimer's disease (AD) is the most common form of dementia, affecting millions of people worldwide (1). As a neurodegenerative brain disease, its effects begin with memory defects and eventually progress to loss of mental function as the disease develops (2). However, there is no effective treatment to cure AD, and the symptoms can only be relieved by drugs or other interventions in the early stage (3,4). Therefore, the early diagnosis of AD is crucial for the timely improvement of patient care (5,6). Mild cognitive impairment (MCI) is a transitional state between normal aging and AD. According to the criteria from fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (7), patients with MCI who progress to AD within the next 3 years can be divided into those with progressive MCI (pMCI) or those with stable MCI (sMCI). Distinguishing between pMCI and sMCI plays an important role in the early diagnosis of AD. However, due to the more subtle differences in cognition and brain structure in patients with MCI (8), the classification of pMCI and sMCI is challenging. Further investigation into distinguishing between these types is highly significant for the study of disease progression and early intervention.

Genetic testing and neuroimaging examination, among other approaches (9), are the common methods for AD diagnosis. Among these, structural magnetic resonance imaging (sMRI), as a noninvasive brain morphometrics method, can capture changes in the brain's anatomical structure and morphological atrophy due to its high contrast and high spatial resolution of soft tissues (10,11). T1-weighted imaging (T1WI), as the most common sequence in sMRI, is relatively easy and cost-effective for acquiring compared to other advanced sequences, such as functional MRI (fMRI) and diffusion tensor imaging (DTI). Therefore, exploring brain disease mechanisms based solely on sMRI may be particularly productive.

In recent years, convolutional neural networks (CNNs) have become a powerful tool in extracting sMRI features for disease diagnosis. One study (12) developed a novel patch-based deep learning network for AD diagnosis which could identify discriminative pathological locations effectively from sMRI. Hu *et al.* (13) proposed a visual geometry group transformer (VGG-Transformer) model that could capture brain atrophy progression features from longitudinal sMRI images which could improve its diagnostic efficacy for MCI. Zhu *et al.* (14) extracted discriminative features from sMRI image blocks and diagnosed AD based on combined feature representation for the whole brain structure, thereby improving diagnostic performance by identifying discriminative pathological locations in sMRI scans. However, these studies mainly extracted atrophic structural features from images in the Euclidean space, ignoring non-Euclidean spatial features in brain networks.

Morphological brain networks (MBNs) characterize the similarity of morphological features derived from sMRI between brain regions, and the robustness and biological basis of a radiomics-based morphology brain network have been demonstrated. One study (15) employed a regional radiomics similarity network (R2SN) to identify the subtypes of MCI, and the associated stratification provided new insights into risk assessment for patients with MCI. Yu *et al.* (16) captured morphological connectivity changes in patients with AD, patients with MCI, and normal control (NC) using individual regional mean connectivity strength (RMCS) from a regional radiomics similarity network. Given that radiomics morphology brain networks can provide novel insights into the mechanisms of brain disease, we speculated that combining Euclidean features of brain atrophy with images with non-Euclidean spatial features in the brain network could achieve superior performance in diagnosing early AD.

The study aimed to verify whether multispatial representational information extracted from gray-matter volume (GMV) in the Euclidean space and radiomics-based morphology brain network (radMBN) in the non-Euclidean space could provide improved performance in distinguishing patients with AD from NC. For this purpose, we developed

Table 1 Demographic information of the participants from the ADNI, AIBL, EDSD, and OASIS databases

Cohort	Group	Age (years)	Sex (M/F)	MMSE
ADNI (N=1,655)	NC (n=603)	73.46±6.16	277/326	29.08±1.10
	MCI (n=770)	72.98±7.68	447/323	27.56±1.81
	AD (n=282)	74.91±7.69	151/131	23.18±2.13
	P value	<0.001	<0.001	<0.001
AIBL (N=557)	NC (n=478)	73.10±6.10	203/275	28.70±1.24
	AD (n=79)	74.16±7.84	33/46	20.41±5.49
	P value	0.172	0.908	<0.001
EDSD (N=388)	NC (n=230)	68.76±6.14	108/122	28.58±2.97
	AD (n=158)	75.54±8.10	66/92	20.89±5.12
	P value	<0.001	0.072	<0.001
OASIS (N=783)	NC (n=594)	67.07±8.70	240/354	29.06±1.54
	AD (n=189)	75.04±7.70	95/94	24.47±4.14
	P value	<0.001	<0.05	<0.001

Significant difference at P<0.001. MMSE and age are expressed as the mean ± standard deviation. ADNI, The Alzheimer's Disease Neuroimaging Initiative; AIBL, Australian Imaging Biomarkers and Lifestyle; EDSD, The European DTI Study on Dementia database; OASIS, The Open Access Series of Imaging Studies; AD, Alzheimer's disease; MCI, mild cognitive impairment; NC, normal control; MMSE, Mini-Mental State Examination; DTI, diffusion tensor imaging; M/F, male/female.

a multispatial information representation model (MSRNet) integrating a Euclidian representation channel and a non-Euclidian representation channel, and validated it on 3,383 participants from four publicly available databases. We then further investigated whether the decision score generated from the MSRNet model could reflect disease progression and contribute to diagnostic performance when combined with clinical indicators.

Methods

Participants

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and included 3,383 participants from four databases: the Alzheimer's Disease Neuroimaging Initiative (ADNI; n=1,655) database (http://adni.loni.usc.edu), the Australian Imaging Biomarkers and Lifestyle (AIBL; n=557) database (http://aibl.csiro.au), the European DTI Study on Dementia (EDSD; n=388) database (http://neugrid4you.eu), and the Open Access Series of Imaging Studies (OASIS; n=783) database (http://oasis-brains.org). In the ADNI cohort, there were significant differences in Mini-Mental State

Examination (MMSE) scores between the NC, MCI, and AD groups [P<0.001, analysis of variance (ANOVA) test]. Similarly, a significant difference in the MMSE score was also observed between the NC and AD groups in the AIBL, EDSD, and OASIS databases (P<0.001, t test). The detailed clinical information is shown in *Table 1*.

Data preprocessing and radMBN construction

Previous studies have shown that gray matter is relevant to the diagnosis of AD. Based on the T1WI images of each participants, we used the CAT12 toolkit (http://dbm.neuro.uni-jena.de/cat/) to segment whole-brain GMV maps (17). The specific segmentation process is described in Appendix 1. All gray-matter images (181×217×181 voxels) were resliced to a size of (91×109×91 voxels) with 2 mm³ isotropic voxels. In this study, the radMBN (18) was constructed on the GMV as follows: (I) we first computed a set of well-defined radiomics features ($N_{radiomics}$=25) for each region of interest (ROI) ($N_{brainregion}$=90) defined by automatic anatomical labeling (AAL), including intensity features and texture features (19). (II) A minimum–maximum (min-max) normalization method was then used to normalize all the radiomics features for each brain region to obtain a 90×25
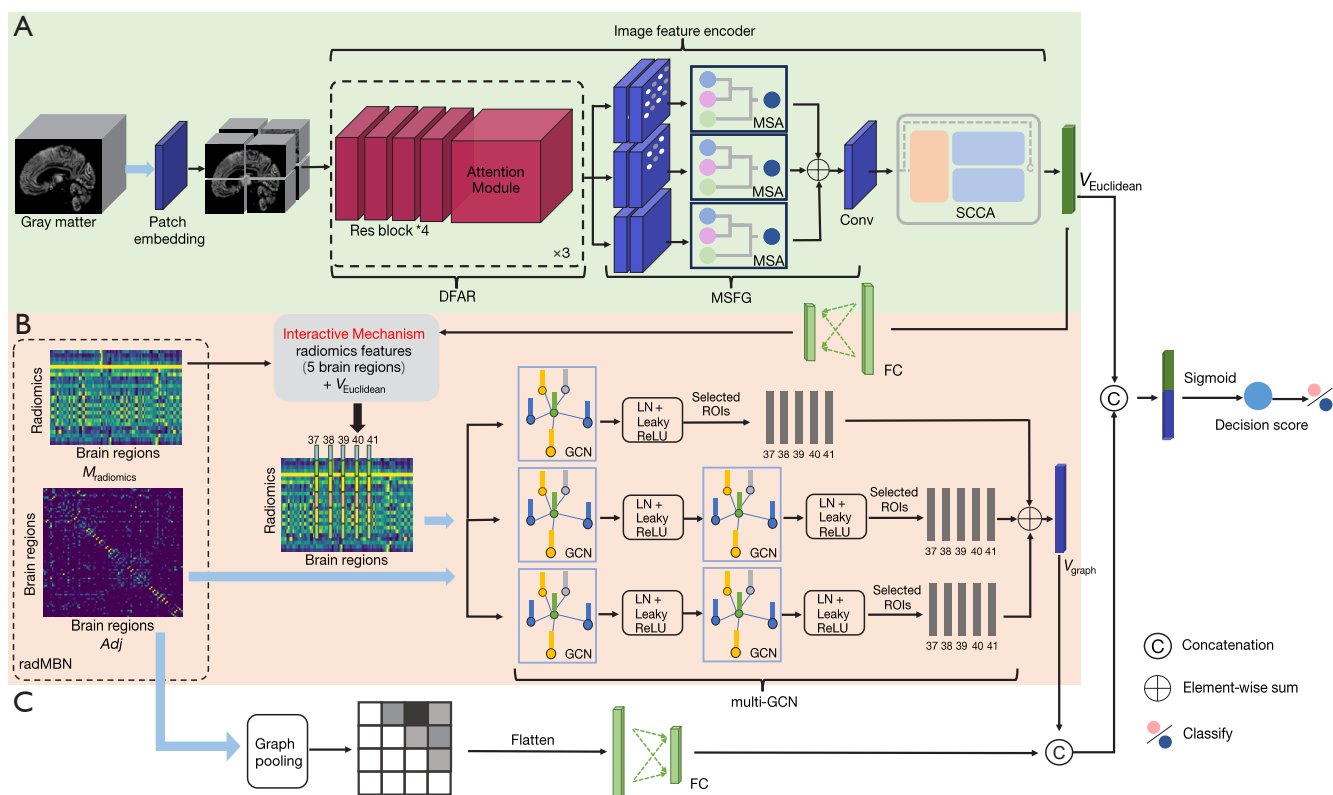
Figure 1 Architecture and explanation of the proposed MSRNet. (A) The Euclidean representation channel based on the gray-matter volume map. (B) The multi-GCN branch of the graph representation channel based on the radMBN. (C) The fully connected layer branch of the graph representation channel. The interactive mechanism connects the Euclidean and graph representation channels in five disease-related brain regions, as identified in *Figure 2*. The blue arrows represent the inputs to the model. DFAR, discriminative features by focusing on atrophic regions; MSA, multiscale convolution self-attention module; MSFG, multiscale module that extracts fine-grained information from images; Conv, convolution layer; SCCA, A feature representation enhancement module that integrated spatial attention, channel attention, and coordinate attention; FC, fully connected layers; radMBN, radiomics-based morphology brain network; ROI, region of interest; LN, layer normalization; ReLU, rectified linear unit; Res block, residual block; GCN, graph convolutional neural network; MSRNet, multispatial information representation model; multi-GCN, multilayer graph convolutional neural network.

matrix of nodal features. (III) The brain network matrix was constructed by calculating the Pearson correlation coefficient (PCC) between the radiomics features between each pair of brain regions. The $l_{2,1}-\text{norm}$ regularization was used to remove redundant connections and unify the network topology during the construction of the brain network matrix, and a 90×90 brain network radMBN was finally obtained for each participant. The details of the radiomics features and radMBN construction are provided in Appendix 2 and Appendix 3.

The proposed multispatial information representation model

In this section, a multispatial information representation

model (MSRNet) for AD diagnosis with sMRI is described. The MSRNet was constructed with a Euclidean representation channel (*Figure 1A*) integrating a multiscale module and a feature enhancement module and with a graph (non-Euclidean) representation channel (*Figure 1B,1C*) integrating a node (brain region) feature aggregation mechanism. This was accomplished via the adoption of a multilayer graph convolutional neural network (multi-GCN) (*Figure 1B*) and a node connectivity aggregation mechanism through the use of fully connected layers (*Figure 1C*). In addition to the multispatial information representation in MSRNet, an interactive mechanism was developed to connect the Euclidean and graph representation channels through five disease-related brain regions which were
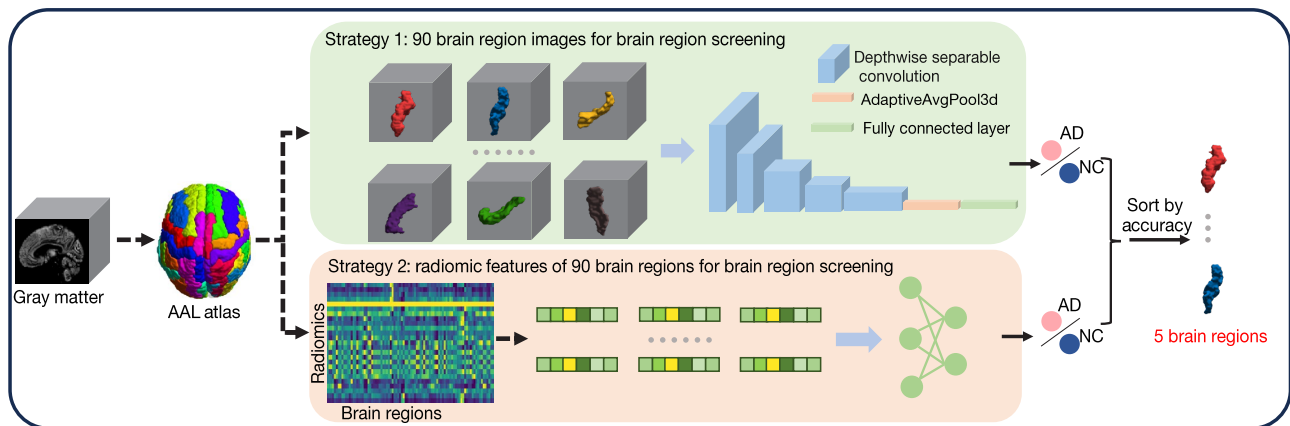
8572

Nan et al. Multispatial model emphasizing key brain regions for AD

Figure 2 Two brain region identification strategies. Strategy 1: the classifier is fed with voxel intensities of the gray-matter volume map. Strategy 2: the classifier is fed with radiomics features. AAL, automatic anatomical labeling. AD, Alzheimer's disease; NC, normal control.
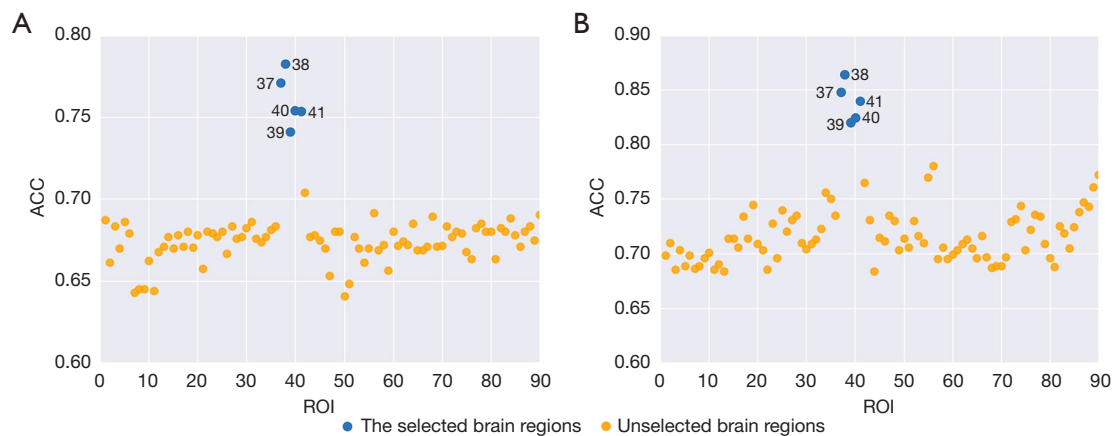

Figure 3 The classification accuracy visualization of the 90 brain regions between the two strategies of (A) voxel intensities and (B) radiomic features. In both strategies, the same five brain regions (blue dots) including the left hippocampus [37], right hippocampus [38], left parahippocampal gyrus [39], right parahippocampal gyrus [40], and left amygdala [41] exhibited superior performance compared to the other methods regardless the classification accuracy threshold. ROI, region of interest; ACC, accuracy.

identified based on a classifier which operated on voxel intensities and radiomics features (*Figures 2,3*).

The Euclidean representation channel in MSRNet

In the Euclidean representation channel, a deep feature encoder was developed for the learning of salient features related to cognitive impairment from the GMV maps. As shown in *Figure 1A*, the image feature encoder is mainly composed of three parts: a feature extraction module for learning discriminative features by focusing on atrophic regions (DFAR), a multiscale module that extracts fine-grained information from images (MSFG), and a feature

representation enhancement module that integrates spatial attention, channel attention, and coordinate attention (SCCA) (20) to enhance the feature representation. The input of the Euclidean representation channel was the GMV map.

The DFAR module for learning discriminative features

The proposed DFAR module in the study is composed of four residual blocks and one attention module. Three DFAR modules are concatenated to better learn disease-related discriminative features from the patches. In contrast to other operations that divide images into patches in advance, we first put the images into patch embedding
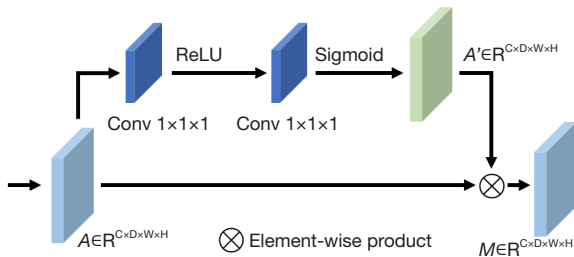
Figure 4 Attention module composition. Conv, convolution layer; ReLU, rectified linear unit.
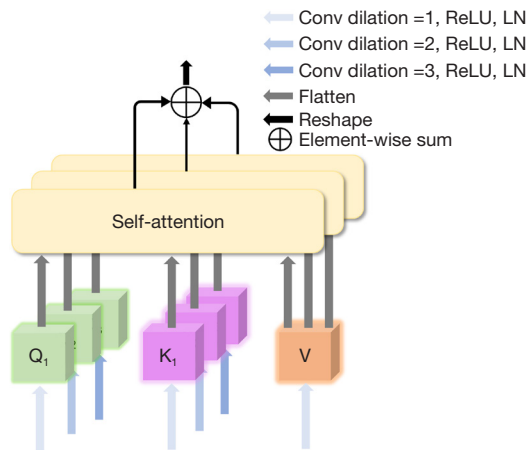

Figure 5 The detailed composition of the multiscale convolution self-attention module (MSA). Conv, convolution layer; ReLU, rectified linear unit; LN, layer normalization.

with patch size p, embedding dimension d via a simple convolution with c input channels and d output channels, kernel size p, and stride p (21); this process is represented as follows:

$$X_0 = F_{patch}(X) = BN\big(\sigma\big(Conv(X, stride = p, kernel_{size} = p)\big)\big) \quad [1]$$

The number of output channels of the residual block in the DFAR module is 128, 128, 256, and 256. The attention module is presented in the *Figure 4*. Specifically, the attention map obtained by the attention module can be formulated by Eq. [2]. With the input of the 1×1×1 convolutional layer being set as $A \in R^{C \times D \times W \times H}$, the calculated attention map $A' \in R^{C \times D \times W \times H}$ can be expressed as follows:

$$A' = Sigmoid\big(f\big(\sigma\big(f(A, \omega_1)\big), \omega_2\big)\big) \quad [2]$$

where $A \in R^{C \times D \times W \times H}$ is the input of the attention module, $A' \in R^{C \times D \times W \times H}$ is the calculated attention feature map,

$Sigmoid(\bullet)$ is the sigmoid function, $f(\bullet)$ is the convolution function, $\sigma(\bullet)$ is the activation function, and ω_1 and ω_2 are the parameters of the two-layer convolution. The output $M \in R^{C \times D \times W \times H}$ of the attention module can be obtained by applying attention to the feature map A via $M = A \otimes A'$, where $\otimes$ is the element-wise product.

The MSFG module for extracting fine-grained information
To integrate multiscale features, an MSFG module was developed, which aggregates a multibranch dilation convolution and a multiscale convolution self-attention module (MSA). The structure of the MSA is shown in *Figure 5*. Due to the fixed size of convolutional kernels and the limited receptive field, traditional convolution fails to capture multiscale feature information. Therefore, we replaced traditional spatial convolutions with dilated convolutions. By setting different dilation rates to expand the convolution kernel's receptive field, the MSA module can integrate multiscale contextual information into the feature map. Subsequently, the multiscale feature maps produced by the MSA module across multiple branches are consolidated and merged, followed by feature aggregation through the spatial convolution layer, and the maps are then inputted to the subsequent feature enhancement module. This process is represented by Eqs. [3-5]:

$$Q_i, K_i = Flatten\big(Conv3d(x, s, r_i)\big) \quad [3]$$

$$V = Flatten\big(Conv3d(x, s)\big) \quad [4]$$

$$Attention(Q_i, K_i, V) = \sum_{i=1}^{n} softmax\left(\frac{Q_i K_i^{T}}{\sqrt{d}}\right)V \quad [5]$$

where x is the input feature map of MSA; $Conv3d(\bullet)$ is the 3D convolution operator; $softmax(\bullet)$ is the softmax function; the step size s of the convolution is set to 1; r_i is the dilated rate, which is set to be 1, 2, or 3, respectively; n is the number of query tokens or key tokens obtained at different scales and is set to 3; and d is the sequence dimension of the input feature map after flattening, which is used to scale the dot product to prevent the size of the dot product from having a large impact on the attention distribution at higher dimensions.

The SCCA module for feature representation enhancement
Directly integrating multiscale features can introduce irrelevant background information, thereby reducing feature representation capability and affecting the accuracy of classification results. Therefore, it is crucial to enhance or suppress the output multiscale feature map regionally. Thus, a feature enhancement module SCCA integrating channel attention, spatial attention, and coordinate attention was
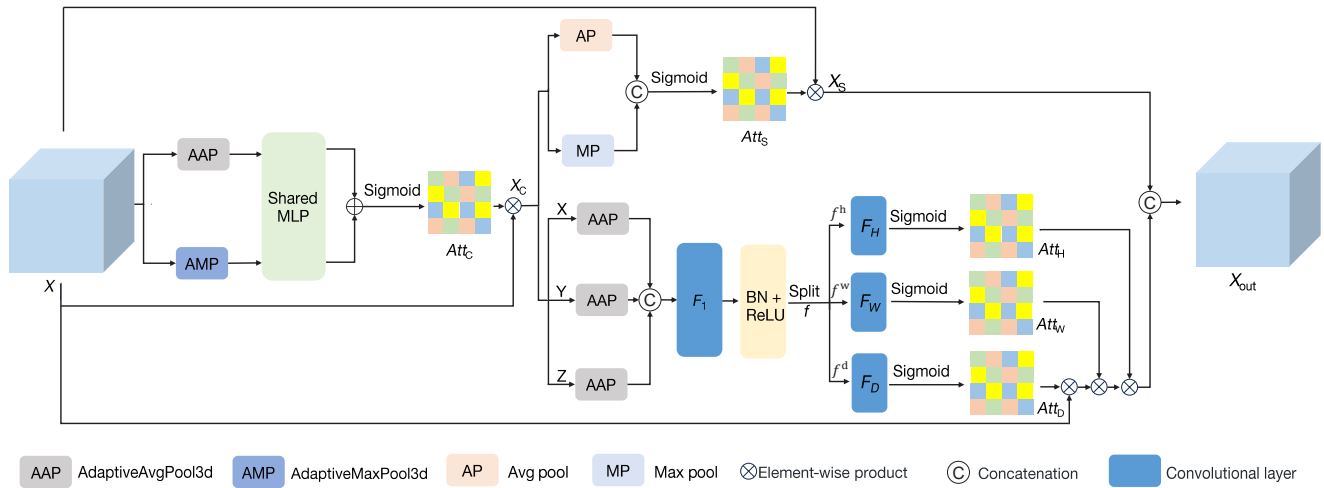
8574

Nan et al. Multispatial model emphasizing key brain regions for AD

Figure 6 The detailed composition of the SCCA. SCCA, A feature representation enhancement module that integrated spatial attention, channel attention, and coordinate attention; MLP, multilayer perceptron; BN, batch normalization; ReLU, rectified linear unit.

included to improve the expression ability of features and suppress unnecessary information. The SCCA module is displayed in *Figure 6*.

Channel attention is adopted via the learning of weights between channel feature maps, with adjustments adaptively being made for the importance among different channels. As shown in *Figure 6*, $X \in R^{C \times D \times W \times H}$ is the input feature map of the channel attention in the module, and its output is X_C. This process is represented by Equation:

$$X_C = Sigmoid\big(MLP\big(AMP(X)\big) + MLP\big(AAP(X)\big)\big) \otimes X \quad [6]$$

where $AMP(\bullet)$ and $AAP(\bullet)$ are the AdaptiveMaxPool3d and AdaptiveAvgPool3d layers, respectively.

The spatial attention mechanism effectively captures important information from different locations in the image. In *Figure 6*, X_C is the channel attention output serving as the input to the spatial attention, and the output of the spatial attention is X_S. This process can be represented by Equation [7]:

$$X_S = Sigmoid\big(Conv\big(Concat\big(AP(X_C), MP(X_C)\big)\big)\big) \otimes X \quad [7]$$

where, $AP(\bullet)$ and $MP(\bullet)$ are the average pooling and maximum (max) pooling, respectively; $Concat(\bullet)$ is the concatenation function; and $Conv(\bullet)$ is the 7×7×7 convolutional layer.

Coordinate attention can enhance or suppress feature representations at specific positions based on pixels. The application of adaptive average pooling along three directions to the output X_C of channel attention and the

passing of the result through convolution and normalization operations can be represented as follows:

$$f = ReLU\Big(BN\Big(F_1\big(concat\big(z^h, z^w, z^d\big)\big)\Big)\Big) \quad [8]$$

where, $F_1(\bullet)$ is the convolution function, $BN(\bullet)$ is the batch normalization layer, $f \in R^{\frac{C}{r \times 1 \times 1 \times (D+W+H)}}$ is the intermediate feature map encoding the spatial information, and r is the reduction ratio for controlling the size of the channel dimension. Along the spatial dimension, f is then split into three independent tensors, f^d, f^w, and f^h, and transformed into tensors of the same channel as X_C with three independent 1 × 1 × 1 convolutions, F_D, F_W, F_H, yielding the following:

$$Att_D, Att_W, Att_H = \sigma\big(F_D\big(f^d\big)\big), \sigma\big(F_W\big(f^w\big)\big), \sigma\big(F_H\big(f^h\big)\big) \quad [9]$$

where $\sigma(\bullet)$ is the sigmoid activation function, and the outputs Att_D, Att_W, and Att_H are the attention weight maps as follow:

$$f_{out} = X \times Att_D \times Att_W \times Att_H \quad [10]$$

where f_{out} is the output with coordinate attention applied to X, and the output X_{out} of SCCA is obtained by concatenating f_{out} and X_S. The image feature encoder ultimately maps the learned Euclidean space features into a one-dimensional feature vector $V_{Euclidean}$.

The graph (non-Euclidean) representation channel in MSRNet

The graph representation channel was implemented by

integrating a node feature aggregation mechanism by adopting a multi-GCN and a node connectivity aggregation mechanism that uses fully connected layers. The input of the graph representation channel is the brain network radMBN and includes a radiomics feature matrix for all brain regions and a connectivity matrix between brain regions. Each node represents a brain region with radiomics feature dimension n=25. The connectivity indicates the edges in the brain connectivity map but weights with a strength value between brain regions that play a crucial role in interregional communication within the brain.

Regarding the node feature aggregation, the multi-GCN was implemented with three different GCN settings. The first channel employs one GCN layer, mapping the node feature dimension from 25 to 64. The second channel employs two GCN layers, mapping the node feature dimension from 25 through 50 to 64. The third channel also employs two GCN layers but maps the node feature dimension from 25 through 12 to 64. Each GCN layer is followed by a layer normalization layer and a leaky rectified linear unit (Leaky_ReLU) layer.

Regarding the node connectivity aggregation, a pooling layer is first applied to filter out nodes with relatively small degrees and the associated connections, and then the upper triangular elements of the pooled adjacency matrix are flattened into a one-dimensional vector and sent to a fully connected layer. The fully connected layer maps the vector into a relatively low-dimensional representation.

The interactive mechanism between Euclidean and graph representations in MSRNet

In addition to the multispatial information representation in MSRNet, an interactive mechanism was also developed to further enhance the feature representation via the connection of the Euclidean and graph representation branches. The interactive mechanism was established based on five disease-related brain regions which were identified based on a classifier that operates on two feature strategies: voxel intensities and radiomic features (*Figures 2,3*). In the following section, the method for identifying the five disease-related brain regions is described, which is followed by and explanation of the mechanism of interaction between the Euclidean and graph representations.

The identification of the five disease-related brain regions

In this study, the brain region identification was implemented based on an AD *vs.* NC classifier which was fed with each brain region for classification accuracy comparisons. Specifically, a total of 90 brain regions were

fed into the AD *vs.* NC classifier based on the AAL atlas, and two different feature strategies were adopted for each brain region in the study. The first strategy is based on the voxel intensities of the GMV map patch, with each patch corresponding to a specific brain region, while the second strategy involves calculating the radiomics features for each brain region.

In strategy 1, to reduce the influence of spatial location and redundant background knowledge, we translated the segmented brain regions to the middle of the image and cutoff the surrounding redundant voxels. Given the input image, $X \in R^{1 \times D \times W \times H}$, strategy 1 could be formulated as follows,

$$R_i = Crop\left(Translate\left(X \odot Mask_i\right)\right), \quad i = 1, 2, \cdots, N \qquad [11]$$

$$m_i = Classify\left(R_i\right), \quad i = 1, 2, \cdots, N \qquad [12]$$

where $\odot$ is the Hadamard product, $Msak_i$ is the binarization mask of the ith brain region obtained from the AAL atlas, $R_i \in R^{1 \times P \times P \times P}$ has P=64, N is the number of brain regions. As shown in *Figure 2*, $Classify(\bullet)$ is a simple classifier consisting of five depth-wise separable convolutional layers, one AdaptiveAvgPool3d layer, and one fully connected layer. m_i were considered to be the classification accuracy value for each brain region.

In strategy 2, the radiomic features of each brain region were fed into the classifier which was established by a linear layer.

It should be noted that each feature strategy generated 90 classification accuracy values. The classification accuracy values were then ranked to identify the most relevant brain regions. *Figure 3* displays the classification accuracy of the 90 brain regions for the two strategies of voxel intensities and radiomic features. In both strategies, the same five brain regions (number 37, 38, 39, 40, and 41, indicated with blue dots in *Figure 3*) of the left hippocampus, right hippocampus, left parahippocampal gyrus, right parahippocampal gyrus, and left amygdala exhibited outstanding performance compared to other regions regardless of the classification accuracy threshold.

The interactive mechanism between the Euclidean and graph representations

In our study, the Euclidean representation channel and the graph representation channel were not separate but interactive. The Euclidean representation channel (*Figure 1A*) was fed with the GMV map of the T1WI image and generated a feature vector $V_{\text{Euclidean}}$ (*Figure 1A*). The graph representation channel was fed with the brain network radMBN and

8576

Nan et al. Multispatial model emphasizing key brain regions for AD

generated a feature vector V_{graph} (*Figure 1B*). The brain network could be represented by a radiomic feature matrix $M_{\text{radiomics}}$ for all nodes and a connectivity matrix between all nodes.

The interactive mechanism was based on the five brain regions identified above. Specifically, the feature vector $V_{\text{Euclidean}}$ (*Figure 1A*) from the image encoder branch was used to enhance the node features of the brain network radMBN that fed into the GCN branch. Given each node represents a brain region, the feature vector $V_{\text{Euclidean}}$ was only used to enhance the features of the five identified brain regions by the following additional operation $M_{\text{radiomics}}(node\_i) = M_{\text{radiomics}}(node\_i) + V_{\text{Euclidean}}$, where $node\_i$ is from the five brain regions.

In addition to the interactive mechanism described above, the feature vector V_{graph} (*Figure 1B*), which was the output of the GCN branch after graph convolution, was constructed based solely on the node features from the five identified brain regions, not from all the brain regions.

Experimental settings

Comparison of diagnostic performance of MSRNet with that of other methods

We validated the superiority of the proposed MSRNet through a series of intra- and inter-database experiments on AD *vs.* NC classification, late MCI (LMCI) *vs.* early MCI (EMCI) classification, and pMCI *vs.* sMCI classification. The methods under comparison included attention-based graph neural network (AGNN) (22), graph attention network (GAT) (23), GraphSAGE (24), hypergraph neural network (HGNN) (25), and Graph U-Net (26) with radMBN as input and ConvMixer (21), 3D visual geometry group 16 (3D VGG-16) (27), and 3D attention network (3DAN) (28) with GMV map as input. In the comparison experiments, 10-fold cross-validation was adopted.

Analysis of progress trajectories of decision scores and clinical indicators

To validate the sensitivity of the proposed MSRNet to MCI progression, a longitudinal trajectory analyses of MMSE, Alzheimer's Disease Assessment Scale 13 (ADAS13), and the decision score which was generated by the MSRNet were performed on individuals with sMCI or pMCI from the ADNI cohort. For participants with sMCI, visiting status remained stable over the time points, so the baseline was set as the origin of time to progression. Specifically, the sMCI and pMCI data were fed separately into the AD and NC

classifiers trained in the ADNI cohort, and a decision score was generated for each participant. The decision scores, MMSE, and ADAS13 were normalized using the max-min normalization method. The "regplot" function of the Seaborn library (29) was used to visualize the progression trajectory. A linear regression model was used to fit the data. The reliability of the results was ensured by adding a 95% confidence interval (CI, an explanation of the confidence intervals is available in Appendix 4).

Contribution of the decision score to clinical indicators in diagnostic performance

To validate the contribution of decision scores to clinical indicators in the diagnosis of AD, four clinical indicators were considered in this experiment including age, apolipoprotein E4 (APOE4), ADAS13, and MMSE. Due to the limited availability of clinical indicators, 269 individuals with AD and 411 NC were selected in the ADNI cohort. To minimize performance bias due to imbalanced category data, instances were synthesized using the synthetic minority oversampling technique (SMOTE) (30) algorithm, which balanced the amount of data across categories in a 1:1 ratio.

For the AD *vs.* NC classification, we enumerated all the clinical indicator combinations and added the decision score separately as input. The area under the curve (AUC) values were obtained to evaluate the classification performance based on six machine learning methods including Bayesian network, support vector machine, logistic regression, k-nearest neighbor, decision tree, and random forest. A 10-fold cross-validation was adopted in experiments.

Ablation study

An ablation study was also implemented to validate the superiority of the multispatial information representation over single-Euclidean or single-graph representation and to validate the effectiveness of the interaction mechanism between the Euclidean and graph representation channels in improving the AD *vs.* NC classification.

Experimental hyperparameter settings

All experiments were performed with Python 3.9 (Python Software Foundation, Wilmington, DE, USA) and Pytorch 1.7.1. The specific parameters were as follows: the dropout rate was 0.3, the batch size was set to 2 to fit the graphics processing unit (GPU) memory, the initial learning rate was set to 5×10^{-6}, and the number of epochs was set to 300. To prevent overfitting during training, we employed early

stopping with a patience value of 10, which terminated training if the validation performance did not improve for 10 consecutive epochs.

Evaluation metrics

In order to validate the performance of the proposed

MSRNet, multiple evaluation metrics, based on a 10-fold cross-validation strategy, were adopted for evaluation, including accuracy (ACC), sensitivity (SEN), specificity (SPE), and AUC.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad [13]$$

$$Sensitivity = \frac{TP}{TP + FN} \quad [14]$$

$$Specificity = \frac{TN}{TN + FP} \quad [15]$$

where TP, TN, FP, and FN are true positive, true negative, false positive, and false negative, respectively.

Results

Comparison of diagnostic performance

Classification of AD vs. NC

We compared the classification performance of the proposed MSRNet with other existing methods for AD vs. NC within the ADNI database [*Table 2* and *Figure 7A*; receiver operating characteristic (ROC) curve plots of all methods under comparison are provided in Appendix 5 and Figure S1A]. When comparing single spatial information methods, we exclusively used either radMBN or GMV as the input. *Table 2* summarizes the experimental results for the classification of AD vs. NC. The results of the image-based methods (21,27,28) were better than those of the brain network-based methods
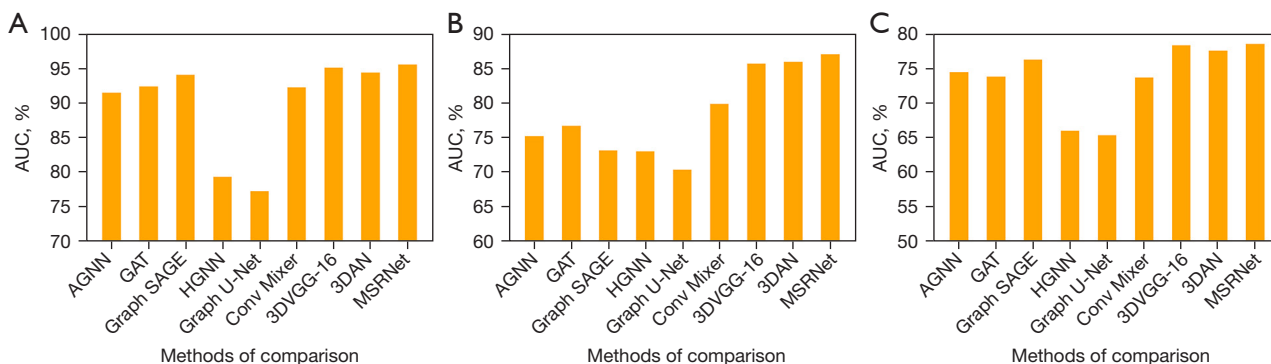
Table 2 Comparison of AD and NC classification performance between the proposed MSRNet and other methods in the ADNI cohort

Method	ACC (%)	SEN (%)	SPE (%)	AUC (%)
AGNN	88.1	75.0	94.1	91.5
GAT	86.5	71.1	93.7	92.4
GraphSAGE	88.9	78.3	93.8	94.1
HGNN	74.9	34.2	94.0	79.3
Graph U-Net	73.8	38.5	90.3	77.2
ConvMixer	86.5	76.6	91.1	92.3
3D VGG-16	90.6	81.8	94.7	95.1
3DAN	88.9	79.1	93.1	94.4
Proposed MSRNet	92.8	88.2	95.0	95.6

AD, Alzheimer's disease; NC, normal control; ACC, accuracy; SEN, sensitivity; SPE, specificity; AUC, area under the curve; MSRNet, multispatial information representation model; ADNI, Alzheimer's Disease Neuroimaging Initiative database; AGNN, attention-based graph neural network; GAT, graph attention network; HGNN, hypergraph neural network; 3D VGG-16, 3D visual geometry group 16; 3DAN, 3D attention network.

Figure 7 AUC histograms based on the comparison methods for the three classification 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 the AD *vs.* NC classifier. AUC, area under the curve; AGNN, attention-based graph neural network; GAT, graph attention network; 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.

8578

Nan et al. Multispatial model emphasizing key brain regions for AD

Table 3 Performance comparison of the proposed MSRNet with other models in AD and NC classification based on inter-database cross-validation with the ADNI, AIBL, EDSD, and OASIS databases

Method	ACC (%)	SEN (%)	SPE (%)	AUC (%)
AGNN	85.2	65.6	92.8	89.7
GAT	84.0	61.9	92.6	88.5
GraphSAGE	87.3	72.9	92.9	91.7
HGNN	77.5	35.8	93.9	72.7
Graph U-Net	77.7	29.5	96.5	71.1
ConvMixer	86.2	78.1	88.8	91.1
3D VGG-16	88.8	79.2	91.4	94.6
3DAN	88.3	73.1	94.2	91.7
Proposed MSRNet	90.6	82.0	93.6	93.9

AD, Alzheimer's disease; NC, normal control; ACC, accuracy; SEN, sensitivity; SPE, specificity; AUC, area under the curve; MSRNet, multispatial information representation model; ADNI, Alzheimer's Disease Neuroimaging Initiative database; AIBL, Australian Imaging Biomarkers and Lifestyle; EDSD, European DTI Study on Dementia database; OASIS, Open Access Series of Imaging Studies; DTI, diffusion tensor imaging; AGNN, attention-based graph neural network; GAT, graph attention network; HGNN, hypergraph neural network; 3D VGG-16, 3D visual geometry group 16; 3DAN, 3D attention network.

Table 4 Comparison of LMCI and EMCI classification performance in the ADNI cohort between the proposed MSRNet and other methods

Method	ACC (%)	SEN (%)	SPE (%)	AUC (%)
AGNN	67.2	84.5	42.1	75.2
GAT	71.0	81.6	55.1	76.7
GraphSAGE	68.9	75.1	59.8	73.1
HGNN	65.9	85.7	36.9	73.0
Graph U-Net	58.7	57.3	60.4	70.3
ConvMixer	70.7	81.2	55.4	79.9
3D VGG-16	77.2	82.6	69.2	85.7
3DAN	77.3	75.9	79.4	86.0
Proposed MSRNet	79.8	80.4	78.8	87.1

LMCI, late mild cognitive impairment; EMCI, early mild cognitive impairment; ACC, accuracy; SEN, sensitivity; SPE, specificity; AUC, area under the curve; MSRNet, multispatial information representation model; ADNI, Alzheimer's Disease Neuroimaging Initiative database; AGNN, attention-based graph neural network; GAT, graph attention network; HGNN, hypergraph neural network; 3D VGG-16, 3D visual geometry group 16; 3DAN, 3D attention network.

(22-26) due to the detectable atrophy in the lesion area. Our method integrated morphological variation features extracted from images with non-Euclidean spatial features from brain networks, thereby enhancing the performance (ACC =92.8%, SEN =88.2%, SPE =95.0%, and AUC =0.956).

We further verified the superiority of our method in the ADNI, AIBL, EDSD, and OASIS databases (*Table 3*). *Table 3* lists the inter-database cross-validation results for the AD diagnosis task of our method compared with those of other methods. The average performance of our method on the four databases was as follows: ACC =90.6%, SEN =82.0%, SPE =93.6%, and AUC =0.939.

Classification of LMCI *vs.* EMCI

As a prodromal symptom of AD, MCI involves a less-apparent lesion, making the differentiation between LMCI and EMCI challenging. We selected 356 patients with LMCI and 240 patients with EMCI from the ADNI cohort for experiments. As shown in *Table 4*, the proposed MSRNet achieved and ACC of 79.8% and AUC of 0.871, which were better than those of the other methods (*Table 4* and *Figure 7B*; the ROC plots of all methods are provided in Appendix 5 and Figure S1B). Specifically, the ACC of the proposed MSRNet was 8.8% higher than that of the methods using only radMBN as the input (22-26) and 2.5% higher than that of the methods of using only GMV as the input (21,27,28).

Classification of pMCI *vs.* sMCI

According to the DSM-5 criteria (7), patients with MCI can be classified as progressive MCI if they progress to AD within three years and as stable MCI if they do not. Discriminating between pMCI and sMCI is also challenging due to the short time span that distinguishes them. We selected 226 patients with pMCI and 538 patients with sMCI from the ADNI cohort. By using the AD *vs.* NC classifier (not trained with MCI participants) in the ADNI cohort to distinguish between pMCI and sMCI, we found that the combined representation of Euclidean and non-Euclidean spatial information achieved an ACC of 73.4% (SEN =63.7%, SPE =77.5%, AUC =0.786), which was superior to that of other methods (*Table 5*, *Figure 7C*; the ROC plots of the methods under comparison are provided in Appendix 5 and Figure S1C). As shown in *Table 5*, although our method could outperform the other methods, the improvement was relatively modest when compared with that for the AD *vs.* NC classification (*Tables 2,3*) and that for the LMCI *vs.* EMCI classification (*Table 4*). The reason for

Table 5 Comparison of pMCI and sMCI classification performance in the ADNI cohort between the proposed MSRNet and other methods

Method	ACC (%)	SEN (%)	SPE (%)	AUC (%)
AGNN	72.0	45.7	83.1	74.5
GAT	70.2	46.1	80.3	73.8
GraphSAGE	72.8	57.5	79.3	76.3
HGNN	70.4	22.2	90.6	66.0
Graph U-Net	70.5	23.0	90.2	65.3
ConvMixer	69.3	50.7	77.1	73.7
3D VGG-16	72.1	63.1	75.8	78.4
3DAN	72.5	62.4	76.7	77.6
Proposed MSRNet	73.4	63.7	77.5	78.6

pMCI, progressive mild cognitive impairment; sMCI, stable mild cognitive impairment; ACC, accuracy; SEN, sensitivity; SPE, specificity; AUC, area under the curve; MSRNet, multispatial information representation model; ADNI, Alzheimer's Disease Neuroimaging Initiative database; AGNN, attention-based graph neural network; GAT, graph attention network; HGNN, hypergraph neural network; 3D VGG-16, 3D visual geometry group 16; 3DAN, 3D attention network.

this is that pMCI and sMCI involve more subtle changes in brain structure, rendering it difficult to capture their differences.

Progression trajectories of the groups in relation to decision score, MMSE, and ADAS13

The decision score is the probability and subsequent category that the model assigns to input samples in a classification task. A sigmoid function is used in the model to convert these scores into a probability distribution (31) and ultimately select the category with the highest score as the prediction. Decision scores have been used to demonstrate the clinical biology underlying deep learning models (15,28).

In *Figure 8A*, the data of patients pMCI who had recently progressed to AD are presented as the baseline data (represented as "0" in *Figure 8A*), and the data from participants at 12, 24, and 36 months before and after conversion are also provided. Meanwhile, the corresponding indices of MMSE and ADAS13 are shown in *Figure 8B,8C*.
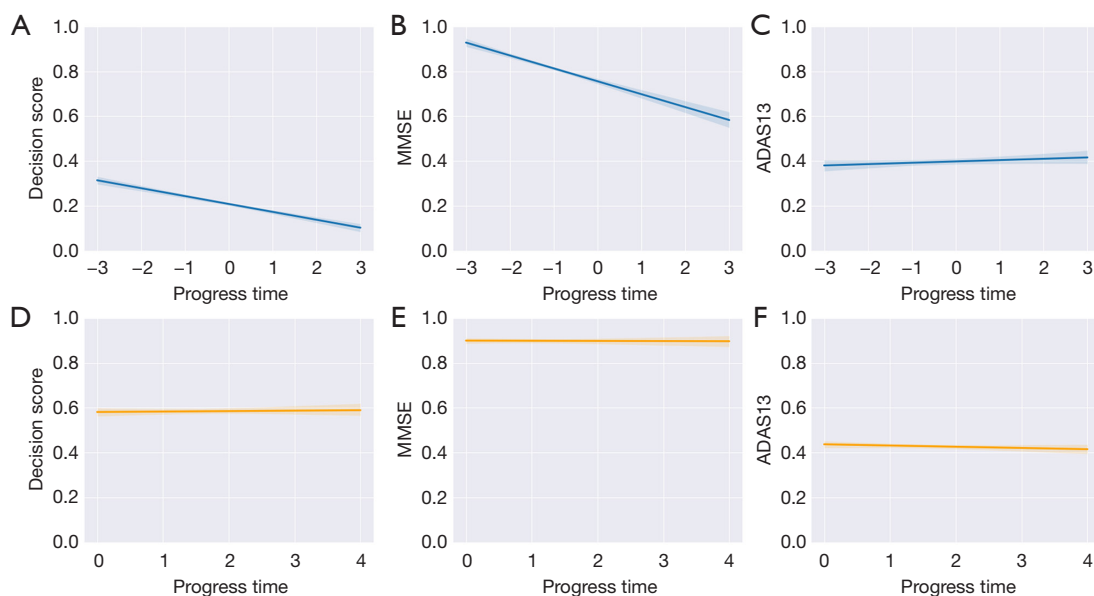

Figure 8 Longitudinal trajectory analysis of MCI progression in the ADNI cohort. (A-C) pMCI progression trajectories of the decision score, MMSE, and ADAS13. (D-F) sMCI progression trajectories of the decision scores, MMSE, and ADAS13. (A-C) In the pMCI group, the decreased confidence of the decision score and MMSE or the increased confidence ADAS13 in pMCI was equivalent to the increase in confidence of AD prediction. (D-F) In the sMCI group, the decision scores, MMSE, and ADAS13 remained relatively stable. The "0" in (A-C) represents the time point of AD onset, "–1", "–2", "–3", "1", "2", and "3" represent data from participants 12, 24, and 36 months before and after onset. The "0", "1", "2", "3", and "4" in (D-F) represent sMCI baseline data and participant data at 12, 24, 36, and 48 months after diagnosis of sMCI. ADAS13, Alzheimer's Disease Assessment Scale 13; MMSE, Mini-Mental State Examination; MCI, mild cognitive impairment; AD, Alzheimer's disease; ADNI, Alzheimer's Disease Neuroimaging Initiative; pMCI, progressive mild cognitive impairment; sMCI, stable mild cognitive impairment.

8580

Nan et al. Multispatial model emphasizing key brain regions for AD

The baseline data of sMCI (represented as "0" in *Figure 8D*) and the data of participants at 12, 24, 36, and 48 months after the diagnosis of sMCI are presented in *Figure 8D*. The corresponding indices of MMSE and ADAS13 are provided in *Figure 8E,8F*.

Clinically, as AD progresses, the patient's cognitive ability (as indicated by measures such as MMSE score) tends to decline (32) while the ADAS13 score tends to increase (33). Based on the data from participants at 12, 24, and 36 months before and after conversion, the trend of declining decision scores aligns with the progression of patients from pMCI to AD. This indicates that changes in decision scores can reflect changes in the patients' condition while being highly consistent with the trajectory of progression of the MMSE and ADAS13. As shown in *Figure 8A*, since the progression trajectory of the decision scores generated by the MSRNet for the pMCI group decreased gradually over time (which was highly consistent with the progression trajectory of the MMSE), the decreased confidence of model prediction in pMCI was equivalent to the increased confidence of model prediction for AD.

Overall, in the pMCI group, the decreased confidence in the decision score and MMSE or the increased confidence in ADAS13 for pMCI was equivalent to the increase in confidence for AD prediction (*Figure 8A-8C*). For the sMCI group, the decision scores, MMSE, and ADAS13 remained relatively stable (*Figure 8D-8F*).

Contribution of decision scores to clinical biomarkers

To further verify whether the decision score obtained by the proposed MSRNet can be used as a general metric to enhance the potential performance in AD diagnosis, we combined four clinical measures, AGE, APOE4, ADAS13, and MMSE, with the decision score and evaluated their combined ability diagnose AD. The experiments were implemented using six machine learning methods including Bayesian network, support vector machine, logistic regression, k-nearest neighbor, decision tree, and random forest. The histogram plots in *Figure 9* indicate that the decision score could provide additional value to the clinical measures of AGE, APOE4, ADAS13, and MMSE in diagnosing AD. The improvement in performance was more pronounced when combining fewer clinical measures. Studies have shown that APOE4 is not only an important mediator of AD susceptibility but may also confer specific phenotypic heterogeneity in AD manifestations (34), making it a promising therapeutic target for AD (35). Based

on the results in *Figure 9*, we observed similar performance when the other features remain unchanged but AOPE4 is replaced by the decision score. These results suggest that the decision score can enhance the diagnosis of AD and may be an easily accessible and universal indicator.

Ablation study

We conducted ablation studies on the AD *vs.* NC classification task to evaluate the effectiveness of the proposed multispatial representation strategy. As shown in *Table 6*, the feature extractor designed to extract Euclidean spatial information from GMV in the MSRNet model achieved promising results on the AD classification task (ACC =91.8%, SEN =87.5%, SPE =93.8%, and AUC =0.955). Compared with other methods that only used GMV as input (*Tables 2,6*), the proposed feature extractor was highly competitive, thereby confirmed the efficacy of the proposed module. The non-Euclidean spatial information extracted from radMBN using multi-GCN achieved an ACC of 89.1%, an SEN of 76.5%, an SPE of 95.0%, and an AUC of 0.949 in the AD *vs.* NC classification task, outperforming other GCN methods (*Tables 2,6*). In addition, the classification performance was improved when further connectivity information between brain regions was added. This suggests that alterations in the strength of connectivity between brain regions are nonnegligible pathological information for disease diagnosis. Subsequently, we introduced the interaction mechanism into the MSRNet model, resulting in performance improvements of 0.6%, 1.7%, and 0.003 for ACC, SEN, and AUC, respectively, as compared to the radMBN + GMV + topology strategy without the interaction mechanism (*Table 6*). This suggests that the interaction mechanism could help the MSRNet integrate multispatial information while focusing on important brain regions, thus improving model performance.

We evaluated the impact of the number of DFAR modules in the Euclidean representation channel and the graph pooling rate in the fully connected channel on model performance. The number of DFAR usage modules was determined by performing the AD *vs.* NC classification task and using only the GMV map as input. The best performance was achieved when the number of DFAR modules was 3, as shown in *Figure 10A*. Moreover, for the AD *vs.* NC classification task, the best results were achieved with a pooling rate of 0.4, with the pooling rate being adjusted from 0.1 to 0.7 (*Figure 10B*).
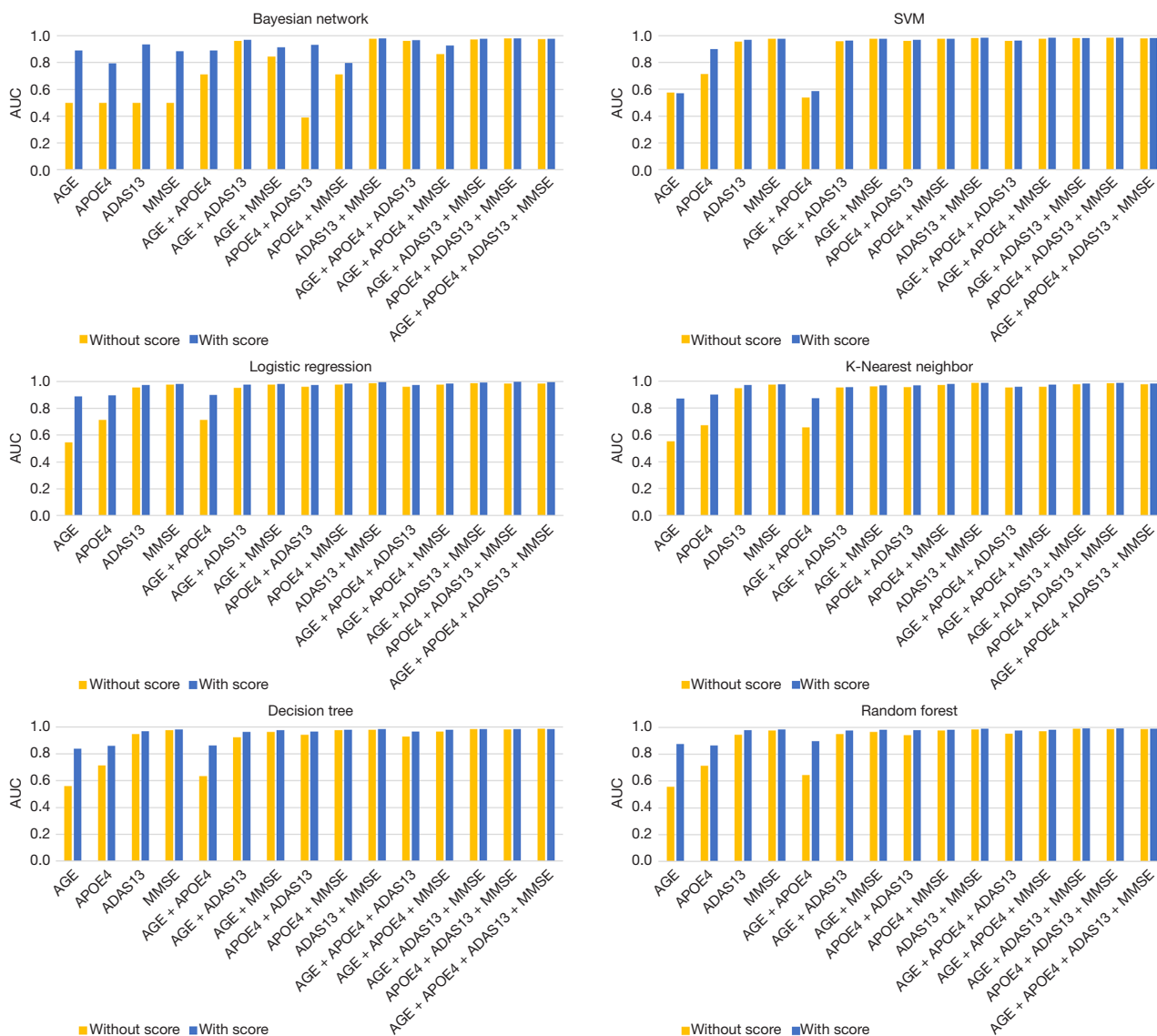
Figure 9 Histogram of the impact of decision scores on the diagnostic performance of clinical indicators. The feature vectors consisting of AGE, ADAS13, MMSE, and APOE4 are based on the performance histogram of six machine learning models for AD diagnosis. Orange represents the decision score not considered, and the blue represents the decision score considered. AUC, area under the curve; SVM, support vector machine; SVM, support vector machine; AD, Alzheimer's disease; ADAS13, Alzheimer's Disease Assessment Scale 13; MMSE, Mini-Mental State Examination; APOE4, apolipoprotein E4.

Discussion

In this study, we developed a deep learning model MSRNet that integrates multispatial information by combining Euclidean spatial information from T1WI with non-Euclidean spatial information from the brain network of radMBN. This study was carried out based on T1WI, which is a simple and easily accessible imaging modality.

The MSRNet was validated on four databases including ADNI, AIBL, EDSD, and OASIS. As shown in *Table 2*, in the intra-database experiments on the ADNI database, the performance for AD diagnosis was as follows: ACC =92.8%, SEN =88.2%, SPE =95.0%, and AUC =0.956. As shown in *Table 3*, in the inter-database experiments on the ADNI, AIBL, EDSD, and OASIS databases, the performance for AD diagnosis was as follows: ACC =90.6%, SEN =82.0%,

Quant Imaging Med Surg 2024;14(12):8568-8585 | https://dx.doi.org/10.21037/qims-24-584

8582

Nan et al. Multispatial model emphasizing key brain regions for AD

Table 6 Ablation of single-spatial information and multispatial information based on AD *vs.* NC classification task in the ADNI cohort

Model input	Interactive mechanism	ACC (%)	SEN (%)	SPE (%)	AUC (%)
radMBN		89.1	76.5	95.0	94.9
GMV		91.8	87.5	93.8	95.5
radMBN + GMV	×	92.1	86.1	95.0	95.2
radMBN + GMV + topology	×	92.2	86.5	94.8	95.3
radMBN + GMV + topology (proposed MSRNet)	√	92.8	88.2	95.0	95.6

The column titled "Interactive mechanism" indicates whether the image information is integrated into important brain regions, with "√" denoting yes and "×" denoting no. AD, Alzheimer's disease; NC, normal control; ADNI, Alzheimer's Disease Neuroimaging Initiative database; ACC, accuracy; SEN, sensitivity; SPE, specificity; AUC, area under the curve; radMBN, radiomics-based morphology brain network; GMV, gray-matter volume.
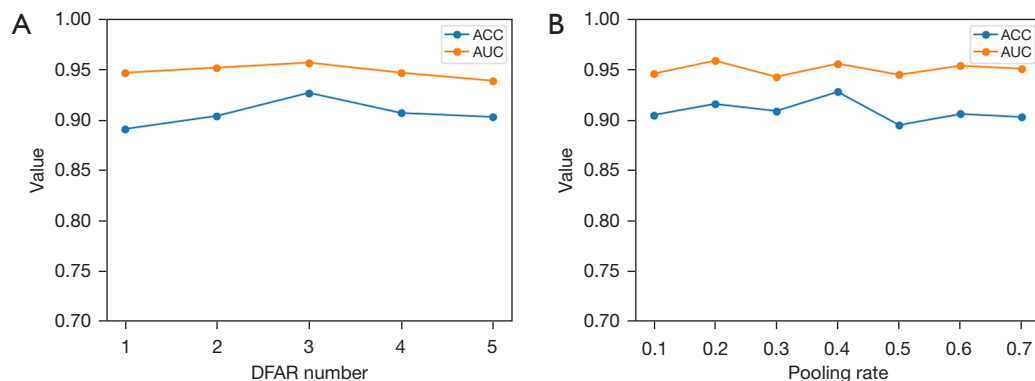

Figure 10 Effect of number of DFAR modules and graph pooling rate on performance. (A) The model achieved optimal performance when the number of DFAR modules was set to 3. (B) The model achieved optimal performance when the graph pooling rate was set to 0.4. ACC, accuracy; AUC, area under the curve; DFAR, discriminative features by focusing on atrophic regions.

SPE =93.6%, and AUC =0.939. The experimental results in *Table 2* and *Table 3* indicate that integrating multispatial information from Euclidean and non-Euclidean spaces for AD diagnosis is more beneficial than is adopting spatial information alone. The MSRNet also achieved satisfactory performance in the AD prodromal stage diagnostic tasks, including for LMCI *vs.* EMCI (*Table 4*) and pMCI *vs.* sMCI classification (*Table 5*).

In the proposed MSRNet, an interactive mechanism was established based on five disease-related brain regions which were identified based on a classifier based on two feature strategies: voxel intensities and radiomic features. The five disease-associated brain regions identified by the two strategies included the left hippocampus, right hippocampus, left parahippocampal gyrus, right parahippocampal gyrus, and left amygdala, which is consistent with the findings of previous studies (36-38). As demonstrated in *Table 6*, the interactive mechanism could enhance the feature representation by connecting the Euclidean and graph

representation branches in the MSRNet model.

Table 7 is a summary of several studies (39-44) related to early diagnosis of AD. Notably, single-modality methods (39-41) achieved relatively superior performance on the AD *vs.* NC diagnostic task. Specifically, Zhao *et al.* (39) optimized the Transformer structure and proposed an inheritable 3D deformable self-attention module which could locate important regions in sMRI. Zhang *et al.* (40) proposed an interpretable 3D residual attention deep neural network that could help locate and visualize important regions in sMRI. Wang *et al.* (41) proposed a multitask-trained dynamic multi-task graph isomorphism network (DMT-GIN) based on fMRI, which could capture spatial information and topological structure in fMRI images. Other studies on multimodal approaches (42-44) reported equally competitive results in the diagnosis of early AD. Kong *et al.* (42) obtained richer multimodal feature information for the early diagnosis of AD by fusing sMRI images and positron emission tomography (PET) images. Zhang *et al.* (43)

Table 7 Comparison of state-of-the-art methods in AD and NC classification

Study	Modality	Participants (AD/NC)	ACC (%)	SEN (%)	SPE (%)	AUC (%)
Zhao *et al.* (39)	sMRI	419/832	92.7	91.9	94.6	97.2
Zhang *et al.* (40)	sMRI	353/650	91.3	91.0	91.9	98.4
Wang *et al.* (41)	fMRI	118/185	90.4	95.9	83.2	89.1
Kong *et al.* (42)	sMRI + PET	111/130	93.2	91.4	95.4	–
Zhang *et al.* (43)	sMRI + PET	129/110	91.1	91.0	91.0	94.1
Tian *et al.* (44)	sMRI + DTI + fMRI	191/167	88.7	86.8	90.9	88.8
Proposed MSRNet	sMRI	282/603	92.8	88.2	95.0	95.6

–, the cited paper does not provide data. AD, Alzheimer's disease; NC, normal control; ACC, accuracy; SEN, sensitivity; SPE, specificity; AUC, area under the curve; sMRI, structural magnetic resonance imaging; PET, positron emission tomography; DTI, diffusion tensor imaging; fMRI, functional magnetic resonance imaging.

proposed a multimodal cross-attention AD diagnosis (MCAD) framework that integrated multimodal data, including sMRI, fluorodeoxyglucose PET (FDG-PET), and cerebrospinal fluid (CSF) biomarkers, to learn the interactions among these modalities and enhance their complementary roles in the diagnosis of AD. Tian *et al.* (44) extracted image features from multimodal MRI data and constructed a GCN based on brain ROIs to extract structural and functional connectivity features between different brain ROIs. In the above studies, those approaches based on single modality extracted only Euclidean spatial information from images or only non-Euclidean spatial information from brain networks for AD diagnosis. In contrast, based on the Euclidean representation channels and non-Euclidean representation channels present in our MSRNet model, multispatial information could be effectively integrated to enhance the performance in diagnosing AD. As shown in *Table 7*, integrating multimodal feature information could improve the diagnostic performance for AD. Notably, PET scans are long, invasive, and expensive, which makes data acquisition more difficult and less unaffordable for patients. Notably, the MSRNet model is based only on sMRI, which has the advantage of being noninvasive and low cost and can ensure diagnostic efficacy.

Several few limitations of this study that need to addressed in future work should be noted. First, the key brain regions were identified solely based on AD *vs.* NC classification task. However, MCI has more subtle differences in cognition and brain structure, and the brain regions identified by the AD *vs.* NC classification may not fully cover the brain regions that could distinguish MCI. As a result, the MCI classification task has more room for improvement as compared to the AD classification task. In future work, we will test various brain region screening methods to make

the model more sensitive to the subtle brain structural changes that are characteristic to MCI.

Conclusions

We developed a multispatial information representation model, MSRNet, for learning multidimensional features from GMV in Euclidean space and radMBN in non-Euclidean space. The experimental results based on four AD databases showed that MSRNet achieved superior performance compared to approaches using single-modality spatial information. In addition, the proposed interaction mechanism was confirmed to be a capable of enhancing the feature representation by connecting the Euclidean and graph representation branches. The longitudinal trajectory study of MCI progression and the combined application with clinical indicators demonstrated that the decision score generated by the MSRNet contributes to the diagnostic capability.

Acknowledgments

Footnote

Conflicts of Interest: All authors have completed the ICMJE

uniform disclosure form (available at https://qims. amegroups.com/article/view/10.21037/qims-24-584/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. This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

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. Tiwari S, Atluri V, Kaushik A, Yndart A, Nair M. Alzheimer's disease: pathogenesis, diagnostics, and therapeutics. Int J Nanomedicine 2019;14:5541-54.

2. Jagust W. Vulnerable neural systems and the borderland of brain aging and neurodegeneration. Neuron 2013;77:219-34.

3. Alzheimer's Association. 2019 Alzheimer's disease facts and figures. Alzheimer's & Dementia 2019;15:321-87.

4. Cummings J, Lee G, Ritter A, Sabbagh M, Zhong K. Alzheimer's disease drug development pipeline: 2019. Alzheimers Dement (N Y) 2019;5:272-293.

5. Ben Ahmed O, Mizotin M, Benois-Pineau J, Allard M, Catheline G, Ben Amar C; Alzheimer's Disease Neuroimaging Initiative. Alzheimer's disease diagnosis on structural MR images using circular harmonic functions descriptors on hippocampus and posterior cingulate cortex. Comput Med Imaging Graph 2015;44:13-25.

6. Atri A. Current and Future Treatments in Alzheimer's Disease. Semin Neurol. 2019;39:227-40.

7. American Psychiatric Association, DSM-5 Task Force. Diagnostic and statistical manual of mental disorders: DSM-5™ (5th ed.). American Psychiatric Publishing, Inc. 2013.

8. Ebrahimighahnavieh MA, Luo S, Chiong R. Deep learning to detect Alzheimer's disease from neuroimaging: A systematic literature review. Computer Methods and Programs in Biomedicine 2020;187:105242.

9. Zhang F, Li Z, Zhang B, Du H, Wang B, Zhang X. Multi-modal deep learning model for auxiliary diagnosis of Alzheimer's disease. Neurocomputing 2019;361:185-95.

10. Cherubini A, Caligiuri ME, Peran P, Sabatini U, Cosentino C, Amato F. Importance of Multimodal MRI in Characterizing Brain Tissue and Its Potential Application for Individual Age Prediction. IEEE J Biomed Health Inform 2016;20:1232-9.

11. Frisoni GB, Fox NC, Jack CR Jr, Scheltens P, Thompson PM. The clinical use of structural MRI in Alzheimer disease. Nat Rev Neurol 2010;6:67-77.

12. Zhang X, Han L, Han L, Chen H, Dancey D, Zhang D. sMRI-PatchNet: A novel efficient explainable patch-based deep learning network for Alzheimer's disease diagnosis with Structural MRI. IEEE Access 2023;11:108603-16.

13. Hu Z, Wang Z, Jin Y, Hou W. VGG-TSwinformer: Transformer-based deep learning model for early Alzheimer's disease prediction. Computer Methods Programs in Biomedicine 2023;229:107291.

14. Zhu W, Sun L, Huang J, Han L, Zhang D. Dual attention multi-instance deep learning for Alzheimer's disease diagnosis with structural MRI. IEEE Transactions on Medical Imaging 2021;40:2354-66.

15. Zhao K, Zheng Q, Dyrba M, Rittman T, Li A, Che T, Chen P, Sun Y, Kang X, Li Q, Liu B, Liu Y, Li S; for the Alzheimer's Disease Neuroimaging Initiative. Regional Radiomics Similarity Networks Reveal Distinct Subtypes and Abnormality Patterns in Mild Cognitive Impairment. Adv Sci (Weinh) 2022;9:e2104538.

16. Yu H, Ding Y, Wei Y, Dyrba M, Wang D, Kang X, Xu W, Zhao K, Liu Y; Alzheimer's Disease Neuroimaging Initiative. Morphological connectivity differences in Alzheimer's disease correlate with gene transcription and cell-type. Hum Brain Mapp 2023;44:6364-74.

17. Gaser C, Dahnke R, Thompson PM, Kurth F, Luders E, The Alzheimer's Disease Neuroimaging Initiative. CAT: a computational anatomy toolbox for the analysis of structural MRI data. Gigascience 2024;13:giae049.

18. Zhu C, Song Z, Wang Y, Jiang M, Song L, Zheng Q, editors. Group Sparse Radiomics Representation Network for Diagnosis of Alzheimer's Disease Using Triple Graph Convolutional Neural Network. 2023 6th International Conference on Software Engineering and Computer Science (CSECS); 2023. doi: 10.1109/CSECS60003.2023.10428290; Chengdu, China: IEEE.

19. Zhao K, Zheng Q, Che T, Dyrba M, Li Q, Ding Y, Zheng Y, Liu Y, Li S. Regional radiomics similarity networks (R2SNs) in the human brain: Reproducibility, small-world properties and a biological basis. Netw Neurosci

2021;5:783-97.

20. Hou Q, Zhou D, Feng J, editors. Coordinate attention for efficient mobile network design. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2921:13708-17.

21. Trockman A, Kolter JZ. Patches are all you need? arXiv:2201.09792.

22. Thekumparampil KK, Wang C, Oh S, Li LJ. Attention-based graph neural network for semi-supervised learning. arXiv:1803.03735.

23. Velickovic P, Cucurull G, Casanova A, Romero A, Lio P, Bengio Y. Graph attention networks. arXiv:1710.10903.

24. Hamilton WL, Ying R, Leskovec J. Inductive representation learning on large graphs. Advances in Neural Information Processing Systems 2017;30.

25. Jiang J, Wei Y, Feng Y, Cao J, Gao Y, editors. Dynamic Hypergraph Neural Networks. IJCAI 2019:2635-2641.

26. Gao H, Ji S. Graph U-Nets. IEEE Trans Pattern Anal Mach Intell 2022;44:4948-60.

27. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv: 2014 Available online: https://doi.org/10.48550/arXiv.1409.1556.

28. Jin D, Zhou B, Han Y, Ren J, Han T, Liu B, Lu J, Song C, Wang P, Wang D. Generalizable, reproducible, and neuroscientifically interpretable imaging biomarkers for Alzheimer's disease. Advanced Science 2020;7:2000675.

29. Waskom ML. Seaborn: statistical data visualization. Journal of Open Source Software 2021;6:3021.

30. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: synthetic minority over-sampling technique. Journal of Artificial Intelligence Research 2002;16:321-57.

31. Zhang Y, Li H, Zheng Q. A comprehensive characterization of hippocampal feature ensemble serves as individualized brain signature for Alzheimer's disease: deep learning analysis in 3238 participants worldwide. Eur Radiol 2023;33:5385-97.

32. Stanley K, Whitfield T, Kuchenbaecker K, Sanders O, Stevens T, Walker Z. Rate of Cognitive Decline in Alzheimer's Disease Stratified by Age. J Alzheimers Dis 2019;69:1153-60.

33. Ge XY, Cui K, Liu L, Qin Y, Cui J, Han HJ, Luo YH, Yu HM. Screening and predicting progression from high-risk mild cognitive impairment to Alzheimer's disease. Sci Rep 2021;11:17558.

34. Emrani S, Arain HA, DeMarshall C, Nuriel T. APOE4 is associated with cognitive and pathological heterogeneity in patients with Alzheimer's disease: a systematic review. Alzheimer's Research Therapy 2020;12:141.

35. Safieh M, Korczyn AD, Michaelson DM. ApoE4: an emerging therapeutic target for Alzheimer's disease. BMC Med 2019;17:64.

36. Park C, Jung W, Suk HI. Deep joint learning of pathological region localization and Alzheimer's disease diagnosis. Sci Rep 2023;13:11664.

37. Prestia A, Boccardi M, Galluzzi S, Cavedo E, Adorni A, Soricelli A, Bonetti M, Geroldi C, Giannakopoulos P, Thompson P, Frisoni G. Hippocampal and amygdalar volume changes in elderly patients with Alzheimer's disease and schizophrenia. Psychiatry Res 2011;192:77-83.

38. Kenny ER, Blamire AM, Firbank MJ, O'Brien JT. Functional connectivity in cortical regions in dementia with Lewy bodies and Alzheimer's disease. Brain 2012;135:569-81.

39. Zhao Q, Huang G, Xu P, Chen Z, Li W, Yuan X, Zhong G, Pun C-M, Huang Z. IDA-Net: Inheritable Deformable Attention Network of structural MRI for Alzheimer's Disease Diagnosis. Biomedical Signal Processing and Control 2023;84:104787.

40. Zhang X, Han L, Zhu W, Sun L, Zhang D. An explainable 3D residual self-attention deep neural network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. IEEE Journal of Biomedical and Health Informatics 2022;26:5289-97.

41. Wang Z, Lin Z, Li S, Wang Y, Zhong W, Wang X, Xin J. Dynamic Multi-Task Graph Isomorphism Network for Classification of Alzheimer's Disease. Applied Sciences 2023;13:8433.

42. Kong Z, Zhang M, Zhu W, Yi Y, Wang T, Zhang B. Multi-modal data Alzheimer's disease detection based on 3D convolution. Biomedical Signal Processing and Control 2022;75:103565.

43. Zhang J, He X, Liu Y, Cai Q, Chen H, Qing L. Multi-modal cross-attention network for Alzheimer's disease diagnosis with multi-modality data. Computers in Biology Medicine 2023;162:107050.

44. Tian X, Liu Y, Wang L, Zeng X, Huang Y, Wang Z. An extensible hierarchical graph convolutional network for early Alzheimer's disease identification. Computer Methods Programs in Biomedicine 2023;238:107597.