The Journal of Supercomputing
https://doi.org/10.1007/s11227-023-05655-9

Research of spatial context convolutional neural networks for early diagnosis of Alzheimer's disease

Yinsheng Tong[1,2] · Zuoyong Li[3] · Hui Huang[1,2] · Libin Gao[1,2] · Minghai Xu[2,4] · Zhongyi Hu[1,2]

Accepted: 4 September 2023

Abstract

The early and effective diagnosis of Alzheimer's disease (AD) and mild cognitive impairment (MCI) has received increasing attention in recent years. However, currently available deep learning methods often ignore the contextual spatial information contained in structural MRI images used for early diagnosis and classification of Alzheimer's disease. This may lead us to miss important structural details by failing to adequately capture the potential connections between each slice and its neighboring slices. This lack of contextual information may cause the accuracy of the network model to suffer, which in turn affects its generalization ability and application in real-life scenarios. To explore deeper the connection between spatial context slices, this research is designed to develop a new network model to effectively detect or predict AD by digging into the deeper spatial contextual structural information. In this paper, we design a spatial context network based on 3D convolutional neural network to learn the multi-level structural features of brain MRI images for AD classification. The experimental results show that the model has good stability, accuracy and generalization ability. Our experimental method had a classification accuracy of 92.6% in the AD/CN comparison, 74.9% in the AD/MCI comparison, and 76.3% in the MCI/CN comparison. In addition, this paper demonstrates the effectiveness of the proposed network model through ablation experiments.

Keywords Alzheimer's disease · Structural magnetic resonance image · Spatial context · Classification

Zuoyong Li, Hui Huang, Libin Gao, Minghai Xu and Zhongyi Hu have contributed equally to this work.

Extended author information available on the last page of the article

Published online: 26 September 2023

1 Introduction

Alzheimer's disease is a neurodegenerative disease, and the cause of the disease is still not clear, and there is no effective treatment for AD so far [1]. It affects brain cells and can lead to memory loss, reduced thinking ability, and eventually loss of self-care [2, 3]. According to the 2021 World Alzheimer's Disease Report, more than 55 million people worldwide currently have Alzheimer's disease, and it is predicted to reach 78 million by 2030. In China, there are currently about 10 million people with AD, and one in 20 seniors has AD. It is expected that by 2050, China will have more than 40 million people with AD, which is more than the total population of Canada. Current medical research indicates that AD is irreversible and can only be treated to improve symptoms and control the progression of the disease. If AD can be accurately diagnosed and intervened at an early stage, this will result in a significant improvement in the quality of life of AD patients. Therefore, an early and accurate analysis of AD prediction is necessary [4]. Early diagnosis of AD remains challenging, and convolutional neural networks(CNN) [5],which do not require manual feature extraction, can increase efficiency even further and have been very successful in early diagnosis [6–15].

Currently, neuroimaging has been widely used as an important biomarker for AD diagnosis. Magnetic resonance imaging (MRI) is a non-invasive, low-cost imaging technique that can clearly display the three-dimensional anatomy of the human brain and has been widely studied. However, magnetic resonance imaging is often disturbed by random noise, so denoising in magnetic resonance imaging is an indispensable process [16]. Methods for diagnosing AD by MRI images can be broadly divided into two categories, including methods based on two-dimensional views and methods based on three-dimensional attempts. In most studies based on two-dimensional views, methods mostly select two-dimensional slices from each subject and classify these coronal, sagittal or horizontal brain images as a whole. Jain et al. [17] used the most informative set of 2D slices and classified medical images by transfer learning and achieved 95.73% of the results . Khuzaie et al. [18] used 2D slices as input to a convolutional neural net and trained the AlzNet convolutional neural network structure using 2D slices, achieving a 99.3% correct rate.

Many 3D view-based methods also work well for classifying AD. Zhang et al. [19] proposed a densely connected convolutional neural network with a connected intelligent attention mechanism for learning multi-level features of brain MR images for AD classification. The convolutional operation is extended to 3D to capture the spatial information of MRI. The features extracted from each 3D convolutional layer were combined with features from all previous layers with different attention, and the study showed significant experimental results . Hu et al. [20] proposed a VGG-TSwinformer model based on Convolutional Neural Network (CNN) and Transformer for short-term longitudinal study of MCI. Comparative results are obtained on the ADNI dataset . Liu et al. [21] developed a new method based on 3D deep convolutional neural networks to accurately distinguish mild Alzheimer's disease dementia from mildly cognitively impaired

and cognitively normal individuals using structural MRI. The experiment yielded good results and the model can also be used to predict progression.

Vaithinathan et al. designed a new magnetic resonance T1-weighted texture extraction technique for Alzheimer's disease classification for early AD diagnosis, starting from the texture of the brain, and achieved good classification results [22]. In addition to this, the literature combines 2D images and 3D images to perform studies on the early diagnostic classification of AD [23, 24]. Bi et al. [25] designed two deep learning methods for functional brain network classification. The convolutional learning method learns deep region connectivity features and the recursive learning method learns deep adjacency location features, in addition, an Extreme Learning Machine (ELM) enhancement structure was implemented to further improve the learning capability, and the results show that the proposed method provides satisfactory learning capability in AD detection applications. Oh et al. [26] used unsupervised learning based on convolutional autoencoders (CAE) to solve AD and NC classification tasks and supervised transfer learning to solve pMCI and sMCI classification tasks . In addition, there is the application of autoencoders in combination with long- and short-term memory networks to resting-state functional magnetic resonance imaging for the early diagnosis of Alzheimer's disease [27].

The papers mentioned above all achieve relatively superior performance in terms of algorithm and performance. However, 3D structural magnetic resonance imaging (SMRI) carries information that is not limited to surface features; it also hides higher-order and richer spatial structural information. In Alzheimer's disease research, atrophy of localized brain regions is of great concern, and thus only a few regions show significant structural changes in SMRI scans, which are highly correlated with pathological features. Nevertheless, currently available deep learning methods tend to ignore the underlying contextual spatial information contained in structural MRI images used for early diagnosis and classification of Alzheimer's disease. Traditional three-dimensional convolutional neural networks (3D CNNs) tend to focus only on local feature extraction. This may lead us to miss important structural details by failing to adequately capture the potential connections between each slice and its neighboring slices. This lack of spatial context information may cause the accuracy of the network model to suffer, which in turn affects its generalization ability and application in real scenarios. To address this problem, we need to develop a more comprehensive deep learning model that can better integrate the structural features of each slice and simultaneously consider the spatial contextual relationships between the slices. In this way, we can enhance the model's understanding of the correlations between different brain regions and between different slices while maintaining high accuracy. This will help improve the accuracy of early diagnosis and classification of Alzheimer's disease.

To address the previously mentioned problems, this study employs an innovative approach to construct a spatial context-associative convolutional neural network for capturing strong spatial correlations between slices in 3D brain images. This network model was specifically designed for early Alzheimer's disease detection and classification tasks, while incorporating a 3D convolutional neural network for Alzheimer's disease classification. The main contributions of this paper are reflected in the following aspects: First, we propose and design a novel network structure called

Spatial Contextual Network (SCN), which aims to mine the higher-order spatially correlated structural information latent in 3D structural magnetic resonance imaging (SMRI). With SCN, we were able to more fully utilize the spatial relationships between slices and significantly improve the classification accuracy. The results of ablation experiments further validate the significant role of SCN in model performance enhancement. Second, we propose a new approach that goes beyond the use of a single feature and combines the original features with the mined spatial context-related features. This combination allows the model to better capture overall spatial structural information and thus diagnose Alzheimer's disease more accurately. Finally, in this paper, we construct an innovative deep learning experimental framework that fuses a 3D convolutional neural network with a spatial context-dependent network specifically for the task of early Alzheimer's disease classification. By combining these two networks, we expect to further improve the performance of the classification model and provide more accurate and reliable support for the early diagnosis of Alzheimer's disease. In summary, this study brings new ideas and approaches to the early detection and classification task of Alzheimer's disease by introducing spatial context-associative convolutional neural networks and an innovative feature combination approach. We believe that this research is of great significance in promoting the development of the medical image analysis field.

2 Data sets and pre-processing

The data used in this paper are from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database [28]. ADNI is a longitudinal multi-center research to develop clinical, imaging, genetic and biochemical biomarkers for the early detection and follow-up of AD. Since its launch more than a decade ago, this landmark public-private partnership has made significant contributions to AD research, enabling the sharing of data between researchers around the world. In this work, more than 1,000 participants collected, validated and used MRI and PET images, genetics, cognitive tests, cerebrospinal fluid and blood biometrics as predictors of AD [29]. There are currently four types of ADNI databases, of which ADNI 1 sets uniform standards for obtaining longitudinal, multi-locus MRI and PET data in AD, MCI and elderly control patients, ADNI GO defines and describes the AD spectrum stage prior to MCI by recruiting 200 EMCIs, ADNI 2 performs longitudinal clinical, cognitive, MRI, and PET (18F-Florbetapir and FDG) and blood and cerebrospinal fluid biomarker studies in 550 newly enrolled subjects and continued these studies for 5 years in approximately 700 subjects in ADNI1 and ADNI GO,ADNI3 achieved prediction of cognitive decline by correlating results with standard clinical measures and pathology and validating baseline and longitudinally obtained biomarker measures.

2.1 Data set source and situation

The data used in this study were obtained from the ADNI database of 416 subjects. The main tags set when acquiring the data are data modality as MRI, image type as

Table 1 Basic information of experimental data

Group	Number of people	Age	Gender(Male/Female)
AD	104	75.5±7.5	50/54
MCI	208	76.0±6.8	97/107
CN	204	76.4±6.5	58/46

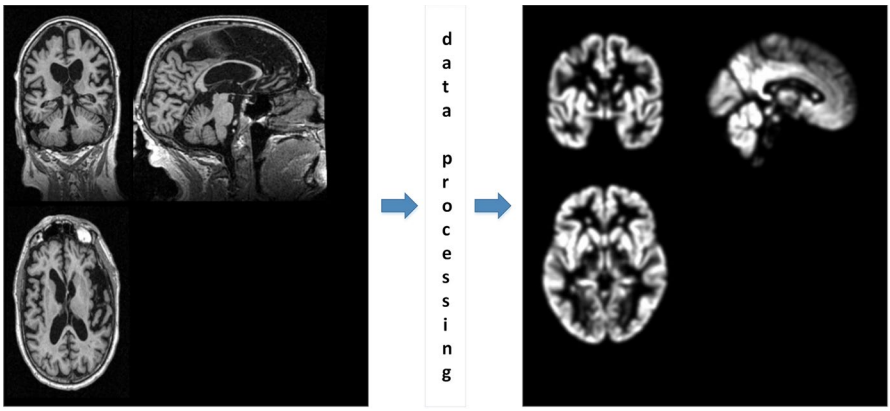

Fig. 1 The processed 3DMRI image

3D, data type as Original and T1 weighting term, respectively. The subjects were more evenly split between men and women and were between 50 and 90 years old. The dataset contains neuroimages assigned with three labels: control (CN) indicates no impairment; mild cognitive impairment (MCI) and Alzheimer's disease (AD). Table 1 shows the approximate status of all data

2.2 Data pre-processing

The original MRI images contain non-brain structures, such as skulls, which increase the computational effort and thus affect the experimental results. Data pre-processing is used for skull stripping, removing noise and artifacts from the data to better highlight the texture of the image and improve the quality of the image for feature extraction. Since the precise separation rate of each neural image is slightly different in the ADNI library, the data were segmented and aligned from the SMRI using the VBM method in order to reduce anatomical differences between individuals [30]. The dataset in this paper uses Cat12, an extended toolkit for Spm12, to first remove non-brain tissue and segment the data into gray matter (GM), white matter (WM), and cerebrospinal fluid. After acquiring the gray matter data, the smoothing operation is performed using the smoothing module in SPM12 with FMWH=4 mm, and the final output image size is 121×145×121. The processed 3DMRI image is shown in Fig. 1.

3 Methods and experiments

Due to the limited number of data sets, this experiment uses the double cross-validation method to get the final experimental results, first, in the case of ensuring that the data are not leaked, the data are divided into 10 equal parts, and one part is taken out as the test set and the other 9 parts are taken as the training set, then the training set is divided into the training set and the validation set according to the ratio of 8:2 and trained using the fivefold cross-validation, after this operation once, the experimental results are obtained once. As a result, the data in the training set will get five models after the fivefold cross-validation method, and then the data in the test set will be put into the five models for prediction, and the average value will be calculated to get the final test set results. Through the above steps, the experimental results of the test set are averaged again after 10 times of testing with non-repeated test data sets, and the final accuracy is obtained. This experimental method perfectly solves the problems of small data sets and easy over-fitting and improves the generalization ability of the model.

3.1 3D convolutional neural network

The convolutional neural network proposed by LeCun et al. [31] is very effective in processing two-dimensional flat images, and the constructed network is also able to automatically extract information such as color and texture shape of the image . The three-dimensional convolution used in this study adds one dimension over the two-dimensional convolution. It is worth noting that the dimension is not the channel dimension, but the spatial dimension of the sliding convolution kernel. The basic definition, principle and working equations of 3D Convolution are very similar to those of 2D Convolution, except that an extra depth dimension is added to the computational equations of 2D Convolution. The formula of 3D convolution is shown in Eq. 1.

$$M^l(w,h,d) = \sum_a \sum_b \sum_c K_{a,b,c} y^{l-1}{}_{(w+a)(h+b)(d+c)} + m \tag{1}$$

where the 3D data of the input M are $x * y * z$, w,h,d are denoted as iterators, K is a convolution kernel of dimension $n1 * n2 * n3$ and a,b,c are denoted as iterators, denotes the output of the first level when $l = 1$ and the output of the last level when $l = L$, y^{l-1} means the output of the previous layer, and m is the bias amount. The cell output of each convolutional layer is calculated as shown in Eq. 2.

$$y^l{}_{w,h,d} = f(M^l{}_{w,h,d}) \tag{2}$$

This work incorporates a 3D convolutional neural network, which is designed to compress and extract the image information of the brain by first extracting the learning features through 3D convolution, and the 3D convolutional neural network can effectively analyze the overall structure of the brain without easily losing the Region

of interest(ROI). Moreover, it can well capture the spatial contextual structural feature information of the data and highlight the spatial attributes, which can effectively improve the accuracy and precision of disease classification. Therefore, this study designs a 3D CNN-SCN model to correlate the original MRI spatial contextual structural information features with the extracted spatial contextual structural information features for classification.

3.2 Spatial contextual association network

This experiment focuses on applying the spatial up-down structural properties mined from 3D SMRI images, combined with the correlation features between spatial up-down slices in the image on slices, so a spatial context network is designed and built. The network combines a 3D convolutional neural network, where the original 3D image is first extracted by the 3D convolutional neural network to learn the original features and compress the neural image, and then fed into a spatial context network, which combines the previously extracted spatial features with the spatial features of the original input image of each patient, and finally connects a fully-connected layer to convert the multi-dimensional output into a single probability between 0 and 1

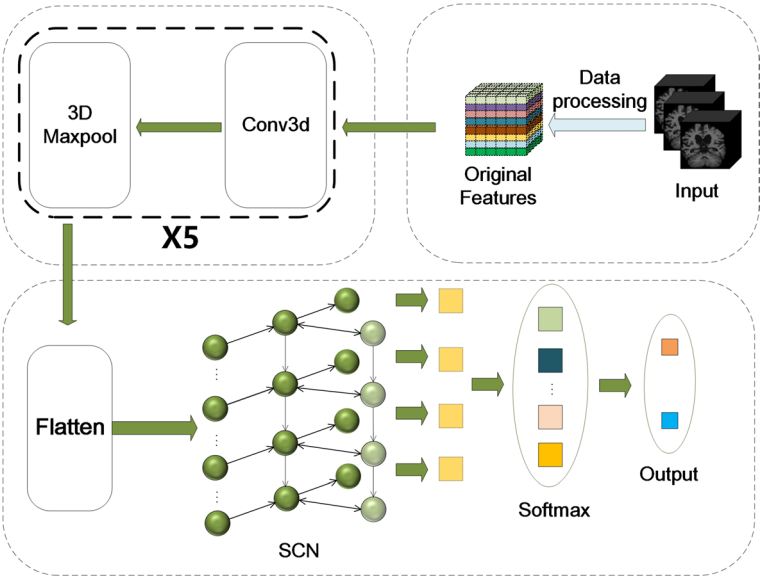

Fig. 2 This is the flowchart of spatial context network structure based on 3D CNN, firstly the preprocessed image will go through 3D convolution and pooling to get the first extracted features, which is used to reduce the feature dimensions, reduce the computation amount, and keep the important spatial information. Then the first extracted features will be sent into SCN as original features after flatten operation, in SCN, we will apply 1D convolution operation to the first extracted features to extract features in a targeted way to capture higher order spatial context information. These features are interactively fused to interact with the original features to obtain richer information. Finally, the obtained fused features are put into a classifier for classification

and output the final prediction. The structure of the network is shown in Fig. 2. The entire network is divided into two learning stages. The first stage extracts spatial features by 3D CNN, and the second stage mines more potential connections in spatial context by linking the original features with the extracted features by SCN. The first stage includes SMRI data pre-processing to extract spatial features. The second stage includes combining and classifying the original features and the extracted features using SCN network.

Spatial contextual association network is a neural network for processing 3D data tensor, which is able to capture more changes between structural spatial up and down slices compared to general neural networks. For example, the information contained in the upper and lower structures of the hippocampal region of the brain in 3D MRI brain images, the hippocampal structure of patients with different degrees of disease will change differently, and the spatial upper and lower domain association network will be able to capture these detailed changes and use the captured feature information as a basis for classification.

The specific structure of the spatial upper and lower domain association network is shown in Fig. 3, where sigmoid denotes the use of sigmoid activation function and tanh denotes the use of tanh activation function.

Firstly 3D CNN extracts features, which is local and contextual image information, by independently performing feature extraction on each voxel of the image. The CNN convolutional neural network used in our study consists of five layers and uses $2 * 2 * 2$ convolutional kernels for feature extraction. The small convolutional kernels enable efficient density inference, which allows capturing more detailed features of spatial structure information. In addition, the small convolution kernel increases the nonlinear fitting ability and reduces the network parameters and computation. In this study, the original baseline $5 * 5 * 5$ convolution is replaced by a $2 * 2 * 2$ convolution, which reduces the network parameters and computation by 94%. In the 3D convolution process, we use only one-element

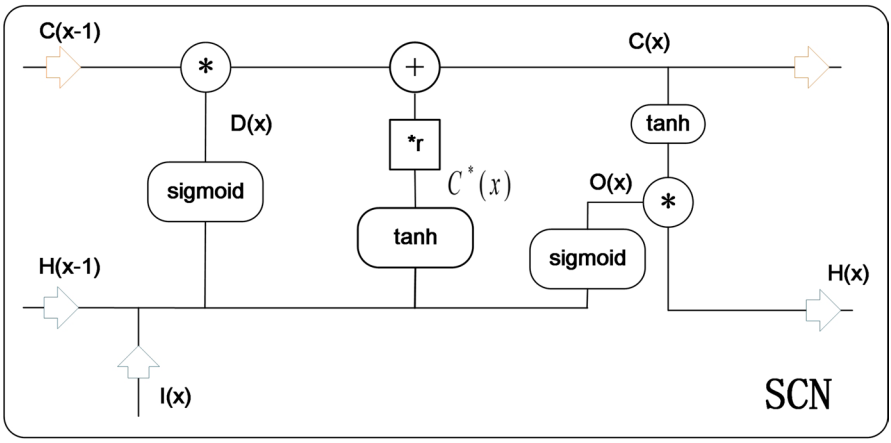

Fig. 3 Flowchart of the spatial context network. This diagram is used to illustrate the specific calculation process of the spatial context network, accurately explaining the calculation order of each formula

steps, because larger steps may miss closely connected spatial structural information features. The density of the spatial up-and-down structural information extracted by the network is ensured by using small convolution kernels and one-dimensional step operations.

The spatial context network combines the feature encoding of 3D CNN with the original features through the input gate, and generates new hidden features $C^*(x)$ using the input data $I(x)$ and the extracted features $H(x-1)$ from the previous layer. The formula for calculating the input gate is shown in Eq. 3. It is important to note that W^C represents the $C(x)$ weights and U^C represents the parameters of $C(x)$.

$$(b)C^*(x) = \tanh(W^C I(x) + U^C H(x-1)) \tag{3}$$

The input gate uses the spatially explicit feature data extracted from the previous layer and the current original spatial feature data to generate new feature data and determine which features are more valuable and should be retained.

Similar to the input gate, the discard gate uses the spatially explicit feature data that has been extracted from the previous layer and the current original spatial feature data, and calculates whether the features from the previous layer have some value. The formula for the discard gate is shown in Eq. 4. It is important to note that W^D represents the $D(x)$ weights and U^D represents the parameters of $D(x)$.

$$D(x) = \text{sigmoid}(W^D I(x) + U^D H(x-1)) \tag{4}$$

The new features obtained can be expressed as shown in Eq. 5 where $C(x-1)$ is the output hidden feature of the previous layer, where r is the hyperparameter.$C^*(x)$ is calculated as shown in Eq. 3.

$$C(x) = D(x)C(x-1) + \mathrm{r}C^*(x) \tag{5}$$

Finally there is an output gate, which determines the final output by hiding the feature state. The formulas are shown in Eqs. 6, 7 where $H(x-1)$ is the output feature of the previous layer. It is important to note that W^O represents the $O(x)$ weights and U^O represents the parameters of $O(x)$ and b is the bias amount.

$$O(x) = \text{sigmoid}\left(W^O I(x) + U^O H(x-1) + b\right) \tag{6}$$

$$H(x) = O(x) * \tanh\left(C(x)\right) \tag{7}$$

4 Experimental results and comparison

Four commonly used performance metrics were used in this experiment, namely area under curve, accuracy, sensitivity, and specificity. The calculation formula is shown in Eqs. 8, 9, 10 where TP indicates true positive, FP indicates false positive, TN indicates true negative, and FN indicates false negative. Population standard deviation (PSD) is used to measure the dispersion or volatility of experimental data.

Table 2 Overall experimental data and average results of AD/CN test set

BATCH	AUC	ACC	SEN	SPE
1	0.9692	0.9130	1.0000	0.8461
2	0.8615	0.7826	0.6154	1.0000
3	0.9206	0.9130	0.7778	1.0000
4	0.9848	0.9565	0.9091	1.0000
5	0.8824	0.9130	0.6667	1.0000
6	0.9230	0.9565	0.9231	1.0000
7	1.0000	1.0000	1.0000	1.0000
8	0.9762	0.9565	1.0000	0.8889
9	0.9848	0.9565	1.0000	0.9091
10	0.9242	0.9130	1.0000	0.8333
Mean	0.9427	0.9261	0.8892	0.9477
PSD	0.0451	0.0571	0.1414	0.0669

Table 3 Overall experimental data and average results of the AD/MCI test set

BATCH	AUC	ACC	SEN	SPE
1	0.6955	0.6875	0.5909	0.9000
2	0.6250	0.7188	0.8000	0.5833
3	0.6857	0.8438	0.9200	0.5714
4	0.8136	0.8438	0.9091	0.7000
5	0.8000	0.7188	0.6800	0.8571
6	0.6389	0.6563	0.5000	0.8571
7	0.7421	0.7500	0.7222	0.7857
8	0.7969	0.8438	0.8750	0.7450
9	0.8178	0.8125	0.9474	0.6154
10	0.7576	0.6563	0.4762	1.0000
Mean	0.7373	0.7532	0.7421	0.7615
PSD	0.0685	0.0732	0.1671	0.1343

$$\mathrm{ACC} = \frac{\mathrm{TP+TN}}{\mathrm{TP+FN+TN+FP}} \tag{8}$$

$$\mathrm{SEN} = \frac{\mathrm{TP}}{\mathrm{TP+FN}} \tag{9}$$

$$\mathrm{SPE} = \frac{\mathrm{TN}}{\mathrm{TN+FP}} \tag{10}$$

Table 2 shows the results and average results of the 10 test set experiments for AD/CN. Table 3 shows the results and average results of the 10-test set experiments for AD/MCI. Table 4 shows the results and average results of the 10-test set experiment for MCI/CN. The accuracy rate was 92.6% in AD patients versus healthy controls,

Table 4 Overall experimental data and average results of MCI/CN test set

BATCH	AUC	ACC	SEN	SPE
1	0.8583	0.7813	0.6842	0.9231
2	0.6812	0.7188	0.6956	0.7778
3	0.8229	0.8438	0.8333	0.8750
4	0.8874	0.9063	0.9524	0.8182
5	0.6270	0.6563	0.7222	0.5714
6	0.7292	0.7500	0.9000	0.5000
7	0.8261	0.7500	0.6522	1.0000
8	0.7273	0.7500	0.7619	0.7273
9	0.6927	0.7813	0.8750	0.5000
10	0.7662	0.6875	0.5238	1.0000
Mean	0.7618	0.7625	0.7601	0.7693
PSD	0.0804	0.0688	0.1241	0.1816

74.9% in AD patients versus MCI patients, and 76.3% in MCI patients versus healthy controls.

To demonstrate the effectiveness of joining the spatial upper and lower domain correlation network, the corresponding ablation experiments are done in this paper. The parameters of the 3D convolutional neural network used in this ablation experiment remain unchanged, and a small convolutional kernel of $2 * 2 * 2$ is also used to extract features, and a one-dimensional step size is used, while the final accuracy is also derived according to the operation of the double cross-validation method. Table 5 shows the experimental results and the average results of 10 test sets of the comparison experiment AD/CN. Table 6 shows the experimental results and average results of the 10 test sets of the comparison experiment AD/MCI. Table 7 shows the experimental results and average results of 10 test sets of the comparison

Table 5 Overall experimental data and average results of the AD/CN test set (comparison experiments)

BATCH	AUC	ACC	SEN	SPE
1	1.0000	1.0000	1.0000	1.0000
2	0.7154	0.7391	0.6154	0.9000
3	0.8968	0.8261	0.8889	0.7857
4	0.9242	0.9565	0.9091	1.0000
5	0.8529	0.8696	0.6667	0.9412
6	0.9077	0.8696	0.8461	0.9000
7	0.9750	0.9565	0.9333	1.0000
8	0.9286	0.8696	0.9286	0.7778
9	1.0000	1.0000	1.0000	1.0000
10	0.8258	0.8696	0.9091	0.8333
Mean	0.9026	0.8957	0.8697	0.9138
PSD	0.0830	0.0783	0.1230	0.0850

Table 6 Overall experimental data and average results for the AD/MCI test set (comparison experiments)

BATCH	AUC	ACC	SEN	SPE
1	0.6455	0.5938	0.4091	1.0000
2	0.7208	0.8125	0.9500	0.5833
3	0.7086	0.5938	0.4800	1.0000
4	0.7909	0.7188	0.6364	0.9000
5	0.7143	0.7500	0.7600	0.7143
6	0.6111	0.5938	0.3333	0.9286
7	0.7460	0.7500	0.7222	0.7857
8	0.8594	0.7813	0.7500	0.8750
9	0.7247	0.6878	0.4737	1.0000
10	0.8658	0.7813	0.6667	1.0000
Mean	0.7387	0.7063	0.6181	0.8787
PSD	0.0778	0.0805	0.1805	0.1359

Table 7 Overall experimental data and average results of MCI/CN test set (comparison experiments)

BATCH	AUC	ACC	SEN	SPE
1	0.7206	0.6875	0.4737	1.0000
2	0.6522	0.6250	0.5217	0.8889
3	0.8802	0.7500	0.6667	1.0000
4	0.8225	0.6875	0.5238	1.0000
5	0.7341	0.7500	0.6667	0.8571
6	0.7667	0.7500	0.7500	0.7500
7	0.7343	0.7500	0.6522	1.0000
8	0.7100	0.7500	0.7619	0.7273
9	0.7500	0.6875	0.6250	0.8750
10	0.6883	0.7500	0.8095	0.6364
Mean	0.7459	0.7188	0.6451	0.8735
PSD	0.0622	0.0419	0.1060	0.1321

experiment MCI/CN. The accuracy rate was 89.6% in AD patients versus healthy controls, 70.6% in AD patients versus MCI patients, and 71.9% in MCI patients versus healthy controls.

From the overall comparative experiments, it is obvious that most of the single experimental indexes have been improved among the ten experiments, and the average results show that the accuracy of all three types of tasks has increased by 3%−4.4%. The results of multiple experiments show that the performance of the spatial upper and lower domain correlation network is very obvious for the network architecture of this experiment, which can make the model more stable and further improve the accuracy. The results show that the performance of the spatial context network is significantly improved for the network architecture of this experiment, which makes the model more stable and further improves the accuracy.

In order to deeply evaluate the effectiveness of spatial context networks (SCNs) in the early detection and classification of Alzheimer's disease, we conducted a series of ablation experiments, with a special focus on the performance of SCNs compared to traditional 2D convolutional neural networks (2D CNNs). These ablation experiments aim to reveal the contribution of SCNs in improving model performance from a spatial context perspective. In the ablation experiments, we first implemented a 2D CNN-based model as a control group. The model accepts 2D slices as input, undergoes operations such as convolution and pooling for feature extraction, and then connects fully connected layers for classification. We compare the 2D CNN model with the full model that introduces SCN to evaluate the effectiveness of SCN in terms of whether it is able to extract spatial context information between slices. The experimental results are shown in Table 8.

In addition, to prove the effectiveness of spatial contextual association networks, a comparison is made with the results of other papers, including these papers in addition to those mentioned above [32–41]. The comparison results are shown in Table 9 and Fig. 4.

From Table 9, it can be seen that the spatial context network proposed in this study has some improvement in performance compared with other methods, which proves that the spatial context network can further improve the classification accuracy and stability.

This research adopts the experimental method of double cross-validation, which solves the problem of insufficient sample size and makes the use of data more efficient. Through a large number of experiments to ensure the reliability of the results and greatly achieve the reduction of chance, this study divides the data into ten parts and obtains ten models by training under the exclusion of data leakage, and sends the ten test sets into the corresponding training models to obtain the experimental results, and finally averages the experimental results to obtain the final experimental results.

In our cross-validation experiments, we embed the dropout layer into the architecture of the deep learning model, which is based on the principle of randomly deactivating a portion of neurons, i.e., setting their outputs to zero, during the training process of the model. This randomization helps to reduce the strong dependencies between neurons, and thus reduces the risk of overfitting the model to a particular training sample. By introducing a dropout layer, the model is forced to consider combinations of multiple submodels during training, which enhances its generalization ability and allows it to better fit unknown data. Multiple training and testing are performed through the cross-validation method. The results show that after the introduction of the dropout layer, the performance of the model on different data

Table 8 SCN and traditional 2D convolutional neural network (2D CNN) ablation experiment

Task	2D slice CNN	Ours
AD/CN	0.8750	0.9261
AD/MCI	0.6722	0.7532
MCI/CN	0.6802	0.7625

Table 9 Comparison of experimental results of this thesis with other papers (comparison experiments)

References	Years	Experimental method	ACC	Task
Vaithinathan et al. [22]	2019	KNN	0.8739	AD/CN
Oh et al. [26]	2019	3D CNN+Transfer learning	0.8600	AD/CN
Xuet al. [32]	2018	SVM	0.8565	AD/CN
Wegmayr et al. [33]	2018	Deep 3D CNN	0.8600	AD/CN
Jie et al. [35]	2020	Wck-CNN	0.8800	AD/CN
Mofrad et al. [36]	2019	DT	0.8600	AD/CN
B et al. [37]	2020	CERF+SVM	0.8620	AD/CN
Gao et al. [41]	2022	CBAM	0.8830	AD/CN
Manhua et al. [38]	2020	3D DenseNet	0.8890	AD/CN
Ours	2022	CNN-SCN	0.9260	AD/CN
Wegmayr et al. [33]	2018	Deep 3D CNN	0.7200	AD/MCI
Aderghal et al. [39]	2017	2D Slice CNN	0.6328	AD/MCI
Aderghal et al. [40]	2018	ROI-based+transfer learning	0.7250	AD/MCI
Ours	2022	CNN-SCN	0.7490	AD/MCI
Wegmayr et al. [33]	2018	Deep 3D CNN	0.6700	MCI/CN
Manhua et al. [38]	2020	3D DenseNet	0.7620	MCI/CN
Aderghal et al. [39]	2017	ROI-based	0.6561	MCI/CN
Aderghal et al. [40]	2018	ROI-based+Transfer learning	0.7250	MCI/CN
Ours	2022	CNN-SCN	0.7630	MCI/CN

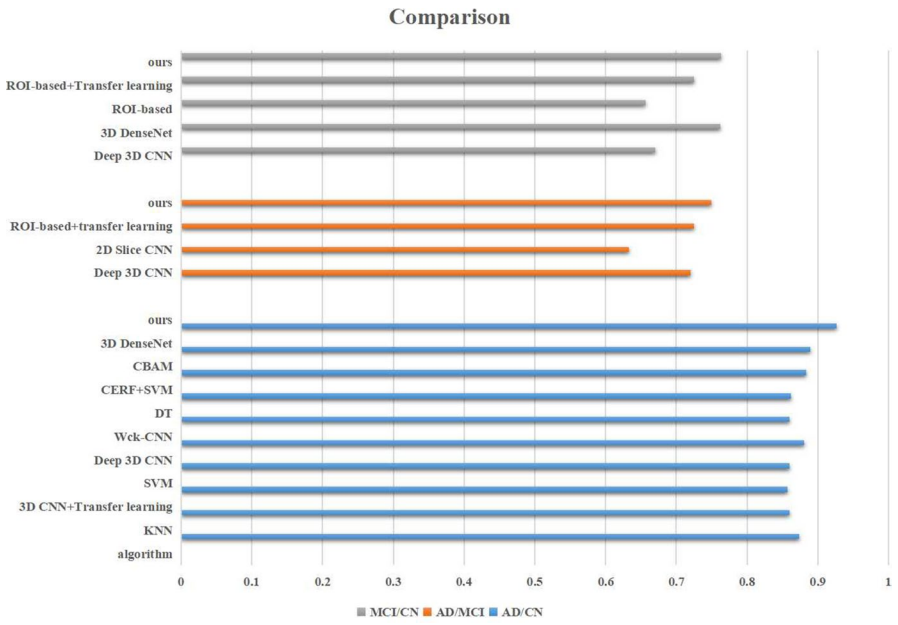

Fig. 4 Comparison of experimental results of this thesis with other papers (comparison experiments)

subsets of cross-validation is more stable, and the generalization error is effectively controlled. This implies that the dropout layer suppresses the overfitting phenomenon to a certain extent and improves the generalization ability of the model, so that it can still maintain a better performance on unseen data.

Overall, in this study we introduced spatial context networks (SCN) as an innovative core methodology designed to address the challenges in 3D brain image analysis and apply it to the task of early detection and classification of Alzheimer's disease. The introduction of SCN has multiple novelties, which bring significant enhancement to the scientific value and practical application contributions of this study. First, SCN introduces a context-aware feature extraction method by emphasizing the spatial contextual relationships between slices. Compared with the traditional 3D convolutional neural network, SCN is able to better capture the higher-order structural information in 3D images, which enhances the feature representation. Secondly, SCN introduces a feature interaction fusion mechanism, which fuses the original features extracted by 3D CNN and the context-aware features extracted by SCN with each other. It brings new ideas for solving the early diagnosis problem and makes the model better in capturing the correlations and details in the image. In addition, we successfully apply SCN to the early detection and classification task of Alzheimer's disease, which is another novel attempt in this field. By combining SCN with 3D convolutional neural networks, we constructed an innovative experimental framework that effectively improves the performance and accuracy of the model in Alzheimer's disease classification. All in all, this study demonstrates unique innovativeness at multiple levels, including the introduction of spatial context networks, feature extraction and interaction, and application domain selection.

Although our method achieves relatively good results in classification. However, the present study still has the following limitations: This study only classifies and diagnoses from structural magnetic resonance imaging data. In fact, there are other data modalities containing spatial attributes that can also be used as information supplements, such as positron emission computed tomography (PET). In future work, we will use data from multiple modalities to conduct the study.

5 Conclusion

In this paper, we propose a deep learning-based spatial upper and lower domain association network combined with 3D convolutional neural network for the classification of early Alzheimer's disease, and demonstrate the classification effect based on ADNI dataset. From the results, the network increased the accuracy over the 3D CNN method by 3% on AD/CN, 4.3% on AD/MCI, and 4.4% on CN/MCI. The network improved the accuracy over the 2D slice method by 5.1% on AD/CN, 8.1% on AD/MCI, and 8.2% on CN/MCI. The accuracy of our method was 92.6% in AD patients versus healthy controls, 74.9% in AD patients versus MCI patients, and 76.3% in MCI patients versus healthy controls. A large number of experiments have demonstrated that spatial context networks can effectively improve the classification accuracy, which indicates that spatial context networks can well mine the information of structural features above and below the space, and the network has

better results for the cases that need to go for spatial structure contextual association features. In the future, the application of deep learning methods in disease diagnosis is worth investigating.

Acknowledgements This work was supported by the Key Project of Zhejiang Provincial Natural Science Foundation under Grant LD21F020001, Grant LSZ19F020001, Grant Z20F020022, and the National Natural Science Foundation of China under Grant U1809209, Grant 62072340, Grant 61972187, Grant 61772254, and the Major Project of Wenzhou Natural Science Foundation under Grant ZY2019020, and the Natural Science Foundation of Fujian Province under Grant 2020J02024. We acknowledge the efforts and constructive comments of respected editors and anonymous reviewers.

Author contributions Conceptualization was performed by Yinsheng Tong. Yinsheng Tong contributed to the development. Yinsheng Tong contributed to the testing/formal analysis. Zuoyong Li, Hui Huang, Minghai XU and Zhongyi Hu acquired the funding. Yinsheng Tong and Zhongyi Hu contributed to the investigation. Yinsheng Tong assisted in the investigation. Zhongyi Hu was involved in the validation. Yinsheng Tong contributed to writing—original draft. Zhongyi Hu and Libin Gao contributed to writing—review and editing. All authors read and approved the final manuscript.

Funding This study was supported by the Key Project of Zhejiang Provincial Natural Science Foundation under Grant LD21F020001, Grant LSZ19F020001, Grant Z20F020022, and the National Natural Science Foundation of China under Grant U1809209, Grant 62072340, Grant 61972187, Grant 61772254, and the Major Project of Wenzhou Natural Science Foundation under Grant ZY2019020, and the Natural Science Foundation of Fujian Province under Grant 2020J02024.

Data availability The datasets generated and/or analyzed during this study are available in the Alzheimer's Disease Neuroimaging Initiative (ADNI) database publicly available at [https://adni.loni.usc.edu/].

Declarations

Conflict of interest The authors declare that they have no competing interests.

Ethics approval and consent to participate Not applicable.

Consent for publication Not applicable.

References

1. Jo T, Nho K, Saykin AJ (2019) Deep learning in Alzheimer's disease: diagnostic classification and prognostic prediction using neuroimaging data. Front Aging Neurosci 11:220. https://doi.org/10.3389/fnagi.2019.00220
2. Wen J, Thibeau-Sutre E, Diaz-Melo M, Samper-González J, Routier A, Bottani S, Dormont D, Durrleman S, Burgos N, Colliot O et al (2020) Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation. Med Image Anal 63:101694. https://doi.org/10.1016/j.media.2020.101694
3. Altinkaya E, Polat K, Barakli B (2020) Detection of Alzheimer's disease and dementia states based on deep learning from MRI images: a comprehensive review. J Inst Electron Comput 1(1):39–53. https://doi.org/10.33969/JIEC.2019.11005
4. Murugan S, Venkatesan C, Sumithra M, Gao X-Z, Elakkiya B, Akila M, Manoharan S (2021) DEMNET: a deep learning model for early diagnosis of Alzheimer diseases and dementia from MR images. IEEE Access 9:90319–90329. https://doi.org/10.1109/ACCESS.2021.3090474.
5. Bhatt D, Patel C, Talsania H, Patel J, Vaghela R, Pandya S, Modi K, Ghayvat H (2021) CNN variants for computer vision: History, architecture, application, challenges and future scope. Electronics 10(20):2470. https://doi.org/10.3390/electronics10202470

6. Noor MBT, Zenia NZ, Kaiser MS, Mamun SA, Mahmud M (2020) Application of deep learning in detecting neurological disorders from magnetic resonance images: a survey on the detection of Alzheimer's disease, Parkinson's disease and schizophrenia. Brain Inform 7(1):1–21. https://doi.org/10.1186/s40708-020-00112-2
7. AbdulAzeem Y, Bahgat WM, Badawy M (2021) A CNN based framework for classification of Alzheimer's disease. Neural Comput Appl 33(16):10415–10428. https://doi.org/10.1007/s00521-021-05799-w
8. Li J, Zou B, Xu Z, Liu Q (2021) Multi-branch multi-task 3D-CNN for Alzheimer's disease detection. In: Chinese Conference on Pattern Recognition and Computer Vision (PRCV). Springer, pp 618–629. https://doi.org/10.1007/978-3-030-88010-1_52.
9. Liu J, Li M, Luo Y, Yang S, Li W, Bi Y (2021) Alzheimer's disease detection using depthwise separable convolutional neural networks. Comput Methods Progr Biomed 203:106032. https://doi.org/10.1016/j.cmpb.2021.106032
10. Khagi B, Kwon G-R (2020) 3D CNN design for the classification of Alzheimer's disease using brain MRI and pet. IEEE Access 8:217830–217847. https://doi.org/10.1109/ACCESS.2020.3040486
11. Salehi AW, Baglat P, Sharma BB, Gupta G, Upadhya A (2020) A CNN model: earlier diagnosis and classification of Alzheimer disease using MRI. In: 2020 International Conference on Smart Electronics and Communication (ICOSEC). IEEE, pp 156–161. https://doi.org/10.1109/ICOSEC49089.2020.9215402.
12. Folego G, Weiler M, Casseb RF, Pires R, Rocha A (2020) Alzheimer's disease detection through whole-brain 3D-CNN MRI. Front Bioeng Biotechnol 8:534592. https://doi.org/10.3389/fbioe.2020.534592
13. Tripathi PC, Bag S (2020) CNN-DMRI: a convolutional neural network for denoising of magnetic resonance images. Pattern Recogn Lett 135:57–63. https://doi.org/10.1016/j.patrec.2020.03.036.
14. Tripathi PC, Bag S (2022) A computer-aided grading of glioma tumor using deep residual networks fusion. Comput Methods Progr Biomed 215:106597. https://doi.org/10.1016/j.cmpb.2021.106597.
15. Angkoso CV, Tjahyaningtijas HPA, Purnomo M, Purnama I (2022) Multiplane convolutional neural network (Mp-CNN) for Alzheimer's disease classification. Int J Intell Eng Syst 15(1):329–340. https://doi.org/10.22266/ijies2022.0228.30
16. Tripathi PC, Bag S (2021) A dilated convolution-based denoising network for magnetic resonance images. In: 2021 International Joint Conference on Neural Networks (IJCNN). IEEE, pp 1–8. https://doi.org/10.1109/IJCNN52387.2021.9533653.
17. Jain R, Jain N, Aggarwal A, Hemanth DJ (2019) Convolutional neural network based Alzheimer's disease classification from magnetic resonance brain images. Cogn Syst Res 57:147–159. https://doi.org/10.1016/j.cogsys.2018.12.015
18. Al-Khuzaie FE, Bayat O, Duru AD (2021) Diagnosis of Alzheimer disease using 2D MRI slices by convolutional neural network. Appl Bionics Biomech. https://doi.org/10.1155/2021/6690539
19. Zhang J, Zheng B, Gao A, Feng X, Liang D, Long X (2021) A 3D densely connected convolution neural network with connection-wise attention mechanism for Alzheimer's disease classification. Magn Reson Imaging 78:119–126. https://doi.org/10.1016/j.mri.2021.02.001
20. Hu Z, Wang Z, Jin Y, Hou W (2023) VGG-TSwinformer: transformer-based deep learning model for early Alzheimer's disease prediction. Comput Methods Progr Biomed 229:107291. https://doi.org/10.1016/j.cmpb.2022.107291
21. Liu S, Masurkar AV, Rusinek H, Chen J, Zhang B, Zhu W, Fernandez-Granda C, Razavian N (2022) Generalizable deep learning model for early Alzheimer's disease detection from structural MRIs. Sci Rep 12(1):17106. https://doi.org/10.1038/s41598-022-20674-x
22. Vaithinathan K, Parthiban L, Initiative ADN et al (2019) A novel texture extraction technique with t1 weighted mri for the classification ofalzheimer's disease. J Neurosci Methods 318:84–99. https://doi.org/10.1016/j.jneumeth.2019.01.011
23. Helaly HA, Badawy M, Haikal AY (2022) Deep learning approach for early detection of Alzheimer's disease. Cogn Comput 14:1711–1727. https://doi.org/10.1007/s12559-021-09946-2
24. Xing X, Liang G, Blanton H, Rafique MU, Wang C, Lin A-L, Jacobs N (2020) Dynamic image for 3D MRI image Alzheimer's disease classification. In: European Conference on Computer Vision. Springer, pp 355–364. https://doi.org/10.1007/978-3-030-66415-2_23.
25. Bi X, Zhao X, Huang H, Chen D, Ma Y (2020) Functional brain network classification for Alzheimer's disease detection with deep features and extreme learning machine. Cogn Comput 12(3):513–527. https://doi.org/10.1007/s12559-019-09688-2

26. Oh K, Chung Y-C, Kim KW, Kim W-S, Oh I-S (2019) Classification and visualization of Alzheimer's disease using volumetric convolutional neural network and transfer learning. Sci Rep 9(1):1–16. https://doi.org/10.1038/s41598-019-54548-6
27. Gao L, Hu Z, Li R, Lu X, Li Z, Zhang X, Xu S (2022) Multi-perspective feature extraction and fusion based on deep latent space for diagnosis of Alzheimer's diseases. Brain Sci 12(10):1348. https://doi.org/10.3390/brainsci12101348
28. Jack CR Jr, Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D, Borowski B, Britson PJ, Whitwell JL, Ward C et al (2008) The Alzheimer's disease neuroimaging initiative (ADNI): MRI methods. J Magn Reson Imaging Off J Int Soc Magn Reson Med 27(4):685–691. https://doi.org/10.1002/jmri.21049
29. Petersen RC, Aisen P, Beckett LA, Donohue M, Gamst A, Harvey DJ, Jack C, Jagust W, Shaw L, Toga A et al (2010) Alzheimer's disease neuroimaging initiative (ADNI): clinical characterization. Neurology 74(3):201–209. https://doi.org/10.1212/WNL.0b013e3181cb3e25
30. Ashburner J, Friston KJ (2005) Unified segmentation. Neuroimage 26(3):839–851. https://doi.org/10.1016/j.neuroimage.2005.02.018
31. LeCun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. Proc IEEE 86(11):2278–2324. https://doi.org/10.1109/5.726791
32. Xu L, Liang G, Liao C, Chen G-D, Chang C-C (2018) An efficient classifier for Alzheimer's disease genes identification. Molecules 23(12):3140. https://doi.org/10.3390/molecules23123140
33. Wegmayr V, Aitharaju S, Buhmann J (2018) Classification of brain MRI with big data and deep 3D convolutional neural networks. In: Medical Imaging 2018: Computer-Aided Diagnosis, vol 10575. SPIE, pp 406–412. https://doi.org/10.1117/12.2293719.
34. Tufail AB, Ma Y-K, Zhang Q-N (2020) Binary classification of Alzheimer's disease using sMRI imaging modality and deep learning. J Digit Imaging 33(5):1073–1090. https://doi.org/10.1007/s10278-019-00265-5
35. Jie B, Liu M, Lian C, Shi F, Shen D (2020) Designing weighted correlation kernels in convolutional neural networks for functional connectivity based brain disease diagnosis. Med Image Anal 63:101709. https://doi.org/10.1016/j.media.2020.101709
36. Mofrad RB, Schoonenboom NS, Tijms BM, Scheltens P, Visser PJ, van der Flier WM, Teunissen CE (2019) Decision tree supports the interpretation of CSF biomarkers in Alzheimer's disease. Alzheimer's Dement Diagn Assess Dis Monit 11:1–9. https://doi.org/10.1016/j.dadm.2018.10.004
37. Bi X-A, Hu X, Wu H, Wang Y (2020) Multimodal data analysis of Alzheimer's disease based on clustering evolutionary random forest. IEEE J Biomed Health Inform 24(10):2973–2983. https://doi.org/10.1109/JBHI.2020.2973324
38. Yu G, Liu Y, Shen D (2016) Graph-guided joint prediction of class label and clinical scores for the Alzheimer's disease. Brain Struct Funct 221(7):3787–3801. https://doi.org/10.1007/s00429-015-1132-6
39. Aderghal K, Benois-Pineau J, Afdel K, Gwenaëlle C (2017) FuseMe: Classification of sMRI images by fusion of deep CNNs in 2D+ ε projections. In: Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing, pp 1–7. https://doi.org/10.1145/3095713.3095749.
40. Liu M, Li F, Yan H, Wang K, Ma Y, Shen L, Xu M (2020) A multi-model deep convolutional neural network for automatic hippocampus segmentation and classification in Alzheimer's disease. Neuroimage 208:116459. https://doi.org/10.1016/j.neuroimage.2019.116459.
41. Gao L, Hu Z, Li Z, Lu X, Li R, Tong Y, Lin D (2022) Alzheimer's disease diagnosis based on collaborative learning augmented algorithms. In: Chinese Intelligent Systems Conference. Springer, pp 364–373. https://doi.org/10.1007/978-981-19-6203-5_34.

Authors and Affiliations

Yinsheng Tong[1,2] · Zuoyong Li[3] · Hui Huang[1,2] · Libin Gao[1,2] · Minghai Xu[2,4] · Zhongyi Hu[1,2]

✉ Minghai Xu
xmhemail@126.com

✉ Zhongyi Hu
hujunyi@163.com

Yinsheng Tong
tys4022@163.com

Zuoyong Li
fzulzytdq@126.com

Hui Huang
huanghui@wzu.edu.cn

Libin Gao
gao571470893@foxmail.com

[1] College of Computer Science and Artificial Intelligence, Wenzhou University, Wenzhou 325035, Zhejiang, China

[2] Key Laboratory of Intelligent Image Processing and Analysis, Wenzhou 325035, Zhejiang, China

[3] College of Computer and Control Engineering, Minjiang University, Fuzhou 350108, Fujian, China

[4] College of Intelligent Manufacturing and Electronic Engineering, Wenzhou University of Technology, Wenzhou 325035, Zhejiang, China