



Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

Identification of Mild cognitive impairment based on quadruple GCN model constructed with multiple features from higher-order brain connectivity

Yuan Li^{a,1,*}, Ying Zou^{d,1}, Hanning Guo^{b,c}, Yongqing Yang^a, Na Li^e, Linhao Li^a, Feng Zhao^{e,*}

^a School of Management Science and Engineering, Shandong Technology and Business University, Yantai, China

^b Department of Computer Science, RWTH Aachen University, Aachen, Germany

^c Institute of Neuroscience and Medicine, Medical Imaging Physics (INM-4), Forschungszentrum Jülich, Jülich, Germany

^d Department of Information Engineering, Yantai Vocational College, Yantai, China

^e School of Computer Science and Technology, Shandong Technology and Business University, Yantai, China

ARTICLE INFO

Keywords:

Protein features
Graph convolutional network
Pooling operation
Quadruple Siamese network

ABSTRACT

Mild cognitive impairment (MCI) is the early stage of Alzheimer's disease, which is associated with abnormal brain proteins, the recognition of MCI being a challenging task. Recent studies have shown that the performance of MCI identification can be improved by combining protein features captured in Positron Emission Computed Tomography (PET). Nevertheless, there are still great challenges in extracting effective features from the vast amount of information. Most brain networks only considered the unilateral features of nodes or edges, ignored the interactions between them. In response to this problem, our study proposed to combine the quadruple Siamese network and GCN with self-attention pooling (QS-SAGCN) for MCI identification. In detail, we constructed the multiple protein features network (MPN) and higher-order MPN (MPHN) by PET images to promote the MCI identification. Furthermore, a pooling operation with self-attention mechanism was incorporated into GCN (SAGCN), which considered the node characteristics and topology in the graph network to facilitate the acquisition of robust biomarkers, simultaneously. Additionally we combined quadruple Siamese network with SAGCN as classification framework to improve the identification accuracy. Our proposed MCI identification method was evaluated on 230 subjects (including 117 MCI subjects, 113 normal control subjects) with both 18F-AV-1451 PET and 18F-AV-45 PET data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Experimental results showed that 1) QS-SAGCN enhanced the ability of feature identification, laying the foundation for obtaining more effective biomarkers for MCI patients; 2) The MCI identification accuracy (93.5%) was obtained by combining QS-SAGCN and higher-order network, indicating that the framework had advantages in mental disorders recognition. Finally, through comparison, the accuracy of our proposed MCI recognition method was superior to some of the existing state-of-the-art methods. Overall, the MCI identification method in this study was effective and promising to assist in the diagnosis of MCI in future clinical practice.

1. Introduction

Alzheimer's disease (AD) is considered as a cognitive dysfunction and neurodegenerative disease (de Mendonça & Ferrari, 2023). The clinical symptoms of AD patients are thinking confusion, language disorder and loss of cognitive function. Mild cognitive impairment (MCI) is a brain dysfunction caused by neurodegeneration (death or loss of nerve cells in the brain). It is also a transitional state between normal aging and AD, which has special biomarkers about abnormal brain proteins precipitation (Kantarci et al., 2021). Early diagnosis and treatment of

MCI patients would take care of great help to the prevention of AD. $A\beta$ protein and Tau protein are important observed proteins in MCI, and the interaction of these two proteins causes damage to nerve cells and leads to impairment of brain function (Massa et al., 2022). Although both $A\beta$ and Tau are important disease-related factors, the etiological basis and neural substrate remain incompletely understood. It is necessary to mine the relationship between the two proteins for MCI diagnostic analysis (Li et al., 2019), deep learning algorithms provided a good tool. Positron Emission Computed Tomography (PET) images has played an important role in the diagnosis of different mental disorders, it could track protein

* Corresponding authors.

E-mail addresses: liy20@sdtbu.edu.cn (Y. Li), Zhaofeng1016@126.com (F. Zhao).

¹ These authors contributed equally to this work.

<https://doi.org/10.1016/j.eswa.2023.120575>

Received 4 January 2023; Received in revised form 4 April 2023; Accepted 27 May 2023

Available online 2 June 2023

0957-4174/© 2023 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

precipitation of the brain. So how to combine deep learning algorithms with PET image features for diagnosing of MCI was our focus in this study.

PET is an advanced clinical imaging technology in nuclear medicine. The general method of PET is to inject glucose, protein, nucleic acid, fatty acid, labeled short-lived radionuclide and other substances into the human body to detect the metabolic activity and protein precipitation of human (Dani et al., 2016). 18F-AV-1451 PET (*Tau*-PET), a novel PHF-*Tau* tracer, is demonstrated to detect tangle pathology in recent years, and it provides a new method to measure *Tau* neuronal tangles in the brain by vivo neuroimaging. Meanwhile 18F-AV-45 PET could also catch the $A\beta$ protein (Chandra et al., 2019). Lian et al. constructed the individual metabolic network of FDG-PET and contributed to the prediction of AD from the horizontal and vertical perspectives respectively (Lian et al., 2018). Duan constructed the *A β* protein network, analyzed the pathological significance of the *A β* protein network, and calculated individual network properties as features to assist in the diagnosis of diseases (Duan et al., 2017). It is still unknown whether the abnormal topological structure in the protein network can help the early diagnosis of disease.

With the development of deep learning algorithms (Song et al., 2022, Wee et al., 2019), more and more researchers utilized these data-driven methods to identify potential neuroimaging biomarkers for computer-aided diagnosis of brain diseases. Neuroscience has always faced the situation of high feature similarity and small sample size. Optimizing the pooling layer is a concern in the feature extraction stage. The Siamese network could also help solve feature similarity and small sample scenarios (Bai et al., 2020). Graph convolutional network (GCN) is able to mine useful brain network patterns in disease classification by explicitly capturing network topology information (Jiang et al., 2020, Ktena et al., 2018). Specifically, GCN has the ability to transmit, aggregate and update the node information in the graph. It could catch the relationship of the nodes in graph to enhance its characteristics, so as to obtain the nodes information and topology of graph network for disease classification (Di Martino et al., 2014). However, the existing GCN based brain connectivity network analysis methods have at least two disadvantages.

GCN may not work well for high-dimensional small samples. In order to avoid overfitting, training GCN often requires a large number of training samples, which is difficult to meet in medical image research. Some studies have tried to combine data from different sites for experiments (Kam et al., 2019), but the data parameters of multiple sites are different, which may affect the performance of GCN.

GCN mainly aims at the information of nodes in brain network, but little is known about the whole network topology. Therefore, the network topology is often ignored during the GCN operation, lacking of effective pooling operations.

In view of the above problems, we proposed the quadruple Siamese network and GCN with self-attention pooling (QS-SAGCN) model based on PET individual higher-order brain connectivity network for identification of MCI patients. First we constructed a multiple protein features network for each individual using a sparse linear regression and calculated the higher-order (second/third-level) brain networks, which was drew on the idea of other reference (Yang et al., 2020). In this way, higher-order networks could reveal higher level and more complex interaction relationships than traditional networks. Then, the self-attention in GCN could take into account node features and network topology, which contributed to graph pooling operations and also filtered out useless information, leaving more discriminative node features. Finally, we developed a quadruple Siamese network GCN model to learn the feature representation of brain connectivity networks, and explicitly incorporated quadruple relationships between subjects into the learning process.

The remainder of this paper is organized as follows. In the Related

Work section, we presented relevant research about deep learning, GCN and Siamese network. In the third section, we mainly introduced the detailed description of our method, including network construction, pooling operation improved by self-attention and the QS-SAGCN framework. In Section IV, we presented experimental results, explored different feature evaluation methods and compared with other methods. In Section V, we discussed the impact of several key components in the proposed method.

2. Related work

2.1. Deep learning for brain disorder identification

With the development of neuroimaging and artificial intelligence technology, many deep learning algorithms have been proposed to explore the abnormalities of human brain connectome and distinguish patients from healthy subjects (De & Chowdhury, 2021, Behrad & Abadeh, 2022). Researchers showed the progress of artificial intelligence and the emergence of deep learning technology provide a promising method for better interpretation of brain image data. Deep learning algorithm has become the most advanced technology for analyzing fMRI data sets, improving performance in various fMRI applications (Shao et al., 2021). Kam et al (Kam et al., 2019) proposed a novel CNN framework that learned embedded features of whole-brain functional connectivity, which could be used for MCI diagnosis, simultaneously. This framework provided intuitive method for better personalized diagnosis of various neurological diseases by taking full advantage of the deep embedded diagnostic features using rs-fMRI. Zhang et al (Zhang et al., 2021b) exploited a FDG-PET cognitive feature based on deep learning, which was applicable to the recognition of Parkinson's disease and Alzheimer's disease. In this framework, the feature maps of the deep CNN model were used to visualize the cognitive dysfunction-related regions, which could provide quantitative biomarkers of cognitive dysfunction in various neurodegenerative diseases. However, the existing deep learning frameworks often ignored the deep-seated topological information of the brain network, which may cause the sub-optimal performance in brain disorder identification.

2.2. Graph convolutional network for brain connectivity analysis

In recent years, research on extending deep learning methods to graph data has received increasing attention. Graphs were widely used as computational model to capture interactions between individual elements represented as nodes in the graph (Yue et al., 2020). Network embedding is a method of transforming nodes into lower dimensional representations while maximally preserving the network structure (Li et al., 2022b). Embedding brain networks into meaningful low-dimensional representations could improve classification performance for disease diagnosis. GCN have been proposed as an effective embedding model in several researches [18; 19] (Zeng et al., 2019), which naturally combine structural information and node features in the learning process. Structural or functional connections of the human brain network represent the information interaction patterns between different regions, and these connections could serve as graph architectures representing the whole-brain scale of each subject. In the field of graph learning, the topological properties of brain network could better describe the structural or functional characteristics. Recent studies have shown that GCN is more effective than other methods in learning representations (Ma et al., 2019). Parisot et al (Parisot et al., 2018) used the GCN method to classify Autism spectrum disorder. Combining imaging and non-imaging data in a single framework, they achieved significant improvements in classification accuracy through precise graph structure. Song et al (Song et al., 2022) introduced a classification framework combining GCN and machine learning. They used a dual-modal fusion brain connectivity network of resting-state fMRI and diffusion tensor imaging to identify significant memory concern and MCI. Moreover,

some studies learned network representations through Siamese graph convolutions, and the final category of test subject is determined according to the similarity between test subject and known subject. Ktena et al (Ktena et al., 2018) applied the Siamese GCN in autism classification, using two GCNs with weight-shared to measure the difference between pairwise functional networks. What the attention needs is that the existing GCN can not be directly used for our task. Specifically, (1) GCN generally requires a high sample size, while our sample is limited; (2) The traditional pooling method has high spatial complexity. In view of the above problems, we give corresponding solutions in the proposed methods.

2.3. The application of Siamese networks in the diagnosis of mental diseases.

Siamese networks have outstanding performance in solving the problem of feature similarity. Reihaneh et al proposed a novel approach, which was a deep Siamese-based 3D CNN. The trained model could use paired networks as input to accurately distinguish subjects, and it could be extended to situations that are not visible (Hassanzadeh and Calhoun, 2020). Ma et al built the end-to-end graph similarity learning framework named Siamese GCN for fMRI analysis. This method learned brain network representations via a supervised metric-based approach with Siamese neural networks using two graph convolutional networks as the Siamese networks. The experiment used a supervised metric based pattern for learning brain network representations, where the Siamese network owned two GCNs as dual networks. (Ma et al., 2019). Yang et al improved the Siamese network architecture equipped with a graph convolutional encoder to learn about graph (or topic) level embedding, which preserved application dependent similarity metrics between brain networks. This algorithm increased the number of training samples and introduced the flexibility of incorporating additional prior information through a specified target graph level distance (Li & Mateos, 2022).

3. Materials and method

3.1. Subjects and image Pre-Processing

All the PET data we used in this paper are from ADNI (Alzheimer's Disease Neuroimaging Initiative: <http://www.loni.ucla.edu/ADNI/>) Database. The ADNI database collects data from 50 sites around the world to help researchers use these data for scientific research. In the ADNI database, the progression of AD is studied by assessing the state of change in brain function and structure.

In this research we selected 230 subjects, including those with MCI and normal controls (NC). MCI patients had reported the subjective memory concern. However, no significant level of impairment was present in other cognitive domains; essentially, their activities of daily living were preserved and no signs of dementia existed. Table 1 lists the demographic data of all subjects.

All PET image data in this study were collected in the resting state. The following steps were mainly used to process PET images: (1) PET images were registered and averaged; (2) The images were normalized, and then smoothed to produce a uniform resolution; (3) We registered

Table 1
Demographic and clinical characteristics of subjects.

	MCI	NC	<i>p</i>
N (total N = 230)	117	113	–
Age	72.37 ± 5.21	75.67 ± 6.22	0.3451 ^b
Gender (male:female)	64:53	60:53	0.383 ^a
FAQ	3.4 ± 2.9	0.1 ± 0.6	<0.001 ^a
MMSE	26.1 ± 2.3	27.8 ± 1.6	<0.001 ^a
CDR	0.5	0	<0.001 ^a

Data are presented as a mean ± standard deviations. *p* was obtained using ^athe two-sample *t* test and ^bthe chi-square test.

the images into a unified MNI(Montreal Institute of Neurology) template.

In the subsequent network construction, we will use the 246 template (Brainnetome Atlas)(Fan et al., 2016) to divide brain regions. In addition, linear regression is utilized before formal calculations to eliminate the effects of gender and age level on the measurements.

3.2. Proposed methods

In order to make the algorithm more consistent with the purpose and data of this study, we proposed a feature extraction framework of GCN based on self-attention pooling mechanism(SAGCN). Further, we drew the framework from the Siamese network idea to construct QS-SAGCN. The overall research idea consists of the following:

Multiple protein features network(MPN): Since only considering the pairwise correlation between brain regions may ignore the potential impact of other regions. This study used sparse linear regression to quantify the cross-region relationships of protein features, and considered the aforementioned regional features as regression target and prediction vector. Besides, existing research has mainly focused on modeling low-order relationships between nodes in brain networks(Yin et al., 2022), lacking the exploration of the underlying higher-order associations that exist in the network itself. Furthermore, higher-order networks based on MPN is considered to describe the state of information exchange between brain regions. **Higher-order MPN(MPHN):** For overcoming the obstacles of dimensionality and making up for the limited amount of data, we calculated MPHN on the basis of MPN. In MPN, each vertex corresponds to multiple pairs of brain regions, and each edge represents how several pairs of brain regions interact. In this way, MPHN can reveal higher levels and more complex interactions than MPN, and enrich the amount of data.

SAGCN: Attention mechanisms have been widely used in deep learning research. This mechanism can be applied to GCN for paying more attention to important nodes and less attention to unimportant nodes in MCI patient. In this study, the self-attention mechanism algorithm was applied to the GCN framework to form SAGCN, obtaining discriminative features.

QSN: In addition, this study will refer to the concept of Siamese network in order to improving the recognition rate under the condition of small sample size. Based on the traditional Siamese network, the quadruple Siamese network(QSN) model collocating SAGCN was built by using the image idea to enhance the robustness of the classification method, called QS-SAGCN.

The computational environment of this experiment were conducted on a PC with a 3.7 GHz Intel Core CPU, an NVIDIA GeForce RTX 3080 Ti graphics card, 32 GB DDR4 RAM, and a 1 TB solid-state disk drive. The proposed method is realized using MATLAB R2018a and Python 3.6. The deep learning frameworks are TensorFlowGPU 1.8 and Keras 2.2.

3.2.1. The brain network construction

In this study, we aimed to improve the accuracy of MCI autodiagnosis by using two kinds of PET images, and discovered potential neural mechanisms by constructing the network based on multiple protein features.

We used the linear regression model with ℓ_1 -norm penalty to construct a network based on multiple protein features, which was called the the least absolute shrinkage and selection operator (LASSO) (Zheng et al., 2018). LASSO algorithm could automatically filtered out the regression quantity that was not closely related to the prediction target(Ryali et al., 2012), so it could more accurately describe the nature of the brain connection pattern(Sporns et al., 2004).

For network construction, we defined each subject as $X =$

$[x_1, x_2, x_3, \dots, x_q]^T \in R^{1 \times q}$, where q was the number of brain regions, and x_q was a sequence composed of p protein features of the q -th brain region. Before constructing the network, we normalized the regional value using the whole brain average and the standard deviation of each protein feature. Specifically, the linear regression model was defined as $y = Aw$, where w represented the regression coefficient and y represented the target vector. In the j -th regression, data matrix was $A = [x_1, x_2, x_{j-1}, x_{j+1}, \dots, x_q]$, and $q-1$ brain regions as independent variables could linearly express y , as shown in formula (1).

$$\min_w \|y - Aw\|_F^2 + \lambda \|w\|_1 \tag{1}$$

Specifically, λ represented the sparsity control parameter. We constructed MPN with the varying $\lambda \in 10^{-4}, a \times 10^{-3}, b \times 10^{-2}, a, b \in 1, 2, \dots, 9$ and calculated the classification performance of different λ in MPN. Meanwhile, the coefficient that was not 0 in the regression coefficient matrix was replaced with 1, and then a binary matrix was obtained. 0 and 1 respectively represented the coefficients of similarity between the target brain region and other brain regions. Fig. 1 showed a MPN containing three types protein (Tau protein content, Aβ protein content, and Tau protein and Aβ protein content ratio for each region). The regression coefficient of each row in the matrix indicated the brain region represented by this row was linearly expressed through other brain regions. The coefficient 1 indicated that the brain region represented by this row is highly similar to the target brain region, while the coefficient 0 indicated that the similarity was low. Because each line of the MPN solution process was independent, the resulting MPN was an asymmetric network. Notably, asymmetric networks may only represent the similarity between brain proteins and cannot reflect other information such as transmission mechanism.

Inspired by previous research, we utilized Pearson coefficient to build a higher-order network (MPHN) on basis of the MPN. Then calculated the higher-order coefficients between multiple regions to represent the information interaction of brain regions, as shown in Fig. 1. On one hand, using this method can make up for the shortcomings of limited data, and more effective features could be extracted through brain networks of different scales; on the other hand, using high-order brain networks could explore more additional information about the brain. Significantly different from lower-level networks, each vertex of higher-order networks corresponded to a pair (or even multiple pairs) of

brain regions, and each edge characterizes how pairs of brain regions interacted. The MPHN is as follows:

$$C_{ij}^{(l)}(k) = \text{corr}(x_i^{(l)}(k), x_j^{(l)}(k)) \tag{2}$$

In addition, we also used the method in formula (2) to calculate the third-level higher-order network on the basis of $\text{MPHN}_{(H+)}$, which is defined as $\text{MPHN}_{(H++)}$.

3.2.2. Graph convolutional network

The strength of deep learning model lies in their ability to automatically discover latent or abstract information from high-dimensional neuroimaging data, which could be an important step toward understanding complex mental disorders. GCN owned many advantages in processing non-Euclidean spatial data in form of the graph data (He et al., 2018). GCN automatically extracted the features of brain network in an end-to-end manner for recognition (Ghorbani et al., 2022). GCN enhanced node features by transmitting and updating node information in the graph, and utilized the obtained node information to mine useful brain connection network patterns for disease classification (Jiang et al., 2020). Capable of capturing network topology, GCN have been proven to own advantages in learning network representations tailored to identify specific brain diseases (Grattarola et al., 2021).

The classical GCN could be thought of as a Laplacian smoothing operator for node features on a graph structure. GCN is consisted by a series of convolutional layers, and each layer in the architecture is followed by the Rectified Linear Unit (ReLU) activation function to increase nonlinearity (Ghorbani et al., 2022). In the application of image data, in order to reduce complexity of the model, it was often necessary to replace the feature resolution with the spatial resolution. To achieve this, statistics were computed in local neighborhoods using pooling layers (Roy et al., 2019). The specific details are shown in formula (Jiang et al., 2020).

$$E^{(l+1)} = \text{Relu}(\text{Pool}(\tilde{D}^{-1/2} \tilde{A} D^{-1/2} E^{(l)} W^{(l)})) \tag{3}$$

Formally, $A \in R^{n \times n}$ was an adjacency matrix, which defines whether there was a connection between two nodes. $\tilde{D}_{ii} = \sum_j \tilde{A}_{ij}$. W was a trainable weight matrix; $E^{(l+1)}$ is the feature matrix (i.e. features computed after 1 steps of the GCN, and node embeddings $E^{(l)}$ generated from previous feature transfer steps). After L iterations of training, a complete

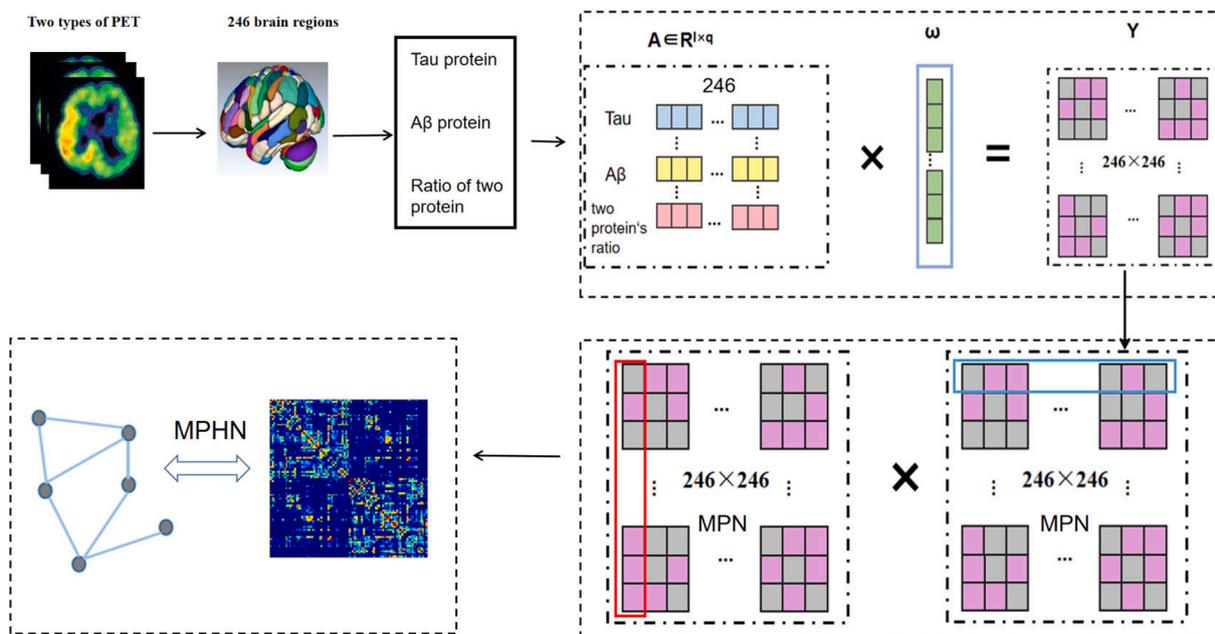


Fig. 1. Construction of multi-level protein features network.

GCN could be obtained, as shown in Fig. 2.

In traditional GCN, feature extraction could be carried out through the relationship between nodes and edges. However, in order to make GCN more in line with the characteristics of data, the following improvements will be made in this study. The traditional pooling method simply considered the topological form of the graph, and the spatial resolution was high. Therefore, attention mechanism was considered to be added to the GCN to reduce the learned parameters and enhance the effectiveness of features. In addition, in order to obtain higher classification accuracy in limited data, this study utilized GCN to extract data features, and added feature similarity measurement mechanism to improve the accuracy of data recognition.

3.2.3. Self-attention pooling of GCN

For the purpose of better reflecting the hierarchical structure of input data, pooling layer effectively reduced the number of parameters and improved computational efficiency. Self-attention pooling could use features and topology to produce hierarchical representations with reasonable temporal and spatial complexity. The GCN based on self-attention mechanism(SAGCN) was proposed in this study. The network frame is shown in Fig. 3. The update formula of node characteristic matrix and adjacency matrix is as formula (4):

$$Z = \sigma(\tilde{D}^{1/2} \tilde{A} \tilde{D}^{1/2} X \Theta_{att}) \quad (4)$$

Where $\Theta_{att} \in R^{F \times 1}$ was the only parameter in the Self-attention Pooling layer. Since the Self-attention score was obtained by the convolution of the graph combining A and X in formula (4), the pooling result could reflect the characteristics and topological structure of the graph.

The following operations were required to update the adjacency matrix and node characteristics by using the pooling layer. According to the importance score and topology of the nodes, the pooling operation could be performed, and some less important nodes could be discarded to form a new graph structure. As shown in formula (5), where KN meant

to select the proportion of nodes to keep according to the attention score value and i indicated an index operation.

$$i = top_rank(attention_score, KN) \quad (5)$$

3.2.4. The modified Siamese network in classification

In this section, Siamese network was incorporated into the classification framework, and the original Siamese network performs classification by measuring how similar two inputs were. Through the calculation of Loss, the similarity of the two inputs was evaluated(Roy et al., 2019). However, the original Siamese network framework only simulated the pairwise relationship of the samples, ignoring its potential complex (such as quadruple) associations. In this work, we adopted a quadruple learning framework in order to mine the quadruple similarity between subjects. New graph embedding based on SAGCN learning was used to measure the similarity between subjects. The quadruple Siamese network(QSN) contained four identical networks (with common parameters), and a bridge was built between them by constructing a loss function.

Specifically, as could be seen from Fig. 4, the T -th QSN input was $\{X_p^t, X_{a1}^t, X_n^t\}$ and $\{X_p^t, X_{a2}^t, X_n^t\}$, where X_p^t and X_n^t were positive and negative samples, and X_{a1}^t and X_{a2}^t were anchor point sample 1 and anchor point sample 2 respectively. Where X represented the brain map of the subject, which contained a set of features F and an adjacency matrix A . That was, positive sample and anchor sample 1 belong to the same category. The negative sample and anchor sample 2 belong to the same categories. In this work, based on the graph representations they learned, each QSN encouraged samples belonging to the same class to be similar, and representations from different classes encouraged to be different.

$$\min_w \frac{\lambda}{2} \|w\|_2 + \sum_{i=1}^N \max[0, (d^+ + \alpha - d^-)] \quad (6)$$

$$d^+ = |f_r(X_p^t) - f_r(X_{ai}^t)|, (i = 1 or 2) \quad (7)$$

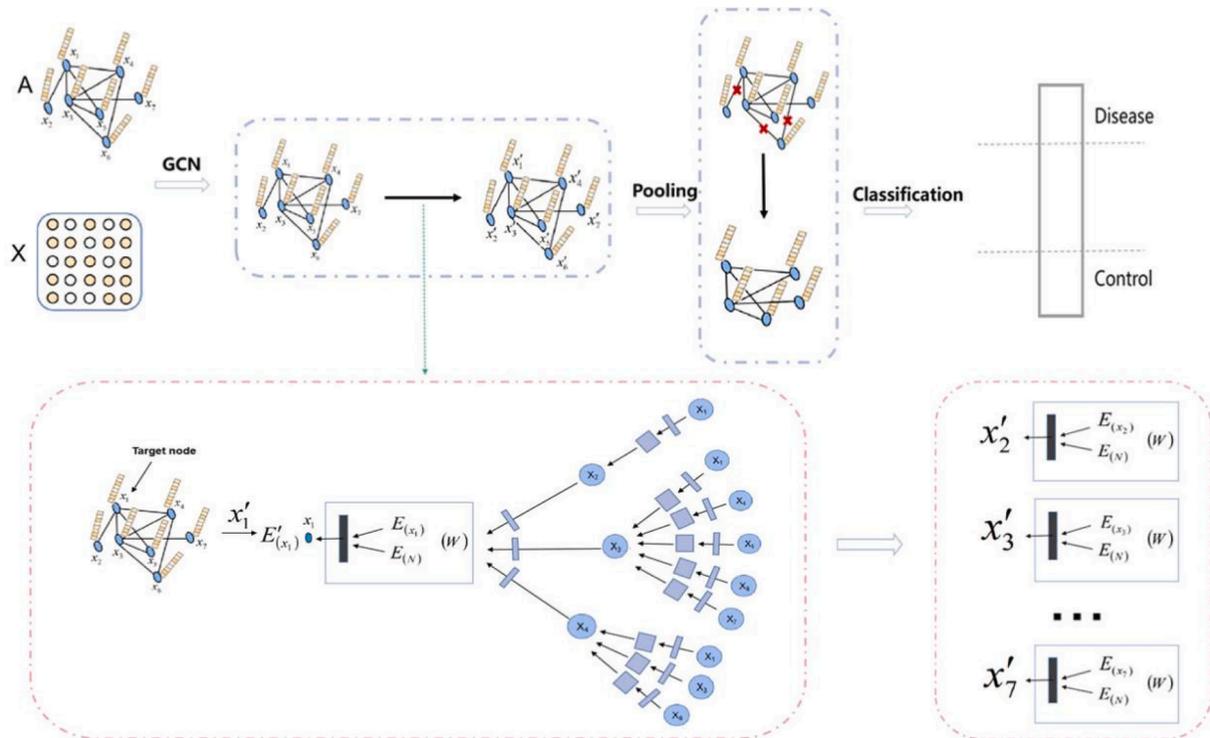


Fig.2. Overall frame diagram of GCN. Where x_i represented the feature vector of the node, x'_i represented the feature vector of the new node obtained after the graph convolution. In the pooling operation, the red fork sign indicated that the pooling operation can discard some less important nodes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

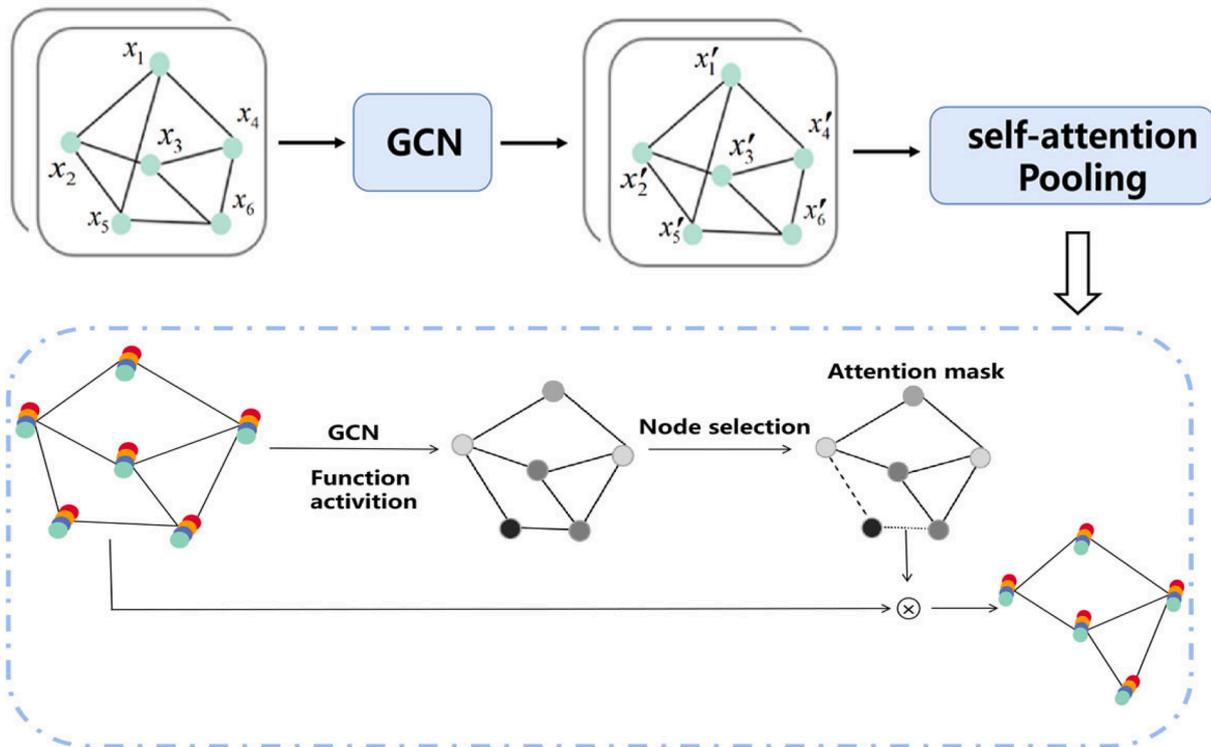


Fig. 3. Pooling operation. After node selection, the dotted lines in the graph connect the node with low attention scores, and then remove it to get a new graph.

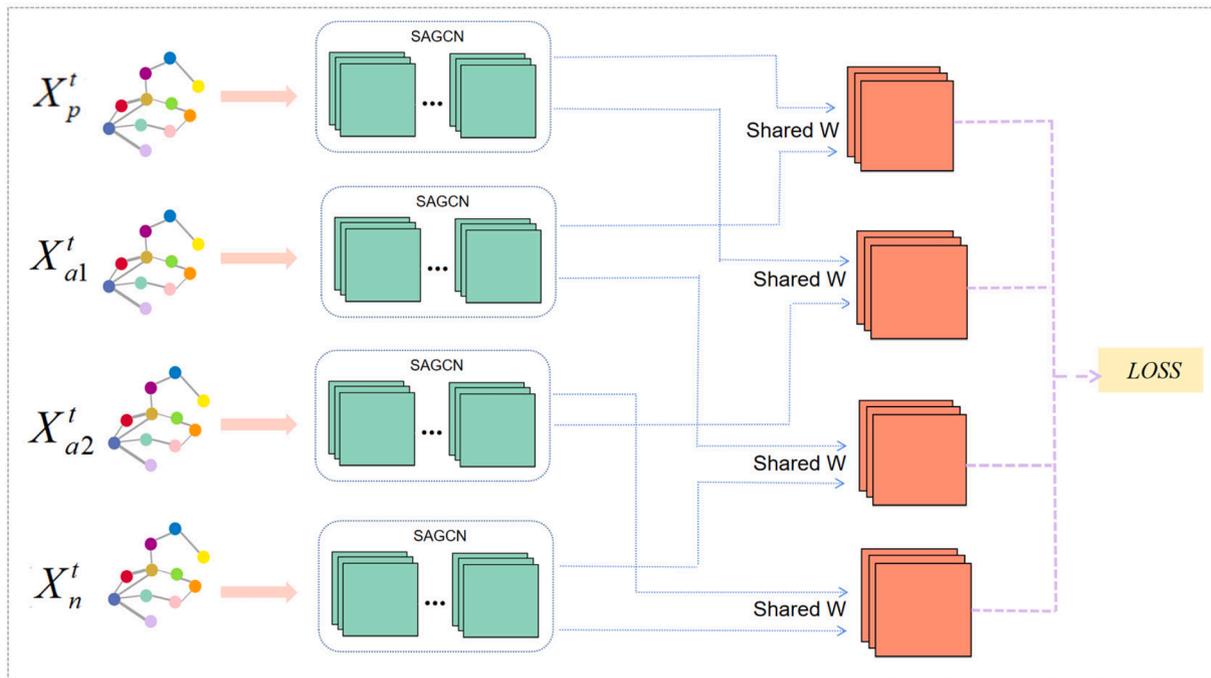


Fig. 4. The quadruple Siamese network.

$$d^- = |f(X_{ai}^t) - f(X_n^t)|, (i = 1 \text{ or } 2) \tag{8}$$

Where $f(\bullet)$ represented the convolution process of SAGCN. In the process of calculating different d , the same set of SAGCN parameters were used. The loss function was shown in formula (6), where a was a super parameter used to distinguish whether the T -th identification was valid. In the process of parameter debugging, when referring to the positive anchor point, make d^+ as small as possible and d^- as large as

possible. Similarly, the negative anchor point was taken as the reverse.

4. Experiment analysis

4.1. The most distinguishing features in MCI diagnosis

To further analyze whether the network construction method proposed in this study has physiological significance. We put interest in the

third-level higher-order network(MPHN_(H+++)) properties with the higher final accuracy. Our network construction method mainly based on the interactive relationship between different brain regions. Therefore, in this section, we focused on the connections between brain regions at the MPH_(H+++). The final results were shown in Table 2 and Fig. 5.

We found that all different connections were concentrated in the Temporal lobe, Frontal lobe, Amygdala and Parahippocampal gyrus. These regions mainly colocalized with DMN. It has sufficient evidence that the changes of brain structure, function and cognition in MCI patients are related to the abnormalities in brain functional subnetworks. Previous studies indicate that Aβ protein aggregation may be driven by the total flow of neuronal activity(Zhang et al., 2021a). Tau protein aggregation may be driven by cross neuronal diffusion, generating a neurodegenerative pattern consistent with a specific functional network and ultimately leading to a specific clinical phenotype(Ismail et al., 2020). These conclusions are consistent with our findings. Meanwhile, the pathological changes regions of MCI patients were mainly concentrated in above regions(as shown in Table 2), which indicated that our results are robust.

4.2. Comparison for MCI diagnosis using different feature extraction

In order to verify effectiveness of the proposed method, we conducted extensive experimental comparison based on the following six networks. The classification performance of above six networks were listed in Table 3 and Fig. 6. We compared the classification effects of different types of networks from the classification accuracy(ACC), sensitivity(SEN), specificity(SPE), and area under the curve(AUC).

Among them, the MPN represented the original regression network, and MPH_(H+) represented the second-level higher-order network of MPN, where MPH_(H+)-GCN represented the combination of the constructed MPH_(H+) and GCN network architecture. Similarly, MPH_(H+++) represented the third-level higher-order network of MPN, where MPH_(H+++)-GCN represented the combination of the constructed MPH_(H+++) and GCN network architecture. SAGCN + MPH_(H+) and SAGCN + MPH_(H+++) represented the combination of the different constructed MPH_(H+) and GCN network architecture with with self-attention pooling operation, and other representations were similar.

Three conclusions could be drawn from Table 3: (1) adopting MPH_(H+) features is superior to MPH_(H+++) features, which indicated that higher-order brain network can obtain more discriminative features by describing more comprehensive information transmission of brain

Table 2

Brain regions with significant group effect in the connections between brain regions for MCI patients and HC.

Region 1	Region 2	P value	Region 1	Region 2	P value
SFG_L_7_1	Hipp_L_2_2	0.00003	SFG_R_7_6	Hipp_L_2_2	0.0072
SFG_L_7_1	Hipp_R_2_2	0.00006	SFG_R_7_6	Hipp_R_2_2	0.0059
SFG_R_7_1	Hipp_L_2_2	0.0001	LOC_R_2_2	MFG_L_7_1	0.0050
SFG_R_7_1	Hipp_R_2_2	0.00006	LOC_R_2_2	MFG_R_7_1	0.0043
SFG_L_7_2	Hipp_L_2_2	0.0012	LOC_R_2_2	MFG_L_7_2	0.0030
SFG_L_7_2	Hipp_R_2_2	0.0013	BG_R_6_5	ITG_L_7_6	0.0003
PoG_L_4_3	FuG_L_3_2	0.0030	BG_R_6_5	ITG_R_7_6	0.0109
PoG_L_4_3	FuG_R_3_2	0.0118	SFG_L_7_1	PhG_L_6_3	0.0042
INS_L_6_2	ITG_R_7_6	0.0107	SFG_L_7_1	PhG_R_6_3	0.0042
CG_R_7_2	FuG_L_3_2	0.0308	Tha_R_8_7	MTG_L_4_1	0.0055
CG_R_7_2	FuG_R_3_2	0.0130	Tha_R_8_7	MTG_R_4_1	0.0299
LOC_L_4_3	MFG_R_7_1	0.0072	BG_R_6_6	IPL_L_6_6	0.0037
LOC_L_4_3	MFG_L_7_2	0.0074	BG_R_6_6	IPL_R_6_6	0.0005

Abbreviations: SFG: Superior Frontal Gyrus; PoG: Postcentral Gyrus; INS: Insular Gyrus; CG: Cingulate Gyrus; LOC: lateral Occipital Cortex; Hipp: Hippocampus; FuG: Fusiform Gyrus; ITG: Inferior Temporal Gyrus; MFG: Middle Frontal Gyrus; SFG: Middle Frontal Gyrus; BG: Basal Ganglia; Tha: Thalamus; ITG:Inferior Temporal Gyrus; PhG: Parahippocampal Gyrus; MTG:Middle Temporal Gyrus; IPL: Inferior Parietal Lobule; L: left; R: right.

regions; (2) The GCN with self-attention mechanism for pool operation can take into account the node characteristics and network topology, and extract more discrimination features; (3) For MPH_(H+) and MPH_(H+++), the performance of feature extraction and feature layer fusion based on SAGCN was the best, which proved the effectiveness of feature fusion.

4.3. Comparison of the different types of Siamese networks

We compared the effect of three different Siamese Networks on the performance of SAGCN classifier. The previous experimental results showed that the third-level higher-order network(MPH_(H+++)) was the brain network with the best discrimination performance, so the analysis in this section was only carried out on the basis of MPH_(H+++). The original Siamese networks was used to deal with the case where two inputs are “relatively similar”(Hoffer & Ailon, 2015). The experiment data were applied to the original Siamese Networks composed of the same scale network, to compare the performance difference between QSN and the original Siamese network. In addition, to further verify the performance of the QSN, we added the triplet Siamese network for comparison(Zhu et al., 2019), which consisted of three instances of the same feed-forward network (with shared parameters). When inputting three samples, the network outputted two intermediate values, namely the distance between the embedding representation of two inputs and the representation (anchor) of the third input.

4.4. Comparisons with state-of-the-art methods

In order to demonstrate the advantages of the MCI classification method proposed in this study, we compared the existing state-of-the-art methods. Zhang et al proposed a new method for data feature representation combined with multi-view information enhancement, in which an improved multi-task feature selection method was used to select the optimal image features for MCI classification. Finally, using the multi-kernel SVM classification method, the classification accuracy of 88.5% was obtained(Liu et al., 2020). Ashtari-Majlan et al built a multi-stream deep convolutional neural network to fed with features-based imaging data to classify MCI patients. The experimental results on the ADNI dataset showed an accuracy of 85.96%(Ashtari-Majlan et al., 2022). Li et al constructed a new adaptive dynamic functional connectivity model, supplemented by a deep spatiotemporal feature fusion method, for the recognition of mild cognitive impairment (MCI). The experimental results showed that an accuracy of 87.7% is achieved (Li et al., 2020). Therefore, it could be seen that our proposed MCI classification method is significantly superior to three existing state-of-the-art methods.

5. Discussion

5.1. The effectiveness of network construction method

In this study, multiple regression model(MPN) was employed to construct the original brain network, taking into account the influence of all brain regions. MCI was accompanied by abnormalities that affected multiple brain regions (i.e. large-scale networks) rather than a single isolated region(Du et al., 2007, Stam et al., 2007). Previous study has shown that Aβ protein and Tau protein well reflect the pathological changes(Firouzi et al., 2018). Besides, Tau protein and Aβ protein played an important role in the development of MCI. The atrophy of cerebral cortex and brain function damage of MCI patients were closely related to both proteins, the distribution of these two proteins might have some certain correlations(Da et al., 2014). Therefore, this study attempted to regard the ratio in each brain region of two proteins as the new feature(Santangelo et al., 2020).

Furthermore, researchers found that higher-order method was generally more sensitive to subtle changes in signals than low-order

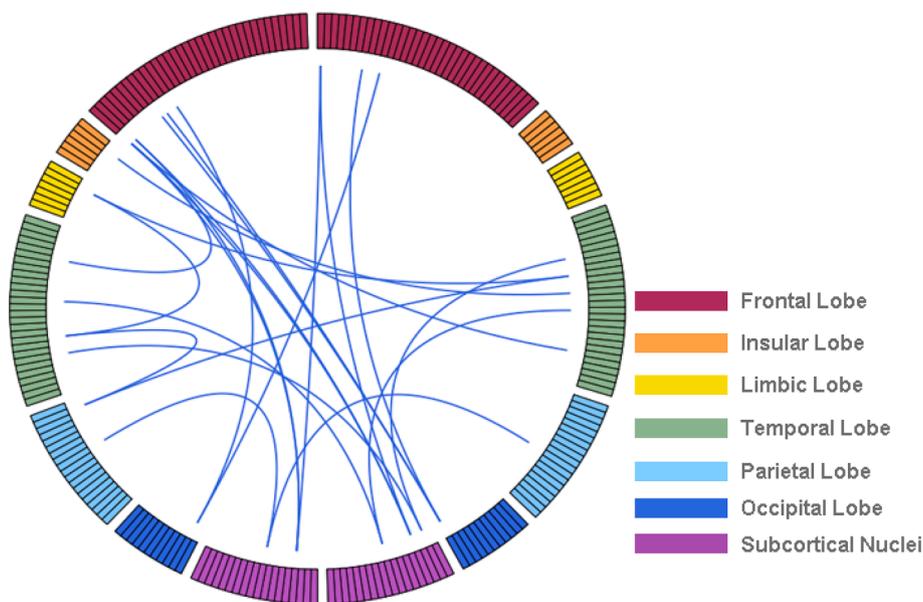


Fig. 5. Nodes with differences in MPH_(H++).

Table 3
MCI classification results with different feature strategies.

	ACC	SEN	SPE	AUC
MPHN _(H+)	0.714 ± 0.012	0.722 ± 0.024	0.713 ± 0.024	0.730 ± 0.032
MPHN _(H++)	0.762 ± 0.013	0.783 ± 0.031	0.755 ± 0.016	0.780 ± 0.033
MPHN _(H+) -GCN	0.812 ± 0.031	0.801 ± 0.021	0.825 ± 0.026	0.815 ± 0.021
MPHN _(H++) -GCN	0.865 ± 0.021	0.854 ± 0.012	0.844 ± 0.021	0.847 ± 0.031
MPHN _(H+) -SAGCN	0.905 ± 0.013	0.916 ± 0.023	0.892 ± 0.043	0.899 ± 0.033
MPHN _(H++) -SAGCN	0.935 ± 0.022	0.903 ± 0.025	0.895 ± 0.032	0.922 ± 0.021

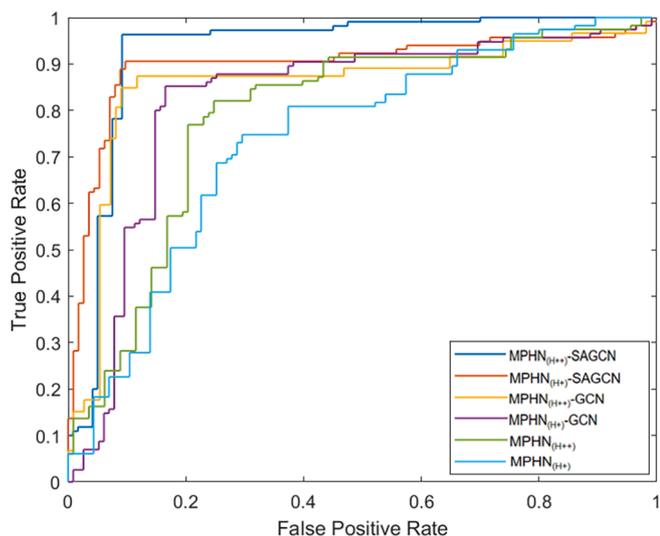


Fig. 6. ROC curve.

method(Jung et al., 2017). In order to mine higher-order information among brain regions, we constructed the two type of higher-order networks. Results in Table 3 showed that the classification accuracy of the

MPHN_(H+) and the MPH_(H++) methods were 71.4% and 76.2% respectively. It could be seen that the higher-order network model improved the accuracy of disease classification because it considered the dynamic higher-order characteristics of brain network. Zhou et al defined higher-order brain networks among three brain regions. Results showed that such higher-order networks made full use of the complementary information in different brain networks and could obtain more discriminative network attributes(Zhou et al., 2018). Researches indicated some higher-order statistics (e.g., correlation between different edges) may also provide additional and useful information for functional connectivity analysis(Fan et al., 2020, Feng et al., 2021). Therefore, in the study, the higher-order information between brain regions were embedded in the brain network and applied to identify MCI patients, which improved the recognition accuracy. In general, one brain region usually interacted with multiple brain regions physiologically, while low-order methods tended to ignore the relationship between multiple brain regions. It is necessary to study effective higher-order brain network construction methods to reveal the correlation of ordered patterns among multiple brain regions from the perspective of protein networks.

5.2. The effectiveness of Self-attention pooling

In recent years, it has been proposed to extend convolutional neural network to graph data, including redefining convolution and graph pooling. Therefore, the application of pooling in graph significantly improved the effectiveness of feature extraction. Previous studies have adopted pooling methods that only consider graph topology(Rhee et al., 2017). With the growing interest in graph pooling, several improved methods have been proposed(Zhang et al., 2018, Yao et al., 2021), which used the characteristics of nodes to obtain smaller graph representation. In this study, a graph pooling method based on self-attention was applied GCN. The node characteristics and the graph topology could be fully considered to further improve the accuracy of classification. In previous study, by introducing a structured multi-head self-attention mechanism, a new architecture was developed to extract effective representations of graph. Three forms of attention mechanisms were employed, namely node attention, layer attention, and graph attention (Gao and Ji, 2019). The self-attention pooling is the hierarchical pooling method, which considered node characteristics, graph topology and reduced computational complexity(Lee et al., 2019). Li et al. showed

that the self-attention pooling could measure the importance of features in classification. Then the attention scores were used to weight features for highlighting the contribution of these discriminative features in classification model(Li et al., 2022a). The self-attention pooling mechanism achieved feature optimization and parameter reduction while extracting more accurate global features. The self-attention pooling algorithm could adaptively extract key features that affected classification results, and considered the contribution of each feature to the final result (Ihalage and Hao, 2022). Moreover, Zhao(Zhao et al., 2022) proposed a self-attention coherent clustering electroencephalogram(EEG) emotion recognition model based on multi pool graph convolutional network in the research, and obtained good recognition results.

5.3. The effectiveness of different types of Siamese networks

We used the quadruple Siamese networks strategy with SAGCN to simulate the relationship between subjects, and verifying its effectiveness in experiments. In this study, we compared the original Siamese Networks and triple Siamese Networks with the experimental model (QSN), and the results were shown in Table 4. In comparison, QSN was superior to the other networks in ACC and AUC values. This may be due to the higher-order relationship between subjects captured by multiple comparisons, which was helpful to discover the real data structure and improve the classification results. Furthermore, considering that the number of subjects in neuroimaging datasets were usually limited (such as tens or hundreds), multiple comparison combined with cross validation strategy could be used as a flexible data enhancement approach (Song et al., 2019). In previous studies(LaFerla & Oddo, 2005; Wang et al., 2018), improved models have been proposed by using the idea of Siamese Network, mainly by comparing the similarities between the identified samples and the reference samples. The results showed that these methods had strong generalization ability and superiority in the task of sample recognition. Moreover, the study proposed a framework based on Siamese networks to extract information features from the differences between the two hemispheres of brain. Using this method, the differences between dopaminergic regions in two hemispheres were found to be an effective biomarker for identifying Parkinson's disease patients(Arco et al., 2022). Alaverdyan et al. employed the Siamese network composed of convolutional self autoencoders as subnetworks, and combined with features extracted from MRI images to identify subtle lesions in epileptic patients(Alaverdyan et al., 2020).

6. Conclusion

The proposed framework was consisted of the GCN combined with self-attention pooling(SAGCN) and collaborated with the quadruple Siamese network(QSN) for MCI classification, which achieved the best MCI classification accuracy (93.5%). Specifically, the third-level higher-order network(MPHN(H++)) of PET images could provided important information for the classification of MCI patients and NCs, and revealed cross-regional alterations in the protein content properties of each patient. Meanwhile, we combined SAGCN and QSN as QS-SAGCN into the classification model, which could utilize the topological structure of the MPHN(H++) to generate a feature representation with reasonable spatial complexity and classify diseases. The excellent classification performance achieved by the experimental framework in this study strongly demonstrated its effectiveness in identifying NC and MCI patients, and may have potential clinical applications in the automatic diagnosis of neuropsychiatric disorders.

7. Limitation and Future work

In order to further improve the current framework, the following limitations need to be considered. (1) The recognition of progressive MCI and stable MCI are not considered in the method. In future study, the classification of progressive MCI and stable MCI will be included in

Table 4

MCI classification results with different Siamese network.

	ACC	SEN	SPE	AUC
Original Siamese network	0.852 ± 0.021	0.860 ± 0.014	0.845 ± 0.016	0.848 ± 0.032
Triplet Siamese network	0.889 ± 0.013	0.883 ± 0.021	0.875 ± 0.026	0.890 ± 0.013
Quadruple Siamese network	0.935 ± 0.022	0.903 ± 0.025	0.895 ± 0.032	0.922 ± 0.021

the method analysis, and the effectiveness of the method will be verified using the classification results of different types of MCI. (2) In the current study, only a single model data is used, without considering the potential relationships between different modal data. Therefore, it is hoped that multimodal joint analysis can be included in future work. (3) Only one basic way is used to construct higher-order networks in this article. In future research, we can use the various ways to build different higher-order brain networks to help us better understand the pathophysiological characteristics of brain diseases. (4) The amount of data in this study is limited, and data from different sites will be considered in the following research to further verify the effectiveness of the experimental framework.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data used in the experiment was the ADNI public database.

Acknowledgments

The Alzheimer's Disease Neuroimaging Initiative (ADNI; National Institutes of Health Grant U01 AG024904) funded the collection and sharing of data for this project. This work was supported in part by Doctoral Scientific Research Foundation of Shandong Technology and Business University (No. BS202016) and the National Natural Science Foundation of China (62176140, 82001775, 61772319, 61873177, 61972235, 61976125 and 61976124); in part of Horizontal project of Yantai Vocational College (No. HX202245); in part of Shandong Provincial College Youth Innovation Introduction and Education Program Team(Intelligent Analysis and Emergency Management of Security Big Data).

References

- Alaverdyan, Z., Jung, J., Bouet, R., et al. (2020). Regularized Siamese neural network for unsupervised outlier detection on brain multiparametric magnetic resonance imaging: Application to epilepsy lesion screening. *Medical Image Analysis [J]*, 60, Article 101618.
- Arco, J. E., Ortiz, A., Castillo-Barnes, D., et al. (2022). Quantifying inter-hemispheric differences in Parkinson's disease using Siamese networks [C] // *Springer; City*, 156–165.
- Ashtari-Majlan, M., Seifi, A., & Dehshibi m m. (2022). A multi-stream convolutional neural network for classification of progressive mci in alzheimer's disease using structural mri images. *IEEE Journal of Biomedical and Health Informatics [J]*, 26, 3918–3926.
- Bai, Y., Hu, J., Su, J., et al. (2020). Pyramid pooling module-based semi-Siamese network: A benchmark model for assessing building damage from xBD satellite imagery datasets. *Remote Sensing [J]*, 12, 4055.
- Behrad, Fatemeh, & Abadeh, Mohammad Saniee (2022). An overview of deep learning methods for multimodal medical data mining. *Expert Systems with Applications, Article 117006*.
- Chandra, A., Valkimadi, P. E., Pagano, G., et al. (2019). Applications of amyloid, Tau, and neuroinflammation PET imaging to Alzheimer's disease and mild cognitive impairment. *Human brain mapping [J]*, 40.

- Da, X., Toledo, J. B., Zee, J., et al. (2014). Integration and relative value of biomarkers for prediction of MCI to AD progression: Spatial patterns of brain atrophy, cognitive scores, APOE genotype and CSF biomarkers. *NeuroImage: Clinical* [J], 4.
- Dani, M., Brooks, D. J., & Edison p. (2016). Tau imaging in neurodegenerative diseases. *European Journal of Nuclear Medicine and Molecular Imaging* [J], 43.
- De, Arijit, & Chowdhury, Ananda S. (2021). DTI based Alzheimer's disease classification with rank modulated fusion of CNNs and random forest. *Expert Systems with Applications*, 169, Article 114338.
- de Mendonça, L. J. C., & Ferrari, R. J. (2023). Alzheimer's Disease Neuroimaging Initiative. Alzheimer's disease classification based on graph kernel SVMs constructed with 3D texture features extracted from MR images[J]. *Expert Systems with Applications*, 211, Article 118633.
- DI MARTINO A, YAN C-G, LI Q, et al. 2014. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry* [J], 19: 659-667 %@ 1476-5578.
- Du, An-Tao, et al. (2007). Different regional patterns of cortical thinning in Alzheimer's disease and frontotemporal dementia. *Brain*, 130(4), 1159–1166.
- Duan, H., Jiang, J., Xu, J., et al. (2017). Differences in β brain networks in Alzheimer's disease and healthy controls. *Brain Research* [J], 1655.
- Fan, L., Li, H., Zhuo, J., et al. (2016). The human brainnetome atlas: A new brain atlas based on connectonal architecture. *Cerebral cortex* [J], 26.
- Fan, X., Gong, M., Xie, Y., et al. (2020). Structured self-attention architecture for graph-level representation learning. *Pattern Recognition* [J], 100.
- Feng, C., Zhu, Z., Cui, Z., et al. (2021). Prediction of trust propensity from intrinsic brain morphology and functional connectome. *Human brain mapping* [J], 42.
- Firozian, A., Whittington, A., Searle, G. E., et al. (2018). Imaging β and Tau in early stage Alzheimer's disease with [18F] AV45 and [18F] AV1451. *EJNMMI research* [J], 8.
- Gao, Hongyang, & Ji, Shuiwang. Graph u-nets. International conference on machine learning. PMLR, 2019.
- Ghorbani, M., Kazi, A., Baghshah, M. S., et al. (2022). Ra-gcn: Graph convolutional network for disease prediction problems with imbalanced data. *Medical Image Analysis* [J], 75.
- GRATTAROLA D, LIVI L, ALIPPI C, et al. 2021. Seizure localisation with attention-based graph neural networks. bioRxiv [J]: 2020.2012. 2003.409979.
- Hassanzadeh, Reihaneh, & Calhoun, Vince D. (2020). Individualized Prediction of Brain Network Interactions using Deep Siamese Networks. 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE). IEEE.
- He, A., Luo, C., Tian, X., et al. (2018). A twofold Siamese network for real-time object tracking [C] // . City., 4834–4843.
- Hoffer, Elad, & Ailon, Nir. (2015). Deep metric learning using triplet network. Similarity-Based Pattern Recognition: Third International Workshop, SIMBAD 2015, Copenhagen, Denmark, October 12-14, 2015. Proceedings 3. Springer International Publishing.
- Ihalag, A., & Hao, Y. (2022). Formula Graph Self-Attention Network for Representation-Domain Independent Materials Discovery. *Advanced Science*, 9(18), 2200164.
- Ismail, R., Parbo, P., Madsen, L. S., et al. (2020). The relationships between neuroinflammation, beta-amyloid and Tau deposition in Alzheimer's disease: A longitudinal PET study. *Journal of neuroinflammation* [J], 17.
- Jiang, H., Cao, P., Xu, M., et al. (2020). Hi-GCN: A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction. *Computers in Biology and Medicine* [J], 127.
- Jung, J., Cloutman, L. L., Binney, R. J., et al. (2017). The structural connectivity of higher order association cortices reflects human functional brain networks. *cortex* [J], 97, 221–239.
- Kam, T. E., Zhang, H., Jiao, Z., et al. (2019). Deep learning of static and dynamic brain functional networks for early MCI detection[J]. *IEEE Transactions on Medical Imaging*, 39(2), 478–487.
- Kantarci, K., Boeve, B. F., Przybelski, S. A., et al. (2021). FDG PET metabolic signatures distinguishing prodromal DLB and prodromal AD. *NeuroImage: Clinical* [J], 31, 102754.
- Ktena, S. I., Parisot, S., Ferrante, E., et al. (2018). Metric learning with spectral graph convolutions on brain connectivity networks. *NeuroImage* [J], 169.
- LaFerla, Frank M., & Oddo, Salvatore (2005). Alzheimer's disease: β , tau and synaptic dysfunction. *Trends in Molecular Medicine*, 11(4), 170–176.
- Lee, Junhyun, Lee, Inyeop, & Kang, Jaewoo. (2019). Self-attention graph pooling. International conference on machine learning. PMLR.
- LI H, HUANG J, ZHOU M, et al. 2022a. Self-attention pooling-based long-term temporal network for action recognition. IEEE Transactions on Cognitive and Developmental Systems %@ 2379-8920 [J].
- LI W, ZHANG H, LI M, et al. 2022b. MGEGFP: a multi-view graph embedding method for gene function prediction based on adaptive estimation with GCN. Briefings in Bioinformatics [J], 23.
- Li, Y., Liu, J., Tang, Z., et al. (2020). Deep spatial-temporal feature fusion from adaptive dynamic functional connectivity for MCI identification. *IEEE transactions on medical imaging* [J], 39, 2818–2830.
- Li, Yang, & Mateos, Gonzalo. (2022). Learning Similarity-Preserving Representations of Brain Structure-Function Coupling. 2022 30th European Signal Processing Conference (EUSIPCO). IEEE.
- Li, Y., Yao, Z., Yu, Y., et al. (2019). Brain network alterations in individuals with and without mild cognitive impairment: Parallel independent component analysis of AV1451 and AV45 positron emission tomography. *BMC psychiatry* [J], 19.
- LIAN C, LIU M, ZHANG J, et al. 2018. Hierarchical fully convolutional network for joint atrophy localization and Alzheimer's disease diagnosis using structural MRI. IEEE transactions on pattern analysis and machine intelligence [J], 42: 880-893 %@ 0162-8828.
- LIU J, PAN Y, WU F-X, et al. 2020. Enhancing the feature representation of multi-modal MRI data by combining multi-view information for MCI classification. *Neurocomputing* [J], 400: 322-332.
- Ma, G., Ahmed, N. K., Willke, T. L., et al. (2019). Deep graph similarity learning for brain data analysis [C] // . City., 2743–2751.
- Massa, F., Chincarini, A., Bauckneht, M., et al. (2022). Added value of semiquantitative analysis of brain FDG-PET for the differentiation between MCI-Lewy bodies and MCI due to Alzheimer's disease. *European Journal of Nuclear Medicine and Molecular Imaging* [J], 49.
- Parisot, S., Ktena, S. I., Ferrante, E., et al. (2018). Disease prediction using graph convolutional networks: Application to autism spectrum disorder and Alzheimer's disease. *Medical Image Analysis* [J], 48.
- Rhee, Sungmin, Seo, Seokjun, & Kim, Sun. (2017). Hybrid approach of relation network and localized graph convolutional filtering for breast cancer subtype classification. arXiv preprint arXiv:1711.05859.
- Roy, S. K., Harandi, M., Nock, R., et al. (2019). Siamese networks: The tale of two manifolds [C] // . City., 3046–3055.
- Ryali, S., Chen, T., Supekar, K., et al. (2012). Estimation of functional connectivity in fMRI data using stability selection-based sparse partial correlation with elastic net penalty. *NeuroImage* [J], 59.
- Santangelo, R., Masserini, F., Agosta, F., et al. (2020). CSF p-Tau/A β 42 ratio and brain FDG-PET may reliably detect MCI "imminent" converters to AD. *European Journal of Nuclear Medicine and Molecular Imaging* [J], 47.
- Shao, L., Fu, C., You, Y., et al. (2021). Classification of ASD based on fMRI data with deep learning. *Cognitive Neurodynamics* [J], 15.
- Song, L., Gong, D., Li, Z., et al. (2019). Occlusion robust face recognition based on mask learning with pairwise differential Siamese network [C] // . City., 773–782.
- Song, X., Zhou, F., Frangi, A. F., et al. (2022). Multi-center and multi-channel pooling GCN for early AD diagnosis based on dual-modality fused brain network. *IEEE Transactions on Medical Imaging* %@.
- Sporns, O., Chialvo, D. R., Kaiser, M., et al. (2004). Organization, development and function of complex brain networks. *Trends in cognitive sciences* [J], 8.
- Stam, C. J., Jones, B. F., Nolte, G., et al. (2007). Small-world networks and functional connectivity in Alzheimer's disease. *Cerebral cortex* [J], 17.
- Wang, Q., Teng, Z., Xing, J., et al. (2018). Learning attentions: Residual attentional Siamese network for high performance online visual tracking [C] // . City., 4854–4863.
- Wee, C. Y., Liu, C., Lee, A., et al. (2019). Cortical graph neural network for AD and MCI diagnosis and transfer learning across populations[J]. *NeuroImage: Clinical*, 23, Article 101929.
- Yang, Y., Gong, Y., Land, L. P. W., et al. (2020). Understanding the effects of physical experience and information integration on consumer use of online to offline commerce. *International Journal of Information Management* [J], 51, Article 102046.
- Yao, D., Sui, J., Wang, M., et al. (2021). A mutual multi-scale triplet graph convolutional network for classification of brain disorders using functional or structural connectivity. *IEEE transactions on medical imaging* [J], 40.
- YIN W, LI L, WU F-X 2022. Deep learning for brain disorder diagnosis based on fMRI images. *Neurocomputing* [J], 469: 332-345 %@ 0925-2312.
- Yue, X., Wang, Z., Huang, J., et al. (2020). Graph embedding on biomedical networks: Methods, applications and evaluations. *Bioinformatics* [J], 36.
- ZENG H, ZHOU H, SRIVASTAVA A, et al. 2019. Accurate, efficient and scalable graph embedding [C] // , IEEE; City. 462-471 %@ 172811246X.
- ZHANG K, MIZUMA H, ZHANG X, et al. 2021a. PET imaging of neural activity, β -amyloid, and Tau in normal brain aging. *European Journal of Nuclear Medicine and Molecular Imaging* [J], 48: 3859-3871 %@ 1619-7089.
- Zhang, M., Cui, Z., Neumann, M., et al. (2018). An end-to-end deep learning architecture for graph classification [C] // . City.
- ZHANG X, YAO L, WANG X, et al. 2021b. A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers. *Journal of neural engineering* [J], 18: 031002 %@ 031741-032552.
- Zhao, H., Liu, J., Shen, Z., et al. (2022). SCC-MPGCN: Self-attention coherence clustering based on multi-pooling graph convolutional network for EEG emotion recognition. *Journal of neural engineering* [J], 19.
- Zheng, W., Yao, Z., Xie, Y., et al. (2018). Identification of Alzheimer's disease and mild cognitive impairment using networks constructed based on multiple morphological brain features. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* [J], 3.
- Zhou, Y., Qiao, L., Li, W., et al. (2018). Simultaneous estimation of low-and high-order functional connectivity for identifying mild cognitive impairment. *Frontiers in neuroinformatics* [J], 12.
- Zhu, Q., Li, H., Huang, J., et al. (2019). Hybrid functional brain network with first-order and second-order information for computer-aided diagnosis of schizophrenia. *Frontiers in neuroscience* [J], 13.