

METHODOLOGY

Open Access



# Identification of early mild cognitive impairment using multi-modal data and graph convolutional networks

Jin Liu<sup>1</sup>, Guanxin Tan<sup>1</sup>, Wei Lan<sup>2</sup> and Jianxin Wang<sup>1\*</sup>

From 15th International Symposium on Bioinformatics Research and Applications (ISBRA'19) Barcelona, Spain. 3-6 June 2019

\*Correspondence:

[jxwang@mail.csu.edu.cn](mailto:jxwang@mail.csu.edu.cn)

<sup>1</sup>Hunan Provincial Key Lab on Bioinformatics, School of Computer Science and Engineering, Central South University, 932 Lushan South Road, 410083, Changsha, China  
Full list of author information is available at the end of the article

## Abstract

**Background:** The identification of early mild cognitive impairment (EMCI), which is an early stage of Alzheimer's disease (AD) and is associated with brain structural and functional changes, is still a challenging task. Recent studies show great promises for improving the performance of EMCI identification by combining multiple structural and functional features, such as grey matter volume and shortest path length. However, extracting which features and how to combine multiple features to improve the performance of EMCI identification have always been a challenging problem. To address this problem, in this study we propose a new EMCI identification framework using multi-modal data and graph convolutional networks (GCNs). Firstly, we extract grey matter volume and shortest path length of each brain region based on automated anatomical labeling (AAL) atlas as feature representation from T1w MRI and rs-fMRI data of each subject, respectively. Then, in order to obtain features that are more helpful in identifying EMCI, a common multi-task feature selection method is applied. Afterwards, we construct a non-fully labelled subject graph using imaging and non-imaging phenotypic measures of each subject. Finally, a GCN model is adopted to perform the EMCI identification task.

**Results:** Our proposed EMCI identification method is evaluated on 210 subjects, including 105 subjects with EMCI and 105 normal controls (NCs), with both T1w MRI and rs-fMRI data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database. Experimental results show that our proposed framework achieves an accuracy of 84.1% and an area under the receiver operating characteristic (ROC) curve (AUC) of 0.856 for EMCI/NC classification. In addition, by comparison, the accuracy and AUC values of our proposed framework are better than those of some existing methods in EMCI identification.

**Conclusion:** Our proposed EMCI identification framework is effective and promising for automatic diagnosis of EMCI in clinical practice.

**Keywords:** Early mild cognitive impairment, Multi-modal MRI data, Graph convolutional networks, Identification



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

## Background

Alzheimer's disease (AD) is a common neurodegenerative disease accompanied by memory, cognitive and motor disorders. As of 2006, there are 26.6 million AD patients worldwide, and by 2050, one in every 85 people worldwide is expected to have AD [1]. As the world is developing into an aging society, the negative impact of AD on families and society will become more and more significant. Mild cognitive impairment (MCI) is an intermediate process in the conversion of normal people to AD, with up to 15% of people with MCI being converted to AD each year [2]. At present, there is no accurate diagnosis and effective treatment for AD. Most researchers hope that patients can be diagnosed in time when they are in the stage of MCI, and then take effective measures to prevent further deterioration of the disease. Therefore, accurate identification of early MCI (EMCI) is critical to human health.

Since magnetic resonance imaging (MRI) can noninvasively measure brain structural and functional changes related to brain disorder development *in vivo*, in recent years it has been widely used in the study of brain disorders [3], such as AD/MCI [4, 5], schizophrenia [6, 7] and autism [8]. Therefore, MRI can provide phenotypes that can be used to diagnose such disorders. MRI falls into two broad categories: structural MRI (such as T1 MRI, and T2 MRI) and functional MRI (such as rs-fMRI and ts-fMRI). Brain structure is typically measured using structural MRI, which can provide relatively high-definition brain structure in grey matter and white matter. There are many metrics to measure brain structure, and most of them have been widely applied in the study of MCI identification, such as grey matter volume, cortical thickness, texture properties and so on [9–12]. Brain function is typically measured using functional MRI, which can provide changes in hemodynamics caused by neuronal activity. Functional connectivity between brain regions is a common measure of brain function. Also, brain networks based on brain regions and functional connectivity between brain regions have been widely used for feature representation in the study of various brain disorders. In the past years, brain function analysis based on graph theory has shown a powerful role in exploring functional impairment of brain disorders, and has been widely used for MCI identification [13–16].

In the past decade, whether structural MRI-based brain structure metrics or functional MRI-based brain function metrics, these metrics were mainly used separately in the studies with MCI. For example, Karas et al. [9] found that the MCI subjects showed a decrease in grey matter volume in the medial temporal lobe. Wang et al. [15] constructed functional brain networks of MCI subjects and found that the length of the shortest path increased in MCI subjects compared with NCs; Zhang et al. [17] first extracted functional connectivity between brain regions from functional MRI data of each subject as feature representation, and then trained a L2-regularized logistic regression classifier based on these functional connectivity features to perform MCI identification. Therefore, many researchers believe that different metrics may contain different-yet-complementary information, and combinations of these metrics may improve MCI classification performance over separate metrics. In fact, recent studies have also been show great promises for improving the accuracy of MCI identification by combining multiple structural and functional metrics, such as grey matter volume (GMV) and shortest path length (SPL). For example, Wee et al. [18] first used both structural MRI and functional MRI data of each subject to construct multiple brain networks for each subject, and then extracted local clustering coefficient from each brain network of each subject as feature representation to perform

the MCI identification task by using a multi-kernel learning algorithm; De Marco et al. [19] used multiple machine learning models based on different metrics from both structural MRI and functional MRI data to investigate the performance of MCI identification; Tripathi et al. [20] proposed an unsupervised framework for the classification of EMCI and LMCI by combining shape and voxel-based features from 12 brain regions; Jie et al. [21] proposed a feature combination framework to combine both temporal and spatial features of dynamic functional networks to perform automatic identification of EMCI and LMCI. So far, although some results have been achieved for the identification of MCI subjects based on structural and functional MRI data, extracting which features and how to combine multiple features to improve MCI identification accuracy have always been a difficult problem.

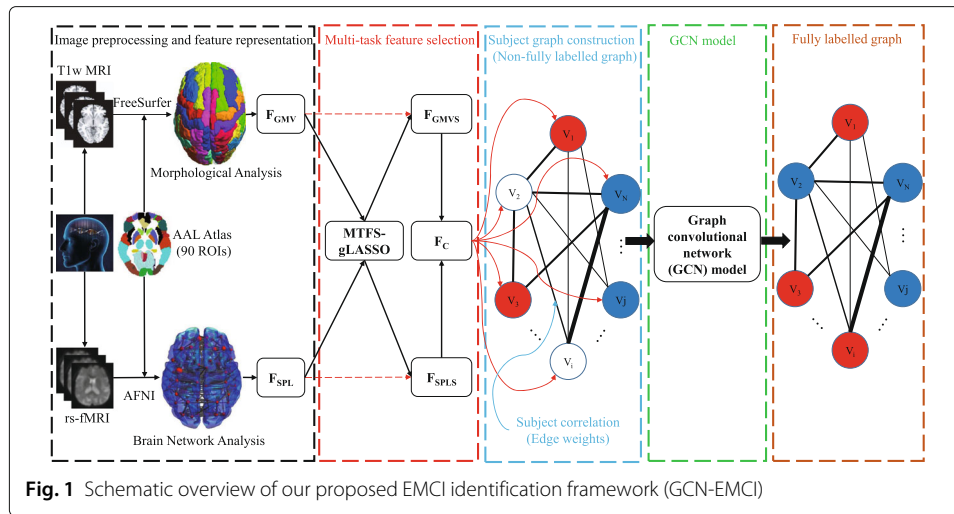
Recently, deep learning models have been widely applied in the fields of medical health [22–25]. Since spectral graph-based convolutional neural network (GCN) models [26, 27] can process irregular graph structures using computational harmonic analysis, many researchers in medical health adopt spectral GCN models to perform various applications, especially at a subject level [28–32]. For example, Anirudh et al. [28] proposed a bootstrapping strategy-based spectral GCN model to perform autism spectrum disorder classification using rs-fMRI data. Guo et al. [29] proposed a spectral GCN model that integrates brain connectivity information to predict visual tasks using MEG data. Ktena et al. [30] proposed a siamese GCN model to learn a graph similarity metric to perform autism and sex classification using rs-fMRI data. In the field of MCI identification, the spectral GCN model is also applied. For example, Parisot et al. [31] proposed a spectral GCN model by combining imaging and non-imaging information to distinguish EMCI from late MCI. However, Parisot et al. [31] only extract GMV of each brain region as imaging features from T1w MRI data, and the Mini Mental State Examination (MMSE), a common scale for AD in clinical practice, has not been taken into account.

Taking the above-mentioned into consideration, in this study we propose a new EMCI identification framework using multi-modal data and graph convolutional networks, which is denoted as GCN-EMCI and shown in Fig. 1. Firstly, we extract GMV and SPL of each brain region based on automated anatomical labeling (AAL) atlas [33] as feature representation from T1w MRI and rs-fMRI data of each subject, respectively. Then, in order to obtain features that are more helpful in identifying EMCI, a common multi-task feature selection method is applied. Afterwards, we construct a non-fully labelled subject graph using imaging and non-imaging phenotypic measures of each subject. Finally, a recent GCN model is adopted to perform the EMCI identification task. The GCN-EMCI is evaluated on 210 subjects (including 105 subjects with EMCI and 105 NCs) with T1w MRI and rs-fMRI data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (<http://adni.loni.usc.edu/>).

## Materials and methods

### Data

A subset of the Alzheimer's Disease Neuroimaging Initiative (ADNI) [34] is used to evaluate our proposed EMCI identification method. This subset includes 210 subjects with both T1w MRI and rs-fMRI data, which are composed of 105 normal controls (NCs) and 105 subjects with early mild cognitive impairment (EMCI). All T1w MRI and rs-fMRI



data are acquired on 3.0 Tesla Philips medical system scanners at multiple sites, and the slice thickness of T1w MRI data and rs-fMRI data is 1.2 mm and 3.0 mm, respectively. Furthermore, rs-fMRI data of each subject contain 140 volumes. Table 1 presents a brief demographic information of the subjects involved in this study. In Table 1, the front and back of  $\pm$  represent mean and standard deviation, respectively. For more details with these subjects from ADNI, please see <http://adni.loni.usc.edu/>.

Prior to EMCI identification using the subjects, Chi-square test [35] is implemented to test the differences in gender, and *t*-test [35] is implemented to test the differences in age and MMSE. As can be seen from Table 1, no significant differences ( $p > 0.05$ ) are found between each of the two groups in gender, age and MMSE.

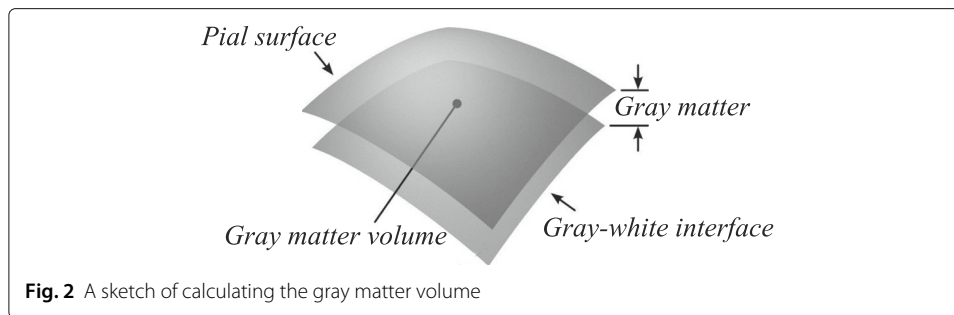
### Image preprocessing and feature representation

As can be seen from Fig. 1, the procedures of image preprocessing and feature representation in our work mainly include two aspects: T1w MRI data preprocessing and feature representation, and rs-fMRI data preprocessing and feature representation. These two aspects are briefly introduced as follows.

Firstly, a standard preprocessing procedure is applied to T1w MRI data of each subject using a standard FreeSurfer pipeline (<https://surfer.nmr.mgh.harvard.edu>) [36], including motion correction, non-uniform intensity normalization, talairach transform computation, skull removal, volumetric segmentation, cortical surface reconstruction and so on. After this standard preprocessing procedure, we can obtain the gray matter (GM) map, which lies between the gray-white interface and the pial surface, and has been widely used to investigate AD/MCI in the literatures [37, 38]. Since EMCI is accompanied by brain atrophy, we suspect that the gray matter volume (GMV) is also accompanied by a decrease. The gray matter volume is defined as the amount of gray matter, and uses the surface-based volume calculation as shown in Fig. 2. For this reason, in this study we extract GMV based on each GM region of the automated anatomical labeling (AAL) atlas

**Table 1** Demographic information of the subjects involved in this study

Demographic information	NC	EMCI	<i>p</i> -value
Number (male/female)	105 (54/51)	105 (49/56)	>0.05
Age (year)	77.1 $\pm$ 6.3	76.3 $\pm$ 5.4	>0.05
MMSE	29.1 $\pm$ 1.1	27.5 $\pm$ 1.8	>0.05



[33] as structural feature representation from T1w MRI data for each subject. For more details of the 90 GM regions of the AAL atlas, please see <http://www.gin.cnrs.fr/en/tools/aal-aal2/>. Finally, we can obtain GWV of each GM region from T1w MRI data for each subject, which is denoted as  $F_{GMV}$ . It is worth mentioning that  $F_{GMV}$  is 90-dimensional vectors.

Secondly, a standard preprocessing procedure is also applied to rs-fMRI data of each subject using the pipeline provided by the Analysis of Functional NeuroImages (AFNI) software (<https://afni.nimh.nih.gov/>) [39], including removing the first 10 rs-fMRI volumes, slice timing, head motion corrections, spatial smoothing, band-pass filtering (0.01-0.1 Hz), nuisance signal regression, and Montreal Neurological Institute (MNI) space normalization and so on. After this standard preprocessing procedure, we can obtain the average rs-fMRI time series of each brain region according to the AAL atlas from the resulted rs-fMRI data of each subject. At present, brain network analysis based on graph theory [40, 41] plays an important role in the study of complex brain diseases, and is widely used in various brain diseases [42], such as AD/MCI, Schizophrenia, Parkinson and so on. The shortest path length (SPL) plays an important role in the information transmission of a brain network, and it is a very important metric to describe the internal structure of the brain network. Meanwhile, the SPL is a measure of functional integration, which can measure the ability to rapidly combine pieces of specialized information. The shorter SPL can transmit the information more quickly and reduce brain consumption. Since EMCI is accompanied by brain atrophy, we suspect that the brain information transmission of EMCI is also accompanied by damage. For this reason, in this study we construct an individual brain network for each subject, which consists of 90 brain regions according to the AAL atlas and functional connectivity between each two brain regions. The functional connectivity is calculated by the pairwise Pearson correlation coefficient between the average rs-fMRI time series of each two brain regions. Then, we compute nodal SPL based on individual brain network of each subject as functional feature representation, which is denoted as  $F_{SPL}$ . It is worth mentioning that  $F_{SPL}$  is also 90-dimensional vectors.

With the above analysis, we can obtain two regional feature sets from both T1w MRI and rs-fMRI data for each subject, i.e.,  $F_{GMV}$  and  $F_{SPL}$ . These two regional feature sets are taken as the original feature representation of each subject.

### Multi-task feature selection

Since the two original feature sets:  $F_{GMV}$  and  $F_{SPL}$  are calculated based on a certain rule, these two feature sets may contain irrelevant or redundant features. Therefore, feature

selection is required for these two original feature sets. Before performing feature selection on each feature set, each original feature set were first scaled individually to range  $[-1, +1]$ . Then, every scaled feature was normalized across all training subjects to obtain its standard score (z-value). These steps ensure that each feature set is within the same scale, minimizing possible bias that may occur when performing selection on features with different dynamic ranges.

In this study, to consider the relationship between different feature selection tasks, the different feature selection tasks should be learned jointly, which is often called multi-task feature selection (MTFS). At present, the group LASSO-based MTFS method (denoted as MTFS-gLASSO) [43] is a common feature selection method, and has been widely used in various feature selection tasks. The MTFS-gLASSO method can be formulated as follows,

$$\min_{\mathbf{W}} \left( \sum_{t=1}^T \|\mathbf{y} - \mathbf{X}_t \mathbf{w}_t\|_2^2 + \lambda \|\mathbf{W}\|_{2,1} \right) \tag{1}$$

where  $\mathbf{X}_t = [\mathbf{x}_t^1, \mathbf{x}_t^2, \dots, \mathbf{x}_t^i, \dots, \mathbf{x}_t^N] \in \mathbb{R}^{N \times P}$  denotes all training subjects in the  $t$ -th task,  $P$  denotes the number of features of each training subject,  $\mathbf{y} = [y^1, y^2, \dots, y^i, \dots, y^N] \in \mathbb{R}^N$  denotes the labels of all training subjects,  $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_i, \dots, \mathbf{w}_T] \in \mathbb{R}^{P \times T}$  is a discriminant matrix,  $\|\mathbf{W}\|_{2,1}$  denotes the  $l_{2,1}$ -norm of  $\mathbf{W}$ , and  $\lambda > 0$  is a parameter to balance the loss function (i.e.,  $\sum_{t=1}^T \|\mathbf{y} - \mathbf{X}_t \mathbf{w}_t\|_2^2$ ) and the regularization term (i.e.,  $\|\mathbf{W}\|_{2,1}$ ). The larger the  $\lambda$  value, the greater the penalty for the parameters in the model, resulting in higher model sparsity, that is, more parameters are trained to zero.

Finally, as shown in Fig. 1, we concatenate the two selected feature sets (i.e.,  $F_{GMVS}$  and  $F_{SPLS}$ ) obtained by MTFS-gLASSO, which is denoted as  $F_C$ .

**Subject graph construction**

Before performing EMCI identification using GCN model, we should first construct a graph using all subjects. A graph is typically defined as  $G = (V, E, C)$ , where  $V$  is the set of vertices (or nodes),  $E$  is the set of edges, and  $C$  is the adjacency matrix describing the graph's connectivity. Therefore, to construct the subject graph, we need to determine the definition of the nodes and edges in this graph.

In this study, we define each subject as a node, and the correlation between each two subjects as edges. For a node  $V_i$ , we use  $F_C(V_i)$  to represent it. To compute the adjacency matrix (i.e.,  $C$ ) of the subject graph, we follow the work of Parisot et al. [31]. Considering a set of  $D$  non-imaging phenotypic measures  $H = \{H_d\}$  (such as gender, age or MMSE), the adjacent matrix is defined as follows:

$$C(i, j) = Corr(V_i, V_j) \sum_{d=1}^D \eta(H_d(V_i), H_d(V_j)) \tag{2}$$

$$Corr(V_i, V_j) = \exp\left(-\frac{[\ell(F_C(V_i), F_C(V_j))]^2}{2\sigma^2}\right) \tag{3}$$

$$\eta(H_d(V_i), H_d(V_j)) = \begin{cases} 1 & \text{if } |H_d(V_i) - H_d(V_j)| < \varepsilon \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where  $Corr(V_i, V_j)$  is an imaging phenotypic measure of correlation between subjects,  $\rho(\cdot, \cdot)$  is a correlation distance,  $\sigma > 0$  is a constant parameter, and  $\eta(H_d(V_i), H_d(V_j))$  a non-imaging phenotypic measure of distance between subjects. It is worth mentioning that different non-imaging phenotypic measures correspond to different values of  $\varepsilon$ .

### Classification using GCN model

Following the work of Parisot et al. [31], schematic illustration of the GCN model in this study is shown in Fig. 3. The input layer of the GCN model is a non-fully labelled subject graph, and the output layer of the GCN model is a fully labelled subject graph. Obviously, the training set consists of labelled nodes (such as  $V_1$  and  $V_j$ ) in the non-fully labelled subject graph, and the testing set consists of unlabelled nodes (such as  $V_2$  and  $V_i$ ) in the non-fully labelled subject graph.

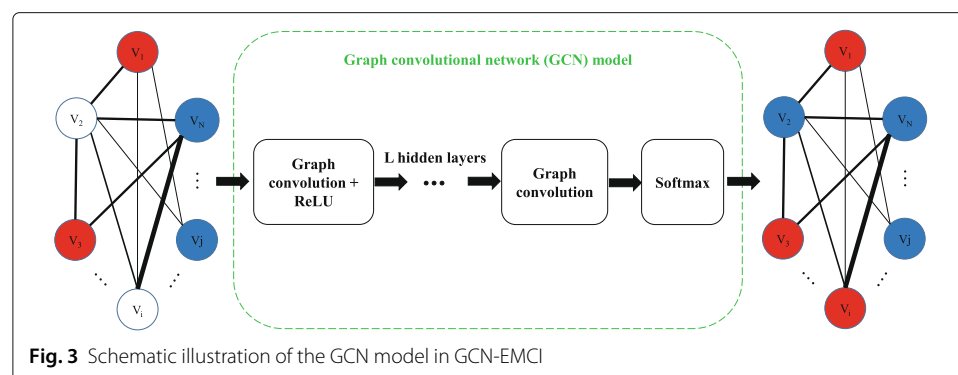
As can be seen from Fig. 3, the GCN model is a semi-supervised classification method and a relatively simple model, which consists of L hidden layers with graph convolution and a softmax layer. The L hidden layers are activated by a rectified linear unit (ReLU) function. After training the GCN model, the softmax function is used in the testing set to assign labels to nodes that have no labels.

## Experiments and results

### Experimental settings

Our proposed EMCI identification framework (GCN-EMCI) is evaluated on 105 subjects with EMCI and 105 NCs via a 5-fold cross-validation strategy. In the multi-task feature selection step, the parameter  $\lambda$  is set to [0, 100] with a step size of 5, and these features with non-zero elements in  $W$  are selected. In the subject graph construction step, the corresponding  $\varepsilon$  values for the non-imaging phenotypic measures: gender, age and MMSE are set to 1, 2 and 2, respectively. In the classification using GCN model step, GCN parameters are similar with [31], we only change: L = 5, dropout rate: 0.01, learning rate: 0.02, epochs: 500, K = 4.

To quantitatively evaluate the classification performance of GCN-EMCI, in this study the three metrics: accuracy (ACC), sensitivity (SEN) and specificity (SPE) are computed. In addition, to quantitatively evaluate the overall performance of GCN-EMCI, the area under receiver operating characteristic (ROC) curve (AUC) value [44–46] is also reported. It is worth mentioning that the greater the values of the four metrics are, the better the classification performance of the method is.



**Table 2** Classification performance of GCN-EMCI based on different subject graphs

Features	ACC(%)	SEN(%)	SPE(%)	AUC
$F_{GMV}$	65.8	69.8	62.7	0.672
$F_{SPL}$	62.7	66.5	59.4	0.637
$F_{GMV} + F_{SPL}$	69.7	71.4	65.6	0.719
$F_{GMVS}$	79.8	83.4	77.1	0.802
$F_{SPLS}$	75.3	78.3	73.2	0.765
$F_C$	81.5	82.7	80.2	0.828
$F_C + H$	84.1	86.5	81.3	0.856

To avoid the bias caused by randomly dividing the subjects in the cross-validation, the 5-fold cross-validation is repeated 50 times in our experiments. In this study we only report the average of 50 repeated experiments.

### Classification performance

In order to verify the effectiveness of GCN-EMCI, we have done a series of experiments based on different subject graphs. First, we only use imaging phenotypic features (i.e.,  $F_{GMV}$ ,  $F_{SPL}$ ,  $F_{GMV} + F_{SPL}$ ,  $F_{GMVS}$ ,  $F_{SPLS}$  and  $F_C$ ) to construct subject graphs, and then we combine imaging and non-imaging phenotypic features (i.e.,  $F_C + H$ ) to construct subject graph. The results of these experiments are shown in Table 2.

As can be seen from Table 2, the classification performance of GCN-EMCI based on subject graphs with original features (i.e.,  $F_{GMV}$ ,  $F_{SPL}$  and  $F_{GMV} + F_{SPL}$ ) are relatively low (such as ACC < 70%), and the classification performance of GCN-EMCI based on subject graphs with selected features (i.e.,  $F_{GMVS}$ ,  $F_{SPLS}$  and  $F_C$ ) are relatively good (such as ACC  $\approx$  80%). However, the classification performance of GCN-EMCI based on subject graph with both imaging features and non-imaging features (i.e.,  $F_C + H$ ) is the best (ACC = 84.1%, SEN = 86.5%, SPE = 81.3%, AUC = 0.856). Experimental results show that GCN-EMCI is effective for EMCI identification.

## Discussion

### Different feature selection methods

To demonstrate the superiority of MTFS-gLASSO in GCN-EMCI, we compare two other common feature selection methods, i.e.,  $t$ -test [35] and LASSO [47]. These two methods are implemented by scipy [48] and scikit-learn [49] packages in Python, respectively. The specific approach of these two methods is to first use the  $t$ -test or LASSO method for  $F_{GMV}$  and  $F_{SPL}$  respectively, and then concatenate  $F_{GMVS}$  and  $F_{SPLS}$  as the feature representation of each subject. It's worth mentioned that when  $t$ -test is used as feature selection method, these features with  $p$ -value less than 0.1 are selected; when LASSO is used as feature selection method, these features whose weights are not equal to 0 are selected. The EMCI/NC classification performance based on GCN-EMCI with different feature selection methods is shown in Table 3.

**Table 3** Comparison with different feature selection methods for EMCI/NC classification

Methods	ACC(%)	SEN(%)	SPE(%)	AUC
$t$ -test	70.9	74.7	68.2	0.728
LASSO	78.5	83.6	76.6	0.798
MTFS-gLASSO	84.1	86.5	81.3	0.856



**Table 4** Comparison with existing methods for EMCI/NC classification

Methods	ACC(%)	SEN(%)	SPE(%)	AUC	<i>p</i> -value
Tripathi et al., 2017 [20]	75.8	74.2	76.7	0.762	<0.01
Jie et al., 2018 [21]	79.5	82.6	77.2	0.801	<0.01
GCN-EMCI	84.1	86.5	81.3	0.856	

As can be seen from Table 3, the EMCI/NC classification performance based on GCN-EMCI with MTFS-gLASSO is best in ACC, SEN, SPE and AUC. This result indicates that MTFS-gLASSO can obtain better feature representation than the other two feature selection methods.

#### Comparison with existing methods

To demonstrate the superiority of GCN-EMCI, we also compare two existing methods [20, 21] in EMCI identification. In our comparative experiments, the existing methods are also repeated 50 times via a 5-fold cross-validation strategy, and the average classification performance is reported in Table 4. In order to statistically verify that the classification performance of GCN-EMCI is better than that of the other two existing methods, we also report the *p*-values of GCN-EMCI and other methods in terms of ACC, which is shown in Table 4.

As can be seen from Table 4, GCN-EMCI obtains the best performance in ACC, SEN, SPE and AUC for EMCI/NC classification. Compared with the two existing methods, GCN-EMCI utilizes the correlation between each two subjects, and uses a GCN model to learn the deep differences between EMCI and NC. In addition, as the statistical *p*-value is less than 0.01, GCN-EMCI is significantly better than the other two existing methods. These results indicate that our proposed method (i.e., GCN-EMCI) is not only effective, but also has a good advantage in EMCI identification.

#### Conclusion

In this study, we propose a new EMCI identification method using multi-modal data and graph convolutional networks. Firstly, we perform image preprocessing and feature representation for both T1w MRI and rs-fMRI data of each subject. Then, in order to obtain features that are more helpful in identifying EMCI, a common multi-task feature selection method is adopted. Afterwards, we construct a subject graph using imaging phenotypic measures and non-imaging phenotypic measures of each subject. Finally, a GCN model is applied to perform the EMCI identification task. Experimental results on 210 subjects from ADNI database demonstrate that our proposed framework is effective for EMCI identification. This method paves the way to discriminative imaging markers for computer-aided identification of EMCI.

#### Acknowledgement

A partial version of this work was presented at the 15th International Symposium on Bioinformatics Research and Applications (ISBRA) in Barcelona, Spain (June 3-6, 2019). We would like to thank the reviewers for their detailed suggestions which greatly improved the quality and readability of this work.

#### Abbreviations

AAL: automated anatomical labeling; ACC: accuracy; AD: Alzheimer's disease; ADNI: Alzheimer's Disease Neuroimaging Initiative; AFNI: Analysis of Functional NeuroImages; AUC: area under ROC curve; EMCI: early mild cognitive impairment; GCN: graph convolutional network; GMV: grey matter volume; LASSO: least absolute shrinkage and selection operator; LMCI: late mild cognitive impairment; MCI: mild cognitive impairment; MEG: magnetoencephalography; MMSE: Mini Mental State Examination; MNI: Montreal Neurological Institute; MRI: Magnetic Resonance Imaging; MTFS-gLASSO:

LASSO-based MTFs; MTFs: multi-task feature selection; NC: normal control; ReLU: rectified linear unit; ROC: the receiver operating characteristic; rs-fMRI: resting state-functional MRI; SEN: sensitivity; SPE: specificity; SPL: shortest path length; ts-fMRI: tasking state-functional MRI

#### About this supplement

This article has been published as part of BMC Bioinformatics Volume 21 Supplement 6, 2020: Selected articles from the 15th International Symposium on Bioinformatics Research and Applications (ISBRA-19): bioinformatics. The full contents of the supplement are available online at <https://bmcbioinformatics.biomedcentral.com/articles/supplements/volume-21-supplement-6>

#### Authors' contributions

JW conceived the project; JL and WL designed the experiments; JL and GT performed the experiments; JL and GT wrote the paper. All authors read and approved the final manuscript.

#### Funding

This work is supported in part by the National Natural Science Foundation of China under Grant No.61802442, No.61702122, No.61877059, the Natural Science Foundation of Hunan Province under Grant No.2019JJ50775, the 111 Project (No.B18059), the Hunan Provincial Science and Technology Program (2018WK4001). Specifically, publication costs are funded by the National Natural Science Foundation of China under grant No.61802442 and the Natural Science Foundation of Hunan Province under Grant No.2019JJ50775.

#### Availability of data and materials

All data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) are available to qualified researchers via the ADNI Access Management System (<http://adni.loni.usc.edu/data-samples/access-data/>). The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Author details

<sup>1</sup>Hunan Provincial Key Lab on Bioinformatics, School of Computer Science and Engineering, Central South University, 932 Lushan South Road, 410083, Changsha, China. <sup>2</sup>School of Computer, Electronics and Information, Guangxi University, 100 Daxue East Road, 530004, Nanning, China.

Received: 26 February 2020 Accepted: 2 March 2020 Published: 18 November 2020

#### References

- Association A. 2016 alzheimer's disease facts and figures. *Alzheimers Dement*. 2016;12(4):459–509.
- Kantarci K, Weigand SD, Przybelski SA, Shiung MM, Whitwell JL, Negash S, Knopman DS, Boeve BF, O'Brien PC, Petersen RC. Risk of dementia in mci: combined effect of cerebrovascular disease, volumetric mri, and 1h mrs. *Neurology*. 2009;72(17):1519–25.
- Liu J, Pan Y, Li M, Chen Z, Tang L, Lu C, Wang J. Applications of deep learning to mri images: a survey. *Big Data Min Analytics*. 2018;1(1):1–18.
- Liu J, Li M, Lan W, Wu F-X, Pan Y, Wang J. Classification of alzheimer's disease using whole brain hierarchical network. *IEEE/ACM Trans Comput Biol Bioinforma*. 2018;15(2):624–32.
- Li H-D, Bai T, Sandford E, Burmeister M, Guan Y. Baihui: cross-species brain-specific network built with hundreds of hand-curated datasets. *Bioinformatics*. 2018;35(14):2486–8.
- Liu J, Li M, Pan Y, Wu F-X, Chen X, Wang J. Classification of schizophrenia based on individual hierarchical brain networks constructed from structural mri images. *IEEE Trans NanoBioscience*. 2017;16(7):600–8.
- Xiang Y, Wang J, Tan G, Wu F-X, Liu J. Schizophrenia identification using multi-view graph measures of functional brain networks. *Front Bioeng Biotechnol*. 2020;7:479.
- Kong Y, Gao J, Xu Y, Pan Y, Wang J, Liu J. Classification of autism spectrum disorder by combining brain connectivity and deep neural network classifier. *Neurocomputing*. 2019;324(9):63–68.
- Karas G, Scheltens P, Rombouts S, Visser P, Van Schijndel R, Fox N, Barkhof F. Global and local gray matter loss in mild cognitive impairment and alzheimer's disease. *Neuroimage*. 2004;23(2):708–16.
- Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, Madsen SK, Parikshak N, Hua X, Toga AW, Jack CR, et al. Validation of a fully automated 3d hippocampal segmentation method using subjects with alzheimer's disease mild cognitive impairment, and elderly controls. *Neuroimage*. 2008;43(1):59–68.
- Liu J, Wang X, Zhang X, Pan Y, Wang X, Wang J. Mmm: classification of schizophrenia using multi-modality multi-atlas feature representation and multi-kernel learning. *Multimed Tools Appl*. 2018;77(22):29651–67.
- Liu J, Wang J, Bin H, Wu F-X, Pan Y. Alzheimer's disease classification based on individual hierarchical networks constructed with 3d texture features. *IEEE Trans NanoBioscience*. 2017;16(6):428–37.

13. Chen G, Ward BD, Xie C, Li W, Wu Z, Jones JL, Franczak M, Antuono P, Li S-J. Classification of alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional mr imaging. *Radiology*. 2011;259(1):213–21.
14. Feng Y, Bai L, Ren Y, Chen S, Wang H, Zhang W, Tian J. Fmri connectivity analysis of acupuncture effects on the whole brain network in mild cognitive impairment patients. *Magn Reson Imaging*. 2012;30(5):672–82.
15. Wang J, Zuo X, Dai Z, Xia M, Zhao Z, Zhao X, Jia J, Han Y, He Y. Disrupted functional brain connectome in individuals at risk for alzheimer's disease. *Biol Psychiatry*. 2013;73(5):472–81.
16. Brier MR, Thomas JB, Fagan AM, Jason H, Holtzman DM, Benzinger TL, Morris JC, Ances BM. Functional connectivity and graph theory in preclinical alzheimer's disease. *Neurobiol Aging*. 2014;35(4):757–68.
17. Zhang X, Hu B, Ma X, Xu L. Resting-state whole-brain functional connectivity networks for mci classification using l2-regularized logistic regression. *IEEE Trans NanoBioscience*. 2015;14(2):237–47.
18. Wee C-Y, Yap P-T, Zhang D, Denny K, Browndyke JN, Potter GG, Welsh-Bohmer KA, Wang L, Shen D. Identification of mci individuals using structural and functional connectivity networks. *Neuroimage*. 2012;59(3):2045–56.
19. De Marco M, Beltrachini L, Biancardi A, Frangi AF, Venneri A. Machine-learning support to individual diagnosis of mild cognitive impairment using multimodal mri and cognitive assessments. *Alzheimer Dis Assoc Disord*. 2017;31(4):278–86.
20. Tripathi S, Nozadi SH, Shakeri M, Kadoury S. Sub-cortical shape morphology and voxel-based features for alzheimer's disease classification. In: *IEEE International Symposium on Biomedical Imaging*. IEEE; 2017. <https://doi.org/10.1109/isbi.2017.7950682>.
21. Jie B, Liu M, Shen D. Integration of temporal and spatial properties of dynamic connectivity networks for automatic diagnosis of brain disease. *Med Image Anal*. 2018;47:81–94.
22. Liu L, Wu F-X, Wang J. Efficient multi-kernel dcnn with pixel dropout for stroke mri segmentation. *Neurocomputing*. 2019;350:117–27.
23. Yu Y, Li M, Liu L, Fei Z, Wu F-X, Wang J. Automatic icd code assignment of chinese clinical notes based on multilayer attention birnn. *J Biomed Inform*. 2019;91:103114.
24. An Y, Huang N, Chen X, Wu F, Wang J. High-risk prediction of cardiovascular diseases via attention-based deep neural networks. *IEEE/ACM Trans Comput Biol Bioinforma*. 2019;PP(99):1–1.
25. Liu L, Chen S, Zhu X, Zhao X-M, Wu F-X, Wang J. Deep convolutional neural network for accurate segmentation and quantification of white matter hyperintensities. *Neurocomputing*. 2020;384:231–42.
26. Zhou J, Cui G, Zhang Z, Yang C, Liu Z, Maosong S. Graph neural networks: A review of methods and applications. 2018. arXiv:1812.08434. <https://arxiv.org/abs/1812.08434>.
27. Zhang Z, Cui P, Zhu W. Deep learning on graphs: A survey. 2018. arXiv:1812.04202. <https://arxiv.org/abs/1812.04202>.
28. Anirudh R, Thiagarajan JJ. Bootstrapping graph convolutional neural networks for autism spectrum disorder classification. 2017. arXiv:1704.07487. <https://arxiv.org/abs/1704.07487>.
29. Guo Y, Nejati H, Cheung NM. Deep neural networks on graph signals for brain imaging analysis. 2017. arXiv:1705.04828. <https://arxiv.org/abs/1705.04828>.
30. Ktena SI, Parisot S, Ferrante E, Rajchl M, Lee M, Glocker B, Rueckert D. Metric learning with spectral graph convolutions on brain connectivity networks. *Neuroimage*. 2018;169:431–42.
31. Parisot S, Ktena SI, Ferrante E, Lee M, Guerrero R, Glocker B, Rueckert D. Disease prediction using graph convolutional networks: Application to autism spectrum disorder and alzheimer's disease. *Med Image Anal*. 2018;48:117–30.
32. Yu Y, Li M, Liu L, Li Y, Wang J. Clinical big data and deep learning: Applications, challenges, and future outlooks. *Big Data Min Analytics*. 2019;2(4):288–305.
33. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B, Joliot M. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*. 2002;15(1):273–89.
34. Carrillo MC, Bain LJ, Frisoni GB, Weiner MW. Worldwide alzheimer's disease neuroimaging initiative. *Alzheimer's Dement*. 2012;8(4):337–42.
35. Balding DJ. A tutorial on statistical methods for population association studies. *Nat Rev Genet*. 2006;7(10):781–91.
36. Fischl B. Freesurfer. *Neuroimage*. 2012;62(2):774–81.
37. Min R, Wu G, Cheng J, Wang Q, Shen D, Initiative ADN. Multi-atlas based representations for alzheimer's disease diagnosis. *Hum Brain Mapp*. 2014;35(10):5052–70.
38. Liu J, Wang J, Tang Z, Hu B, Wu F-X, Pan Y. Improving alzheimer's disease classification by combining multiple measures. *IEEE/ACM Trans Comput Biol Bioinforma*. 2018;15(5):1649–59.
39. Cox RW. Afni: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res*. 1996;29(3):162.
40. Rubinov M, Sporns O. Complex network measures of brain connectivity: uses and interpretations. *NeuroImage*. 2010;52(3):1059–69.
41. Bassett DS, Sporns O. Network neuroscience. *Nat Neurosci*. 2017;20(3):353–64.
42. Liu J, Li M, Pan Y, Lan W, Zheng R, Wu F-X, Wang J. Complex brain network analysis and its applications to brain disorders: a survey. *Complexity*. 2017;2017(8362741):.
43. Yuan M, Lin Y. Model selection and estimation in regression with grouped variables. *J R Stat Soc*. 2006;68(1):49–67.
44. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (roc) curve. *Radiology*. 1982;143(1):29–36.
45. Lan W, Wang J, Li M, Liu J, Wu F-X, Pan Y. Predicting microrna-disease associations based on improved microrna and disease similarities. *IEEE/ACM Trans Comput Biol Bioinforma*. 2018;15(6):1774–82.
46. Chen Q, Lai D, Lan W, Wu X, Chen B, Chen Y-PP, Wang J. lldmsf: Inferring associations between long non-coding rna and disease based on multi-similarity fusion. *IEEE/ACM Trans Comput Biol Bioinforma*. 2019;PP(99):1–1.
47. Tibshirani R. Regression shrinkage and selection via the lasso. *J R Stat Soc Ser B Methodol*. 1996;58(1):267–88.

48. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, et al. Scipy 1.0—fundamental algorithms for scientific computing in python. 2019. arXiv preprint arXiv:1907.10121. <https://arxiv.org/abs/1907.10121>.
49. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, et al. Scikit-learn: Machine learning in python. *J Mach Learn Res.* 2011;12(Oct):2825–30.

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

