Graph Convolutional Network with Morphometric Similarity Networks for Schizophrenia Classification

Hye Won Park, Seo Yeong Kim, and Won Hee Lee(✉)

Kyung Hee University, Yongin, Republic of Korea
whlee@khu.ac.kr

Abstract. There is significant interest in using neuroimaging data for schizophrenia classification. Graph convolutional networks (GCNs) provide great potential to improve schizophrenia classification using brain graphs derived from neuroimaging data. However, accurate classification of schizophrenia is still challenging due to the heterogeneity of schizophrenia and their subtle differences in neuroimaging features. This paper presents a new graph convolutional framework for population-based schizophrenia classification that leverages graph-theoretical measures of morphometric similarity networks inferred from structural MRI scans and incorporates variational edges to reinforce the learning process. Specifically, we construct individual morphometric similarity networks based on inter-regional similarity of multiple morphometric features (cortical thickness, surface area, gray matter volume, mean curvature, and Gaussian curvature) extracted from T1-weighted MRI. We then formulate an adaptive population graph where each node is represented by the topological features of individual morphometric similarity networks and each edge models the similarity between the topological features of the subjects and incorporates the phenotypic information. An encode module is devised to estimate the associations between phenotypic data of the subjects and to adaptively optimize the edge weights. Our proposed method is evaluated on a large dataset collected from nine sites, resulting in a total sample of 366 patients with schizophrenia and 590 healthy individuals. Experimental results demonstrate that our proposed method improves the classification performance over traditional machine learning algorithms, with a mean classification accuracy of 81.8%. The most salient regions contributing to classification are primarily identified in the middle temporal gyrus and superior temporal gyrus.

Keywords: Graph Convolutional Networks · Morphometric Similarity Networks · Schizophrenia

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-43907-0_60.

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14220, pp. 626–636, 2023.
https://doi.org/10.1007/978-3-031-43907-0_60

1 Introduction

Schizophrenia is a severe and chronic psychiatric disorder characterized by psychotic episodes, cognitive impairment, and impaired functioning with high rates of disability [4]. Early diagnosis of schizophrenia is beneficial to patients as it leads to better prognosis and reduces symptoms more effectively [15]. However, early detection of schizophrenia is still challenging since the complex and heterogeneous symptoms are associated with schizophrenia, complicating the optimal treatment of patients and limiting positive outcomes. The current diagnosis of schizophrenia is mainly based on patient's clinical symptoms and clinical interviews. Therefore, there is an urgent need to establish an objective approach for an accurate diagnosis of schizophrenia.

There is an increasing interest in developing accurate and robust techniques to classify subjects into groups using neuroimaging data. Previous studies focused on handcrafted feature-based machine learning approach, which requires the process of feature extraction and selection prior to disease classification [28, 30]. Pre-selected features extracted from neuroimaging data might not be sufficient and generalizable across different disease datasets, thus yielding a substantial variation in classification performance.

With the rapid development of deep learning methods, recent studies have shown great potential for integrating neuroimaging data and deep learning models for an automatic diagnosis of brain diseases [12, 23]. In particular, graph convolutional network (GCN) has become a promising approach for medical image classification such as schizophrenia [16], Alzheimer's disease [10], and autism spectrum disorder [27]. GCN is capable of preserving graph topological properties while automatically learning features to perform various graph-related tasks such as graph classification and graph representation learning [2]. This is especially important for psychiatric disorders like schizophrenia which has complex and subtle differences in neuroimaging features compared to healthy individuals. Previous studies mainly focused on functional MRI data as the input of GCN for disease classification [18, 19]. Parisot *et al.* proposed a population GCN model combining subject-specific imaging and non-imaging information for autism spectrum disorder and Alzheimer's disease classification [18]. Qin *et al.* examined the effectiveness of GCN in distinguishing patients with major depressive disorder from healthy controls [19]. However, most studies adopted previously captured information based on functional brain networks which describe statistical dependence between functional MRI time series instead of modeling cortical networks based on similarity in regional cortical morphology estimated from structural MRI of the brain.

Recent studies have highlighted the potential of cortical networks, so-called morphometric similarity networks (MSNs), to predict individual differences in brain morphometry [13, 14], thereby allowing for their potential utility to capture individual cognitive variation [24]. Using the MSNs, it is also possible to unravel changes in the brain during normal cortical development [26] and psychiatric disorders [17] and to identify patterns of abnormal morphometric similarity that classify autism spectrum disorder [31] and Alzheimer's disease and mild cognitive impairment from healthy individuals [32]. However, much less is known about whether the MSNs are clinically useful phenotype and the ability of MSNs for schizophrenia classification.

To fill these gaps, we developed a generalizable graph convolutional framework for population-based schizophrenia classification using the MSNs inferred from structural MRI data. Our contributions are summarized as follows: 1) We propose a new population graph model for integrating MSN-driven features derived from structural MRI and phenotypic information; 2) A novel feature selection strategy is introduced to leverage graph-theoretical measures of the MSNs, which is new and generalizable for graph neural networks; 3) We validate the feasibility of our proposed method by conducting a comprehensive evaluation on a large schizophrenia dataset, which shows superior performance in classification over traditional machine learning approaches; 4) A complete sensitivity analysis for key parameters in our GCN-based classification framework is performed; 5) The most salient regions contributing to classification are primarily identified in the middle temporal gyrus and superior temporal gyrus.

2 Materials and Methods

Figure 1 shows the overview of our proposed classification framework. We construct individual MSNs based on inter-regional similarity of multiple morphometric features extracted from T1-weighted MRI. We formulate a population graph model where each node is represented by graph-theoretical measures of the MSNs and each edge models the similarity between the topological features of the subjects and incorporates the phenotypic information. We apply a threshold to the population graph to remove spurious connections. The edge weights are adaptively updated by using an MLP-based encoder based on non-imaging data of the subjects. Spectral graph convolutions are applied for learning, followed by the MLP-based classification for schizophrenia.

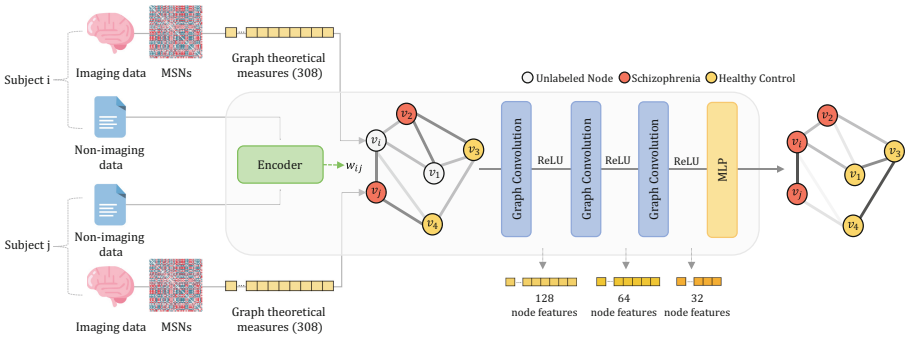

Fig. 1. Overview of our proposed classification framework.

2.1 Datasets

In this study, structural T1-weighted magnetic resonance imaging (MRI) scans were collected from six public databases, including DecNef [25], COBRE [1], CANDI [11],

MCICShare [7], UCLA CNP [3], and CCNMD [20]. A total of 956 subjects comprising 366 patients with schizophrenia and 590 healthy controls were included after quality control (for details, see Supplementary Materials). Ethical approvals and informed consents were obtained locally for each study, covering both participation and subsequent data sharing. Details about sample and demographic information are presented in Supplementary Table S1.

2.2 Morphometric Similarity Networks

The structural MRI scans from all subjects were preprocessed using the recon-all command from FreeSurfer (version 7.2.0) [6]. Cortical parcellation was based on an atlas with 308 cortical regions derived from the Desikan-Killiany atlas [21]. In each subject, this procedure generated 1540 morphometric features (308 regional measures of gray matter volume, surface area, cortical thickness, Gaussian curvature, and mean curvature). Morphometric features for each subject can be expressed as a set of 308 vectors of length 5. Each morphometric feature was standardized across 308 regions so that the data have a mean of zero and a standard deviation of one. A $N \times N$ correlation matrix, representing an individual MSN, was constructed by computing the Pearson's correlation coefficient between the z-normalized morphometric features of region i and region j for all (i, j) of regional properties. This procedure resulted in 308×308 MSN for each subject.

2.3 Graph Construction

We consider a population comprising N subjects; each subject being associated with a set of imaging and phenotypic information. We define the population graph as an undirected weighted graph $G = (V, E, W)$, where V is a set of $|V| = N$ nodes, E is a set of edges. $W \in \mathbb{R}^{N \times N}$ is a weighted adjacency matrix of population graph G. Each node v represents a subject, and the feature vectors for nodes are extracted from imaging data. Each edge is defined as the similarity between nodes. The two main decisions required to build the population model are the definition of the feature vector representing each node and the connectivity of the graph corresponding to its edges and weights.

Feature Vector. The MSN has a high dimensionality with 47,278 features considering only the upper triangle of the MSN. Using the Brain Connectivity Toolbox (BCT) [22], we computed the following graph-theoretical measures: strength, betweenness centrality, and clustering coefficient. These measures were considered as node features for subjects. The strength is the sum of weights of edges connected to the given node. The betweenness centrality is defined as the fraction of all shortest paths in the network that contain a given node. Nodes with high values of betweenness centrality participate in many shortest paths. The clustering coefficient is a measure of network segregation and is defined as the fraction of triangles around the given node.

We investigated two additional feature selection approaches. Firstly, we applied the ridge classifier to perform recursive feature elimination (RFE). RFE iteratively removes the irrelevant features to achieve the desired number of features while maximizing the

ridge classification accuracy. Secondly, we used the k-best selection with ANOVA, a univariate feature selection algorithm that selects the k-best features with the highest F-values.

Graph Edge. The process of computing graph edges involved two steps: the calculation of initial edge weights and adaptive learning through encoders. The initial similarity between nodes were determined using imaging features and categorical information. Considering a set H categorical phenotypic measures $M = \{M_h\}$, such as sex and scan site, the initial edge weight W_I between node v and w are defined as follows:

$$W_I(v, w) = Sim(v, w) \sum_{h=1}^{H} \gamma(M_h(v), M_h(w)) \tag{1}$$

$$Sim(v, w) = exp\left(-\frac{[\rho(x(v), x(w)]^2}{2h^2}\right) \tag{2}$$

$$\gamma(M_h(v), M_h(w)) = \begin{cases} 0, \textit{ if } M_h(v) \neq M_h(w) \\ 1, \textit{ if } M_h(v) = M_h(w) \end{cases} \tag{3}$$

where $Sim(v, w)$ is a measure of similarity between node v and node w in the imaging features. γ is a measure of similarity between node v and node w in the categorical variables. ρ is the Pearson's correlation distance. $x(v)$ and $x(w)$ are the topological feature vectors of node v and w, respectively. h denotes the width of the kernel.

We set a threshold on $W_I(v, w)$ via quantile Q, , and adaptively calculated the edge weights only for the remaining edges. We investigated four different encoders (PAE, EA, L2, and Cosine + Tanh) that were used to determine the edge weights between node v and node w based on phenotypic information (e.g., sex, age, and scan site). The subject's normalized phenotypic inputs were projected into a latent space $\varphi \in \mathbb{R}^{64}$ using MLP. The pairwise association encoder (PAE) [9] scores the association between node v and w as follows: $W(v, w) = 0.5 * (S_C(\varphi_v, \varphi_w) + 1)$, where $S_C(\varphi_v, \varphi_w)$ is the cosine similarity between latent vector φ_v and φ_w. The Edge Adapter (EA) [8] scores the association between node v and w as follows: $W(v, w) = \sigma(\sum_{64} \alpha|\varphi_v - \varphi_w|)$, where $|\varphi_v - \varphi_w|$ is the absolute difference vector of two latent vectors. α is the 64-dimension trainable parameter used in fully connected layer and σ is the sigmoid function. We also designed the L2 encoder as follows: $W(v, w) = \sigma(1/||\varphi_v - \varphi_w||_2^2)$, where $||\varphi_v - \varphi_w||_2^2$ denotes L2 distance of two latent vectors. Finally, we designed the Cosine + Tanh encoder combining Tanh with cosine similarity as follows: $W(v, w) = Tanh(ReLU(S_C(\varphi_v, \varphi_w)))$. The encoders were optimized with graph convolution models using gradient descent algorithm.

2.4 Spectral Graph Convolutions

To extract spatial features from each graph node, we used the spectral graph convolution in the Fourier domain. A spectral convolution of signal x with a filter $g_\theta = diag(\theta)$ defined in the Fourier domain is defined as a multiplication in the Fourier domain: $g_\theta * x = U g_\theta U^T x$, where U is the matrix of the eigenvectors of the normalized Laplacian matrix. The normalized graph Laplacian $\mathcal{L}$ is defined as $\mathcal{L} = I - D^{-1/2} W D^{-1/2}$,

where I and D are the identity matrix and the diagonal degree matrix, respectively. To reduce computational complexity of convolution, we used Chebyshev spectral convolutional network (ChebConv) [5] that approximates spectral graph convolution using Chebyshev polynomials. The Chebyshev polynomials are computed recursively as $T_k\left(\tilde{\mathcal{L}}\right) = 2\tilde{\mathcal{L}}T_{k-1}\left(\tilde{\mathcal{L}}\right) - T_{k-2}\left(\tilde{\mathcal{L}}\right)$, with $T_0\left(\tilde{\mathcal{L}}\right) = 1$ and $T_1\left(\tilde{\mathcal{L}}\right) = \tilde{\mathcal{L}}$. A K-order ChebConv is defined as $g_{\theta'} * x = \sum_{k=0}^{K} T_k(\tilde{\mathcal{L}})\theta'_k x$, where $\tilde{L}$ is the rescaled Laplacian and θ'_k are filter parameters. The output spatial features from l^{th} convolutional layer with input features $H^l \in \mathbb{R}^{N \times C^l}$ can be calculated as: $H^{l+1} = \sum_{k=0}^{K} T_k(\tilde{\mathcal{L}})H^l\Theta_k^l$, where $\Theta_k^l \in \mathbb{R}^{C^l \times C^{l+1}}$ are the trainable weights for the polynomial of order k in the l^{th} layer.

The model consists of a GCN with L hidden layers. Each hidden layer is followed by a ReLU activation function to introduce non-linearity. The output layer is full connected and comprises two convolutional layers with 256 and 2 channels. The model was trained using the entire population graph. During training, the training nodes were labeled and the test nodes were masked. Cross-entropy loss computed on the training nodes was used to train the GCN and edge encoder. The performance of the model was evaluated using the test nodes.

2.5 Interpretability

We identified relevant node features that contribute to the classification using GNNExplainer [29]. Given a trained GCN model and its prediction on a test node, explainer maximizes the mutual information between a GCN's prediction and the distribution of subgraph structure. GNNExplainer identified a compact subgraph together with a small subset of node features that have a decisive role in the GCN's prediction. The node feature importance was measured by computing feature masks. Feature importance was determined as an average value across subjects to represent the contribution of each brain region. Finally, we reported the top 10 node features with the largest average value of feature importance.

3 Experiments and Results

Experimental Settings. We used pooled stratified cross-validation to split the samples into training and testing sets for the evaluation of the GCN. The pooled samples were randomly divided into 5-folds, of which 1-fold served as the testing set, and the remaining 4-folds were used as the training set. This strategy ensures that training and testing sets contain the equivalent proportions of each class. The model performance was evaluated on the testing set in terms of balanced accuracy (BAC), sensitivity (SEN), specificity (SPE), F1-score, and area under the receiver operating characteristic curve (AUC).

In our experiments, we set Chebyshev polynomial order $K = 3$, quantile $Q = 0.5$, and hidden layer $L = 3$. The number of channels was set to [128, 64, 32]. All models were trained using Adam optimizer with an initial learning rate 0.01, weight decay 5×10^{-5}, and dropout rate 0.2 for 300 epochs to avoid overfitting. Our model was implemented with the PyTorch framework and trained on NVIDIA RTX 3090 GPU.

Competing Methods. To validate the superiority of our proposed model, we compared the GCN model with traditional machine learning algorithms including support vector machine (SVM), random forest (RF), and K-nearest neighbor (KNN). The upper triangle of the MSN was used as input features. We tuned hyperparameters using 5-fold cross-validation to find the optimal parameters via grid search.

3.1 Results and Analysis

Results of Schizophrenia Classification. Table 1 presents the performance of our proposed model compared with those obtained by several machine learning models. The GCN model based on clustering coefficient achieved a superior performance compared with other machine learning models with an average accuracy of 80.2%.

Table 1. Classification comparisons between different methods.

Methods	BAC	SEN	SPE	F1-Score	AUC
RF	52.53 ± 1.17	56.37 ± 12.56	10.65 ± 1.56	17.77 ± 2.23	63.04 ± 2.19
SVM	54.03 ± 1.65	55.74 ± 8.62	17.22 ± 4.97	25.69 ± 5.92	61.04 ± 1.48
KNN	56.03 ± 2.81	50.35 ± 4.09	30.87 ± 7.48	37.81 ± 6.95	57.04 ± 2.58
Ours	**80.18 ± 2.85**	**70.32 ± 4.66**	**82.55 ± 8.44**	**75.50 ± 3.24**	**88.50 ± 1.83**

Identification of Discriminating Regions. Figure 2 shows the top 10 brain regions contributing to GCN classification. The most salient regions contributing to classification were primarily identified in the middle temporal gyrus, superior temporal gyrus, and inferior temporal gyrus. These findings suggest that these regions are crucial in distinguishing patients with schizophrenia from healthy controls. Detailed results are provided in Supplementary Table S2.

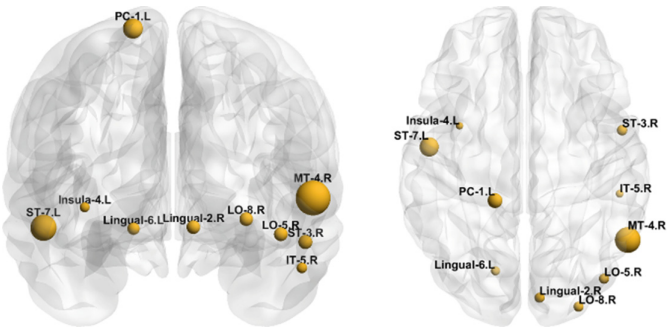

Fig. 2. The top 10 salient brain regions. Abbreviations: MT, middle temporal gyrus; ST, superior temporal gyrus; PC, postcentral gyrus; LO, lateral occipital cortex; IT, inferior temporal gyrus.

3.2 Ablation Studies

We provide detailed investigation of three key components (the density of MSNs, feature selection strategy, and edge encoder) and their influence on classification results.

Influence of the Connection Density of MSNs. We constructed a series of MSNs with different connection densities ranging from 10% to 100% in 10% increments. For each matrix, we computed strength, betweenness centrality, and clustering coefficient to examine their influence on classification performance. In Fig. 3(a), using GCN with strength provides decreased classification performance as the connection density increases. Betweenness centrality shows a lower performance as the density decreases. The use of clustering coefficient is most reliable and provides the highest performance (81.8% accuracy) at connection density of 30%.

Effect of the Feature Selection Strategy. We explored the influence of feature selection strategy on classification performance for the three methods: graph-theoretical measures, RFE, and k-best. The number of features for RFE and k-best considered was C = [308, 500, 1000]. Figure 3(b) shows that using graph-theoretical measures as node features yields improved classification performance. When using clustering coefficient, the GCN model achieved the best performance with 80.18% accuracy. The k-best method shows the lowest performance compared to the other methods.

Influence of the Encoder. We examined the influence of four different types of encoders (PAE, EA, L2, and Cosine + Tanh) on classification performance. Figure 3(c) presents that using PAE as an encoder provides the highest classification performance compared to the other encoders.

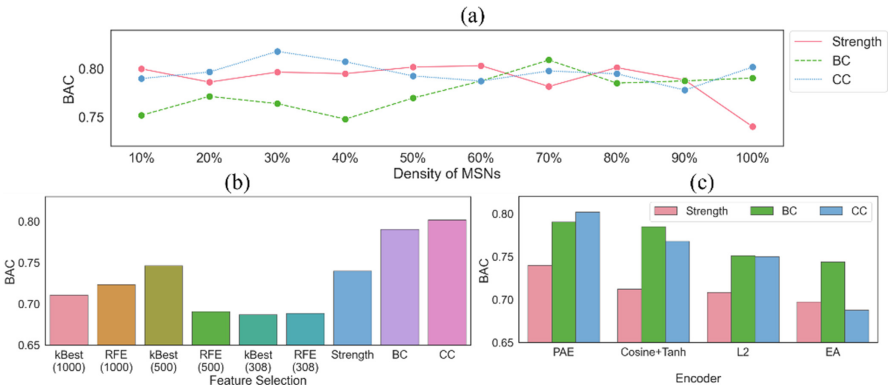

Fig. 3. Ablation studies on three different key components. Classification accuracy (a) for different density of MSNs, (b) for different features selection strategies, and (c) for different types of encoders. Abbreviations: BC, betweenness centrality; CC, clustering coefficient; EA, Edge Adapter.

4 Conclusion

In this study, we proposed an improved GCN structure, which combines graph-theoretical measures of MSNs derived from structural MRI data and non-imaging phenotypic measures for disease classification. This study explored the application value of our proposed GCN model based on cortical networks (MSNs) in differentiating patients with schizophrenia from healthy controls. Using a large multi-site dataset and various validation strategies, reliable and generalizable classification accuracy of 81.8% can be achieved. These results indicate that the MSNs serve as a useful and clinically-relevant phenotype and GCN modeling shows promise in detecting individual patients with schizophrenia. Further, we investigated different graph structures and their influence on classification performance. The GCN model making use of subject-specific clustering coefficient as imaging feature vectors and the PAE encoder performed best. By examining the saliency patterns contributing to GCN classification, we identified the most salient regions in the middle temporal gyrus and superior temporal gyrus. These findings suggest the potential utility of GCN for enhancing our understanding of the underlying neural mechanisms of schizophrenia by identifying clinically-relevant disruptions in brain network topology.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1C1C1009436), the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI) funded by the Ministry of Health and Welfare, Republic of Korea (grant number: HI22C0108), and the Institute for Information and Communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. RS-2022–00155911, Artificial Intelligence Convergence Innovation Human Resources Development (Kyung Hee University)).

References

1. Aine, C., Bockholt, H.J., Bustillo, J.R., et al.: Multimodal neuroimaging in schizophrenia: description and dissemination. Neuroinformatics **15**, 343–364 (2017)
2. Bessadok, A., Mahjoub, M.A., Rekik, I.: Graph neural networks in network neuroscience. IEEE Trans. Pattern Anal. Mach. Intell. **45**(5), 5833−5848 (2022)
3. Bilder, R., Poldrack, R., Cannon, T., et al.: UCLA consortium for neuropsychiatric phenomics LA5c Study. OpenNeuro (2020)
4. Charlson, F.J., Ferrari, A.J., Santomauro, D.F., et al.: Global epidemiology and burden of schizophrenia: findings from the global burden of disease study 2016. Schizophr. Bull. **44**, 1195–1203 (2018)
5. Defferrard, M., Bresson, X., Vandergheynst, P.: Convolutional neural networks on graphs with fast localized spectral filtering. In: Advances in Neural Information Processing Systems, pp. 3844–3852. (2016)
6. Fischl, B.: FreeSurfer. Neuroimage **62**, 774–781 (2012)
7. Gollub, R.L., Shoemaker, J.M., King, M.D., et al.: The MCIC collection: a shared repository of multi-modal, multi-site brain image data from a clinical investigation of schizophrenia. Neuroinformatics **11**, 367–388 (2013)

8. Huang, Y., Albert, C.: Semi-supervised multimodality learning with graph convolutional neural networks for disease diagnosis. In: 2020 IEEE International Conference on Image Processing (ICIP), pp. 2451–2455 (2020)
9. Huang, Y., Chung, A.C.: Edge-variational graph convolutional networks for uncertainty-aware disease prediction. In: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference. Lima, Peru, October 4–8, 2020, Proceedings, Part VII 23, pp. 562–572. Springer, Cham (2020)
10. Jiang, H., Cao, P., Xu, M., Yang, J., Zaiane, O.: Hi-GCN: a hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction. Comput. Biol. Med. **127**, 104096 (2020)
11. Kennedy, D.N., Haselgrove, C., Hodge, S.M., et al.: CANDIShare: a resource for pediatric neuroimaging data. Neuroinformatics **10**, 319–322 (2012)
12. Khodatars, M., Shoeibi, A., Sadeghi, D., et al.: Deep learning for neuroimaging-based diagnosis and rehabilitation of autism spectrum disorder: a review. Comput. Biol. Med. **139**, 104949 (2021)
13. Khundrakpam, B.S., Lewis, J.D., Jeon, S., et al.: Exploring individual brain variability during development based on patterns of maturational coupling of cortical thickness: a longitudinal MRI study. Cereb. Cortex **29**, 178–188 (2019)
14. Kong, R., Li, J., Orban, C., et al.: Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion. Cereb. Cortex **29**, 2533–2551 (2019)
15. Larson, M.K., Walker, E.F., Compton, M.T.: Early signs, diagnosis and therapeutics of the prodromal phase of schizophrenia and related psychotic disorders. Expert Rev. Neurother. **10**, 1347–1359 (2010)
16. Lei, D., Qin, K., Pinaya, W.H., et al.: Graph convolutional networks reveal network-level functional dysconnectivity in schizophrenia. Schizophr. Bull. **48**, 881–892 (2022)
17. Morgan, S.E., Seidlitz, J., Whitaker, K.J., et al.: Cortical patterning of abnormal morphometric similarity in psychosis is associated with brain expression of schizophrenia-related genes. Proc. Natl. Acad. Sci. **116**, 9604–9609 (2019)
18. Parisot, S., Ktena, S.I., Ferrante, E., et al.: Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer's disease. Med. Image Anal. **48**, 117–130 (2018)
19. Qin, K., Lei, D., Pinaya, W.H., et al.: Using graph convolutional network to characterize individuals with major depressive disorder across multiple imaging sites. EBioMedicine **78**, 103977 (2022)
20. Repovš, G., Barch, D.M.: Working memory related brain network connectivity in individuals with schizophrenia and their siblings. Front. Hum. Neurosci. **6**, 137 (2012)
21. Romero-Garcia, R., Atienza, M., Clemmensen, L.H., Cantero, J.L.: Effects of network resolution on topological properties of human neocortex. Neuroimage **59**, 3522–3532 (2012)
22. Rubinov, M., Sporns, O.: Complex network measures of brain connectivity: uses and interpretations. Neuroimage **52**, 1059–1069 (2010)
23. Sadeghi, D., Shoeibi, A., Ghassemi, N., et al.: An overview of artificial intelligence techniques for diagnosis of Schizophrenia based on magnetic resonance imaging modalities: methods, challenges, and future works. Comput. Biol. Med. 105554 (2022)
24. Seidlitz, J., Váša, F., Shinn, M., et al.: Morphometric similarity networks detect microscale cortical organization and predict inter-individual cognitive variation. Neuron **97**, 231–247 (2018)
25. Tanaka, S.C., Yamashita, A., Yahata, N., et al.: A multi-site, multi-disorder resting-state magnetic resonance image database. Scientific data **8**, 227 (2021)
26. Váša, F., Seidlitz, J., Romero-Garcia, R., et al.: Adolescent tuning of association cortex in human structural brain networks. Cereb. Cortex **28**, 281–294 (2018)

27. Wen, G., Cao, P., Bao, H., et al.: MVS-GCN: A prior brain structure learning-guided multi-view graph convolution network for autism spectrum disorder diagnosis. Comput. Biol. Med. **142**, 105239 (2022)
28. Winterburn, J.L., Voineskos, A.N., Devenyi, G.A., et al.: Can we accurately classify schizophrenia patients from healthy controls using magnetic resonance imaging and machine learning? a multi-method and multi-dataset study. Schizophr. Res. **214**, 3–10 (2019)
29. Ying, Z., Bourgeois, D., You, J., Zitnik, M., Leskovec, J.: Gnnexplainer: generating explanations for graph neural networks. Adv. Neural Inf. Process. Syst. 9244–9255 (2019)
30. Zhao, F., Zhang, H., Rekik, I., An, Z., Shen, D.: Diagnosis of autism spectrum disorders using multi-level high-order functional networks derived from resting-state functional MRI. Front. Hum. Neurosci. **12**, 184 (2018)
31. Zheng, W., Eilam-Stock, T., Wu, T., et al.: Multi-feature based network revealing the structural abnormalities in autism spectrum disorder. IEEE Trans. Affect. Comput. **12**, 732–742 (2019)
32. Zheng, W., Yao, Z., Xie, Y., Fan, J., Hu, B.: Identification of Alzheimer's disease and mild cognitive impairment using networks constructed based on multiple morphological brain features. Biol. Psychiatry: Cogn. Neurosci. Neuroimaging **3**, 887–897 (2018)