Classification of Autism Spectrum Disorder Using rs-fMRI data and Graph Convolutional Networks

Tianren Yang
Knight Foundation School of Computing and Information Sciences
Florida International University (FIU)
Miami, Florida
tyang009@fiu.edu

Mai A. Al-Duailij
Princess Nourah Bint Abdul Rahman University
Riyadh, Saudi Arabia
maalduailij@pnu.edu.sa

Serdar Bozdag
Department of Computer Science and Engineering, Department of Mathematics, BioDiscovery Institute
University of North Texas
Denton, Texas
serdar.bozdag@unt.edu

Fahad Saeed
Knight Foundation School of Computing and Information Sciences,
Florida International University (FIU)
Miami, Florida
fsaeed@fiu.edu

2022 IEEE International Conference on Big Data (Big Data) | 978-1-6654-8045-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/BigData55660.2022.10021070

***Abstract*— Autism spectrum disorder (ASD) affects large number of children and adults in the US, and worldwide. Early and quick diagnosis of ASD can improve the quality of life significantly both for patients and their families. Prior research provides strong evidence that structural and functional magnetic resonance imaging (MRI) data collected from individuals with ASD exhibit distinguishing characteristics that differ in local and global, spatial and temporal neural patterns of the brain – and therefore can be used for diagnostic purposes for various mental disorders. However, the data from MRI are high-dimensional and advanced methods are needed to make sense out of these datasets. In this paper, we present a novel model based on graph convolutional network (GCN) that can utilize resting state fMRI (rs-fMRI) data to classify ASD subjects from health controls (HC). In addition to using the graph from traditional correlation matrices, our proposed GCN model incorporates graphlet topological counting as one of the training features. Our results show that graphlets can preserve the topological information of the graphs obtained from fMRI data. Combined with our GCN, the graphlets retain enough topological information to differentiate between the ASD and HC. Our proposed model gives an average accuracy of 64.27% on the whole ABIDE-I data sets (1035 subjects) and highest site-specific accuracy of 75.9%, which is comparable to other state-of-the-art methods – while potentially open to being more interpretable.**

Keywords*— Graph Convolutional Networks, Autism Spectrum disorder, ABIDE I dataset, graphlet

I. INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental syndrome that affects children severely, causing difficulties in social skills and communication [1], [2]. The current clinical diagnosis of ASD is accomplished using behavioral, and cognitive metrics such as developmental, language, behavioral estimations and child's guardians reports in different settings (e.g., home and school) [3]. Early diagnosis for ASD diagnosis remains challenging [4], and may lack specificity, and sensitivity [5], [6] or the rigor required of other diseases such as *diabetes*. Both ADOS and ADR-R tests are prone to false positives. Despite the importance of conferring a diagnosis of ASD early to implement intervention, there are numerous barriers, delays, and conflicting information many families experience during the assessment and diagnosis process [7]. The median age at diagnosis for ASD is 46 - 74 months according to CDC – approx. *diagnostic delay of 2.2 years on average* [8] - and impacts considerable number of children [9]. This delay is consequential considering that the early interventional measures such as applied behavioral analysis (ABA) – in addition to being satisfactory to parents, and medical staff alike – more optimistically, can and does, modify the ASD patterns to improve the health, level of functioning and well-being of the children [10], [11]. These clinical metrics can be imperfect measures as they are subject to high test-retest variability, and are influenced by factors (such as environment, social structure, other mental disorders, etc.) other than changes due to ASD [12]– termed as epistemological uncertainty, and urgent need for *distinctive*, and *reliable* biomarkers is widely accepted.

In the recent decade, advances in neuroimaging technologies have provided a critical step, and have made it possible to measure the pathological, and functional changes associated with ASD in the brain [13], [14]. Resting state functional magnetic resonance imaging (rs-fMRI) – is non-invasive and fast - has the potential to record functional abnormalities in brain networks specific to ASD as compared to the typical brain [15]. For brain disorders, rs-fMRI is a widely used approach as it spontaneously measures low-frequency fluctuations in the blood-oxygen-level-dependent (BOLD) signal to investigate the functional architecture of the brain. However, the subtle changes in functional brain networks are impossible to be used as biomarker using conventional radiological, or computational methods [16]. Advance machine-learning models applied to fMRI data offer a systematic, and possibly interpretable approach to learn subtle patterns leading to objective classification, and biomarker discovery. Such discoveries offer observation of neurological processes, track progression, evaluate whether a specific approach (pharmacological or otherwise) is helping, and offer differentiation power to classify people with typical, ASD, and other mental disorders.

In this paper, we designed and developed a graph neural network [17]–[19] (GNN) based machine-learning model that can utilize rs-fMRI data to classify ASD subjects from health

control (HC). Our method exploits the topological properties of the brain network to distinguish between ASD and HC. In addition to lowering the dimensionality of the input rs-fMRI features, we developed a technique to extract graph motifs (or graphlets) from the regions of interest (ROIs) from the fMRI connectivity matrix, leading to a novel feature that have the distinguishing power to classify between ASD and healthy controls. While other machine-learning models preprocess and transform the connectivity matrix into a set of flattened vectors leading to losing the spatial and temporal information of the brain network; our proposed GCN preserves complex topological information, and hence may be open to interpretation. Our extensive experiments using Autism Brain Imaging Data Exchange (ABIDE-I) benchmarking datasets [20], [21] demonstrate that our model can reach an average accuracy of 64.27% (across whole data set), and exhibits highest accuracy of 75.9% on some of the sites. These results are comparable to other state-of-the-art methods such as ASD-DiagNet [16].

The rest of this paper is structured as follows: in section 2, we discuss the related state-of-the-art tools and approaches of ASD classification. In section 3, we discuss the detail of the design of the machine-learning algorithm. The experimental result and discussion will be presented in Section 4, and conclusions and future work are in Section 5.

II. Background and Related Work

The Autism Brain Imaging Data Exchange (ABIDE) initiative has collaborated with 17 international imaging sites to provide the functional and structural brain imaging benchmarking datasets [13], [14], [20], [22]–[24]. These data sets have been used to develop, test and train machine-learning models. Some of these include probabilistic neural network [13] to classify resting state fMRI data of subjects under age of 20. Another model BrainNetCNN [25] with an element-wise layer achieved an accuracy of 68.7%. Traditional machine-learning models such as Support Vector Machines (SVM) and random forests (RF) are explored in several studies [23], [26], [27].

More advance machine-learning models such as deep learning methods, Autoencoders, Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN) are also gaining popularity due to superior performances in disease diagnosis and classification [18], [22], [27]–[29]. A deep learning method [20] that consists of two stacked autoencoders, and a multi-layer classifier achieved an average of 70% accuracy (across all sites). Another model used extraction of spatiotemporal features from the full 4-D data using 3-D CNNs and 3-D Convolutional LSTMs achieves state-of-the-art results with F1-scores of 0.78 and 0.7 on NYU and UM sites, respectively [30]. Our previously proposed and highly popular method, ASD-DiagNet [22], consists of an autoencoder and a single layer perceptron (SLP) for classification decision and achieved 63.4% accuracy using the complete ABIDE dataset. The maximum accuracy exhibited by ASD-DiagNet is 82%. However, all these traditional and deep-learning methods are not amenable to interpretation, which may be essential for clinical purposes.

In the recent years, Graph Neural Networks (GNNs) have gained a lot of attention because these networks can analyze data that cannot be represented in a grid-like structure i.e. graphs. Application areas include social networks, biological networks, or brain connectomes, which are traditionally represented in the form of graphs. GNN as a deep learning method can learn the spatial patterns in the graphs and has achieved success in disease diagnosis [18], [31], [32]. Some of these models include graph similarity metric using a Siamese graph convolutional neural network (s-GCN) for pair and subject classification [33]. Another hierarchical graph convolution network [18] framework was applied to ASD and Alzheimer's disease. Another more recent GNN proposed regularized pooling layer [34] to highlight significant regions of interest (ROIs) for inferring neuroimaging-derived biomarkers for ASD and improving the accuracy of ASD classification. While these methods have been successful in achieving results for classification of brain data for ASD and Alzheimer's, the impact (of usage) of brain network topology used for achieving the classification is unclear. By extension, the utility of these methods for interpretation is going to be limited.

Another limitation of these methods is selective reporting of subset of ABIDE results. Some studies have chosen to drop the subjects that are under or above a specific age [13], [35] while others have tested them on data from a subset of sites [30]. ASD-DiagNet is the state-of-the-art method in this filed, and has reported results on all 1035 subjects in ABIDE I dataset. These results will be used as a baseline to compare the results of our proposed GCN method in this paper for comprehensive and complete analysis.

III. Materials and Methods

A. Graph Convolutional Network

We use the GCN model proposed by Kipf et al. [36] as the basis of our model. We define an undirected graph for each subject $G = \{\mathcal{R}_i, A_i\}$, where $\mathcal{R}_i = \{r_i^1, r_i^2 \dots, r_i^N\}$ is the set of N nodes, and $A_i \in R^{N\times N}$ is the adjacency matrix of the i-th subject [36]. For CC200 functional parcellation, N equals to 200 and is the number of regions of interests (ROIs).

Each of the convolutional layer can be defined as:

$$H^{l+1} = \sigma\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}H^l W^l\right) \quad (1)$$

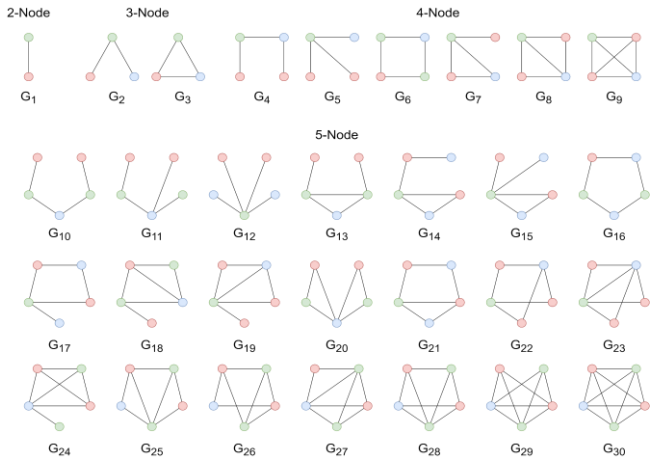

Fig. 1. 30 isomorphic types of undirected subgraphs between size 2 and 5.

Where $\tilde{A} = A + I$, I is the identity matrix and A is the adjacency matrix of the undirected graph G. $\tilde{D}$ is the diagonal node degree matrix of $\tilde{A}$, and W is a trainable weight matrix. Here σ is the activation function and we choose $ReLu(\cdot) = \max(0, \cdot)$.

Graph convolutional layer is used to predict the structure of the graph by using edge features. For example, the weights and the Pearson's correlation coefficient are some of the representations of the relationship between different ROIs. We can define the f_i as the features for the nodes in the l^{th} layer. Then the propagation mode for the update of node representation is given by [31]:

$$f_i = \sigma\left(W_0^l f_i + \sum_{j \in \mathcal{N}(i)} \varphi\left(w_1^{l-1} f_j^{l-1}, e_{ij}\right)\right) \quad (2)$$

Where $N(i)$ denotes the set of indices of neighboring nodes, e_{ij} denotes the features related to the edge of node v_i and v_j. σ is the activation function, and φ is a function used to embed the neighboring nodes' features. In this paper, we choose the ReLu as the activation function for each hidden layer.

B. Node and Edge Feature Construction

1) Pearson's correlation coefficient: The characteristics of nodes were collected from imaging data. After the Craddock template *[37]* was used to partition the ABIDE I preprocessed fMRI data *[38]* into 200 brain areas (CC-200 parcellation), for each of the subject, we construct a graph with 200 nodes using these ROIs. Then, average time series were computed in the brain and the Pearson correlation coefficient between two brain regions was calculated to obtain each subject's whole functional connectivity matrix and this was done for all pairs.

The following equation was used to obtain the coefficient between each region x, and y of length of n.

Algorithms 1. Calculating the graphlet counting vector

Input: an $N \times N$ correlation matrix A_{co}, where in this case $N = 200$. A vector V_g of size 1×30 initialized. 30 types of graphlets G_1 to G_{30}. A given threshold θ.

1: create a graph G with edge weights given by A_{co}

2: remove self-loop edge in G

3: **for** each value a_{ij} in A_{co} (edge in G):

4: remove edge where $|A_{co}| < \theta$. (we have $G \rightharpoonup \hat{G}$)

5: end for

6: **for** $i = 0$ *to* 29 **do**:

7: calculate isomorphism count in $G' \rightarrow V_g[i]$

8: end for

$$p_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (3)$$

Where $\bar{x}$, and $\bar{y}$ are the mean of x and y vectors, respectively. With the 200 regions provided, the computing of all pairwise correlations yield a matrix of 200×200. The functional connectivity matrix is a real symmetric matrix, and the diagonal elements represent the correlation between each region and themselves, so we only keep the upper right triangle of the matrix, which is a widely-adopted method as in [3], [15], [20], [22]. These pairs give $n(n-1)/2 = 19900$ features. The feature chosen for nodes still have a high dimension. Therefore, to avoid the redundancy of the features and possible overfitting problem, we reduce the feature vectors with Recursive Feature Elimination [39] (RFE) with a fixed number of 5000. For our proposed method, the Pearson's correlation coefficient is also used for the graphlet counting vector calculation and is described below.

2) Graphlet counting vector: We propose graphlet counting vector as a feature for preserving the topological information of the fMRI data.

With the correlation matrix calculated, we construct a complete graph with 200 nodes by it. For each pair of the nodes in this graph, there is an edge containing the coefficient value. All the nodes also have a self-loop which is corresponding to the diagonal value of the correlation matrix. For each of the subjects, we construct a vector of 30 dimensions to represent the number of 30 different graphlets with 2 to 5 nodes [40]. A subgraph G' of a graph G is a graph such that $V(G') \subseteq V(G)$ and $E(G') \subseteq$

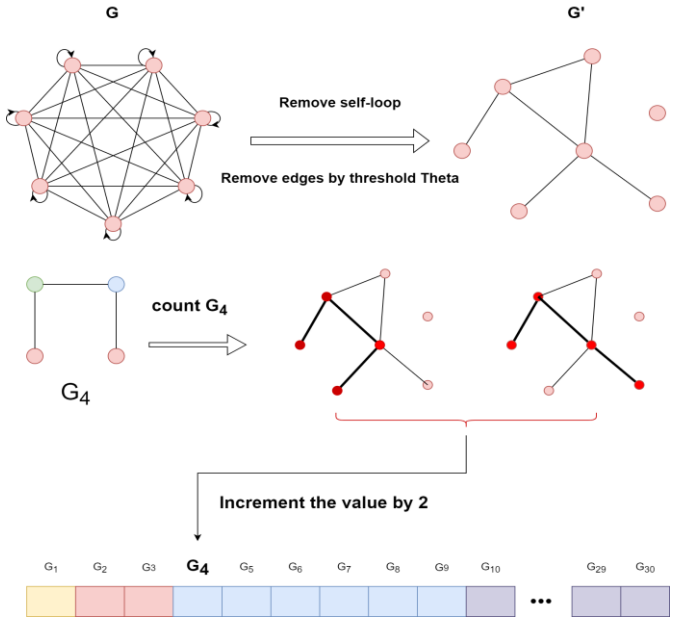

Fig. 2. An example of counting the graphlet type 4 in a 7-node graph.

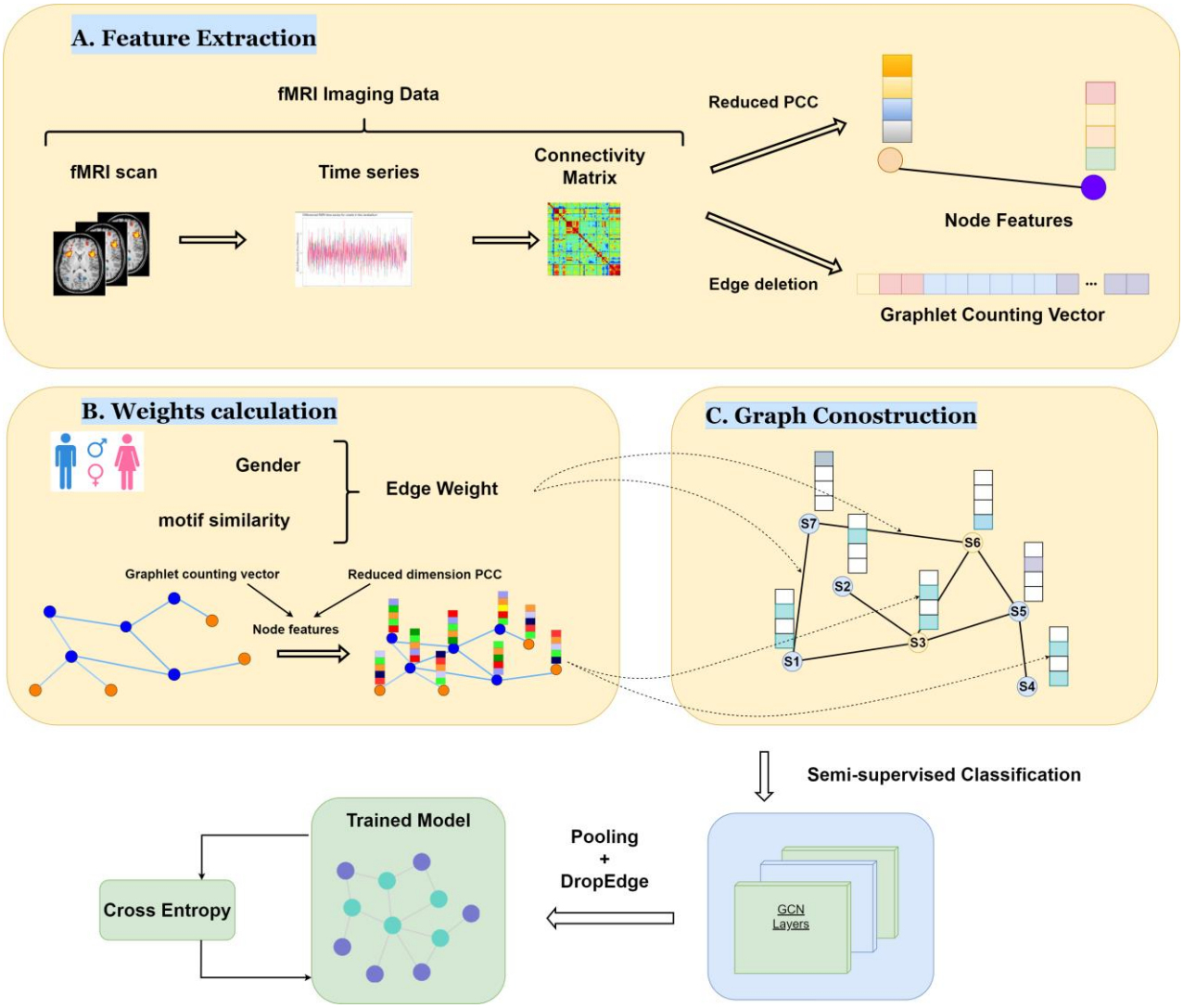

Fig. 3. An overview of the pipeline used for building and training the model of classifying ASD

E(G). And a graphlet is defined as a small, connected, non-isomorphic subgraph. From the 2 nodes to 5 nodes, there are a total of 30 different undirected subgraphs named G_1 to G_{30} as shown in Fig. 1. Like the process we proposed in previous section, we use the 200 regions provided to compute all pairwise correlations and yield a matrix of 200×200. In this case, we have a complete graph G_{pcc} that has 200 nodes, and between each pair of nodes there exists an edge with coefficient value. This graph also contains 200 self-loop edges which will be removed in the next stage. All these values are in range $[-1, 1]$ and with the threshold θ_g we choose, we can get a new graph G'_{pcc}. After all the self-loops and isolated nodes are removed, we use Networkx Python library to count the number of 30 types of graphlets and store the value into a $V_{gc} = (n_{G_0}, n_{G_1}, \dots, n_{G_{29}})$. The whole process is followed by the Algorithm 1 and illustrated in Fig. 2.

3) Edge Weight Construction: The relationship between nodes in the graph in the form of edges is also an important factor to be considered in modeling. We calculate the weight of each pair of the nodes by defining a similarity variable between them. This similarity variable value is consist of two parts – the first one is the phenotypic information, i.e. the gender of the subjects, and the second is the graph motif similarity. It is defined as:

$$Co(M_x, M_y) = Sim(V_{gc}x, V_{gc}y) \quad (4)$$

Where $Sim(V_{gc}x, V_{gc}y)$ represents the similarity of the two graphlet counting vectors (GCV). In this paper, we use the Cosine Similarity to measure it. Given two GCVs x and y with dimension of 30, we define their similarity value as:

$$Sim(V_{gc}x, V_{gc}y) = \frac{x \cdot y}{\|x\|\|y\|} = \frac{\sum_{i=1}^{30} x_i y_i}{\sqrt{\sum_{i=1}^{30} x_i^2}\sqrt{\sum_{i=1}^{30} y_i^2}} \quad (5)$$

Where $\|x\|$ is the Euclidean norm of vector $\vec{x}$. This measure computes the cosine of the angle between $\vec{x}$ and $\vec{y}$. A cosine value of 0 means the two vectors have no match and the closer the cosine value to 1, the smaller the angle between them and the greater the match between vectors.

For the category information gender, φ is defined as:

$$\varphi\left(A_x, A_y\right) = \begin{cases} 0, & if\ A_x \neq A_y \\ 1, & if\ A_x = A_y \end{cases} \quad (6)$$

At last, we can compute the weight for each pair of the edges:

$$W(u, v) = Co(M_u, M_v) + \varphi\left(A_u, A_v\right) \quad (7)$$

TABLE I. LEAVE-ONE-SITE-OUT RESUTLES

Site	Subject Count	Acc.	Sen.	Spe.
CALTECH	37	66.3	53.6	79.7
CMU	27	53.8	68.5	37.9
KKI	48	64.9	62.8	66.4
LEUVEN	63	58.7	52.4	64.1
MAXMUN	52	53	49.7	55.8
NYU	175	68.1	53.9	78.75
OHSU	26	67.6	59.9	74.2
OLIN	34	**75.9**	67.7	**86.3**
PITT	56	65.8	**72.8**	58.3
SBL	30	48.7	66.5	30.9
SDSU	36	50.1	36.7	58.6
STANFORD	39	55.4	59.9	51.1
TRINITY	47	63.6	55.9	70.4
UCLA	98	69.5	57.3	84.5
UM	140	66.2	54.1	77
USM	71	67.4	64.2	73.3
YALE	56	68.3	68.7	67.9
Mean	**1035 (total)**	**64.27**	**57.95**	**69.97**

C. GCN structures

Graph convolutional neural network cannot achieve satisfactory performance by simply stacking graph convolutional modules to deepen the network, and vanishing gradient problem will result in network over-smoothing [41]. Our proposed model, shown in Fig. 3, has one input layer, several hidden layers, and an output layer in this model. Input layer and hidden layers were followed by node pooling layer, and output layer was followed by SoftMax.

The node pooling layer is used to reduce the redundant features of the graph [17]. The methods include aggregating the neighboring nodes to one node and pruning the original graph G to a smaller subgraph with only some important nodes. In order to classify at the graph level, we need a dimensionality reduction layer to bring the number of nodes and the feature dimension of each node down to a suitable size. Previous research results [42] have shown that some parts of the brain (ROIs) hold more weight in prediction and classification than other regions, so they must be retained in the pooling process.

To avoid overfitting, we applied DropEdge [43] to the feature graph when the hidden layer was convolved. The idea of DropEdge is to randomly remove a certain number of edges from the input graph at each training epoch, which can be treated as a data augmenter. In our experiment, the super-parameter value of drop rate was set to 0.2, which means edges were deleted with a probability of 20%.

IV. EXPERIMENT RESULT AND DISCUSSION

The purpose of this paper is to propose a functional connectivity based graph neural networks model that using fMRI time series and the topological information (graphlet count) within ROIs to classify and assist ASD diagnosis.

A. Data Acquisition

In this paper, we used the ABIDE I data, which is a collection of 1112 subjects composed of structural and rs-fMRI data including 539 people with ASD and 573 healthy people. Because various locations have different data collecting equipment, parameters setting, diagnosis processes, and assessment techniques, ABIDE dataset I is a highly heterogeneous database [17], [31]. After filtering the incomplete and corrupted subjects data, our experiment dataset consists of 505 subjects with ASD and 530 healthy controls from all the 17 sites [20], [44]. The data was preprocessed using a widely adopted pipeline called Configurable Pipeline for the Analysis of Connectomes (C-PAC) and parcellated the brain into 200 region of interests (ROIs) using Craddock 200 [37] (CC200) functional parcellation.

Fig. 4. Comparison results with the existing methods

B. Implementation details

For all the experiments reported in this section, we run them on a Linux server running Ubuntu Operating System. The server contains two Intel Xeon (Skylake) Processors at 2.1 GHz with a total 48 GBs of RAM. The system contains an NVIDIA Tesla K-40c GPU with 2,880 CUDA cores and 12 GBs of RAM. CUDA version 11.4 and PyTorch library were used for conducting the experiments. In the training process, learning rate was set to 0.005, and dropout rate was set to 0.15. Each model was iteratively trained 200 times.

We implemented a GCN-based ASD classification using fMRI data with the pipeline of PyTorch framework [45]. As mentioned above, the nodes of the graph were obtained from the classification of brain regions as well as the graphlet count vector, while the edges feature of the graph were constructed by the weight between pairs of the ROIs with age information and graph motif similarities. The potential representations of the graph were learned by the GCN layer, and the salient features were fed to the classification layer for autism label prediction [19]. The readout layer was used to aggregate the node feature vectors into a single vector which was then used as input to the final classifier. With a limited number of data topics available from the dataset, this model was evaluated using a k-fold cross-validation technique to overcome possible overfitting problems. The dataset was randomly partitioned into k groups of equal size, one of which was used to obtain the classification performance and this process was repeated k times. Due to the subject size limitation of the ABIDE dataset, when evaluating the results site by site, we use k=5 to avoid substantial bias in the results for a particular data site due to the insufficient number of the subjects [22].

C. Evaluation

To evaluate the proposed model, the entire data set was divided into two sets of subjects: 80% training data (828 subjects) and 20% testing data (207 subjects). Comparing to other neural networks, graph networks are a more natural way to represent brain network since graphs can inherently capture relationships and complex interactions between ROIs, and thus possess the power of interpreting relational information in human brains. The improved performance of our proposed GCN model in classifying functional connectivity networks comes from graph convolutional layers for node and edge representation.

Additionally, as an evaluation of the classifier performance across sites, we performed a leave-one-site-out cross validation process. This process excluded all the subjects from one site from the training process and used that data as the test set to evaluate the model. We repeated this process for all the 17 sites and the results of these analyses are reported in Table I.

For a whole data sites evaluation, we performed 10-fold cross-validation on all 1,035 subjects using CC-200 atlas. We compared accuracy, sensitivity, and specificity of our approach with [21], [22]. As shown in Fig 4, our method achieves 64.27% accuracy as compared to 60% and 63.2% for Neilsen et. al and ASD-DiagNet. We were able to perform well for sensitivity 57.95% as compared to ASD-DiagNet which exhibited 53.67% but fall short when compared with Neilsen et. al. (62%). We outperformed the other two leading methods in specificity with our GCN model achieving 69.97% as compared to 58% and 68.4% for Neilsen et. al. and ASD-DiagNet, respectively. These are encouraging early results for GCN based classification of ASD vs. HC.

V. Conclusions and Discussions

In this paper, we proposed a novel graph convolutional network (GCN) method to classify between autistic brain scans from healthy brain scans. To utilize the spatial structure of the brain network and extract information from the topology of the connectome, we utilized a novel graphlet count vector feature, which were then fed to the GCN layers to perform the classification. We evaluated our developed method using ABIDE-I data set, which contains all the images from 17 different acquisition sites and compared with other state-of-the-art techniques. Our results showed a mean test accuracy of 64.3% in a 10-fold cross validation experiment using the GCN algorithm and Drop Edge strategy for whole ABIDE-I data set with comparable sensitivity and superior specificity as compared to the state-of-the-art methods. Our method provides a novel architecture and method for classifying, characterizing ASD biomarkers and can potentially be adaptable for diagnosis of other mental illnesses such as Alzheimer's and ADHD disease. Further investigation will be needed to interpret the GCN models specific to ASD classification and our unique graph motif based model will be more amenable to explanations, which will be essential for clinical adaptations.

There are few limitations to our work presented in this paper. One is that the ABIDE dataset size is small and the network training stops relatively fast before it is overfitted – a limitation that is shared among all the existing machine-learning models. Another is that the current model is only tested on ABIDE-I dataset. More testing for larger data sets in the future will give us more insights.

Acknowlegements

This research is funded by National Science Foundation (NSF) award # TI-2213951. In addition, part of this research is funded by supplemental grant to NIH NIGMS R01GM134384. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author (s) and do not necessarily reflect the views of the National Science Foundation (NSF) or National Institutes of Health (NIH).

References

[1] L. Shao, C. Fu, Y. You, and D. Fu, "Classification of ASD based on fMRI data with deep learning," *Cogn. Neurodyn.*, vol. 15, no. 6, pp. 961–974, Dec. 2021, doi: 10.1007/s11571-021-09683-0.

[2] T. Eslami, F. Almuqhim, J. S. Raiker, and F. Saeed, "Machine Learning Methods for Diagnosing Autism Spectrum Disorder and Attention- Deficit/Hyperactivity Disorder Using Functional and Structural MRI: A Survey," *Front. Neuroinformatics*, vol. 14, 2021,

Accessed: May 14, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fninf.2020.575999

[3] C. Yang, P. Wang, J. Tan, Q. Liu, and X. Li, "Autism spectrum disorder diagnosis using graph attention network based on spatial-constrained sparse functional brain networks," *Comput. Biol. Med.*, vol. 139, p. 104963, Dec. 2021, doi: 10.1016/j.compbiomed.2021.104963.

[4] C. A. Mazefsky and D. P. Oswald, "The discriminative ability and diagnostic utility of the ADOS-G, ADI-R, and GARS for children in a clinical setting," *Autism*, vol. 10, no. 6, pp. 533–549, Nov. 2006, doi: 10.1177/1362361306068505.

[5] D. H. Geschwind and M. W. State, "Gene hunting in autism spectrum disorder: on the path to precision medicine," *Lancet Neurol.*, vol. 14, no. 11, pp. 1109–1120, Nov. 2015, doi: 10.1016/S1474-4422(15)00044-7.

[6] A. Jahedi, C. A. Nasamran, B. Faires, J. Fan, and R.-A. Müller, "Distributed Intrinsic Functional Connectivity Patterns Predict Diagnostic Status in Large Autism Cohort," *Brain Connect.*, vol. 7, no. 8, pp. 515–525, Oct. 2017, doi: 10.1089/brain.2017.0496.

[7] M. Lappé, L. Lau, R. N. Dudovitz, B. B. Nelson, E. A. Karp, and A. A. Kuo, "The Diagnostic Odyssey of Autism Spectrum Disorder," *Pediatrics*, vol. 141, no. Supplement 4, pp. S272–S279, Apr. 2018, doi: 10.1542/peds.2016-4300C.

[8] K. Zuckerman, O. J. Lindly, and A. E. Chavez, "Timeliness of Autism Spectrum Disorder Diagnosis and Use of Services Among U.S. Elementary School–Aged Children," *Psychiatr. Serv.*, vol. 68, no. 1, pp. 33–40, Jan. 2017, doi: 10.1176/appi.ps.201500549.

[9] V. Hus and C. Lord, "The autism diagnostic observation schedule, module 4: revised algorithm and standardized severity scores," *J. Autism Dev. Disord.*, vol. 44, no. 8, pp. 1996–2012, 2014.

[10] M. J. Maenner *et al.*, "Prevalence and Characteristics of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2018," *MMWR Surveill. Summ.*, vol. 70, no. 11, pp. 1–16, Dec. 2021, doi: 10.15585/mmwr.ss7011a1.

[11] S. L. Hyman *et al.*, "Identification, Evaluation, and Management of Children With Autism Spectrum Disorder," *Pediatrics*, vol. 145, no. 1, p. e20193447, Jan. 2020, doi: 10.1542/peds.2019-3447.

[12] R. Lane, "Expanding boundaries in psychiatry: uncertainty in the context of diagnosis-seeking and negotiation," *Sociol. Health Illn.*, vol. 42, no. S1, pp. 69–83, Aug. 2020, doi: 10.1111/1467-9566.13044.

[13] T. Iidaka, "Resting state functional magnetic resonance imaging and neural network classified autism and control," *Cortex*, vol. 63, pp. 55–67, Feb. 2015, doi: 10.1016/j.cortex.2014.08.011.

[14] H. Chen *et al.*, "Multivariate classification of autism spectrum disorder using frequency-specific resting-state functional connectivity—A multi-center study," *Prog. Neuropsychopharmacol. Biol. Psychiatry*, vol. 64, pp. 1–9, Jan. 2016, doi: 10.1016/j.pnpbp.2015.06.014.

[15] F. Almuqhim and F. Saeed, "ASD-SAENet: A Sparse Autoencoder, and Deep-Neural Network Model for Detecting Autism Spectrum Disorder (ASD) Using fMRI Data," *Front. Comput. Neurosci.*, vol. 15, 2021, Accessed: May 22, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fncom.2021.654315

[16] T. Eslami, V. Mirjalili, A. Fong, A. R. Laird, and F. Saeed, "ASD-DiagNet: A Hybrid Learning Approach for Detection of Autism Spectrum Disorder Using fMRI Data," *Front. Neuroinformatics*, vol. 13, p. 70, Nov. 2019, doi: 10.3389/fninf.2019.00070.

[17] M. Cao *et al.*, "Using DeepGCN to identify the autism spectrum disorder from multi-site resting-state data," *Biomed. Signal Process. Control*, vol. 70, p. 103015, Sep. 2021, doi: 10.1016/j.bspc.2021.103015.

[18] H. Jiang, P. Cao, M. Xu, J. Yang, and O. Zaiane, "Hi-GCN: A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction," *Comput. Biol. Med.*, vol. 127, p. 104096, Dec. 2020, doi: 10.1016/j.compbiomed.2020.104096.

[19] N. S. Dsouza, M. B. Nebel, D. Crocetti, J. Robinson, S. Mostofsky, and A. Venkataraman, "M-GCN: A Multimodal Graph Convolutional Network to Integrate Functional and Structural Connectomics Data to Predict Multidimensional Phenotypic Characterizations," in *Proceedings of the Fourth Conference on Medical Imaging with Deep Learning*, Aug. 2021, pp. 119–130. Accessed: May 26, 2022. [Online]. Available: https://proceedings.mlr.press/v143/dsouza21a.html

[20] A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz, and F. Meneguzzi, "Identification of autism spectrum disorder using deep learning and the ABIDE dataset," *NeuroImage Clin.*, vol. 17, pp. 16–23, Jan. 2018, doi: 10.1016/j.nicl.2017.08.017.

[21] J. A. Nielsen *et al.*, "Multisite functional connectivity MRI classification of autism: ABIDE results," *Front. Hum. Neurosci.*, vol. 7, 2013, doi: 10.3389/fnhum.2013.00599.

[22] T. Eslami, V. Mirjalili, A. Fong, A. R. Laird, and F. Saeed, "ASD-DiagNet: A Hybrid Learning Approach for Detection of Autism Spectrum Disorder Using fMRI Data," *Front. Neuroinformatics*, vol. 13, 2019, Accessed: May 22, 2022. [Online]. Available: https://www.frontiersin.org/article/10.3389/fninf.2019.00070

[23] A. Abraham *et al.*, "Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example," *NeuroImage*, vol. 147, pp. 736–745, Feb. 2017, doi: 10.1016/j.neuroimage.2016.10.045.

[24] S. Itani and D. Thanou, "Combining anatomical and functional networks for neuropathology identification: A case study on autism spectrum disorder," *Med. Image Anal.*, vol. 69, p. 101986, Apr. 2021, doi: 10.1016/j.media.2021.101986.

[25] J. Kawahara *et al.*, "BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment," *NeuroImage*, vol. 146, pp. 1038–1049, Feb. 2017, doi: 10.1016/j.neuroimage.2016.09.046.

[26] V. Subbaraju, M. B. Suresh, S. Sundaram, and S. Narasimhan, "Identifying differences in brain activities and an accurate detection of autism spectrum disorder using resting state functional-magnetic resonance imaging : A spatial filtering approach," *Med. Image Anal.*, vol. 35, pp. 375–389, Jan. 2017, doi: 10.1016/j.media.2016.08.003.

[27] A. R. J. Fredo, A. Jahedi, M. Reiter, and R.-A. Müller, "Diagnostic Classification of Autism Using Resting-State Fmri Data and Conditional Random Forest," p. 5.

[28] S. Parisot *et al.*, "Disease prediction using graph convolutional networks: Application to Autism Spectrum Disorder and Alzheimer's disease," *Med. Image Anal.*, vol. 48, pp. 117–130, Aug. 2018, doi: 10.1016/j.media.2018.06.001.

[29] M. Aledhari, S. Joji, M. Hefeida, and F. Saeed, "Optimized CNN-based Diagnosis System to Detect the Pneumonia from Chest Radiographs," in *2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, San Diego, CA, USA, Nov. 2019, pp. 2405–2412. doi: 10.1109/BIBM47256.2019.8983114.

[30] A. El-Gazzar, M. Quaak, L. Cerliani, P. Bloem, G. van Wingen, and R. Mani Thomas, "A Hybrid 3DCNN and 3DC-LSTM Based Model for 4D Spatio-Temporal fMRI Data: An ABIDE Autism Classification Study," in *OR 2.0 Context-Aware Operating Theaters and Machine Learning in Clinical Neuroimaging*, Cham, 2019, pp. 95–102. doi: 10.1007/978-3-030-32695-1_11.

[31] X. Li *et al.*, "BrainGNN: Interpretable Brain Graph Neural Network for fMRI Analysis," p. 28.

[32] X. Li, N. C. Dvornek, Y. Zhou, J. Zhuang, P. Ventola, and J. S. Duncan, "Graph Neural Network for Interpreting Task-fMRI Biomarkers," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, Cham, 2019, pp. 485–493. doi: 10.1007/978-3-030-32254-0_54.

[33] S. I. Ktena *et al.*, "Metric learning with spectral graph convolutions on brain connectivity networks," *NeuroImage*, vol. 169, pp. 431–442, Apr. 2018, doi: 10.1016/j.neuroimage.2017.12.052.

[34] X. Li *et al.*, "Pooling Regularized Graph Neural Network for fMRI Biomarker Analysis," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2020*, Cham, 2020, pp. 625–635. doi: 10.1007/978-3-030-59728-3_61.

[35] A. Kazeminejad and R. C. Sotero, "Topological Properties of Resting-State fMRI Functional Networks Improve Machine Learning-Based Autism Classification," *Front. Neurosci.*, vol. 12, 2019, Accessed: Aug. 29, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fnins.2018.01018

[36] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks." arXiv, Feb. 22, 2017. Accessed: Aug. 17, 2022. [Online]. Available: http://arxiv.org/abs/1609.02907

[37] R. C. Craddock, G. A. James, P. E. Holtzheimer III, X. P. Hu, and H. S. Mayberg, "A whole brain fMRI atlas generated via spatially constrained spectral clustering," *Hum. Brain Mapp.*, vol. 33, no. 8, pp. 1914–1928, 2012, doi: 10.1002/hbm.21333.

[38] C. Cameron *et al.*, "The Neuro Bureau Preprocessing Initiative: open sharing of preprocessed neuroimaging data and derivatives," *Front. Neuroinformatics*, vol. 7, 2013, doi: 10.3389/conf.fninf.2013.09.00041.

[39] X. Chen and J. C. Jeong, "Enhanced recursive feature elimination," in *Sixth International Conference on Machine Learning and Applications (ICMLA 2007)*, Cincinnati, OH, USA, Dec. 2007, pp. 429–435. doi: 10.1109/ICMLA.2007.35.

[40] P. Ribeiro, P. Paredes, M. E. P. Silva, D. Aparicio, and F. Silva, "A Survey on Subgraph Counting: Concepts, Algorithms, and Applications to Network Motifs and Graphlets," *ACM Comput. Surv.*, vol. 54, no. 2, pp. 1–36, Mar. 2022, doi: 10.1145/3433652.

[41] Y. Wang, J. Liu, Y. Xiang, J. Wang, Q. Chen, and J. Chong, "MAGE: Automatic diagnosis of autism spectrum disorders using multi-atlas graph convolutional networks and ensemble learning," *Neurocomputing*, vol. 469, pp. 346–353, Jan. 2022, doi: 10.1016/j.neucom.2020.06.152.

[42] R. C. Craddock, M. P. Milham, and S. M. LaConte, "Predicting intrinsic brain activity," *NeuroImage*, vol. 82, pp. 127–136, Nov. 2013, doi: 10.1016/j.neuroimage.2013.05.072.

[43] Y. Rong, W. Huang, T. Xu, and J. Huang, "DropEdge: Towards Deep Graph Convolutional Networks on Node Classification." arXiv, Mar. 12, 2020. Accessed: Aug. 22, 2022. [Online]. Available: http://arxiv.org/abs/1907.10903

[44] M. D. Fox, "Clinical applications of resting state functional connectivity," *Front. Syst. Neurosci.*, 2010, doi: 10.3389/fnsys.2010.00019.

[45] M. Fey and J. E. Lenssen, "Fast Graph Representation Learning with PyTorch Geometric," arXiv, arXiv:1903.02428, Apr. 2019. doi: 10.48550/arXiv.1903.02428.