National College of Ireland

Enhancing the accuracy of Autism detection using fMRI images with Graph Autoencoder and Graph Neural Networks

MSc Research Project
Data Analytics

Sumit Rai
Student ID: x20207603

School of Computing
National College of Ireland

Supervisor: Hicham Rifai

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Sumit Rai
Student ID:	x20207603
Programme:	Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Hicham Rifai
Submission Due Date:	15/08/2022
Project Title:	Enhancing the accuracy of Autism detection using fMRI images with Graph Autoencoder and Graph Neural Networks
Word Count:	5604
Page Count:	22

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	30th August 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Enhancing the accuracy of Autism detection using fMRI images with Graph Autoencoder and Graph Neural Networks

Sumit Rai
x20207603

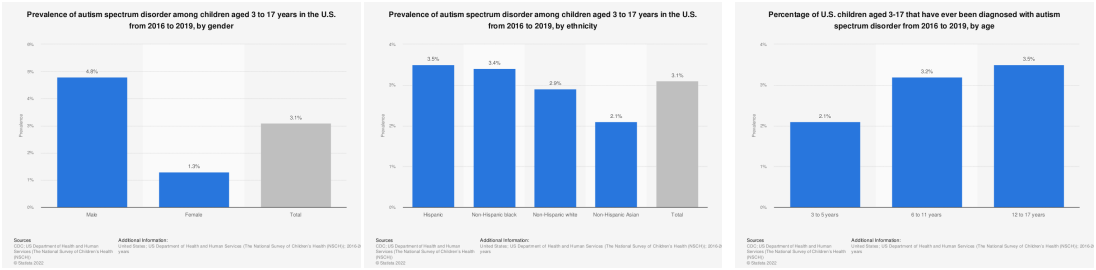
Abstract

Autism Spectral Disorder (ASD) is a neurological developmental disorder. It affects the brain and the symptoms usually manifest as challenges in social skills, restrictive interests and repetitive behaviour. The traditional methods for ASD detection have been based on behavioral observations which are neither efficient nor accurate. Recently, resting-state Functional Magnetic Resonance Imaging (rs-FMRI) has been used to understand the mechanisms of brain disorders such as ASD. In this paper, a graph neural network with graph auto-encoder is proposed and implemented to enhance the accuracy of the ASD detection using rs-FMRI. The correlation between blood oxygen level-dependent (BOLD) signals in the region of interest (ROI) in the brain is used to create functional connectivity matrix. Then, the graph auto-encoder is used for feature representation and a graph neural network is used for the classification task. Both the networks (GAE & GNN) are trained together to tune the latent representation by graph auto-encoder for classifying the subjects as Autistic or non-Autistic. An accuracy of around **55%** is achieved on ABIDE preprocessed dataset, which is less than ideal for medical applications. The results show that the proposed framework should be improved further to achieve an improved classification accuracy over other research which are usually above 78% on ABIDE-I preprocessed dataset.

1 Introduction

Autism Spectral Disorder (ASD) consists of multiple common neurodevelopmental disorders such as Autism and Asperger's Syndrome. The clinical symptoms usually observed are challenged social communication and interaction skills, unusual interests, repetitive behaviours (stereotypical movements) and delayed cognitive skills. According to the World Health Organization (WHO), ASD is very prevalent and 1 in 160 children suffers from it globally. The prevalence of ASD in the USA by gender, ethnicity and age is shown in Figure 1a, Figure 1b and Figure 1c respectively. [1]

[1]The statistics are from Statista.



(a) Prevalence by gender (b) Prevalence by ethnicity (c) Prevalence by age

Figure 1: Prevalence of autism spectrum disorder among children aged 3 to 17 years in the U.S. from 2016 to 2019

Although ASD is a lifelong disorder, but the early detection may help in reducing the severity of the symptoms, better supervision and thereby providing a better quality of life.

The cost of treating an ASD patient is a major financial burden for any economy and is expected to increase. There has been a lot of research going on for ASD treatment and huge amount of money have been getting invested by many countries actively for the neurobiology and neurogenetics research for ASD. However, the etiology of ASD is still not fully known and its pathology is unknown. There has been no conclusive identification of unified structural traits due to multiple varieties of causes of ASD.

The traditional methods for detection of ASD are based on symptoms. These methods are neither efficient nor accurate. The current clinical diagnosis depends on a checklist of criteria available in diagnostic and statistical manuals. Furthermore, the ASD diagnostic metrics and categories suffer from limitations in its observational capabilities. Moreover, the reduction in distinction due to dynamicity in the definition of ASD has further made the situation complex, due to which the chances of misdiagnosis and dual diagnosis have increased.

The current advancement in neural imaging technologies and machine learning algorithms have sparked interest in utilizing deep learning techniques in the detection of ASD. These techniques are non-invasive in nature and aid in investigating the structural and functional patterns in the brain. Specifically, the resting state functional magnetic resonance imaging (rs-FMRI) has been used in the detection of Autism through functional abnormalities in an autistic brain. These abnormalities are based on resting-state blood oxygen level-dependent contrasts. These contrasts represent the energy used by the cells and help in providing the subject's baseline BOLD variance. For these analyses, the brain is parcellated into different regions known as biomarkers. The pairwise interactions between the biomarkers hold the key to deciphering the complex functional interactions to detect and differentiate between ASD patients and normal individuals.

The paper is organized as follows: Section 2 describes the literature review or related work, section 3 descibes the methodology. The design specification is provided in section 4. Section 5 illustrates the implementation. The model evaluation is discusses in section 6 followed by discussion in section 6.3. Finally, the paper is concluded in section 7.

2 Related Work

The spark in the study of mental disorders using functional connectivities as a potential biomarker has led to an increase in the developmental efforts of computer diagnostic mod-

els (CAD). Many studies on mental disorders have utilized these functional connectivities as biomarkers such as mental states classification.(e.g. emotions (Kassam et al.; 2013) and mental disorders (e.g. Schizophrenia, ASD (Yahata et al.; 2017)), semantic categories (O'toole et al.; 2005) and learning (Bauer and Just; 2015)). Few of the relevant research in this field is summarised below:

2.1 Traditional Machine Learning

To reduce the variability in fMRI images collected from multiple sites, data harmonization is suggested in Ingalhalikar et al. (2021). The sources of site variability are differences in scanner, difference in image acquisition parameters, movement of head etc. The *Combating batch effect when combining batches* (ComBat) technique for data harmonization is used. The difference in classification accuracy between harmonized vs. non-harmonized data is compared using ANN, Autoencoders and Random Forests. The dataset used is ABIDE-I preprocessed and the parcellation method used is CC200. Similar to other research, the dataset is used to create an adjacency matrix using Pearson coefficient to calculate co-activation between regions of the brain. The ablation study is also carried out by creating 12 brain sub-networks. The discriminative power of sub-networks on classification accuracy is studied across sites. The highest accuracy of 71.35% is achieved on harmonized data using ANN as a classifier.

2.2 Convolutional Neural Networks

The enhanced convolutional neural network is discussed in Kashef (2022) for the efficient diagnosis of ASD consisting of two temporal convolutional blocks with a kernel of size 4. The filter size of the first block is 32 and then a filter of size 64 is used. The max pooling layer is used to decrease overfitting and for regularization, a dropout layer is used. The ReLU function is used for the activation function with Softmax classifier. This proposed enhanced convolutional neural network achieved an accuracy of 80% on ABIDE-I preprocessed dataset.

An improvement to classical CNN is suggested in Wang (2021). The input is converted from 2-D to 1-D with a 1-D kernel & a pooling filter. The activation function ReLU is replaced with PReLU activation function. The Softmax classifier is used before pre-training and after pre-training, a parameterized SVM is used. The issue of overfitting is avoided by replacing L1 with L2 regularization for cross entropy loss. A dataset containing 416 Autistic and 441 non-Autistic subjects is used. The accuracy is 84.44% using 10-fold cross validation, whereas sensitivity is 85.39% and specificity is 80.57%.

In M and Jaganathan (2021), the summary of sMRI data is combined with rs-FMRI data to get better results. A 3D CNN is used for dimension reduction of summary data while a sparse matrix is created from fMRI images by passing through variational autoencoder. Both of these data are used as input to the graph convolutional network for the classification of Autism. The accuracy achieved is 60.9% - 62.6% on ABIDE preprocessed dataset.

2.3 Siamese Neural Networks

A different approach is provided in Tummala (2021). A deep learning framework which uses T1-weighted MRI images is proposed. The Siamese Neural Nets (SNN) using ResNet50 (pre-trained computer vision residual network model) for the ASD classification is utilized. The preprocessing steps include reorientation, cropping and registration of affine. A validation accuracy of 99% using 5-fold stratified cross-validation is achieved but the test accuracy hasn't been mentioned.

2.4 Federated Learning

The privacy issues and access rights while sharing 3D MRI brain scans is discussed in Fan et al. (2021) which hinders aggregated analysis. The result is low performance and lack of generalizability of local models which are trained using data from a single site due to limited data size. A solution is proposed that uses a guide-weighted federated learning framework of deep learning (3D-GWFDL). The training of local models is perfomed at each site and then the federated model is updated with gradients from each site. There is no sharing of private raw data, but instead, encrypted private site information is shared using differential privacy method. In contrast to other research, the structural MRI data (ABIDE-I and ABIDE-II) is used. An increase in accuracy ranging from 0.92% to 4.02% from local models trained on a single site is seen.

2.5 Graph Neural Networks

A graph-based network using deep belief network (DBN) is proposed by Huang et al. (2021). DBN is a graph extension of K-nearest neighbours. The further refinement is performed by using a restricted path-based depth-first search algorithm. The dataset ABIDE-I is used, having 505 and 530 ASD patients and typical controls respectively. The raw rs-fmri data is preprocessed by extracting the mean time series and converting it into a matrix T of features. The matrix T is a 2 dimensional matrix in which the element T_{ij} is the mean time series of ith region of interest (ROI) and jth timestamp. The graph-based feature selection (GBFS) is utilized to select remarkable functional connectivities using external and internal measures. The level of co-activation between the ROI's time series is measured by using Pearson's coefficient. The remarkable connections are chosen from a total of 19900 [(200 x (200-1)/2 = 19900] possible connections by filtering out those connections for which the following conditions satisfy: $FC_{ASD}(i,j) \geq mean_{ASD} + \alpha * STD_{ASD}$ for ASD subjects, $FC_{ASD}(i,j) < mean_{ASD} + \alpha * STD_{ASD}$ for typical controls where $FC_{ASD}(i,j)$ is connection mean, $mean_{ASD}$ is global mean and STD_{ASD} is global standard deviation. The graph is extended by choosing K nearest neighbours for K=6 selected empirically and exploring it with a restricted path-based depth-first search algorithm implementation (RP-DFS). Finally, A DBN layer consisting of 3 layers is used and an accuracy of 0.764 $\pm$0.022 based on 10-fold cross-validation is achieved with a sensitivity of 0.778 and specificity of 0.750.

In Yin et al. (2021) , a connectivity-based graph attention network is proposed for the detection of ASD from rs-FMRI data. The dataset ABIDE-I is used with Powell atlas as parcellation method. The preprocessing step uses Analysis of Functional Neuroimages

(AFNI) and FMRIB's Software Library (FSL) software packages for fMRI processing. The cleaning of data is performed using the similar steps mentioned in Mostafa et al. (2020) for motion artifacts and noises. A sparse connectivity matrix of size 264 x 264 is generated from the ROI's time series and using Pearson's coefficient for determining connectivity strength. The connectivity matrix is converted into an adjacency matrix representing a graph for each subject.There are 2 groups of features, one related to graph statistics such as centralities, betweeness, eccentricity and others related to statistics of time series such as mean, variance, skewness, and kurtosis. So, the number of dimensions for each node feature vector is 7. The latent representation of the nodes is learnt using a graph attention network and classification is performed using the MLP layer. The adjacency matrix represents a learnt graph and time series characteristics embedding of a single graph. A 5-fold cross validation with 3:1 train-test split is used and a mean accuracy of 82.3% is achieved with a sensitivity of 83.6% and specificity of 78.3%.

A multi-class classification of ASD subtypes is presented in Al-Hiyali et al. (2021). The ABIDE dataset labelled with ASD subtypes such as ASD, APD, PPD-NOS and normal control (NC) are classified using a Convolutional Neural Network with dynamic input functional connectivity. The BOLD signals are extracted using DPARSF Matlab tool. The dynamic functional connectivity between nodes of the brain is quantified by a new metric wavelet coherence (WCF). This gives the overall coherence variability over time. The standard brain atlas, *Automated Anatomical Labelling (AAL)* is used as for brain parcellation. The matrix R has time series as columns of 90 nodes related to cortical regions are selected out of a total 116 nodes. It is then used to create a 2-D matrix of wavelet conherence (WC). The element in the coherence matrix denotes the pairwise node coherence over a period of time at a frequency F. The RMS value over the time points is used to calculate the average over the frequency range. The filtration of the nodes is peformed by graph-based filtering to create a scalogram containing coherent synchronous features. A CNN is used for classification with 10-fold cross-validation and an accuracy of 88.6% is achieved.

In Wang et al. (2022), a multi-altlas graph convolutional networks are combined for the automatic diagnosis of ASD with ensemble learning on preprocessed ABIDE-I dataset having different brain atlases. The Configurable Pipeline for Analysis of Connectomes (CPAC) is used for the preprocessing of the time-series brain regions. The correlation between brain regions are calculated using Pearson coefficient for each subject. The feature selection is performed using recursive feature elimination (SVM-RFE). The relevant features for automatic diagnosis of ASD are selected by the multi-atlas graph convolutional network method(MAGCN). The Ridge classifier is used for the final automatic ASD diagnosis after combining the feature representation with the proposed ensemble learning method. The accuracy achieved is 75.86% while corresponding figures for specificity is 71.53% and sensitivity of 79.24%.

A graph convolutional network is presented in Ma et al. (2021). A region of interest is represented by a node in the graph. The functional connectivities between ROIs is averged over the population and used as feature vectors. The parcellation method divides the brain into 200 regions to create a square matrix of functional connectivity with size 200. This is created for each subject using Pearson's correlations. All these matrices are averaged to calculate a mean matrix from the connectivity matrix of all the subjects.

A backbone adjacency matrix is created from the mean matrix where elements $\mathcal{A}_{i,j} = 1$ if $\overline{A}_{\mathrm{i,j}} \geq \tau$ else 0. A combination of layers (convolutions, pooling, dense) to create different networks is analysed such as as C^2PCP, CPCP, C^2P, C^2P + Dense1 and CP are used to study the effect on classification performance. Additionally, the ablation study is also performed using phenotypic attributes. The best accuracy of around 78% is achieved with a C^2P (GraphConv(20)– GraphConv(10)–GraphPool(10)) model.

The BrainGNN, a graph neural network for the analysis of fMRI image data is presented in Li et al. (2021). It consists of Region of interest (ROI)-aware Graph Convolution Layer (Ra-GConv) for embedding weights learning, ROI-top K pooling layer for reduction of dimesnions and for flatterning, a readout layer. The datasets used are Biopoint and HCP. The BrainGNN achieved an accuracy of 79.8 and 94.4 on Biopoint and HCP datasets respectively.

In Yin et al. (2022), an unsupervised autoencoder is used for learning the latent features from the fMRI data, while a supervised classifier is used for ASD classification. This is different from commonly used two staged networks in which learning takes place separately. Instead, the loss functions of the two stages are combined together so that the learning of the latent features is guided to improve classification accuracy. The unlabeled data is used to experiment with the learning of the encoder part of the autoencoder network and then labeled data to train both the networks jointly. The accuracy using 5 fold cross-validation is 87.2% while sensitivity is 89.9% & specificity is 80.3%.

2.6 Convolution Neural Networks and Transferred Learning

A transfer learning based approach is given in Liang et al. (2021). A convolutional network is used for the extraction of features. The learning of category prototypes takes place automatically. Also, the learning of CNN and prototype happen together for ASD classification using cross-entropy as loss function. The weights are initialised using weights from non-medical applications for achieving early convergence of the training algorithm. The dataset ABIDE-I preprocessed is used. The accuracy using 10-fold cross validation is 77.3%. Also, using all sites except one for training and remaining site for testing gives an inter-site acccuracy of 67.8%.

2.7 ABIDE information

In Di Martino et al. (2014), the information on Autism Brain Imaging Data Exchange (ABIDE) dataset is provided. The process of collection is explained. The ASD subjects are 539 while typical controls are 573, making a total of 1112 fMRI images.

In summary as shown in Table 1, most of the mentioned papers used ABIDE dataset. A correlation matrix of functional connectivities is created from the dataset and then an adjacency matrix is created whose elements are 0 if they are less than cutoff and 1 if greater than or equal to the cutoff value of correlation. The adjacency matrix is used for the classification task. There are two kinds of data representation using graphs - one graph per subject with nodes as regions of interest having phenotypic attributes and

Table 1: Summary of the literature review

Research paper	Dataset	Technique	Accuracy
Ingalhalikar et al. (2021)	ABIDE-I preprocessed	Traditional Machine Learning	71.35%
Kashef (2022)	ABIDE-I preprocessed	Deep Learning	80%
Wang (2021)	Custom dataset	Traditional Machine Learning	84.44%
M and Jaganathan (2021)	ABIDE preprocessed	Deep Learning	60.9% - 62.6%
Tummala (2021)	Custom Dataset	Deep Learning	NA
Fan et al. (2021)	ABIDE-I and ABIDE-II	Deep Learning	0.92% to 4.02% increase
Huang et al. (2021)	ABIDE-I	Deep Learning	76.4%
Yin et al. (2021)	ABIDE-I	Deep Learning	82.3%
Al-Hiyali et al. (2021)	ABIDE	Deep Learning	88.6%
Wang et al. (2022)	ABIDE-I	Deep Learning	75.86%
Ma et al. (2021)	BIDE	Deep Learning	78%
Li et al. (2021)	Biopoint and HCP	Deep Learning	79.8 on Biopoint, 94.4 on HCP dataset
Yin et al. (2022)	Custom dataset	Deep Learning	87.2%
Liang et al. (2021)	ABIDE-I preprocessed	Deep Learning	77.3%

edges as functional connectivities or a single graph with nodes as subjects and functional connectivities as node attributes and edges having phenotypic information.

3 Methodology

3.1 ABIDE dataset

The preprocessed neuoroimaging (FMRI) dataset known as Autism Brain Imaging Data Exchange (ABIDE) is a public open source dataset. It is available at ABIDE - I Preprocessed as part of the *Preprocessed Connectome Project (PCP)*. The dataset is already available as pre-processed using Configurable Pipeline for the Analysis of Connectomes C-PAC pipeline. The dataset is available as a set of subject files along with a csv file having phenotypic attributes of all the subjects along with additional information. In total, there are 1102 files aggregated by the collaboration of 16 international imaging sites. It contains neuroimaging data of 539 individuals with ASD and 573 typical controls. The phenotypic file (Phenotypic_V1_0b_preprocessed1) has information about the site (*SITE_ID*), individuals (*SUB_ID, AGE_AT_SCAN, SEX, HANDEDNESS_CATEGORY, HANDEDNESS_SCORES*), category (*DX_GROUP*), file identifier (*FILE_ID*) along with other information.

A sample of time series is illustrated in fig Figure 2 for the first 10 regions of the brain (CC200 parcellation (Caltech)). The data is a blood-oxygen-level-dependent (BOLD) signal for each small volume (cubic) known as a voxel. The BOLD values of voxels within

a region are averaged to get a representative value for a region for each of 200 brain regions.

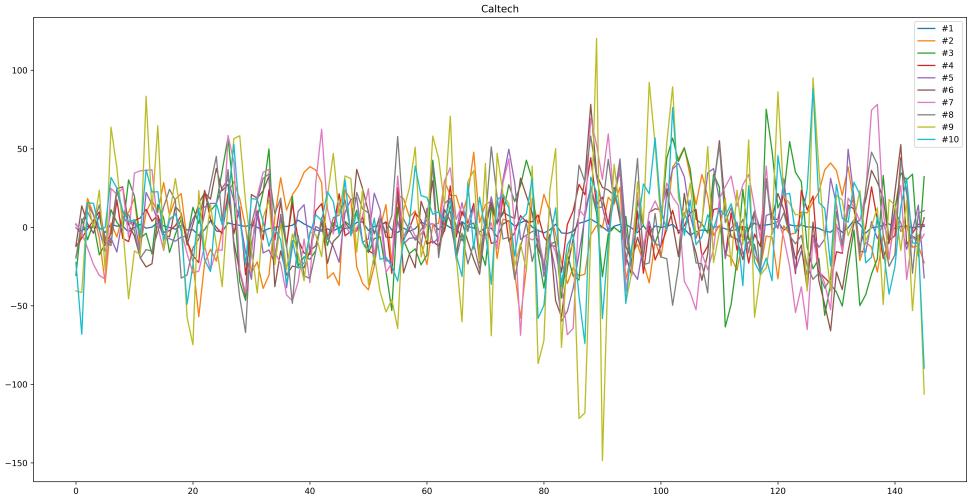


Figure 2: A sample time series containing 146 datapoints for the first 10 regions of the brain

3.1.1 Subject file structure

For each subject, the whole brain region is divided into 200 regions using Craddock 200 (CC200) functional parcellation method. Each subject file contains a time series containing 146 data points from each of the 200 brain regions arranged as 200 columns for that subject. The subject files are available in tab delimited format. The column names are given in the format #<column-position> for e.g. #1, #2 , .. #200. For each of the brain regions, a time series consisting of 146 data points is taken which are the values within a column. The values are numerical (positive and negative) in nature.

3.1.2 Phenotypic file structure

The information about the fields in the ABIDE phenotypic file is summarised in the Table 2. Following the HIPAA guidelines and 1000 Functional Connectomes Project / INDI protocols, the datasets are anonymous and are free from any protected health information.

Table 2: Description of fields

COLUMN LABEL	DATATYPE	DESCRIPTION
SUB_ID	Numeric	ABIDE unique ID number
DX_GROUP	Numeric	Diagnostic Group
DSM_IV_TR	Numeric	DSM_IV_TR Diagnostic Category
AGE_AT_SCAN	Numeric	Age at time of scan in years
SEX	Numeric	Subject Gender
HANDEDNESS_CATEGORY	String	Subject Handedness Category
HANDEDNESS_SCORES	Numeric	Subject Handedness Scores
FIQ	Numeric	FIQ Standard Score
VIQ	Numeric	VIQ Standard Score

Table 2 – continued from previous page

COLUMN LABEL	DATATYPE	DESCRIPTION
PIQ	Numeric	PIQ Standard Score
FIQ_TEST_TYPE	String	IQ Test Administered for full scale IQ
VIQ_TEST_TYPE	String	IQ Test Administered for verbal IQ
PIQ_TEST_TYPE	String	IQ Test Administered for performance IQ
ADI_R_SOCIAL_TOTAL_A	Numeric	Reciprocal Social Interaction Subscore (A) Total for Autism Diagnostic Interview-Revised
ADI_R_VERBAL_TOTAL_BV	Numeric	Abnormalities in Communication Subscore (A) Total for Autism Diagnostic Interview-Revised
ADI_RRB_TOTAL_C	Numeric	Restrictive, Repetitive, and Stereotyped Patterns of Behaviour Subscore (C) Total for Autism Diagnostic Interview-Revised
ADI_R_ONSET_TOTAL_D	Numeric	Abnormality of Development Evident at or Before 36 Months Subscore (D) Total for Autism Diagnostic Interview-Revised
ADI_R_RSRCH_RELIABLE	Numeric	Was ADI scored and administered by research reliable personnel?
ADOS_MODULE	Numeric	Autism Diagnostic Observation Schedule Module
ADOS_TOTAL	Numeric	Classic Total ADOS Score (Communication subscore + Social Interaction subscore)
ADOS_COMM	Numeric	Communication Total Subscore of the Classic ADOS
ADOS_SOCIAL	Numeric	Social Total Subscore of the Classic ADOS
ADOS_STEREO_BEHAV	Numeric	Stereotyped Behaviors and Restricted Interests Total Subscore of the Classic ADOS
ADOS_RSRCH_RELIABLE	Numeric	Was ADOS scored and administered by research reliable personnel?
ADOS_GOTHAM_SOC_AFFECT	Numeric	Social Affect Total Subscore for Gotham Algorithm of the ADOS
ADOS_GOTHAM_RRB	Numeric	Restricted and Repetitive Behaviors Total Subscore for Gotham Algorithm of the ADOS
ADOS_GOTHAM_TOTAL	Numeric	Social Affect Total + Restricted and Repetitive Behaviors Total
ADOS_GOTHAM_SEVERITY	Numeric	Individually Calibrated Severity Score for Gotham Algorithm of the ADOS
SRS_VERSION	Numeric	Social Responsiveness Scale Version
SRS_RAW_TOTAL	Numeric	Total Raw Score the Social Responsiveness Scale
SRS_AWARENESS	Numeric	Social Responsiveness Scale Social Awareness Subscore Raw Total
SRS_COGNITION	Numeric	Social Responsiveness Scale Social Cognition Subscore Raw Total
SRS_COMMUNICATION	Numeric	Social Responsiveness Scale Social Communication Subscore Raw Total
SRS_MOTIVATION	Numeric	Social Responsiveness Scale Social Motivation Subscore Raw Total
SRS_MANNERISMS	Numeric	Social Responsiveness Scale Autistic Mannerisms Subscore Raw Total
SCQ_TOTAL	Numeric	Social Communication Questionnaire Total
AQ_TOTAL	Numeric	Total Raw Score of the Autism Quotient
COMORBIDITY	String	Any other comorbidities?
CURRENT_MED_STATUS	Numeric	Currently Taking Medications?
MEDICATION_NAME	String	Active ingredient of any currrent psychoactive medications
OFF_STIMULANTS_AT_SCAN	Numeric	Off stimulus 24 hours prior to scan?
VINELAND_RECEPTIVE_V_SCALED	Numeric	Vineland Adaptive Behavior Scales Receptive Language V Scaled Score
VINELAND_EXPRESSIVE_V_SCALED	Numeric	Vineland Adaptive Behavior Scales Expressive Language V Scaled Score
VINELAND_WRITTEN_V_SCALED	Numeric	Vineland Adaptive Behavior Scales Written Language V Scaled Score
VINELAND_COMMUNICATION_STANDARD	Numeric	Vineland Adaptive Behavior Scales Communication Standard Score
VINELAND_PERSONAL_V_SCALED	Numeric	Vineland Adaptive Behavior Scales Personal Daily Living Skills V Scaled Score
VINELAND_DOMESTIC_V_SCALED	Numeric	Vineland Adaptive Behavior Scales Domestic Daily Living Skills V Scaled Score
VINELAND_COMMUNITY_V_SCALED	Numeric	Vineland Adaptive Behavior Scales Community Daily Living Skills V Scaled Score
VINELAND_DAILYLVNG_STANDARD	Numeric	Vineland Adaptive Behavior Scales Daily Living Skills Standard Score
VINELAND_INTERPERSONAL_V_SCALED	Numeric	Vineland Adaptive Behavior Scales Interpersonal Relationships V Scales Score
VINELAND_PLAY_V_SCALED	Numeric	Vineland Adaptive Behavior Scales Play and Leisure Time V Scales Score
VINELAND_COPING_V_SCALED	Numeric	Vineland Adaptive Behavior Scales Coping Skills V Scales Score
VINELAND_SOCIAL_STANDARD	Numeric	Vineland Adaptive Behavior Scales Socilaization Standard Score
VINELAND_SUM_SCORES	Numeric	Sum of Vineland Standard Scores (Communication + Daily Living Skills + Socialization)

Table 2 – continued from previous page

COLUMN LABEL	DATATYPE	DESCRIPTION
VINELAND_ABC_STANDARD	Numeric	Vineland Adaptive Behavior Coposite Standard Score
VINELAND_INFORMANT	Numeric	Vineland Adaptive Behavior Scales Informant
WISC_IV_VCI	Numeric	WISC-IV Verbal Comprehension Index
WISC_IV_PRI	Numeric	WISC-IV Perceptual Reasoning Index
WISC_IV_WMI	Numeric	WISC-IV Working Memory Index
WISC_IV_PSI	Numeric	WISC-IV Processing Speed Index
WISC_IV_SIM_SCALED	Numeric	WISC-IV Sim Scales
WISC_IV_VOCAB_SCALED	Numeric	WISC-IV Vocabulary Scaled
WISC_IV_INFO_SCALED	Numeric	WISC-IV Information Scaled
WISC_IV_BLK_DSN_SCALED	Numeric	WISC-IV Block Design Scaled
WISC_IV_PIC_CON_SCALED	Numeric	WISC-IV Picture Concepts Scaled
WISC_IV_MATRIX_SCALED	Numeric	WISC-IV Matrix Reasoning Scaled
WISC_IV_DIGIT_SPAN_SCALED	Numeric	WISC-IV Digit Span Scaled
WISC_IV_LET_NUM_SCALED	Numeric	WISC-IV Letter-Number Sequencing Scaled
WISC_IV_CODING_SCALED	Numeric	WISC-IV Coding Scaled
WISC_IV_SYM_SCALED	Numeric	WISC-IV Symbol Search Scaled
EYE_STATUS_AT_SCAN	Numeric	Eye Status During Rest Scan
AGE_AT_MPRAGE	Numeric	Age at Anatomical Scan in years
BMI	Numeric	Body Mass Index

The phenotypic fields used as node attributes in the current paper are *SITE_ID, AGE_AT_SCAN, SEX, HANDEDNESS_CATEGORY*, while other time-related node attributes such as *mean, variance, skewness and kurtosis* are generated from time-series data from the subject file.

3.2 Functional connectivity

The pairwise correlation is calculated to find the function connectivity between regions of the brain known as regions of interest (200 ROIs). The Pearson coefficient is used to calculate the correlation value whose equation is shown below:

$$\rho_{i,j} = \frac{\sum_{t1}^{tn}(x_{i,t} - \overline{x_i})(x_{j,t} - \overline{x_j})}{\sqrt{\sum_{t1}^{tn}(x_{i,t} - \overline{x_i})}\sqrt{\sum_{t1}^{tn}(x_{j,t} - \overline{x_j})}}$$

A heatmap of the correlations between the first 30 regions of the brain is given in Figure 3

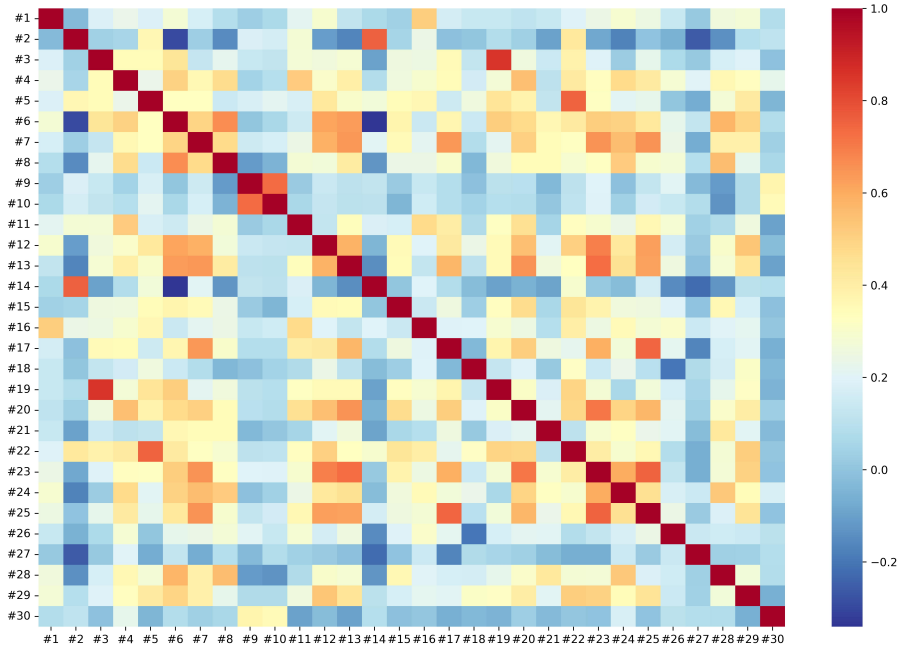


Figure 3: A sample heatmap of correlations between first 30 regions of the brain

3.3 Additional Preprocessing

For each subject, the raw time-series file containing signals from 200 brain regions as per Craddock parcellation file is read and a correlation matrix is created whose elements represent the correlation between brain regions. The correlation matrix is converted into an adjacency matrix by replacing values with 0 which are less than a cutoff value and with 1 which is greater than or equal to the cutoff. Then, the adjaceny matrix is converted into a sparse matrix/edge_index by keeping only non-zero elements. The sparse matrix is converted into a directed matrix by adding a connection (j,i) for each (i,j) in the sparse matrix as required for model training. Also, for each subject, there are two types of node features: time series related and phenotype related. The time series related node features such as mean, variance, skew, kurtosis are calculated from the time series data from the file, whereas the phenotypic attributes such as *SITE_ID", "AGE_AT_SCAN", "SEX", "HANDEDNESS_CATEGORY"* are read from the phenotype file provided as part of the ABIDE dataset. The classification label DX_GROUP is also read from the same phenotypic file. The node features and edge_index along with classification label are used to generate an object (torch_geometric.data) and saved on the disk for later use during model training.

4 Design Specification

The proposed model is shown in Figure 4. It consists of a graph autencoder followed by a graph neural network.

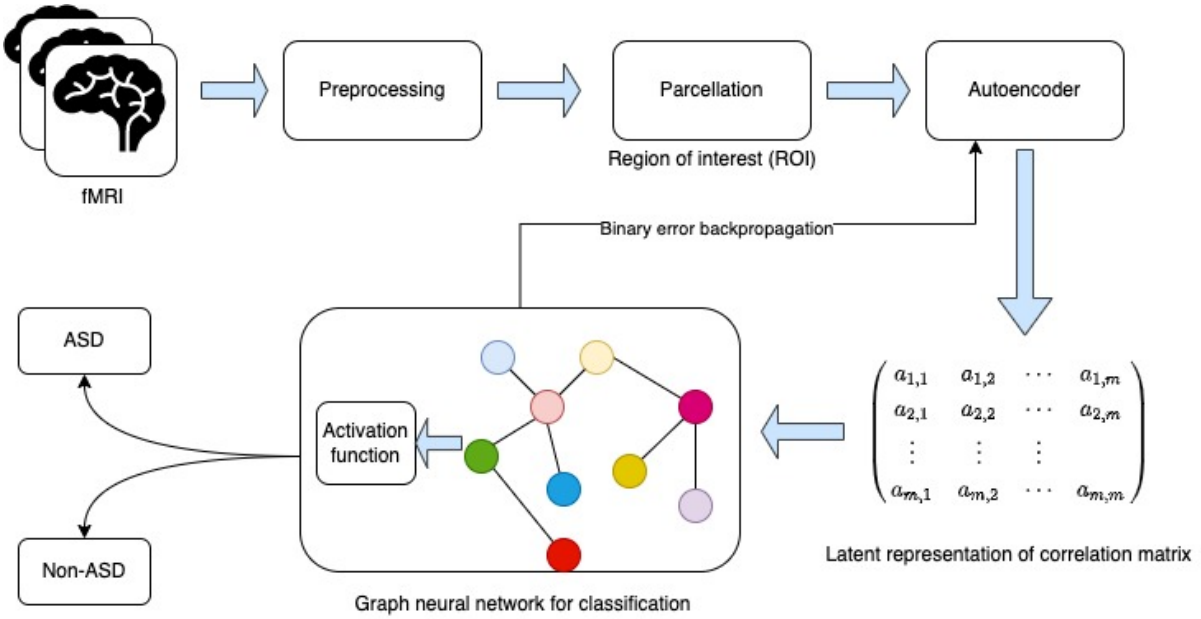


Figure 4: Flow diagram of the proposed novel model

4.1 Graph Autoencoder

The autoencoder network is a neutral network that consists of an encoder, hidden layer and a decoder. The encoder converts the input data from one space into a latent space. Depending on the hidden layer and input dimension, the encoder converts the input to a lower dimension or higher dimension representation.

The decoder part regenerates the original input from this lower or higher dimensional representation. The corresponding version which is used for graphical data is known as graph autencoder. An autoencoder diagram is shown in Figure 5. A graph autoencoder is used to learn the latent representation of the data while discarding the useless redundant information.

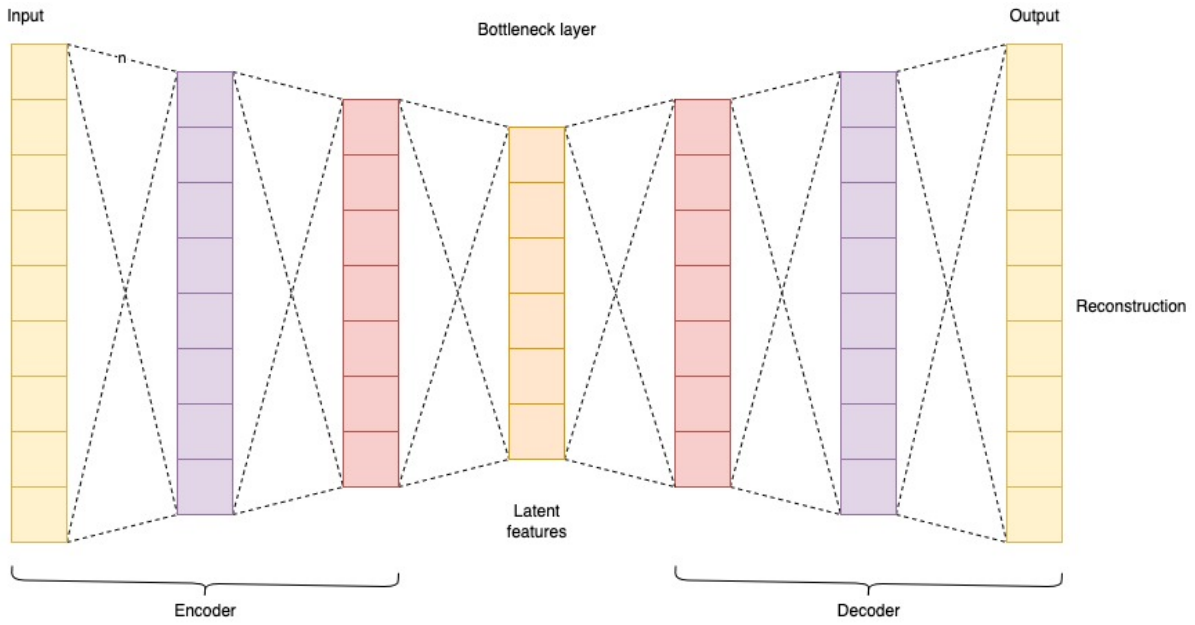


Figure 5: Diagram of an autoencoder

4.2 Graph Neural Network

The fMRI data can be represented efficiently by a graph. A graph is represented by nodes and edges. In the current paper, the regions of interest are represented as nodes with phenotypic attributes as node attributes, while edges represent functional connectivities between regions of interest for each subject. The graph neural network is used to classify the graph as either ASD or non-ASD for each subject.

Figure 6 shows the representation of brain functional connectivities as correlation matrix and graph.

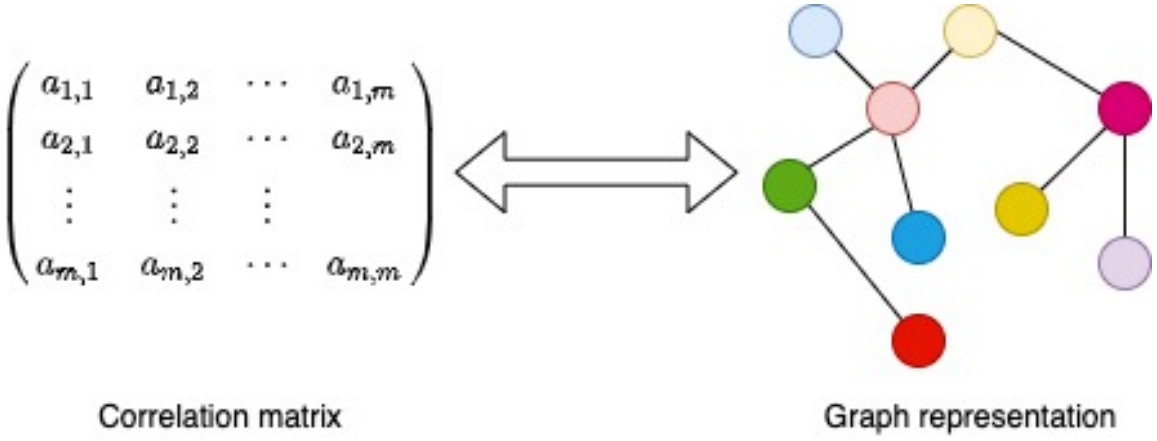


Figure 6: Representation of brain functional connectivities as correlation/adjacency matrix and graph

The loss function of both the graph autoencoder and graph neural network is combined and they are trained jointly to direct the graph autoencoder learning towards enhancing

the classification accuracy by the graph neural network.

5 Implementation

5.1 Preprocessing

The Pytorch library is used for the transformation and modelling the solution. First, for each of the fmri files, it is loaded into a numpy array (ndarray). Then, pairwise correlation is calculated using ndarray *corrcoef* function in numpy for each pair of the possible combination of columns, such as column 1 & 2, column 1 & 3 etc. till column 199 & 200 to create a correlation matrix of size 200 x 200 corresponding to 200 regions of interests in the brain. After that, the correlation matrix is converted into an adjacency matrix by using the *where* function in numpy to convert the values in the correlation matrix by using the formula: 1 if cij >= cutoff_value and 0 for cij < cutoff, where cutoff_value is used as 0.1. Then, the correlation matrix is converted into sparse matix. The sparse matrix is then converted into a directed matrix by adding the corresponding connections i.e. (j, i) for every (i, j) in the sparse matrix.

Then, for each sparse matrix (graph), the DX_GROUP column from the phenotype file is used to get the class of the graph (1 for control group and 2 for Autistic group) and the converting these classes to zero based class i.e. (0 for control and 1 for Autism). After that , the phenotypic features such as *SITE_ID, AGE_AT_SCAN, SEX, HANDEDNESS_CATEGORY* and time related features such as *mean, variance, skewness* and *kurtosis* are calculated and used as node features.

Then, the node features, label and adjacency_matrix are used to construct an object known as Data and saved on the disk for later use during training.

5.2 Modelling

The model consists of an autoencoder and a graph neural network for the task of classification of the graph for Autism detection.

The autoencoder consists of encoder and decoder part. The encoder network encodes the input into another dimension whereas decoder converts it to original dimension.

The graph neural network consists of multiple layers of graph convolutions and activation functions. There are two GCN networks used in the current paper. The layer configurations of the two GCN's are shown in the table Table 3 and 4 below:

The model is implemented using modules *torch_geometric* & *torch* which are available in Pytorch. For the graph autoencoder, the phenotypic attributes are encoded to 21 features from 7 features and then decoded back to 7 features. A new class *GCNEncoder* is coded which is passed as an encoder to the GAE. The encoder uses 2 graph convolutions with a relu activation function as shown in Figure 7.

Table 3: Classification model I

Layer	Name	Parameters
0	GAE	Graph Autoencoder
1	GCNConv	Graph Convolution layer
2	ReLU	Activation
3	GCNConv	Graph Convolution layer
4	Global Mean Pooling	Mean Pooling layer
5	Dropout layer	To prevent over-training
6	Linear Layer	For flattening

Table 4: Classification model II

Layer	Name	Parameters
0	GAE	Graph Autoencoder
1	GCNConv	Graph Convolution layer
2	ReLU	Activation Layer
3	GCNConv	Graph Convolution layer
4	ReLU	Activation Layer
5	GCNConv	Graph Convolution layer
6	Global Mean Pooling	Mean Pooling layer
7	Dropout layer	To prevent over-training
8	Linear Layer	For flattening

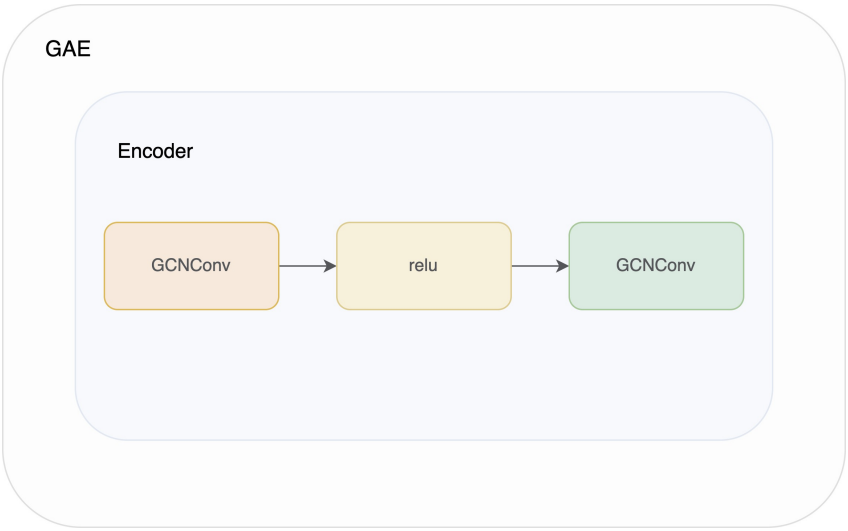
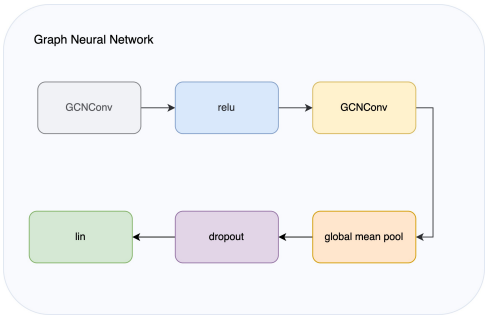
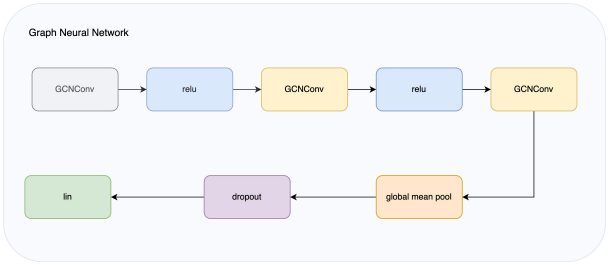


Figure 7: Encoder of the graph autoencoder



(a) Graph Neural Network I



(b) Graph Neural Network II with additional GCNConv & relu layer

Figure 8: Graph Neural Networks

The encoder ouptput is fed to a graph neural network. The graph neural network consists of 2 graph convolutions *GCNConv* with a relu activation in model I and 3 graph convolutions *GCNConv* with two relu activation in model II . For pooling, global_mean_pool is used with a dropout layer to prevent over-training and then finally, a linear layer for flattening in both the models.

6 Evaluation

For the evaluation of the model, the ABIDE-I preprocessed dataset consisting of 1102 subjects after additional preprocessing is divided into 2 sets: 80% training set and 20% testing set. The training set is shuffled before training. A batch size of 256 is used for the training. The results using the two models are discussed below:

6.1 Experiment / Case Study 1

One model as shown in Figure 8a is trained for 500 epochs with a learning rate of 0.0001. The graph of loss vs. epoch is shown in Figure 9 for graph auto-encoder used for latent representation learning and in Figure 10 for graph neural network used for classification.

The two losses are combined during training to direct the latent representation learning towards enhancing ASD classification accuracy.

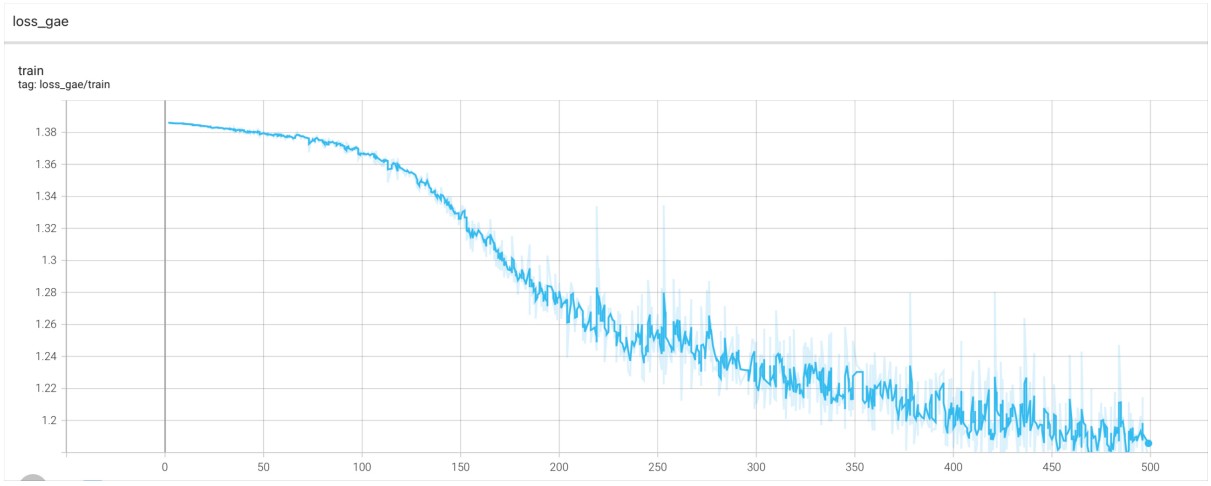


Figure 9: Loss vs. epoch for the graph autoencoder

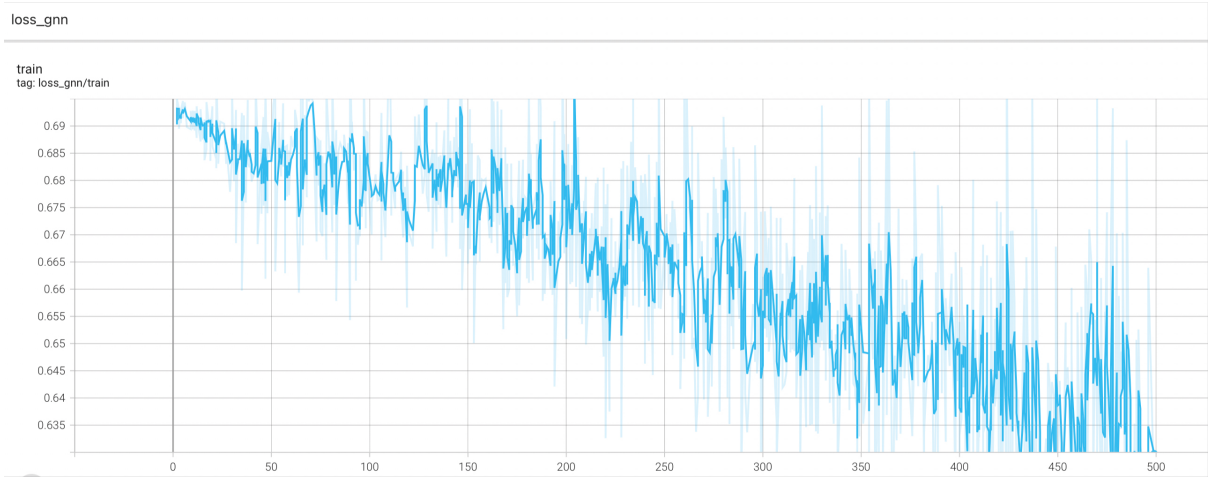


Figure 10: Loss vs. epoch for the classification graph neural network

6.2 Experiment / Case Study 2

The second model as shown in Figure 8b is also trained for 500 epochs. A different learning rate of 0.001 is used for training. The graph of loss vs. epoch is provided in Figure11 and Figure 12 for the graph autoencoder for latent representation and gnn for classification respectively.

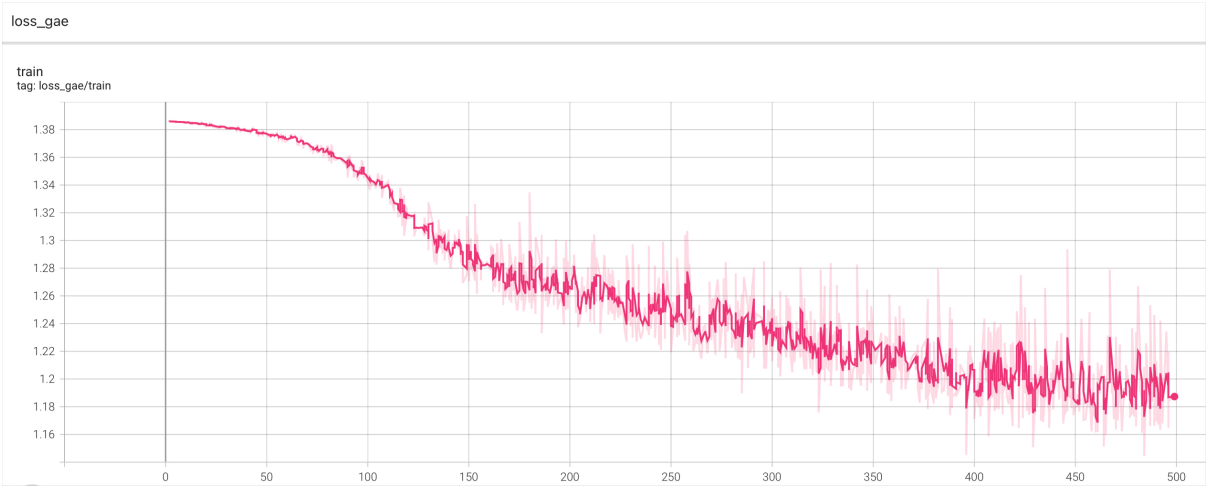


Figure 11: Loss vs. epoch for the graph autoencoder

Figure 12: Loss vs. epoch for the classification graph neural network

The graphs of loss vs epoch for experiment case I and case II is shown together for comparison in Figure 13 for graph autoencoder whereas for graph neural network, it is shown in 14.

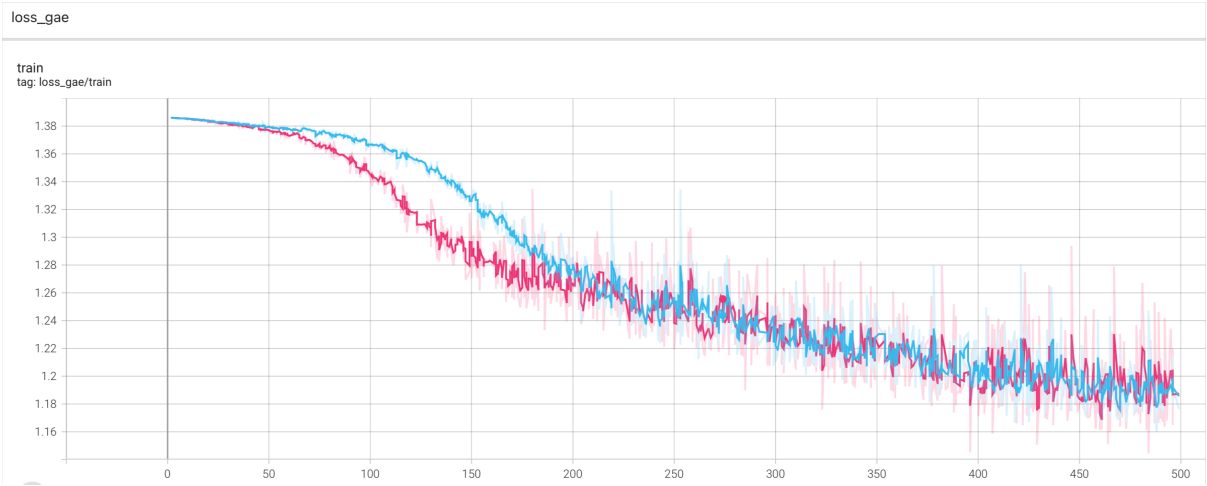


Figure 13: Loss vs. epoch comaprison between model I and II for the graph autoencoder

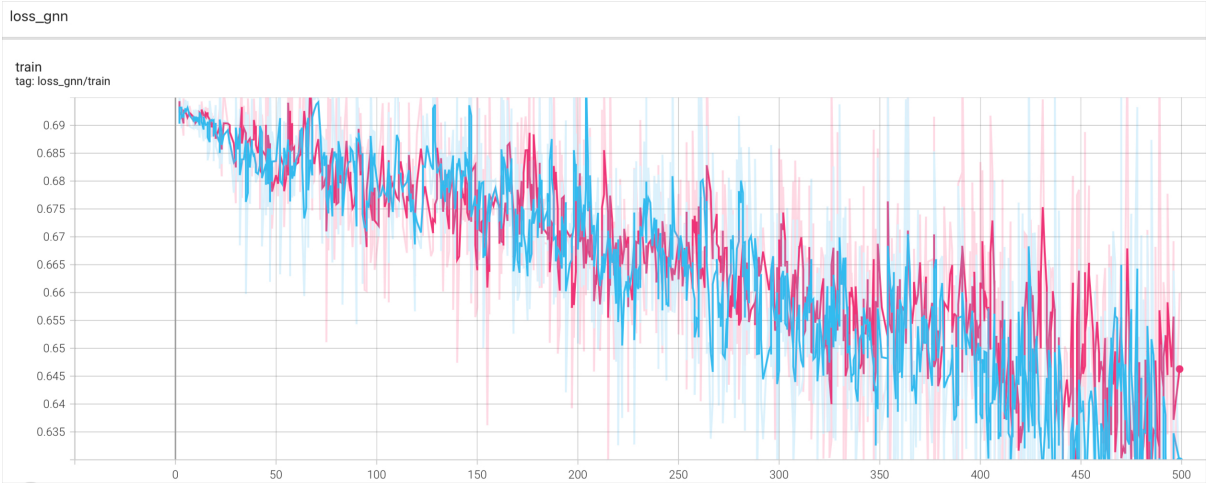


Figure 14: Loss vs. epoch for the classification graph neural network between model I and II

6.3 Discussion

As shown in Figure 9, the loss decreases as epoch progresses and maximum learning take place from epoch 100 to 250 as evident by steep decrease in loss. However, after epoch 250 onwards, the rate of decrease in loss reduces and graph almost becomes asymptotic suggesting learning by the model has almost stopped.

For the GNN as shown in Figure 8a, there is a downward trent but there is more variance. The value keeps on fluctuating around a value. The maximum accuracy achieved is around 55%.

A similar trend can be seen in Figure 11 but maximum learning take place between epoch 75 and 175. For the GNN, the loss value is decreasing but the variance observed in loss is high.

As can be seen from transposition of graphs of model I and model II for graph autoenoder in Figure 13 and graph classification neural network in Figure 14, the values decrease faster in model II as compared to model I for GAE but remains almost similar for GNN. However, the final accuracy observed is still around 55%.

The above discussion suggests that the phenotypic attributes used in the current paper and a cutoff value of 0.1 for functional connectivity are not enough to classify Autism successfully. The additional phenotypic attributes and variation of cutoff values for functional connnectivities might help in making the differentiation between Autistic and non-Autistic subjects which might help in enhancing ASD classification accuracy.

7 Conclusion and Future Work

In this research paper, a novel approach for enhancing the accuracy of the Autism classification by using phenotypic attributes and functional connectivities between brain regions as features is proposed and implemented in Pytorch. A graph autoencoder is trained for learning the latent representation and graph neural network for the classification on ABIDE-I preprocessed dataset. The training is performed jointly to optimize latent representation learning for enhancing classification accuracy. This is in contrast to other methods in which the training is performed separately.

For each subject , the nodes in the graph represents phenotypic attributes while edges represent functional connectivities between brain regions. The whole graph for each subject is classified as Autistic or non-Autistic.

However, the test accuracy has been around 55% which is not ideal for medical applications. A further investigation with other graph convolution layers such as ChebConv, SAGEConv, GraphConv, GATConv, GATv2Conv etc. and other pooling layers such as global_max_pool, SAGPooling etc.) might help in enhancing the accuracy.

The ABIDE-II dataset which is currently available as unprocessed dataset on the web at ABIDE-II can also be used in the future either by pre-processing it locally (out of scope for current research) or when its pre-processed version becomes available (similar to ABIDE-I) in the future.

References

Al-Hiyali, M. I., Yahya, N., Faye, I. and Al-Ezzi, A. (2021). Classification of asd subtypes based on coherence features of bold resting-state fmri signals, *2021 International Conference on Intelligent Cybernetics Technology Applications (ICICyTA)*, pp. 17–22.

Bauer, A. J. and Just, M. A. (2015). Monitoring the growth of the neural representations of new animal concepts, *Human Brain Mapping* **36**(8): 3213–3226.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.22842*

Di Martino, A., Yan, C.-G., Li, Q., Denio, E., Castellanos, F. X., Alaerts, K., Anderson, J. S., Assaf, M., Bookheimer, S. Y., Dapretto, M., Deen, B., Delmonte, S., Dinstein, I., Ertl-Wagner, B., Fair, D. A., Gallagher, L., Kennedy, D. P., Keown, C. L., Keysers, C., Lainhart, J. E., Lord, C., Luna, B., Menon, V., Minshew, N. J., Monk, C. S., Mueller, S., Müller, R.-A., Nebel, M. B., Nigg, J. T., O'Hearn, K., Pelphrey, K. A., Peltier, S. J., Rudie, J. D., Sunaert, S., Thioux, M., Tyszka, J. M., Uddin, L. Q., Verhoeven, J. S., Wenderoth, N., Wiggins, J. L., Mostofsky, S. H. and Milham, M. P. (2014). The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism, *Molecular Psychiatry* **19**(6): 659–667.
URL: *https://doi.org/10.1038/mp.2013.78*

Fan, Z., Su, J., Gao, K., Hu, D. and Zeng, L.-L. (2021). A federated deep learning framework for 3d brain mri images, *2021 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–6.

Huang, Z.-A., Zhu, Z., Yau, C. H. and Tan, K. C. (2021). Identifying autism spectrum disorder from resting-state fmri using deep belief network, *IEEE Transactions on Neural Networks and Learning Systems* **32**(7): 2847–2861.

Ingalhalikar, M., Shinde, S., Karmarkar, A., Rajan, A., Rangaprakash, D. and Deshpande, G. (2021). Functional connectivity-based prediction of autism on site harmonized abide dataset, *IEEE Transactions on Biomedical Engineering* **68**(12): 3628–3637.

Kashef, R. (2022). Ecnn: Enhanced convolutional neural network for efficient diagnosis of autism spectrum disorder, *Cognitive Systems Research* **71**: 41–49.
URL: *https://www.sciencedirect.com/science/article/pii/S1389041721000759*

Kassam, K. S., Markey, A. R., Cherkassky, V. L., Loewenstein, G. and Just, M. A. (2013). Identifying emotions on the basis of neural activation, *PLOS ONE* **8**(6): 1–12.
URL: *https://doi.org/10.1371/journal.pone.0066032*

Li, X., Zhou, Y., Dvornek, N., Zhang, M., Gao, S., Zhuang, J., Scheinost, D., Staib, L. H., Ventola, P. and Duncan, J. S. (2021). Braingnn: Interpretable brain graph neural network for fmri analysis, *Medical Image Analysis* **74**: 102233.
URL: *https://www.sciencedirect.com/science/article/pii/S1361841521002784*

Liang, Y., Liu, B. and Zhang, H. (2021). A convolutional neural network combined with prototype learning framework for brain functional network classification of autism spectrum disorder, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **29**: 2193–2202.

M, K. and Jaganathan, S. (2021). Graph convolutional model to diagnose autism spectrum disorder using rs-fmri data, *2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP)*, pp. 1–5.

Ma, Y., Yan, D., Long, C., Rangaprakash, D. and Deshpande, G. (2021). Predicting autism spectrum disorder from brain imaging data by graph convolutional network, *2021 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–8.

Mostafa, S., Yin, W. and Wu, F.-X. (2020). Autoencoder based methods for diagnosis of autism spectrum disorder, *in* I. Măndoiu, T. M. Murali, G. Narasimhan, S. Rajasekaran, P. Skums and A. Zelikovsky (eds), *Computational Advances in Bio and Medical Sciences*, Springer International Publishing, Cham, pp. 39–51.

O'toole, A. J., Jiang, F., Abdi, H. and Haxby, J. V. (2005). Partially distributed representations of objects and faces in ventral temporal cortex, *J. Cognitive Neuroscience* **17**(4): 580–590.
URL: *https://doi.org/10.1162/0898929053467550*

Tummala, S. (2021). Deep learning framework using siamese neural network for diagnosis of autism from brain magnetic resonance imaging, *2021 6th International Conference for Convergence in Technology (I2CT)*, pp. 1–5.

Wang, C. (2021). Indentification of autism spectrum disorder based on an improved convolutional neural networks, *2021 3rd International Conference on Machine Learning, Big Data and Business Intelligence (MLBDBI)*, pp. 235–239.

Wang, Y., Liu, J., Xiang, Y., Wang, J., Chen, Q. and Chong, J. (2022). Mage: Automatic diagnosis of autism spectrum disorders using multi-atlas graph convolutional networks and ensemble learning, *Neurocomputing* **469**: 346–353.
URL: *https://www.sciencedirect.com/science/article/pii/S0925231221011048*

Yahata, N., Kasai, K. and Kawato, M. (2017). Computational neuroscience approach to biomarkers and treatments for mental disorders., *Psychiatry Clin Neurosci* **71**(4): 215–237.

Yin, W., Li, L. and Wu, F.-X. (2021). A graph attention neural network for diagnosing asd with fmri data, *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pp. 1131–1136.

Yin, W., Li, L. and Wu, F.-X. (2022). A semi-supervised autoencoder for autism disease diagnosis, *Neurocomputing* **483**: 140–147.
URL: *https://www.sciencedirect.com/science/article/pii/S0925231222001606*