T2R2 東京科学大学 リサーチリポジトリ
Science Tokyo Research Repository

論文 / 著書情報
Article / Book Information

題目(和文)	
Title(English)	Improving automated detection of autism spectrum disorder with deep learning based on resting-state and task-based fMRI data
著者(和文)	RAKHIMBERDINAZarina
Author(English)	Zarina Rakhimberdina
出典(和文)	学位:博士(学術), 学位授与機関:東京工業大学, 報告番号:甲第12606号, 授与年月日:2023年9月22日, 学位の種別:課程博士, 審査員:村田 剛志,秋山 泰,岡﨑 直観,DEFAGO XAVIER,石田 貴士
Citation(English)	Degree:Doctor (Academic), Conferring organization: Tokyo Institute of Technology, Report number:甲第12606号, Conferred date:2023/9/22, Degree Type:Course doctor, Examiner:,,,,
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

Powered by T2R2 (Science Tokyo Research Repository)

TOKYO INSTITUTE OF TECHNOLOGY
DOCTORAL THESIS

Improving automated detection of autism spectrum disorder with deep learning based on resting-state and task-based fMRI data

---

Author:
Zarina RAKHIMBERDINA

Supervisor:
Tsuyoshi MURATA

A thesis submitted in fulfillment of the requirements
for the degree of Doctor of Philosophy

in the

School of Science and Computing
Department of Computer Science

Abstract

Deep learning is finding increasing applications in the automated medical detection of neurodevelopmental disorders, such as autism spectrum disorder (ASD). Automated tools aim to overcome the limitations of traditional screening methods, which are based on clinicians' subjective observations and are also costly and time-consuming. With the rapid development of brain fMRI technology, more research is focused on using it to uncover the intrinsic biological markers of ASD. This research aims to improve classification for ASD, both in terms of the deep learning framework and the choice of fMRI imaging data. Specifically, we will focus on two fMRI acquisition paradigms: resting-state and visual task-based approaches. First, the author proposes an improved ASD prediction framework based on a graph neural network and analyzes the contribution of the graph structure. The proposed model is a competitive alternative to the current best deep learning methods for ASD classification in the resting-state paradigm. Visual task-based fMRI is another method used to study brain activity related to visual stimuli, including face recognition and gaze perception relevant to ASD research. By studying ASD with visual task-based fMRI, researchers can gain insight into how the brains of individuals with ASD respond to visual stimuli. Thus, the author further generalizes the proposed graph neural network to visual task-based fMRI. Finally, the author introduces a novel multi-modal framework to incorporate visual stimuli information that may be useful in detecting and treating of ASD.

Acknowledgements

I would like to express my gratitude to Professor Tsuyoshi Murata, my academic supervisor, for his guidance and support throughout my entire graduate studies. Also, I would like to thank my PhD thesis committee members, Prof. Yutaka Akiyama, Prof. Xavier Defago, Prof. Naoaki Okazaki, and Prof. Takashi Ishida, for their dedicated time, consideration, and constructive criticism. Moreover, I wish to express my appreciation to my former and current laboratory colleagues, Xin Liu, Sunil Kumar Maurya, Jodelet Quentin, and Nuttapong Chairatanakul, for their support and enriching research discussions. Their valuable feedback contributed to the enhancement of the quality of my research findings. Lastly, a heartfelt thank you goes to my mother, whose encouragement has been a guiding light throughout my journey in pursuing and completing my graduate studies.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Over the past decade, machine learning methods have been used for automated medical detection of neurodevelopmental disorders, such as autism spectrum disorder (ASD). Automated medical detection aims to overcome the limitations of traditional observational methods, which are based on clinicians' subjective observations and are often times costly and time-consuming. With the rapid development of brain imaging technology, more research is focused on uncovering the intrinsic biological markers of ASD.

This research aims to identify the optimal practices for improving the classification of ASD, both in terms of the data choice and deep learning framework. This chapter will provide an introduction to the thesis by presenting the background of ASD-related research, followed by the introduction of the research problem, the research aim, and objectives.

1.1 Autism Spectrum Disorder (ASD)

Autism spectrum disorder (ASD) refers to a neurodevelopmental disorder that affects peoples' communication, learning ability, behavior, and social interaction. The term "spectrum" refers to a range of levels of disability in functioning that can occur in people with ASD. Both children and adults can be affected by a disease across all socioeconomic and ethnic groups. People with ASD may have repetitious behavioral patterns or narrow interests, which are usually present from childhood and affect everyday functioning. However, not everyone diagnosed with ASD may express those symptoms.

A traditional diagnosis of ASD is based on an assessment of the intellectual disability and language functioning of a patient. A complete evaluation requires a multidisciplinary team, including a neurologist, psychologist, speech therapist, psychiatrist, and other experts who diagnose and treat the disorder. Consequently, such observational tests are time-consuming, costly, and moreover subjective to the experts' evaluation.

ASD-related traditional behavioral markers that have been found include diminished social interaction and attention, atypical visual attention, and non-smooth visual object tracking. Furthermore, children with ASD may exhibit atypical social behaviors such as decreased attention to social scenes, decreased frequency of gaze to faces (Belen et al., 2020). Finding the biomarkers associated with ASD using modern computational methods is particularly valuable in understanding the roots of the neurodevelopmental disorder, which can lead to earlier diagnosis and more targeted treatment (Li et al., 2018). More recent research focuses on the role of visual information processing in patients with ASD by utilizing computer vision methods, such as eye tracking (Greene et al., 2019) and the analysis of behavior to visual stimuli (Belen et al., 2020).

1.2 Brain Imaging

The application of machine learning methods for brain imaging analysis is targeted to develop new ASD detection methods without requiring expert knowledge. Among different brain functional activity measuring techniques, functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) are widely used nowadays (Straaten and Stam, 2013).

Compared to other imaging methods, fMRI is considered to be the most advanced tool of clinical diagnosis due to the high spatial resolution of imaging data (Bullmore and Sporns, 2012). fMRI is used for identifying normal and abnormal brain functioning patterns by measuring so-called blood-oxygen-level-dependent (BOLD) signals. The method is based on the fact that neural activity and blood flow are correlated: the activation of the brain

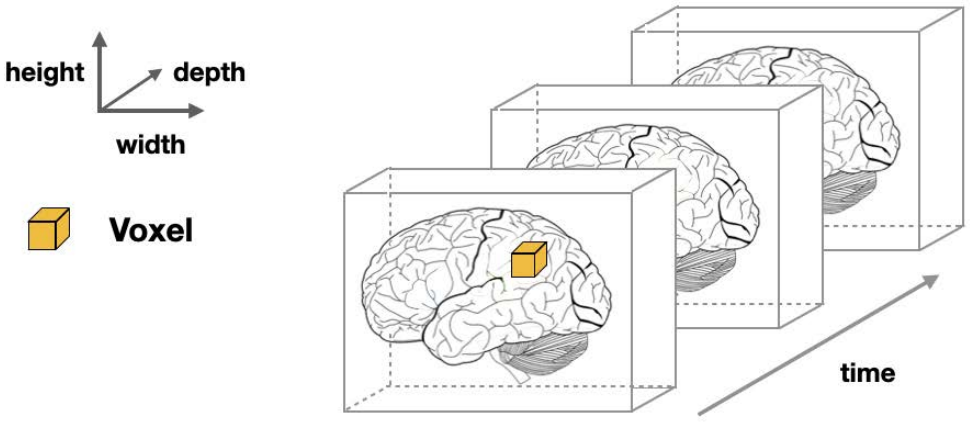


FIGURE 1.1: Graphical illustration of voxel in fMRI data.

region requires an increased supply of blood oxygen to that region (Ogawa et al., 1990). This way, fMRI allows capturing the different patterns of the brain's activations under varying physical factors, such as body movements, presentation of sensor stimuli, or during pathological states (Bassett and Bullmore, 2009; Bullmore and Sporns, 2012).

An fMRI scan can be viewed as a 4-dimensional matrix with three spatial dimensions and one temporal dimension, as illustrated in Figure 1.1. The smallest unit of fMRI that stores a signal value is called a voxel. The amount of whole-brain voxel-wise fMRI signals could be immense: a three-dimensional snapshot in time can contain millions of such voxels per subject. Hence, fMRI is known for its high spacial resolution, which provides more detailed imaging, but at the same time increases difficulty in the analysis. (Zhang et al., 2021b).

1.3 Deep Learning Methods for Brain Imaging Data

The study of the brain has advanced greatly in recent decades, partially due to the rapid growth of technology used in collecting and processing medical imaging data. Modern analysis of brain imaging data is interdisciplinary

research that requires expertise from various technical fields, such as machine learning and image processing. Computational approaches to analyzing brain functioning are important for diagnosing brain disorders or mental illnesses. Optimized computational algorithms enable the efficient processing of large volumes of brain imaging data and provide tools for uncovering important patterns.

Deep learning models provide state-of-the-art computational methods for automated feature extraction mechanisms, thanks to their capability to approximate complex functions (Jaiswal et al., 2018). However, the application of AI in detecting neurodevelopmental disorders using brain imaging information is a challenging task, due to the complexity of the fMRI data, and the oftentimes lack of specific triggers which could characterize the disease. Another big challenge in applying deep learning directly is the scarcity of the available brain imaging data required for making a reliable prediction given the input and output labels.

1.4 fMRI as Brain Imaging Data for Detecting ASD

Two approaches exist for acquiring fMRI data (see Figure 1.2). The first one is resting-state fMRI (rsfMRI), a passive acquisition state in which subjects lie motionless and are not instructed to think of or do anything in particular. The second one is task-based fMRI (tfMRI), which is an externally imposed paradigm that utilizes inputs and outputs from subjects (Zhang et al., 2016). In addition to differences in acquisition, brain fMRI data has multiple ways of its representation. Since fMRI data captures both spacial and time-based patterns of brain activity, it can be represented as 1) BOLD signal feature vectors corresponding to different regions (Kamitani and Tong, 2005); 2) as three-dimensional images; or 3) as a connectivity network between different brain regions. The latter representation is typically used for resting-state data.

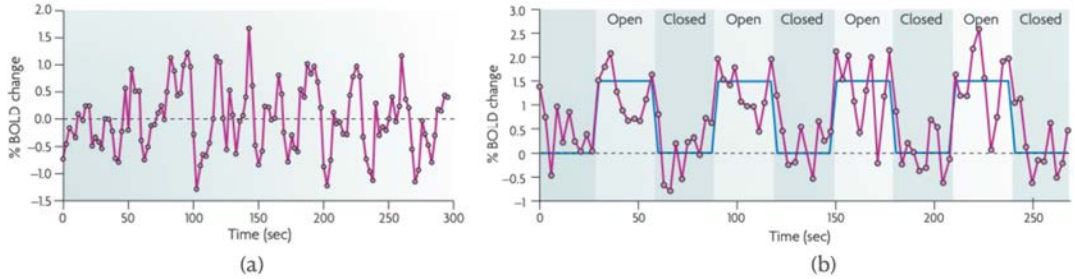


FIGURE 1.2: BOLD signal corresponding acquired through two fMRI paradigms: a) resting-state data recorded during resting; b) task-based data recorded from the primary visual cortex during a simple task paradigm that requires opening and closing eyes. Image adapted from Fox and Raichle, 2007.

1.4.1 Resting-state fMRI

The resting state or task-free fMRI data is a promising tool to investigate functional alterations in the human brain in a relaxed state when a human subject is not engaged in any particular task during scanning session. Despite the absence of tasks or stimuli in a resting state, the human brain is known to exhibit low-frequency fluctuations in blood oxygen level-dependent signal (Ogawa et al., 1990). Studying the brain at rest helped to discover the principles of macro-scale brain organization and shed light on the relationship between spontaneous and evoked activity. Nowadays, resting-state data is actively applied in the study of such neurodevelopmental disorders as attention deficit hyperactivity disorder and autism spectrum disorder with the expectation to improve existing diagnostics (Finn, 2021).

Resting-state brain fMRI signals became the conventional representation of brain activity in detecting neurological disease. Existing diagnostic studies focus on alterations in resting-state connectivity to characterize neural function (Kowalczyk et al., 2021). Therefore, we focus on the analysis of resting-state data in the first part of the thesis.

1.4.2 Task-based fMRI

The primary goal of task-based functional MRI is mapping the response of the brain to perceptual, motor, or cognitive stimulation. By measuring how

BOLD signal changes between task-stimulated states, task-based fMRI is used broadly to identify brain regions that are functionally involved in specific task (Zhang et al., 2016).

Among all sensory perceptions, vision is the human's dominant sense, which also affects cognition (Ripley and Politzer, 2010). For example, vision problems are common in patients who have suffered a traumatic brain injury (Armstrong, 2018). Professor Mriganka Sur of MIT's Department of Brain and Cognitive Sciences and Professor David Williams Director of Rochester's Center for Visual Science suggest that understanding how vision works is key to understanding how the brain works as a whole system (Hagen, 2012).

Consequently, many studies adopt visual task-based fMRI for studying and treatment of such cognitive and neurodevelopmental disorders as attention deficit hyperactivity disorder (Kowalczyk et al., 2021), obsessive-compulsive disorder (Pagliaccio et al., 2019), autism spectrum disorder (Li et al., 2018), because of how differently these patients visually perceive and process the world. Compared to using just resting-state connectivity data, task-based fMRI allows an extended investigation of brain activation patterns specific to different cognitive processes.

New research on neurobehavioral genetics by Gandal et al., 2022 sheds light on sensory hypersensitivity linked to ASD. The authors report that brain changes associated with ASD affect gene expression across cortical regions, with the most gene drop-offs observed in the primary visual cortex. In regards to ASD, atypical visual signal processing has been observed using various stimuli and tasks, with several reports of altered sensitivity (Utzerath et al., 2018; Kolodny et al., 2020)

1.5 Thesis Organization

1.5.1 Motivation

Studying the human brain at rest became very popular over the past decades and accounts for a large percentage of neuroimaging research. Resting-state

research offered the field new research methods which enriched our knowledge of brain functioning. Primarily, the community holds great expectations for the application of deep learning methods for discovering markers of ASD. However, the resting-state paradigm has a number of limitations that restrict the amount of knowledge that we can extract from it. As a result, the existing research in neurological disease might miss out on the prospects of using task-based imaging data.

Both resting-state fMRI and task fMRI are the two modalities that are used to identify individuals with ASD. Given the lack of research analyzing the efficiency and interpretability of the resting-state and task-based fMRI data in DNN-based neurological disease prediction, this study will aim **to propose a framework for classifying Autism Spectrum Disorder (ASD) using both resting-state and task-based fMRI data, considering their unique differences**.

Many studies including this one consider ASD classification into binary classes. This contrasts with the fact that in reality ASD is diagnosed with varying degrees of severity or on a spectrum. However, it is important to note that the end goal of our research is to be able to use an ASD classification framework as a screening tool in identifying individuals who may have ASD and not to provide a definitive diagnosis. A further evaluation by a qualified healthcare professional would be necessary to confirm an ASD diagnosis. In addition, the machine learning-based classification of ASD using non-binary labels is limited by the available annotations in the dataset. For example, the Autism Brain Imaging Data Exchange (ABIDE) dataset, used in this study, focuses on classifying ASD as present or absent and therefore contains binary labels. Autism severity scores are only partially available in ABIDE dataset, and their distribution varies greatly from site to site.

For now, ASD classification into binary classes is aligned with our motivation to use the classification framework during screening, which can classify ASD using non-behavioral markers from fMRI signals. Therefore one of the potential uses is during the early stage of the disease, where behavioral

markers can't be used. Given that the training dataset contains data from patients with varying degrees of disease severity, binary classification of ASD as present or absent is sufficient for this task.

We state the following objectives which will help to achieve the research aim:

1. Following the existing research on the resting-state fMRI dataset, propose a framework for population analysis and identify the importance of imaging resting-state fMRI and non-imaging features in the prediction task.

2. Generalize the proposed framework to visual task-fMRI for ASD classification.

3. Considering the differences between resting-state and task-fMRI data, evaluate the effectiveness of deep learning methods applied to task-based visual fMRI.

With regard to the research objectives stated above, we aim to answer the following research questions:

1. What kind of neural networks can efficiently model the population of subjects using fMRI data containing both imaging and non-imaging information?

2. Can a population-based classification framework help in improving ASD prediction in task-based fMRI?

3. How effective are neural networks with task-based fMRI data compared to resting-state data?

As a solution to research question 1, we present a graph-based neural network that can incorporate brain imaging information from subjects and inter-subject similarity. We explore the contribution of non-imaging features to the improved classification.

To address question 2, we present additional experiments where we integrate task-based visual fMRI data obtained for the study of ASD and the population graph-based framework.

To answer question 3, we consider task-fMRI as a data-rich brain imaging data source from which different neural networks can extract meaningful information for classification. As discussed in Section 2.1.1, resting-state fMRI data has its own limitations. Therefore, we explore different neural network learning strategies to learn better from rich task-based fMRI.

This study will contribute to the body of knowledge on ASD detection by first evaluating the effectiveness of the traditional resting-state research paradigm. Second, it will address the current shortage of research in the task-based paradigm. Finally, this study will present deep learning techniques that can be useful for improving ASD detection in task-based paradigms.

1.5.2 Structure and Organization

The content of this thesis is organized as follows:

1. In Chapter 1, we introduce the context of the study, that is, two main brain fMRI imaging types: resting-state and task-based fMRI. We identify research objectives and questions and emphasize the value of this research.

2. In Chapter 2, we present the related background literature on the state-of-the-art methods in brain disorder classification using Deep Neural Networks (DNNs). We overview the application of DNNs for resting-state and task-based fMRI. First, we present several neural network models popular in the analysis of resting-state data. Then, we present deep learning frameworks for task-based fMRI.

3. Chapter 3 presents our framework for ASD detection using population graphs for resting-state data. We introduce a population graph-based framework and present empirical results to show the effectiveness and limitations of the presented method.

4. Chapter 4 provides a generalization of the proposed framework for task-based fMRI analysis for ASD detection. We discuss the prospects of using task-based fMRI for ASD by unifying the graph-based classification framework introduced in Chapter 3 with visual task-based fMRI. In addition, we explore the existing DNN methods for task-fMRI which can be relevant for disease prediction. We propose a framework that can incorporate brain imaging and task stimuli-related information.

5. Chapter 5 discusses the design considerations and limitations of our proposed framework. We also show the usefulness of combined task-fMRI features and visual stimuli-related features as input. Furthermore, we discuss the implications and limitations of using resting-state fMRI and task-based fMRI, and deep learning techniques that can or cannot be shared with these two fMRI acquisition paradigms.

6. Finally, in Chapter 6, we conclude this thesis and suggest potential future work.

1.6 List of Publications

Here, we provide a list of publications that this thesis is based on.

Journals

- Rakhimberdina, Zarina, Xin Liu, and Tsuyoshi Murata (2020). "Population Graph-Based Multi-Model Ensemble Method for Diagnosing Autism Spectrum Disorder". In: *Sensors* 20.21. Number: 21 Publisher: Multidisciplinary Digital Publishing Institute, p. 6001.

- Rakhimberdina, Zarina, Quentin Jodelet, Xin Liu, and Tsuyoshi Murata (2021). "Natural Image Reconstruction From fMRI Using Deep Learning: A Survey". en. In: *Frontiers in Neuroscience* 15, p. 795488. ISSN: 1662-453X. DOI: 10.3389/fnins.2021.795488.

- Rakhimberdina, Zarina, Xin Liu, and Tsuyoshi Murata (2022). "Strengthening Robustness Under Adversarial Attacks Using Brain Visual Codes". In: *IEEE Access* 10, pp. 96149–96158.

Conference

- Rakhimberdina, Zarina and Tsuyoshi Murata (2020). "Linear Graph Convolutional Model for Diagnosing Brain Disorders". en. In: *Complex Networks and Their Applications VIII*. Ed. by Hocine Cherifi et al. Studies in Computational Intelligence. Cham: Springer International Publishing, pp. 815–826. ISBN: 978-3-030-36683-4.

Chapter 2

Background

In this chapter, we introduce the topics that will build the foundation for discussion in later chapters. Namely, we first describe two types of fMRI data: resting-state fMRI and task-based fMRI used in literature for classifying ASD, and provide a discussion of the benefits of one over the other. Following the description of the fMRI data, this chapter will present the background literature on the state-of-the-art DNNs used for the analysis of resting-state fMRI and task-based fMRI.

2.1 Resting-state and Task-based fMRI

Resting-state functional connectivity fMRI is performed while the participants are at rest. The participants are not engaged in any particular task, and their brain activity is measured in the absence of external stimuli. In contrast, a task-based fMRI is performed when the participant is engaged in a task. For example in a visual task, the participant is typically presented with visual stimuli, and their brain activity is measured in response to the stimuli. The key distinction in the application of resting-state and task-based fMRI to ASD classification lies in the input type, namely, the distinction between uni-modal and multi-modal inputs (as shown in Figure 2.1). Uni-modal input refers to input that comes from a single sensory modality. In case of resting-state fMRI, only fMRI data is used for the analysis and single-input DNNs are sufficient for analysis. On the other hand, task-based fMRI contains additional essential visual stimuli information, and thus contains two modalities:

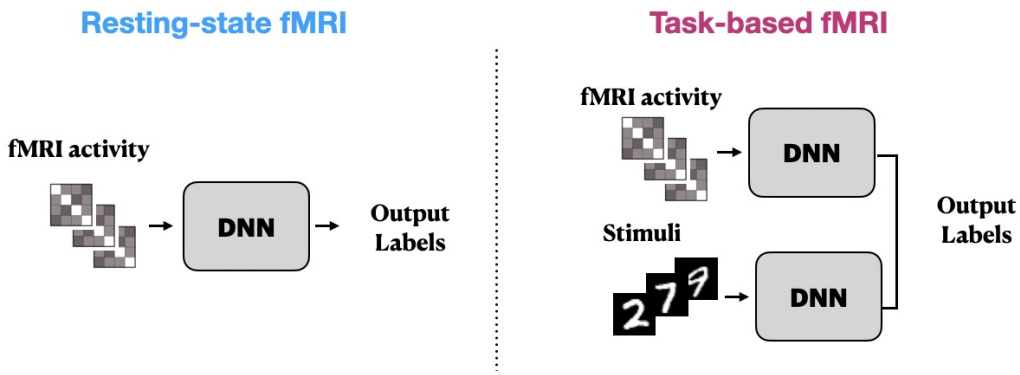


FIGURE 2.1: Distinction between uni-modal and multi-modal inputs.

fMRI signal and stimuli. Multi-modal input can provide more information and context, but it can also be more complex to analyze. To learn the relationship between both modalities DNNs need to process a multi-modal input. In addition, the presence of rich visual stimuli in task-based fMRI enables utilizing modern computer vision models and techniques, such as the use of visual image representation from models trained on large-scale visual datasets.

Both task-based fMRI and resting-state fMRI analysis have their own unique advantages and limitations. Task-based fMRI is useful for identifying the neural networks involved in performing a specific task, while resting-state fMRI is useful for identifying intrinsic functional connectivity patterns in the brain.

2.1.1 Resting-state fMRI

The resting state or task-free fMRI data is a promising tool to investigate functional alterations in the human brain in a relaxed state when the human subject is not engaged in any particular task during the scanning session. Despite the absence of tasks or stimuli in a resting state, the human brain is known to exhibit low-frequency fluctuations in blood oxygen level-dependent signal (Ogawa et al., 1990). Studying the brain at rest helped to discover the principles of macro-scale brain organization and shed light on the relationship between spontaneous and evoked activity. Nowadays, resting-state data is

actively applied in the study of such neurodevelopmental disorders as attention deficit hyperactivity disorder and autism spectrum disorder with the expectation to improve the existing diagnostics (Finn, 2021). The resting state fMRI (rsfMRI) has several advantages over task-fMRI for disease prediction:

- First, the data acquisition is usually shorter and less complex, especially when samples need to be collected from a large group of people (Albuquerque et al., 2021; O'Connor and Zeffiro, 2019).
- Second, due to the absence of the specific task, the data preprocessing and analysis is significantly simplified (Liu et al., 2022). This makes it easy to share the data across the laboratories with the goal of creating a large-scale database.
- Finally, rsfMRI is more practical when individuals, such as children, paralyzed, or cognitively impaired patients, are unable to perform fMRI tasks (O'Connor and Zeffiro, 2019).

Because of these advantages, a number of researchers have been investigating the brain's activation patterns during the resting state and many public datasets are focused on using rsfMRI. On the other hand, a resting-state fMRI data acquisition has limitations in decoding activity patterns due to two main factors:

- The brain activity patterns in the resting state can potentially take a large number of possible configurations. The lack of a stimulus or an experimental task causes much greater variability in activity patterns than expected. A big variation between training and test data samples across a single individual subject makes the task quite challenging for computational models designed for extracting common patterns from the data (Guidotti et al., 2015).
- In addition to the variability within a single-subject scan, the variability of fMRI signals across different subjects can be a source of concern. The evidence shows that it has been particularly challenging to obtain

consistent fMRI activations across different brains and populations due to the huge variability between individuals (Zhang et al., 2016). This makes it challenging to compare the imaging data of different individuals.

Despite the limitations and due to the simplicity and practicality of the experimental setup, resting-state brain fMRI signals became the conventional representation of brain activity in the diagnosis of neurological disease, focusing on alterations in resting-state connectivity to characterize neural function (Kowalczyk et al., 2021). Therefore, we focus on the analysis of rsfMRI data in the first part of the thesis.

2.1.2 Task-based fMRI

The primary goal of task-based functional MRI is mapping the response of the brain to perceptual, motor, or cognitive stimulation. By measuring how BOLD signal changes between task-stimulated states, task-based fMRI is used broadly to identify brain regions that are functionally involved in specific task (Zhang et al., 2016). However, it has been shown by recent works that task-induced brain state manipulation can be used to improve the prediction of individual traits (Greene et al., 2020; Greene et al., 2019). Li et al., 2018 showed that visual regions of the brain can be used to distinguish controls from ASD patients, since controls may attend to the visual features more closely than ASD subjects. Several properties of the visual task-based state fMRI make it appealing for diagnosis and disease prediction (Zhang et al., 2016):

- Data acquired from task-based fMRI has enhanced interpretability and sensitivity to meaningful between-subject variability in healthy adults and patients (Finn, 2021). Even tasks that are seemingly unrelated to the target behavior, such as predicting fluid intelligence from a finger-tapping task, may yield better improved prediction results compared to the resting state. This way, tasks serve as an instrument that one can

use to constrain overall variability between subjects (i.e., reduce noise) and at the same time preserve meaningful differences.

- The presence of the task allows more control over the data processing such as averaging the samples across different runs or scanning sessions and across subjects. Thus, task paradigms enable researchers to parse signals into meaningful and non-meaningful segments and look for generalization patterns in the behavior (Finn, 2021).

- Compared to the resting-state, task-based fMRI analysis is based on more standardized methods, while rsfMRI analysis still lacks a consensus over the best analysis methods, including maximizing inter-subject signal variability (O'Connor and Zeffiro, 2019).

Among all sensory perceptions, vision is the human's dominant sense, which also affects cognition (Ripley and Politzer, 2010). Vision problems are common in patients who have suffered a traumatic brain injury (Armstrong, 2018). Professor Mriganka Sur of MIT's Department of Brain and Cognitive Sciences and Professor David Williams Director of Rochester's Center for Visual Science suggest that understanding how vision works is key to understanding how the brain works as a whole system (Hagen, 2012).

Consequently, many studies adopt visual task-based fMRI for studying and treatment of such cognitive and neurodevelopmental disorders as attention deficit hyperactivity disorder (Kowalczyk et al., 2021), obsessive-compulsive disorder (Pagliaccio et al., 2019), autism spectrum disorder (Li et al., 2018), because of how differently these patients visually perceive and process the world. Compared to using just resting-state connectivity data, task-based fMRI allows an extended investigation of brain activation patterns specific to different cognitive processes.

New research on neurobehavioral genetics by Gandal et al., 2022 sheds light on sensory hypersensitivity linked to ASD. The authors report that brain changes associated with ASD affect gene expression across cortical regions,

with the most gene drop-offs observed in the primary visual cortex. In regards to ASD, atypical visual signal processing has been observed using various stimuli and tasks, with several reports of altered sensitivity (Utzerath et al., 2018; Kolodny et al., 2020)

2.1.3 fMRI Data Representations

Typical fMRI data representations include voxel-based and connectome-based as presented in Figure 2.2. Below, we will discuss the differences between these two representations.

Voxel-based fMRI representation. Voxel-based fMRI captures the activity of individual voxels in the brain. Voxel-based analysis provides high image resolution, meaning it can identify activity in specific brain regions. However, it does not consider the relationship between different regions of the brain.

Connectome-based fMRI representation. On the other hand, connectome-based fMRI representation, captures the functional connectivity between different brain regions. It considers the relationship between voxels, creating a network of connections between them. This approach provides insight into how brains function and how different regions work together. Voxel-based fMRI representation focuses on local activity in the brain, while connectome-based fMRI focuses on the functional connectivity between brain regions.

The choice of representation depends on the research question and the goal of the analysis. Generally, connectome-based representation is useful for reducing complex matrices containing millions of voxels of resting-state fMRI into more easy-to-process connectivity networks. While task-based fMRI usually produces data that is task-specific and focused on a particular brain region, therefore the common representation is the vector of individual voxels.

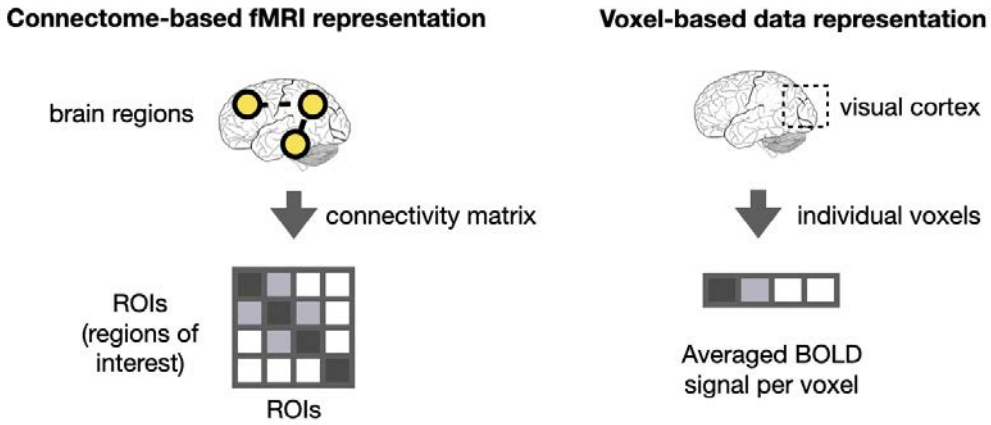


FIGURE 2.2: Voxel-based and connectome-based representations of fMRI data.

2.2 DNN Analysis for Resting-state and Task-based fMRI Data

As mentioned in Section 1.4, the amount of data collected from brain fMRI sessions could be immense, reaching several million data points (voxels) for a single subject. Subsequently, the servers handling such amounts of data are required to have high memory capacity as well as high computing power. Nowadays, deep neural networks (DNNs) are used as a powerful analysis framework, which can satisfy the requirements of large data inputs and complex computations with the data. Deep learning is revolutionizing data analysis tools in many domains, including neuroscience (Jordan and Mitchell, 2015; LeCun, Bengio, and Hinton, 2015). In particular, DNNs have recently gained significant attention in the analysis of brain connectivity patterns acquired through functional magnetic resonance imaging (fMRI).

While both resting-state and task-based acquiring methods use fMRI to measure neural activity, they differ in their approach and the type of information they provide. Resting-state functional connectivity (RSFC) fMRI measures the correlation between different regions of the brain while the subject is at rest. Whereas, task-based fMRI measures the changes in brain signals that occur when a subject performs a specific task.

The differences in the data acquisition between resting-state and task-based fMRI, and the presence of the rich external stimuli in the visual task paradigm, have different implications for the design of DNNs.

RSFC fMRI data is generally static in nature, there are no explicit time frames of interest and the correlations between brain regions remain relatively constant over recording time. Consequently, DNNs designed to process RSFC fMRI data tend to be shallow depth-wise, with fewer network layers and fewer learning parameters. On the other hand, task-based fMRI data is highly dynamic, with the patterns of brain activity, captured by individual voxels, changing more rapidly as the subject goes through different tasks. DNNs designed to analyze task-based fMRI data tend to be deeper and more complex, with more network layers and more learning parameters. In addition, the nature of the stimuli, such as visual images, requires the adoption of the existing visual DNNs to effectively map representations from the brain signal domain to the image domain.

2.2.1 Deep Neural Networks for Resting-state fMRI Analysis

Based on the differences in fMRI data representation and the task, the diversity of DNN models in the literature used for fMRI analysis can be categorized into several groups. The first group of methods relies on simple fully connected DNNs. Such methods have become a popular choice for the analysis of rsfMRI, represented in the form of resting-state functional connectivity networks (RSFC). The second group adapts convolutional neural networks (CNNs) and recurrent neural networks (RNNs) popular in image, speech, and video analysis.

Another group of methods uses a graph structure to represent data samples. Since the task involves a set of human subjects, it is natural to construct a population graph, a structure composed of the entire set of human subjects and representing inter-individual connectivity between them, as shown in Figure 2.3. In this approach, each subject is modeled as a node with corresponding brain-connectivity data, and each edge is defined based on the

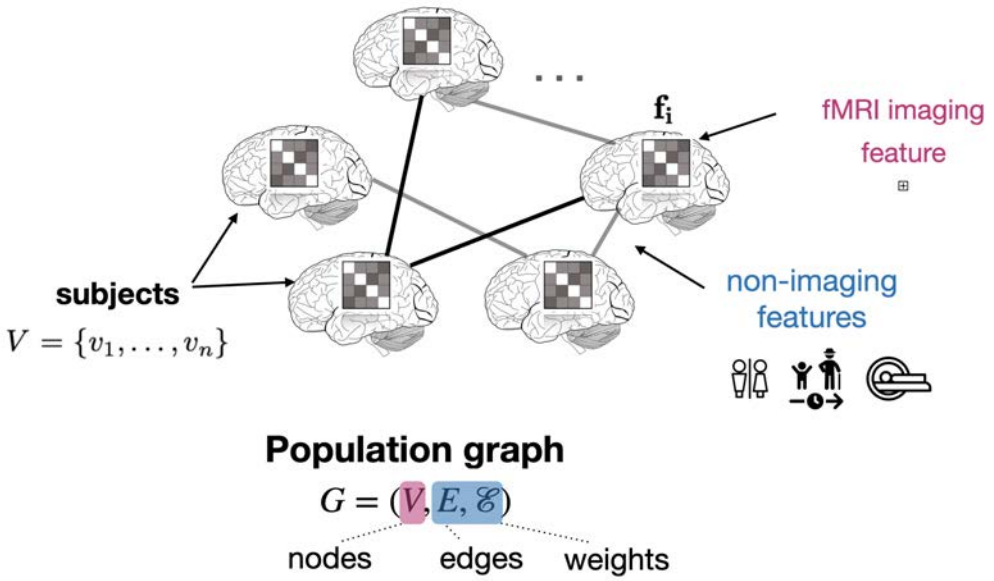


FIGURE 2.3: Representing a population using a graph.

similarity between subjects' phenotypic features (age, gender, handedness, etc.).

Convolutional Neural Networks

CNN architecture has been widely used for the analysis of functional connectivity of resting-state functional magnetic resonance imaging. Using this approach, 2-dimensional brain connectivity matrices, which represent relations between pairs of a brain region, are fed as input to convolutional layers. For example, in the research by Sherkatghanad et al., 2019, the authors focused on the automated detection of autism spectrum disorder (ASD) using CNN with a brain imaging dataset. The authors presented a CNN-based architecture with one convolutional layer, followed by max-pooling and densely connected layers (see Figure 2.4). The hidden layer followed by max-pooling is used to decrease the number of features and avoid overfitting. After the max-pooling layer, a dropout is applied which keeps only 25% of the nodes for training. Finally, the output is concatenated and passed to a dense layer, which is subsequently used for classification (Sherkatghanad et al., 2019).

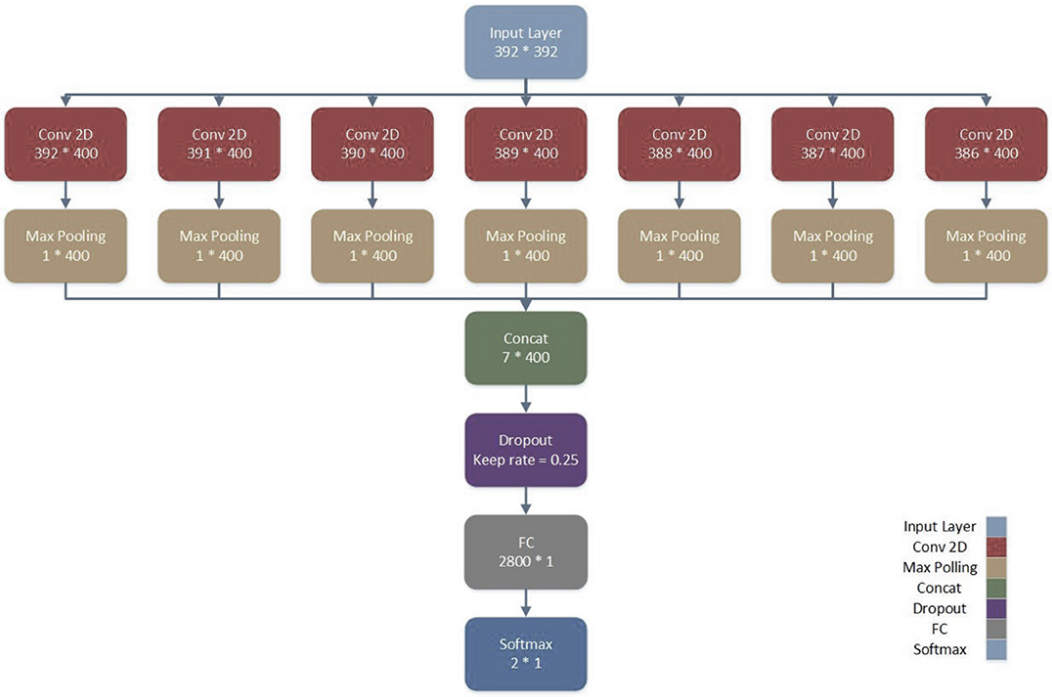


FIGURE 2.4: Proposed CNN architecture by Sherkatghanad et al., 2019 for automated detection of ASD. The figure is adapted from a corresponding paper.

Graph Convolutional Networks

Over the past few years, Graph Convolutional Network (GCN) has been used as a powerful method for processing graph structure. The strength of GCN originates from combining local node-level context and global neighborhood-level context (Defferrard, Bresson, and Vandergheynst, 2016; Kipf and Welling, 2017). Similar to standard two-dimensional image convolution (Krizhevsky, Sutskever, and Hinton, 2012), which uses a rectangular filter over the values of the neighboring pixels, graph convolution aims to aggregate the neighborhood information of nodes in a graph using graph Laplacian matrix (Kipf and Welling, 2017).

In the context of population-based brain disorder detection, Graph Convolutional Network (GCN) takes adjacency matrix representation $\mathbf{A}$ of population graph $G = (V, E, \mathcal{E})$ as an input. The element a_{ij} of symmetric adjacency matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ corresponds to edge weight between nodes v_i and v_j. Some of the popular methods used to determine edge weight include

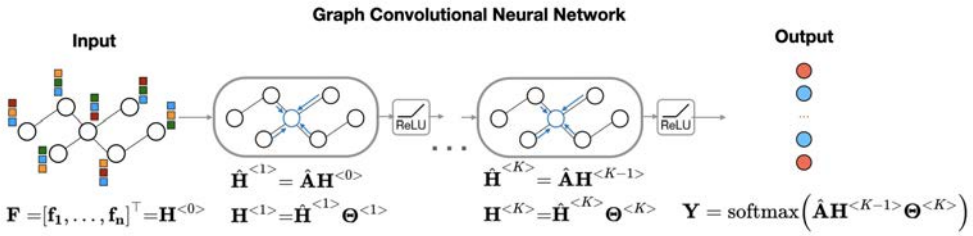


FIGURE 2.5: Schematic pipeline for GCN.

the Pearson correlation, the K-nearest neighbor (KNN), and the distance-based graph Shuman et al., 2013. The graph is annotated with feature matrix $\mathbf{F} = [\mathbf{f_1}, \ldots, \mathbf{f_n}]^\top$, where each row $\mathbf{f_i} \in \mathbb{R}^D$ corresponds to a D-dimensional fMRI features of node v_i. The feature matrix $\mathbf{F}$ is used as an input feature matrix to the first layer of GCN, i.e $\mathbf{H}^{<0>} = \mathbf{F}$. The output of the model is binary matrix $\mathbf{Y} \in \mathbb{R}^{n \times C}$, where each row $\mathbf{y_i}$ denotes the probability of node v_i belonging to one of the $C = 2$ classes, healthy or pathological (see Figure 2.5).

GCN can be classified into spectral-based (Defferrard, Bresson, and Vandergheynst, 2016) and spatial-based (Niepert, Ahmed, and Kutzkov, 2016). Spectral-based GCN is built upon the concept of spectral convolutional neural networks that utilize the concepts of the graph Fourier transform and the normalized graph Laplacian matrix. Spatial-based GCN represents a graph convolution operation using the spatial relationships between the graph nodes. In other words, convolutional operation on a graph structure in layer l is performed by multiplying adjacency matrix $\mathbf{A}$ and the input feature matrix $\mathbf{H}^{<k-1>}$ from the previous layer $k-1$: $\mathbf{H}^{<k>} = \mathbf{A}\mathbf{H}^{<k-1>}$. One multiplication operation $\mathbf{A}\mathbf{H}^{<k>}$ averages feature vectors of nodes within the one-hop neighborhood. By stacking K layers, GCN aims to aggregate K-hop neighborhood information from node features. In addition to performing graph convolution, each layer k linearly transforms hidden feature representation by a weight matrix $\mathbf{\Theta}^{<k>}$. Then, a non-linear activation such as ReLU is applied to each layer, resulting in the following forward propagation formula for layer k:

$$\mathbf{H}^{<k>} = \mathrm{ReLU}(\hat{\mathbf{A}}\mathbf{H}^{<k-1>}\mathbf{\Theta}^{<k>}), \tag{2.1}$$

where $\hat{\mathbf{A}} = \mathbf{D}^{-\frac{1}{2}}(\mathbf{A} + \mathbf{I})\mathbf{D}^{-\frac{1}{2}}$ is normalized adjacency matrix containing self-loops and $\mathbf{D}$ is a degree matrix of $\mathbf{A}$. Finally, class labels for each node in the graph are computed using the final K-th softmax layer:

$$\mathbf{Y}_{GCN} = \text{softmax}(\hat{\mathbf{A}}\mathbf{H}^{<K-1>}\mathbf{\Theta}^{<K>}) \tag{2.2}$$

It is important to note that there are many research works using graph-based deep learning models that aim to map functions to brain regions, model functional connectivity of the brain, and analyze the brain's response to external events. These research works have been used for various tasks, including gender classification, emotion recognition, and brain motor imagery, and have been suggested to have potential clinical applications (Ahmedt-Aristizabal et al., 2021). However, they are not related to detecting or classifying a disease. Thus, their contributions are not presented in this work.

2.2.2 Deep Neural Networks for Task-based fMRI Analysis

The objectives of task-based fMRI can range from the classification of motor functions for prosthetic arm movement control (Cohen et al., 2014) to more complex tasks such as the identification of disease markers in vision, speech, and the motor function (Zhang et al., 2021b). Depending on the research objectives, different DNNs are used for task-based fMRI. The main advantage of using task-based fMRI with DNNs is the presence of additional input modalities such as image, video, or speech, which can be correlated with brain activity. Using a task paradigm gives an opportunity to qualitatively assess the meaningfulness of brain activity patterns captured in an fMRI scan. Moreover, the presence of external input modalities, such as natural images, makes it possible to overcome fMRI data size limitation (thousands of samples) by leveraging the models pre-trained on the relatively abundant (millions of samples) data. In the interest of this thesis, we will focus on the visual task-based fMRI.

Convolutional neural network (CNN) for visual task-based fMRI. In

computer vision, the convolution operation is a popular method for extracting information from images. When applied to task-based fMRI analysis, CNNs usually aim to extract visual feature representation from the stimuli images and map them to the corresponding brain activity. Convolutional layers in CNN perform feature extraction by filtering the information within a neighborhood of pixels in an image. Because of convolutional operation, the CNN has a better feature extraction capability compared to a simpler multilayer feed-forward neural network, which disregards the structural information of input images (LeCun et al., 1989). Stacking convolutional layers on top of each other allows learning hierarchical visual features of input images, known as feature abstraction. The lower CNN layers learn low-level details, whereas the higher CNN layers extract global high-level visual information from images (Mahendran and Vedaldi, 2015). The use of CNNs is ubiquitous in image processing tasks, including image reconstruction. Specifically, encoder-decoder (Beliy et al., 2019; Gaziv et al., 2020), U-Net (Fang, Qi, and Pan, 2020), generative adversarial network (Goodfellow et al., 2014), and variational autoencoder (Kingma and Welling, 2014) are popular architectures that adopt stacked convolutional layers to extract features at multiple levels.

To map brain fMRI activity to the visual stimuli, Shen et al. (Shen et al., 2019a) utilized a pretrained VGG-19-based convolutional DNN to extract hierarchical features from stimuli images (see Figure 2.6 A). The DNN consists of sixteen convolutional layers followed by three fully connected layers. This method was motivated by the finding that hierarchical image representations obtained from different layers of deep neural network correlate with brain activity in the visual cortex (Eickenberg et al., 2017; Horikawa and Kamitani, 2017). Using this fact, one can establish a hierarchical mapping from fMRI signals in the low/high-level areas of visual cortices to the corresponding low/high-level features from the DNN. For this task, the authors implemented a feature decoder D that maps fMRI activity patterns to multilayer DNN features. These decoded fMRI features correspond to the hierarchical image features obtained from DNN. The optimization is performed on the

feature space by minimizing the difference between the hierarchical DNN features of the image and multilayer features decoded from fMRI activity.

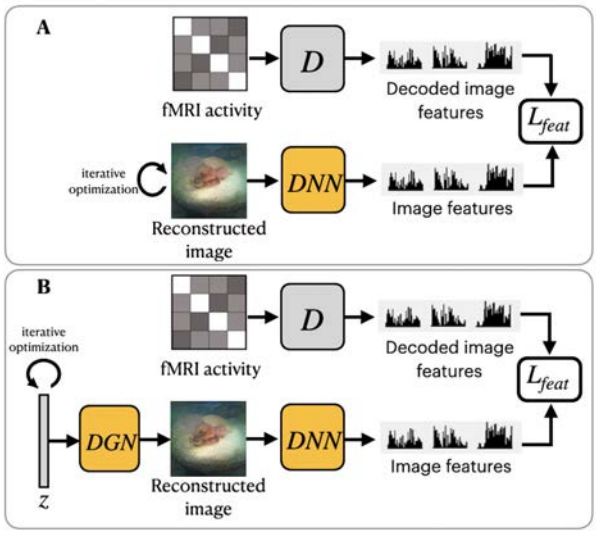


FIGURE 2.6: Overview of two variations of frameworks proposed by Shen et al., 2019a: **(A)** `ShenDNN` and **(B)** `ShenDNN+DGN`. The yellow color denotes the use of pretrained components.

Encoder–decoder methods. In computer vision, encoder–decoder models are multilayer models that incorporate convolutional layers. Encoder–decoder models are widely used in image-to-image translation (Isola et al., 2017) and sequence-to-sequence models (Cho et al., 2014). They learn the mapping from an input domain to an output domain via a two-stage architecture: an encoder E that compresses the input to the latent space representation $\mathbf{z} = E(\mathbf{x})$ and a decoder D that produces the output from the latent representation $\mathbf{y} = D(\mathbf{z})$ (see Figure 2.7 A) (Minaee et al., 2021). The compressed latent representation vector $\mathbf{z}$ serves as a bottleneck, which encodes a low-dimensional representation of the input. The model is trained to minimize the reconstruction error, which is the difference between the reconstructed image and the ground-truth image.

Beliy et al., 2019 presented a CNN-based encoder–decoder model, where the encoder E learns the mapping from stimulus images to the corresponding fMRI activity, and a decoder D learns the mapping from fMRI activity to their corresponding images. The framework of this method, which we refer

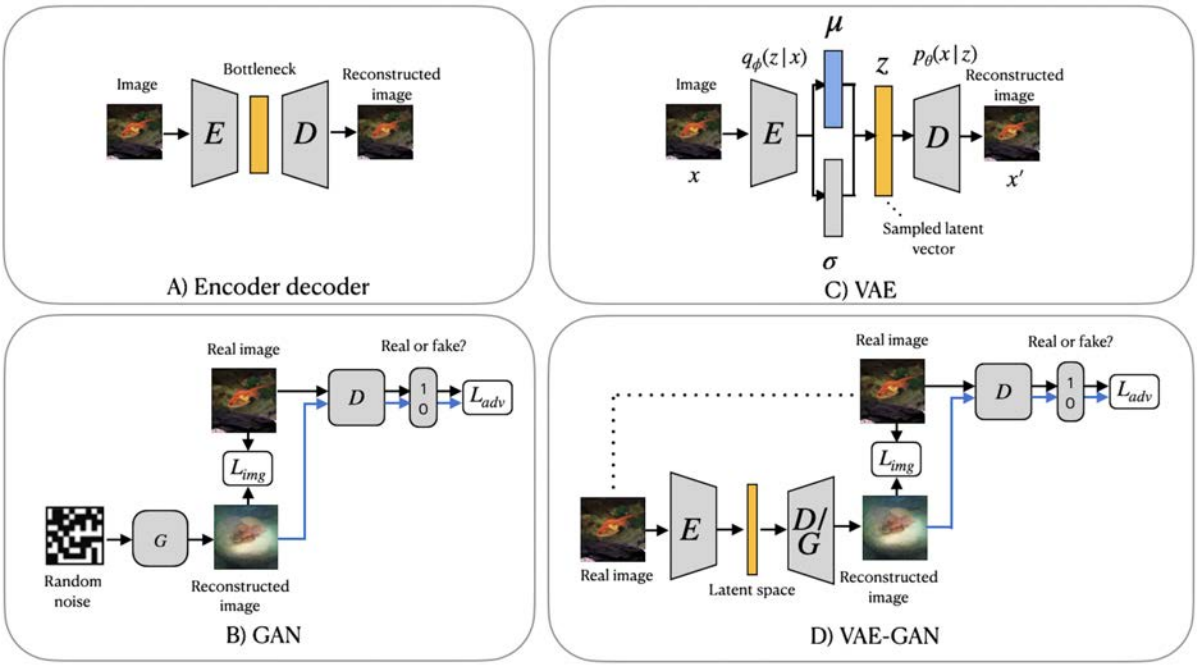


FIGURE 2.7: **(A)** Encoder–decoder architecture. **(B)** VAE. **(C)** GAN. **(D)** VAE-GAN.

to as `BeliyEncDec`, is presented in Figure 2.8. By stacking the encoder and decoder back-to-back, the authors introduced two combined networks *E-D* and *D-E*, whose inputs and outputs are natural images and fMRI recordings, respectively. This allowed the training to be self-supervised on a larger dataset of unlabeled data. Specifically, 50,000 additional images from the ImageNet validation set and test fMRI recordings without stimulus pairs were used as unlabeled natural images and unlabeled fMRI samples. The training was conducted in two steps. In the first step, the encoder *E* builds a mapping from stimulus images to fMRI activity. It utilizes the weights of the first convolutional layer of the pre-trained AlexNet (Krizhevsky, Sutskever, and Hinton, 2012) and is trained in a supervised manner to predict fMRI activity for input images. In the second step, the trained encoder *E* is fixed, and the decoder *D* is jointly trained using labeled and unlabeled data. The entire loss of the model consists of the fMRI loss of the encoder *E* and the Image loss (RGB and features loss) of the decoder *D*.

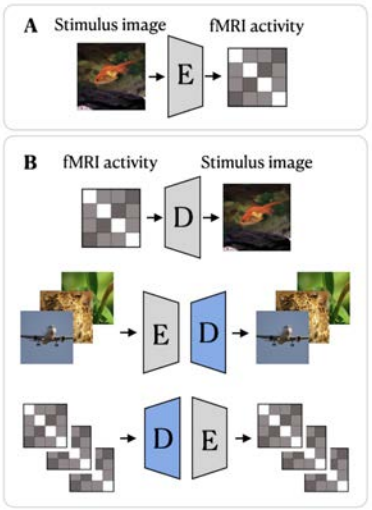


FIGURE 2.8: `BeliyEncDec` framework proposed by Beliy et al., 2019: **(A)** supervised training of the Encoder; **(B)** supervised and self-supervised training of the Decoder. The weights of the Encoder are fixed. The blue color denotes the components of the model trained on external unlabeled data.

2.3 Challenges in Applying DNN Models

There are several challenges in applying DNN to fMRI data. The main challenge is the lack of large annotated datasets. Unlike image data, fMRI datasets are relatively small and expensive to acquire. This limits the amount of data available for training DNN models. Another challenge is the variability of brain activation patterns across individuals, which makes it difficult to generalize DNN models to new subjects.

In the case of resting-state fMRI, the main challenge in applying DNN is the lack of a clear objective to guide the learning of DNN, unlike task-based fMRI, where the cognitive task or stimuli acts as a clear objective for the analysis. This makes it more difficult to define a loss function and train DNN models. Another challenge is the presence of artifacts and noise in the data, which can negatively impact the performance of DNN models.

In the case of task-based fMRI, the presence of time-locked stimuli that are shared across participants allows extending the amount of brain stimuli samples recorded from one brain. In addition, this time-locked stimuli is

shared across participants and this allows to significantly improve the signal-to-noise ratio and the inter-subject correlation. Stimuli information provides ground truth for averaging across scanning sessions and helps parse the complex signal into meaningful chunks. This in turn helps to reveal correlations across subjects, as recent research on visual task-based fMRI shows an improvement in the ability of DNN to generalize to new subjects (Kamitani and Tong, 2005; Beliy et al., 2019). While the presence of cognitive tasks or stimuli acts helps to improve the performance of DNN models, it also poses the challenge of how to design a model to handle multi-modal input, which contains both brain signals and the stimuli information.

Chapter 3

Classifying Autism Spectrum Disorder using Resting-state fMRI

3.1 Introduction

Automated medical detection of autism spectrum disorder (ASD) using deep learning techniques has shown great promise. With the advent of neural networks, the analysis of complex brain imaging data, such as functional magnetic resonance imaging (fMRI) data, has become more feasible. Numerous applications of neural networks have been proposed, including those based on direct image processing of brain data (Horikawa and Kamitani, 2017; Sherkatghanad et al., 2019) and those that utilize phenotypic information of subjects for disease prediction (Khosla et al., 2018; Parisot et al., 2018). In this chapter, we focus on Autism Spectrum Disorders (ASD) prediction using functional connectivity networks from resting-state fMRI.

Resting-state functional connectivity (RSFC) measures the temporal correlation between the blood-oxygen-level-dependent signal in different brain areas during a resting state (Biswal et al., 1995). RSFC can reveal new patterns in the brain network that can lead to neurological disorders and thus are widely used to study brain organization and mental disorder (Bassett and Bullmore, 2009; Hulvershorn et al., 2014; Bullmore and Sporns, 2009). RSFC is computed for each subject and is represented using a square matrix where each entry corresponds to the strength of the functional connectivity between

two regions of the brain. Along with RSFC, fMRI datasets contain subject-level non-imaging phenotypic information. For example, the Autism Brain Imaging Data Exchange (ABIDE) database combines both imaging (RSFC) and phenotypic information, including biological sex, age, and imaging site (Di Martino et al., 2014). Phenotypic information, in particular biological sex, was shown to be useful for the prediction of ASD since the disease affects females less frequently than males (Werling and Geschwind, 2013). The efficient use of both imaging and non-imaging information for neurological disorder prediction has become the focus of many recent works (Parisot et al., 2018; Ktena et al., 2018).

3.1.1 Motivation and Objective

The goal of this chapter can be summarized into the following points:

1. To design a framework that can classify ASD in a population using resting-state fMRI.

2. Analyze the role of graph structure in the classification and identify the role of imaging resting-state fMRI and non-imaging features.

In motivation (1), we build network models using resting-state functional connectivity data and additional non-imaging features corresponding to subjects' phenotypes [1]. In (2), we would like to propose a more general approach, where the informativeness of the features is decided by the model.

3.1.2 Organization

We begin this chapter by introducing the background research in Section 3.2. We then present the process of population graph construction using fMRI data and describe the differences between various edge-defining functions in Section 3.2.2. Section 3.3 describes the dataset. In Section 3.4.1, we present a simplified graph-based framework that can classify ASD using resting-state

[1] By phenotype we imply non-fMRI features, which characterize the person's age, biological sex, handedness, etc.

fMRI. We then analyze the role of graph structure in the classification and identify the role of imaging resting-state fMRI and non-imaging features using methods of graph signal processing (GSP) in Section 3.4.2. We then move on to analyzing multiple population graphs using the tools from GSP in Sections 3.4.3. Section 3.6 shows the detailed results of our proposed multi-model population graph-based ensemble.

3.2 Background

3.2.1 Graphs and Graph Neural Networks.

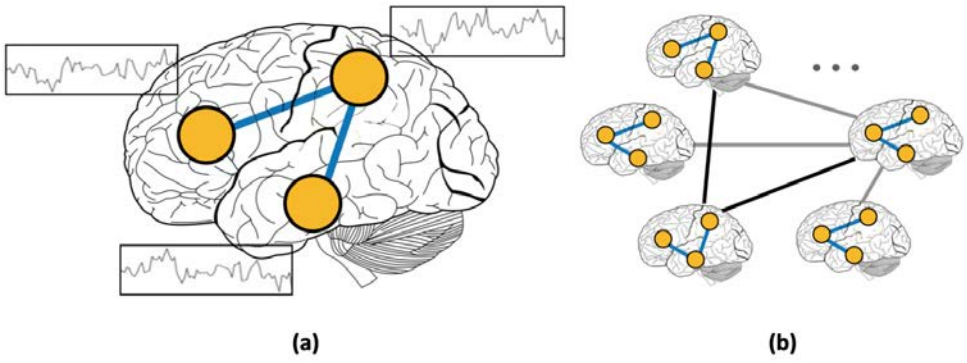


FIGURE 3.1: (a) RSFC graph and (b) population graph

Graphs present a natural way of modeling complex interactions by combining features of different modalities (Leskovec, Huttenlocher, and Kleinberg, 2010; Escala-Garcia et al., 2020). For example, in social networks, nodes represent individuals, and the presence of a link between two nodes signifies the existence of a friendship between the corresponding individuals. Unlike social networks, where links are predefined, graphs constructed from medical data require a more elaborate choice of the subjects' features for defining edges.

The diversity of graph-based models proposed earlier in the literature can be categorized into two classes, based on the way the nodes are defined. The first class of models (Figure 3.1(a)) uses graphs to describe the structural or functional connectivity of the human brain on an individual level (Bullmore

and Sporns, 2009). In other words, nodes represent brain regions and edges represent functional correlations between the time series of those regions (Ji, Maurits, and Roerdink, 2019; Hsieh, Sun, and Liang, 2014; Ventresca, 2019). As a result, each constructed graph corresponds to one subject and further analysis is performed using graph comparison metrics. The second class of models (Figure 3.1(b)) involves the construction of a population graph, a structure composed of the entire set of human subjects and representing inter-individual connectivity between them. In this approach, each subject is modeled as a node with corresponding brain-connectivity data, and each edge is defined based on a specific similarity metric (Parisot et al., 2018). This property of incorporating different types of features made population graph models effective for brain disorder classification (Parisot et al., 2018; Parisot et al., 2017).

Some researchers found incorporating additional non-imaging data useful for defining graph structure (Parisot et al., 2017; Parisot et al., 2018). In this work, we use graphs to represent inter-subject connectivity or population graph composed of the entire set of human subjects. Each subject is modeled as a node with corresponding RSFC data, and each edge is defined based on the similarity between subjects' features (RSFC, age, sex, etc.). Thus, the advantage of using graphs for medical diagnosis is multi-fold. First, modeling the subject's data as a population graph allows incorporating subjects' non-imaging features (Parisot et al., 2018; Parisot et al., 2017). Second, the existing graph-based methods, such as graph neural networks (Kipf and Welling, 2017), allow efficient computations on brain imaging datasets.

One such method is Graph Convolutional Neural Network or GCN. It was introduced as an efficient method for node classification on graphs (Kipf and Welling, 2017). Unlike traditional methods that focus either on the features (regression models, convolutional neural networks, etc.) or on the network structure (community detection algorithms, node embeddings, etc.), GCN provides a computational framework that accounts for both node features and graph structure (Defferrard, Bresson, and Vandergheynst, 2016;

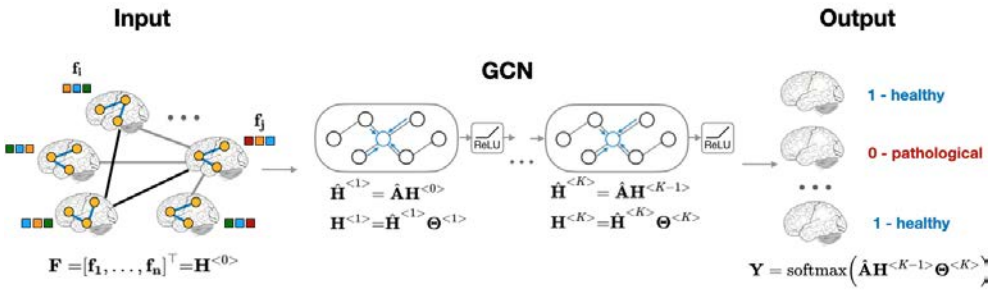


FIGURE 3.2: Overview of subject classification using graph neural networks. The model's inputs are an adjacency matrix representation $\hat{\mathbf{A}}$ of the population graph and a set of subjects' RSFC features $\mathbf{F}$.

Kipf and Welling, 2017). The aggregation of global neighborhood-level information is accomplished via graph convolution operation, which in contrast to standard image convolutions, is performed by matrix multiplication of graph Laplacian (Kipf and Welling, 2017). The computational pipeline of GCN applied to brain imaging data is presented in Figure 3.2.

Related works by Parisot et al., 2018; Anirudh and Thiagarajan, 2019 showed graph-based models to be useful for improving classification performance on brain imaging data. In the study of Autism Spectrum Disorder and Alzheimer's disease, Parisot et al., 2018 defined edges in a population graph based on the similarity between phenotypic features and RSFC patterns of the subjects. To identify the most informative phenotypic features, the authors compared the performance of GCN on the set of population graphs constructed using different combinations of those features. Based on the results, the authors concluded that along with RSFC, the biological sex of the subject and the location of the imaging facility contributed significantly towards the improved prediction of Autism Spectrum Disorder. Whereas for Alzheimer's disease, a different combination of features (sex and genetic information) resulted in the best model's prediction performance. To mitigate the problem of choosing the best-performing graph, the authors in (Anirudh and Thiagarajan, 2019) proposed to use a set of graphs with randomly removed edges. However, a significant issue with the previous works remains.

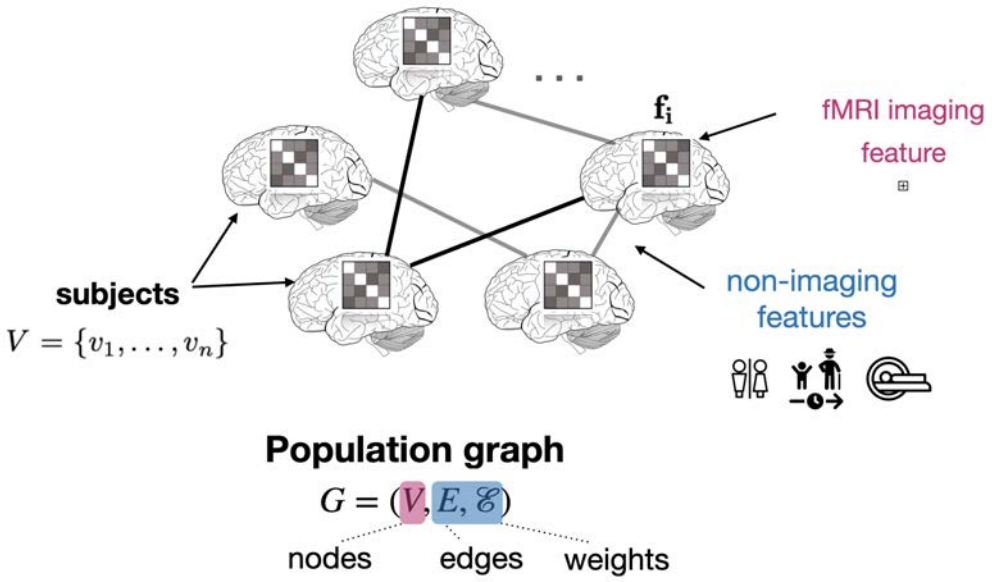


FIGURE 3.3: Representing a population of subjects using a graph.

This is a lack of systematic approach to choosing a particular graph definition over the other. In most cases, the similarity function used to define the edges becomes problem-specific, i.e., the choice of a particular graph definition might work for one dataset but fail or not be applicable for others. Therefore, there is a need for a robust model that is less sensitive to the choice of the underlying graph construction method.

3.2.2 Graph Construction

We describe the process of population graph construction adopted by (Parisot et al., 2018). Using the set of healthy controls and patients with ASD, we define the population graph as follows. The graph nodes represent the subjects from the dataset, and edges connecting the nodes represent the similarity between subjects' imaging and phenotypic features (see Figure 3.3). More explicitly, we construct an undirected weighted graph $G = (V, E, \mathcal{W})$, where the set of nodes $V = \{v_1, \ldots, v_n\}$ corresponds to a set of subjects. Each node v_i is associated with a d-dimensional feature vector $\mathbf{x}_i$ extracted from fMRI imaging data in the form of RSFC. The feature matrix $\mathbf{X} \in \mathbb{R}^{n \times d}$ consists of stacked feature vectors of n nodes in the graph. The set of edges $E \subseteq V \times V$

TABLE 3.1: Statistics of the ABIDE I subset used in this work.

	ABIDE
Total subjects in a subset	871
Patients	403
Healthy controls	468
Female/Male	144/727
Age range	6 - 58
Age mean	16.94
Sites	20

corresponds to links between the nodes, and $\mathcal{W} : E \mapsto \mathbb{R}$ is a function which assigns weight w_{ij} to each edge as follows:

$$\mathcal{W}(i,j) = sim\left(\mathbf{x}_i, \mathbf{x}_j\right) \sum_{h=1}^{H} \mathbf{1}[M_h(v_i) = M_h(v_j)], \tag{3.1}$$

where $sim\left(\mathbf{x}_i, \mathbf{x}_j\right)$ is defined based on correlation distance between lower triangular elements of RSFC matrices, and $M_h(v_i)$ is categorical phenotypic feature value[2] corresponding to node v_i.

3.3 Autism Spectrum Disorder Dataset

In this section, we describe the publicly available fMRI dataset used in our experiments (see details in Tables A.1 and 3.2). Autism Brain Imaging Data Exchange I (ABIDE I)[3] dataset combines structural and functional MRI data of 1,112 subjects from 17 international acquisition sites, which we refer to as imaging sites (Di Martino et al., 2014). Both imaging and non-imaging data are provided in the ABIDE. A single RSFC matrix represents functional connectivity between 111 regions of interest (ROIs) in the brain extracted from the fMRI scan. Non-imaging data correspond to phenotypic features, such as sex, age, and imaging site.

[2]Categorical phenotypic features are sex (male/female), imaging site number, and age (categorical values corresponding to age groups).

[3]http://preprocessed-connectomes-project.org/abide/

TABLE 3.2: Non-imaging information of ABIDE I subsset (871 subjects) for each individual site.

Site	ASD	Healthy control	Male count	Female count
CALTECH	5	10	10	5
CMU	6	5	7	4
KKI	12	21	24	9
LEUVEN	26	30	49	7
MAX_MUN	19	27	42	4
NYU	74	98	136	36
OHSU	12	13	25	0
OLIN	14	14	23	5
PITT	24	26	43	7
SBL	12	14	26	0
SDSU	8	19	21	6
STANFORD	12	13	18	7
TRINITY	19	25	44	0
UCLA	48	37	10	5
UM	47	73	7	4
USM	43	24	24	9
YALE	22	19	49	7

Quality assurance in ABIDE dataset. Seventeen international acquisition sites contributed fMRI to the ABIDE dataset. Upon receiving the data, it was checked for extreme outliers, impossible data entries, and missing values. Each site reviewed and verified the accuracy of the data before it was distributed (Di Martino et al., 2014; Abraham et al., 2017), the chance accuracy level is 54.03%.

The importance of non-fMRI features in ABIDE dataset. Previous studies claim the importance of non-fMRI including site, gender, and age information in the classification model. In the case of site information, the claim is that the same site fMRI contains a smaller variation in signal, thus the variations across different sites can impact the analysis (Parisot et al., 2018; Abraham et al., 2017). Other non-imaging features such as gender and age are also considered in various studies (Parisot et al., 2018; Ktena et al., 2018).

ABIDE subset. Based on the specifics of a research problem, researchers use different subsets of the ABIDE dataset. There are several reasons why researchers use different subsets from ABIDE dataset:

- Depending on the problem and study design, researchers might use different subsets of the dataset. Some studies may focus on influence of specific age range, gender or geographical region, while others may only use brain imaging data. For example, Di Martino et al., 2014 focused on the subset of 763 subjects including the individuals in different age groups. Nielsen et al., 2013 explored the impact of scanning site differences on resting-state functional connectivity and, specifically, examined the influence of site-specific factors, such as scanner manufacturer and different imaging parameters, within subset of ABIDE I dataset (964 subjects and 16 separate international sites).

- The quality of the brain imaging preprocessing techniques also influences the number of subjects. For example, common preprocessing steps include correction for motion, and slice timing, spatial normalization, and smoothing. Other steps may include denoising, artifact correction, and registration to a standard brain template. The differences in these preprocessing steps results in a different number of subjects selected for analysis.

To ensure a fair comparison with other baselines, we utilize the same subset of data in all our analyses. This subset consists of 403 patients with ASD and 468 healthy individuals, and was selected based on quality assurance preprocessing method established by Abraham et al., 2017. This subset has been used in related works by Abraham et al., 2017; Parisot et al., 2018; Sherkatghanad et al., 2019; and Anirudh and Thiagarajan, 2019, allowing for more accurate conclusions when comparing our findings to these baselines.

3.4 Proposed Method

In this section, we first present a simple yet efficient graph-based framework that can classify ASD using resting-state fMRI. We then analyze the role of

graph structure in the classification and identify the role of imaging resting-state fMRI and non-imaging features using methods of graph signal processing.

3.4.1 Simplified Graph Neural Network

In standard GCN, Convolutional operation on a graph structure in layer l is performed by multiplying adjacency matrix $\mathbf{A}$ and the input feature matrix $\mathbf{H}^{<k-1>}$ from the previous layer $k-1$: $\mathbf{H}^{<k>} = \mathbf{A}\mathbf{H}^{<k-1>}$. One multiplication operation $\mathbf{A}\mathbf{H}^{<k>}$ averages feature vectors of nodes within the one-hop neighborhood. By stacking K layers, GCN aims to aggregate K-hop neighborhood information from node features. In addition to performing graph convolution, each layer k linearly transforms hidden feature representation by a weight matrix $\mathbf{\Theta}^{<k>}$. Then, a non-linear activation such as ReLU is applied to each layer, resulting in the following forward propagation formula for layer k:

$$\mathbf{H}^{<k>} = \mathrm{ReLU}(\hat{\mathbf{A}}\mathbf{H}^{<k-1>}\mathbf{\Theta}^{<k>}), \tag{3.2}$$

where $\hat{\mathbf{A}} = \mathbf{D}^{-\frac{1}{2}}(\mathbf{A}+\mathbf{I})\mathbf{D}^{-\frac{1}{2}}$ is normalized adjacency matrix containing self-loops and $\mathbf{D}$ is a degree matrix of $\mathbf{A}$.

Multi-layer Graph Convolutional Network (GCN) is considered to be the state-of-the-art method for learning population graph representations (He et al., 2018; Parisot et al., 2018). Inspired by the convolutional neural network in the image processing domain, GCN adopts similar neural network architecture by stacking feature propagation layers followed by a non-linear activation function. However, the application of a multilayer neural network results in greater complexity due to non-linear transformations (He et al., 2018; Wu et al., 2019).

SGC (Wu et al., 2019), which stands for Simple Graph Convolutions, is the simplest formulation of a graph convolutional model which was introduced to better understand and explain the mechanisms of GCN. The simplicity of the linear graph convolutional model is achieved by removing non-linear activations between layers, and aggregating weight matrices corresponding

to each layer into a single matrix. (Wu et al., 2019) also showed that the multilayer Graph Convolutional Neural (GCN) network primarily benefits from local averaging achieved by a graph convolution operation. Therefore, after removing non-linear layers from the equation (3.2) and collapsing the repeated multiplications with normalized adjacency matrix $\hat{\mathbf{A}}$ into a single matrix $\hat{\mathbf{A}}^K$, the resulting linear model becomes as follows:

$$\mathbf{Y}_{SGC} = softmax(\hat{\mathbf{A}}^K \mathbf{F}\boldsymbol{\Theta}), \tag{3.3}$$

with K-hop convolutional component $\hat{\mathbf{A}}^K\mathbf{F}$ (where $\mathbf{F} = \mathbf{H}^{<0>}$) multiplied by logistic regression classifier parametrized by weight matrix $\boldsymbol{\Theta} = \Pi_{k=1}^{K}\boldsymbol{\Theta}^{<k>}$.

In our work (Rakhimberdina and Murata, 2020), we utilize the linear architecture of SGC for performing efficient computations on the constructed population graphs (see Figure 4.5). For each of the datasets, we first compute normalized adjacency matrix $\hat{\mathbf{A}}$ (with added self-loops) of the population graph $G = (V, E, \mathcal{E})$. We further annotate each node v_i in the graph with feature vector $\mathbf{f_i}$ representing fMRI connectivity. For simplicity of further calculations, we stack individual feature vectors into a single feature matrix $\mathbf{F} = [\mathbf{f_1}, \ldots, \mathbf{f_n}]^\top$. Adjacency matrix $\hat{\mathbf{A}}$ and feature matrix $\mathbf{F}$ become the inputs to SGC model that is further trained to learn output matrix $\mathbf{Y} \in \mathbb{R}^{n\times 2}$. The computational speedup and memory efficiency of SGC over GCN is achieved by precomputing $\hat{\mathbf{A}}^K\mathbf{F}$, which minimizes memory consumption as the model learns only a single weight matrix $\boldsymbol{\Theta}$. The training of the model reduces to binary logistic regression performed with the Adam optimization algorithm.

3.4.2 Population Graph Construction

Constructing a population graph from the set of patients and control subjects is not a straightforward task, as there exist multiple edge definitions that map the data to the graph structure. In a graph setting, after defining subjects as nodes, we are interested in identifying the features that capture the intrinsic relationship between the nodes. The optimal graph structure is

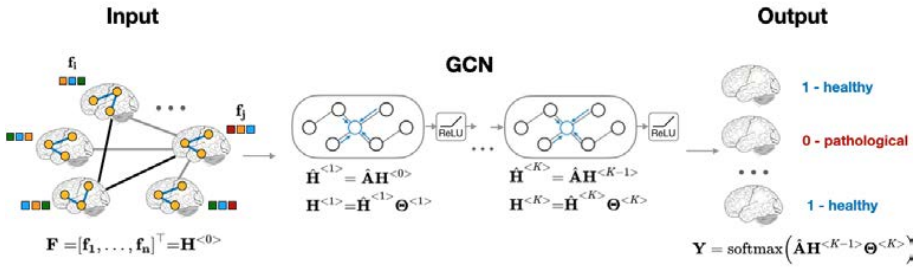


FIGURE 3.4: Overview of classification pipelines based on SGC. SGC takes adjacency matrix representation $\hat{\mathbf{A}}$ of population graph $G = (V, E, \mathcal{E})$ as an input.

considered to be the one in which clusters of patients and healthy subjects can be well separated. Learning the optimal graph structure becomes even more crucial in the later stage as the most optimal graph topology permits subsequent efficient data processing. In the following section, we first describe the process of population graph construction adopted in the literature. We then proceed to the analysis and selection of the best-performing graphs using Graph Signal Processing (GSP) tools. Finally, we present the ensemble of multiple graph-based models on brain imaging data.

From the edge definition in equation (3.1), it is clear that the parameter that affects the graph topology is the edge defining function $\mathcal{W}$. We therefore briefly describe the variety of edge-defining functions proposed in this and earlier works. For clarity, we categorize the resulting graphs into four groups:

1. *sim_RSFC*: a fully connected weighted graph constructed using a correlation between RSFC features. This corresponds to using a weight function $\mathcal{W}(i,j) = sim\left(\mathbf{x}_i, \mathbf{x}_j\right)$.

2. *sim_phenotype* graphs: graphs which are constructed using a combination of phenotypic features (*site, sex, age*). The graph construction corresponds to taking only the second part of the equation (3.1), i.e when $\mathcal{W}(i,j) = \sum_{h=1}^{H} \mathbf{1}[M_h(v_i) = M_h(v_j)]$. Totally there are seven *sim_phenotype* graphs constructed using combinations of three phenotypic features: *sim_site*, *sim_age*, *sim_sex*, *sim_site_age*, *sim_site_sex*, *sim_sex_age*, and *sim_site_age_sex*.

3. *sim_RSFC_phenotype* graphs: graphs that utilize the combination of RSFC features and phenotypic features for edge definition. Similarly, there are seven *sim_RSFC_phenotype* graphs in total: *sim_RSFC_site*, *sim_RSFC_age*, *sim_RSFC_sex*, *sim_RSFC_site_age*, *sim_RSFC_site_sex*, *sim_RSFC_sex_age*, and *sim_RSFC_site_age_sex*.[4]

4. Baseline graphs: a) *FC* – a fully connected graph with edge weight equal to 1, b) *identity* graph – a fully disconnected graph with adjacency matrix equal to the identity matrix, and c) *random* graph, constructed by randomly assigning binary edges in *identity* graph.

Graph Signal Processing

After constructing multiple population graphs based on the different choices of edge-defining functions, we are interested in identifying the ones which capture the most optimal representation of the dataset. For this purpose, we utilize the tools from Graph Signal Processing (GSP) to analyze the underlying structures of the resulting graphs. This section describes the fundamental concepts of GSP, such as graph Fourier Transform and graph filtering. GSP tools help to perform a quantitative comparison between graphs produced using a different choice of edge definition. In particular, based on the characteristics of signal smoothness, we are interested in selecting the most relevant features for population graphs constructed using different edge-defining functions.

Graph Fourier Transform

GSP, in contrast to classical signal processing, analyzes signals that reside on the nodes of graphs. These graph signals can also be referred to as node features. With GSP, we can generalize the signal processing concepts, such as signal smoothness and signal filtering, to a graph domain. We introduce the

[4]Note that when we are not using any features, the graph becomes *sim_RSFC*, which corresponds to the first group.

fundamental concepts of GSP using conventional definitions of the adjacency matrix, graph Laplacian matrix, and graph Fourier transform (Chung, 1996).

Given a graph G with the adjacency matrix $\mathbf{A}$ and degree matrix $\mathbf{D}$, the combinatorial Laplacian matrix of graph G is defined as a difference $\mathbf{L} = \mathbf{D} - \mathbf{A} \in \mathbb{R}^{n \times n}$. A feature vector $\mathbf{x} \in \mathbb{R}^d$ defined on each node is called a graph signal. Since $\mathbf{L}$ is a positive semidefinite matrix, it can be decomposed into a complete set of orthonormal eigenvectors and corresponding eigenvalues:

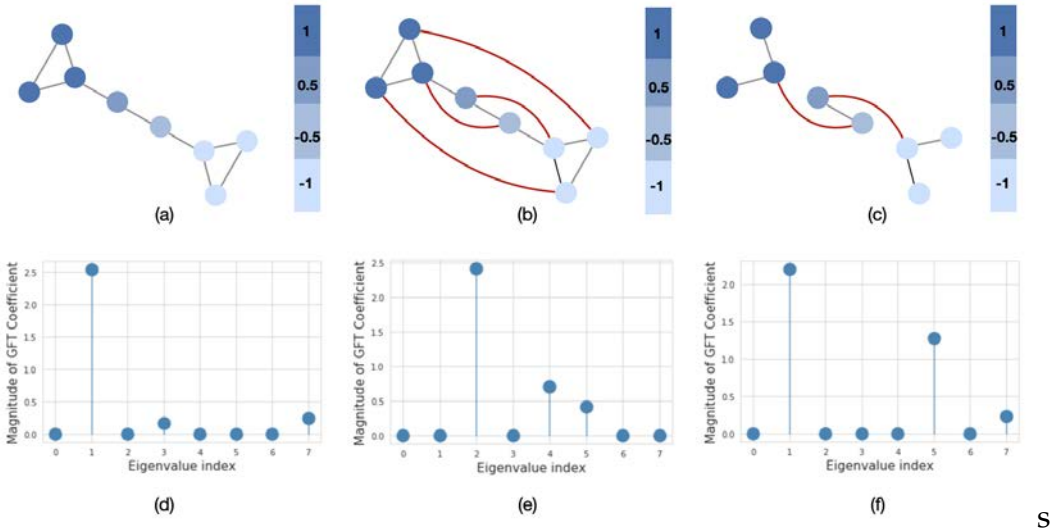
$$\mathbf{L} = \mathbf{U}\Lambda\mathbf{U}^T, \tag{3.4}$$

where $\mathbf{U}$ is a matrix containing eigenvectors as columns, and $\mathbf{\Lambda}$ is a diagonal matrix that contains eigenvalues along the diagonal.

Finally, we introduce the graph Fourier transform (GFT) and its inverse, which is used to represent a graph signal in node and spectral domains. Given a signal $\mathbf{x} \in \mathbb{R}^d$ and eigendecomposition of Laplacian $\mathbf{L}$, the graph Fourier transform of $\mathbf{x}$ is defined as $\hat{\mathbf{x}} = \mathbf{U}^\top \mathbf{x}$ and represents the signal in the graph spectral domain. While the inverse graph Fourier transform of $\hat{\mathbf{x}}$ is defined as $\mathbf{x} = \mathbf{U}\hat{\mathbf{x}}$. $\hat{\mathbf{x}}$ is also referred to as a frequency component of $\mathbf{x}$.

Similarly to time series analysis, the graph frequency components $\hat{\mathbf{x}}$ represent the signal variation with respect to the graph structure. To evaluate how much a signal varies, we use a graph's spectral representation. For example, in Figure 3.5 (a), we plot a barbel graph with two clusters and a fixed signal over its nodes. The signal varies smoothly over the graph since more similar values appear on neighboring nodes (low-frequency signals). The spectral representation of this graph in Figure 3.5 (d) shows that the signal mainly consists of low-frequency components. However, if we rewire or remove some edges from the original graph, i.e., introduce dissimilar signals on neighboring nodes (Figures 3.5 (b) and (c)), we can observe that the signal will have more energy in the higher frequencies (Figures 3.5 (e) and (f)). The last two graphs are referred to as high-frequency graphs. In population graph construction, we are particularly interested in the case when the graph

is constructed in such a way that the neighboring nodes have similar features. This property of the graph is referred to as global smoothness (Dong et al., 2019).



s

FIGURE 3.5: The importance of the graph structure. We plot the graph signal in the node (upper row) and spectral (lower row) domains. The red lines represent the addition of new edges to the original graph in (a). The eigenvalues in (d), (e), and (f) are sorted in order of increasing magnitude.

Graph Filtering

The knowledge about whether the graph is smooth or not helps to perform efficient computations by applying graph filtering (Shuman et al., 2013). It has been shown in (Zhu and Rabbat, 2012) that smooth signals have compressible Fourier coefficients. This is because the sorted magnitude of Fourier coefficients exhibits power-law decay; thus, the largest coefficient can be used to approximate the signal. Using this fact, we can filter the original graph signal $\mathbf{x}$ by extracting signal components corresponding to different frequencies as follows:

$$\tilde{\mathbf{x}} = \mathbf{U}\mathbf{H}\mathbf{U}^{\top}\mathbf{x}, \tag{3.5}$$

where $\mathbf{H}$ is a diagonal filter matrix. Specifically, we can define low-pass filtering for k lowest graph frequencies, by setting diagonal elements of $\mathbf{H}$ to 1 for the first k eigenvalues and 0 otherwise. This is equivalent to using only

the first k eigenvectors:

$$\tilde{\mathbf{x}}_k = \mathbf{U}_{[0:k]}\mathbf{H}\mathbf{U}_{[0:k]}^{\top}\mathbf{x}. \tag{3.6}$$

3.4.3 Analysis of Population Graphs using GSP

Evaluation of Fourier Transform Coefficients

In this section, we investigate four groups of constructed population graphs, defined in Section 3.2.2. Using GFT decomposition, we calculate the magnitude of graph frequency coefficients $\hat{\mathbf{x}}$ to understand which graph frequency components contribute most to the signal $\mathbf{x}$. Figures 3.6 and 3.7 show how the decomposed signals are distributed across each of the following population graphs: *sim_RSFC*, *sim_RSFC_site*, *sim_RSFC_sex*, *FC*, and *random*. We plot the magnitude of GFT coefficients computed for one feature of RSFC-based feature vector $\mathbf{x}$. Clearly, the smoothness of a signal depends on the underlying graph structure. In *sim_RSFC*, *sim_RSFC_site*, *sim_RSFC_sex* (Figure 3.6) the contribution is the highest from low-frequency components. On the other hand, *FC* and *random* graphs are not smooth (Figure 3.7), as the values of frequency components fluctuate around zero, and the contribution from all frequency components is relatively equivalent and small.

Classification using Low Frequency Components

In the previous section, we discovered that some of the graphs exhibit a low-frequency nature. Smooth graph signals decay rapidly and can be closely approximated by graph Fourier coefficients (Shuman, Ricaud, and Vandergheynst, 2016). Also, taking the first k eigenvectors is proven to be efficient for graph clustering purposes (Shuman et al., 2013). Therefore, by incrementally adding more than one smooth eigenvector, we can improve the performance of graph clustering (Ng, Jordan, and Weiss, 2001). We use this fact to test how low and high-frequency components in different graphs contribute towards the accuracy of subject classification. To identify the range of the low best-performing graph frequencies, for each graph configuration, we construct a multi-layer

feedforward neural network (with two hidden layers of size 512 and 64, separated by the ReLU activation function). We train and evaluate the performance of the model on the subject classification task as follows:

1. Using equation (3.6), we filter the graph signal incrementally using the first k-frequency components.

2. We train a multi-layer feedforward neural network on the reconstructed features and report test accuracy.

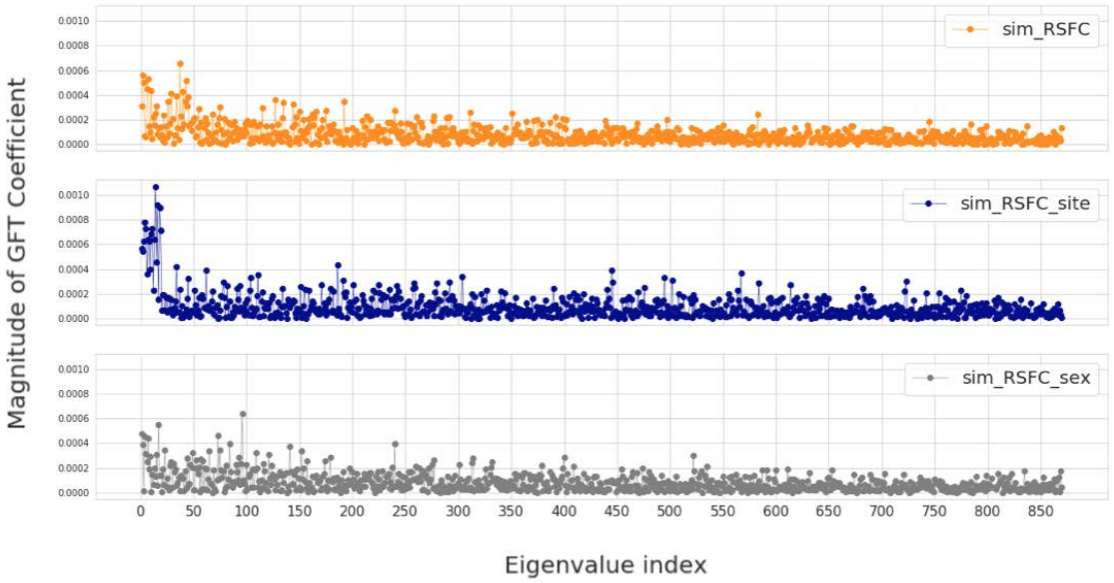


FIGURE 3.6: The magnitude of GFT coefficients for graphs which exhibit low-frequency nature. The eigenvalues are sorted in order of increasing magnitude.

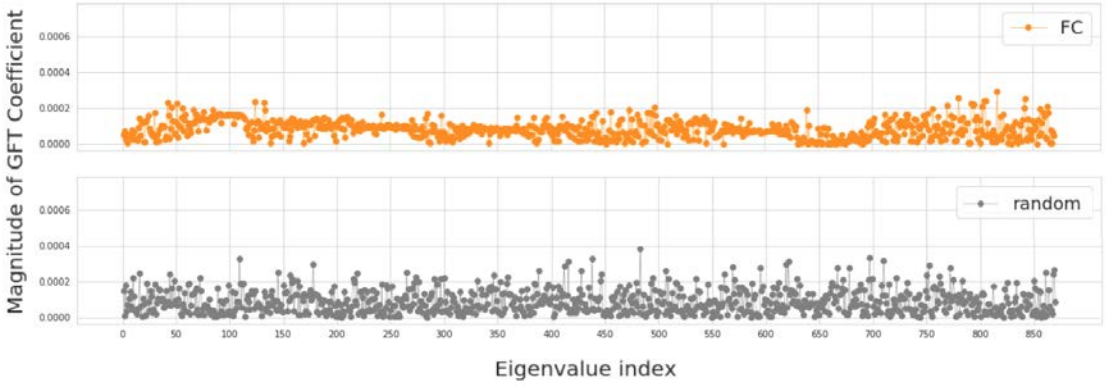


FIGURE 3.7: The magnitude of GFT coefficients for graphs that do not exhibit low-frequency nature. The eigenvalues are sorted in order of increasing magnitude.

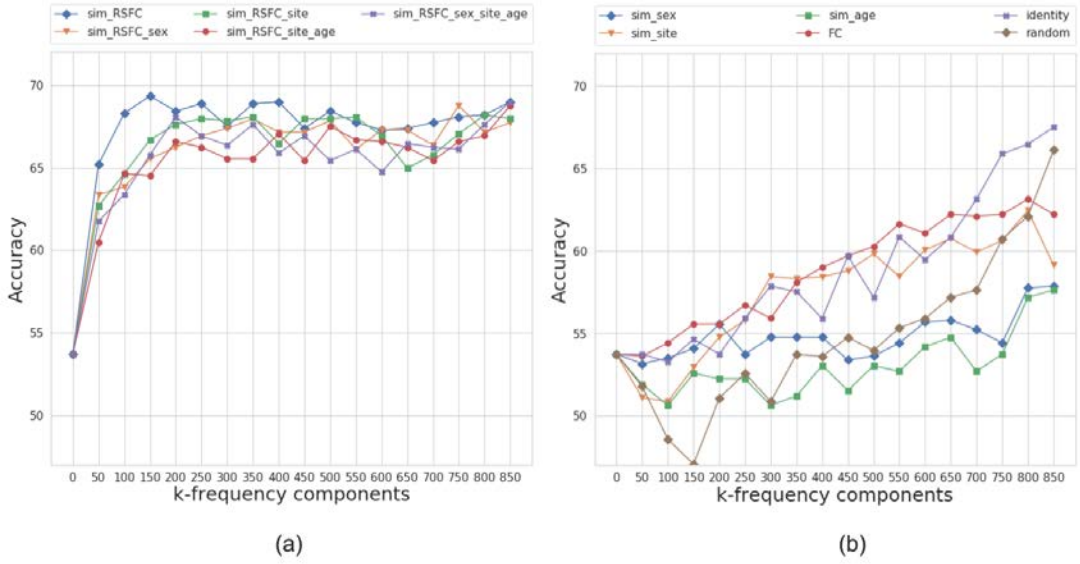


FIGURE 3.8: The average accuracy of the classifier over ten folds on frequency-filtered feature vectors for the ABIDE dataset. The plot shows the cumulative contribution of frequency components to the classification accuracy. Based on the performance of models at different frequency regimes, we split the graphs into two categories, i.e., those that yield higher classification accuracy at (a) low or (b) high frequencies. The eigenvalues corresponding to k-frequency components are sorted in order of increasing magnitude.

Figure 3.8 presents the performance results of the multi-layer feedforward neural network model for different graph configurations. Based on the models' performance at different frequency regimes, we split the graphs into two categories: those that yield higher classification accuracy at either low or high frequencies.

For the ABIDE dataset, the best-performing graph is *sim_RSFC*, i.e., the graph constructed solely using RSFC data. We classify it as a low-frequency graph type since the maximum average accuracy of 69.8% is achieved using only the first 150 frequency components (first 20% of graph spectrum). The graphs constructed using the combination of RSFC, sex, site, and age (*sim_RSFC_site*, *sim_RSFC_sex_site_age*, *sim_RSFC_sex*, and *sim_RSFC_site_age*) also showed low-frequency nature of the graph signal. For all these graphs, the model achieves an accuracy of over 66% on its first 200 frequency components (Figure 3.8 (a)).

On the other hand, the performance of *identity*, *FC*, *random* graphs and

graphs from *sim_phenotype* group exhibits high-frequency nature, as the model requires seeing all graph frequency components to achieve the best performance (see Figure 3.8 (b)). It is worth noting that even after seeing all the frequency components, the top performance of high-frequency graphs is inferior to the performance of low-frequency graphs. We attribute this to the fact that smooth graphs that use RSFC features capture the underlying structure of the population graphs better than those that do not utilize graph structure or rely only on non-fMRI features.

We analyze the role of imaging resting-state fMRI and non-fMRI (phenotype) features on the graph framework. Based on the classification accuracy results presented in Table 3.8 graphs build with different combinations of fMRI and non-fMRI features, we conclude that non-fMRI features alone (shown in the right-hand side by baselines *sim_site*, *sim_age*, and *sim_site*) do not contribute significantly, toward the model performance. For example, non-fMRI features such as site information can improve the accuracy only up to 58%. fMRI features (and combinations using fMRI features) contribute more towards improved accuracy. The most optimal population graphs are the ones that utilize fMRI features or a combination of fMRI and non-fMRI features. These population graphs also performed best at low frequencies, unlike graphs without fMRI features.

Population Graph-based Multi-model Ensemble for ASD Prediction

In many cases, the choice of population graph can be problem-dependent, and due to discrepancies of heterogeneous data (such as coming from different sources or having unique features), no population graph can be considered entirely accurate. Thus, it is reasonable to consider the performance of multiple population graphs. For this purpose, we propose a multi-model ensemble, which integrates multiple best-performing population graphs to yield a better classification performance (Rakhimberdina, Liu, and Murata, 2020). The model consists of two stages. The schematic representation of the ensemble model is shown in Figure 3.9.

In the first stage, we select eight best-performing low-frequency graphs and, along with RSFC features, pass each of them through frequency filtering feedforward neural network described in Section 3.4.3. Each of these networks can be considered as a learning block, which learns a lower-dimensional representation of input graphs.

In the second stage, we create an ensemble by combining the predictions of each of the eight learning blocks. The ensemble model is implemented as a fully-connected layer with a learnable weighting mechanism. It takes as input $n \times 512$ feature vector p, which is composed of eight stacked $n \times 64$ hidden representations p_i from the first stage. For each sample in the test set of size n, the final prediction is a binary $n \times 1$ vector, with 0 representing a healthy subject and 1 representing a patient diagnosed with ASD. We employ a weighting mechanism to assign importance to each individual learning block, corresponding to different input graphs. The classification results of eight models for each subject in the test set are combined by passing them through a softmax layer. For each hidden representation p_i from the first stage, the corresponding weight α_i is calculated according to the following equation, where w is a randomly initialized scoring vector:

$$\alpha_i = \frac{\exp(w_i)}{\sum_j \exp(w_j)}. \tag{3.7}$$

After that, the weighted sum of the individual prediction blocks $H = \sum_i \alpha_i p_i$ is passed into the sigmoid layer:

$$\hat{y} = sigmoid(WH + b), \tag{3.8}$$

where W and b are weight and bias matrices of fully-connected layers, respectively. Finally, the loss is calculated using the cross-entropy error over the labeled samples:

$$\mathcal{L} = -\sum_c Y_c \log\left(\hat{Y}_c\right) \tag{3.9}$$

where Y_c and $\hat{Y}_c$ is the actual and the predicted labels.

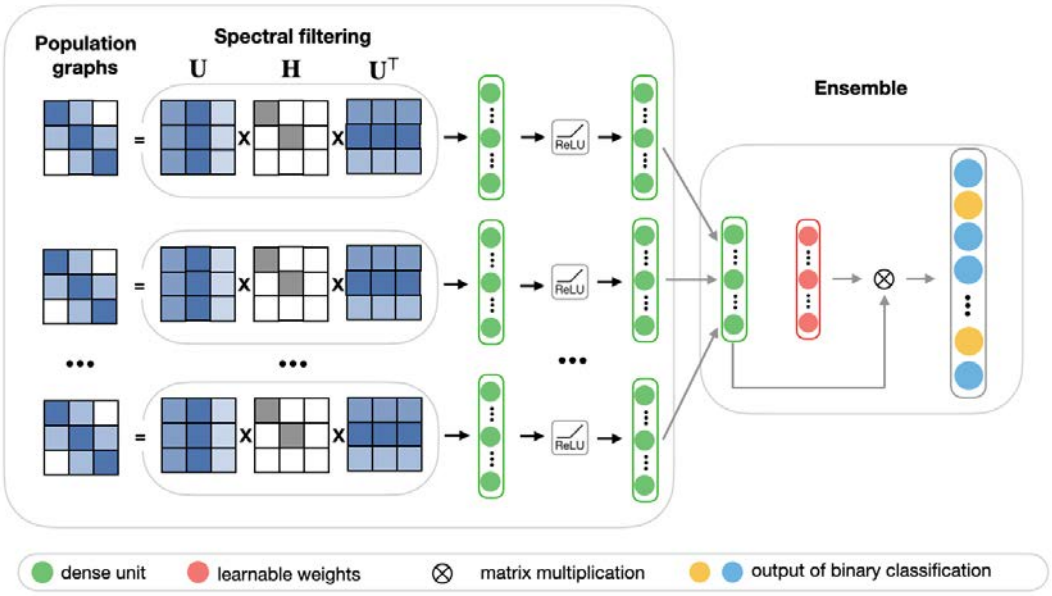


FIGURE 3.9: The schematic representation of the ensemble model.

3.5 Experimental Settings

3.5.1 Baseline Methods

We establish a baseline by comparing our method against several competitor methods developed for ASD detection. These methods are divided into two categories: non-graph methods and graph-based methods.

Non-graph Methods

- **CNN** (Sherkatghanad et al., 2019) proposed a convolutional neural network architecture to classify ASD patients and control subjects using RSFC. The authors reported a classification accuracy of 70.22% and Area Under the Receiver Operating Characteristic Curve (AUC) of 0.75 on a subset of 871 subjects from the ABIDE dataset.

- **ASD-DiagNet** (Eslami et al., 2019) increased the computational speed of autoencoder based DNN (Heinsfeld et al., 2017) by introducing a hybrid learning procedure. Similarly, only RSFC features were used as the input to the model.

- **Ensemble_mv** (Kam, Suk, and Lee, 2017) utilized multiple discriminative restricted Boltzmann machines (DRBM) to classify 263 subjects from the ABIDE dataset. The majority voting strategy was used for combining the prediction outputs of individual DRBMs.

Graph-based Methods

- **GCN** (Parisot et al., 2018) Graph Convolutional Network for classifying ASD (Parisot et al., 2018). The authors were among the first to propose an application of Graph Convolutional Networks (GCN) on population-based brain disorder classification task. In this research, authors combined imaging (fMRI scans) and non-imaging (phenotypic data) data by representing subjects as nodes, and pairwise similarities between them as edges. Further, the authors formulated the task of subject classification as node labeling over the population graph.

- **Ensemble_bootstrap**. Anirudh and Thiagarajan, 2019 proposed a bootstrapping approach by generating twenty randomized graphs from the initial population graph. Each randomized graph was passed through a graph neural network. The final output of the ensemble was determined by averaging the probability estimates of each of the networks.

3.5.2 Training Setup

Due to the limited size of the ABIDE dataset and for the sake of fair comparison with baseline methods, we used ten-fold stratified cross-validation, with the stratification criterion being an equal allocation of the class labels: patients with ASD and healthy subjects. The hyperparameters for each graph-based model in the first stage (the number of layers, the learning rate, dropout rate, etc.) were tuned using the Optuna hyperparameter optimization framework (Akiba et al., 2019). We trained a graph-based model on 200 epochs with Adam (Kingma and Ba, 2015) optimizer and learning rate = 0.01. We fixed the graph filtering frequency threshold to $k = 200$. All models were developed using the open-source machine learning library PyTorch and trained

on NVIDIA GeForce GTX TITAN X GPU with 12GB memory size and CPU Intel(R) Core(TM) i7-7700K CPU 4.20GHz.

3.5.3 Evaluation Metrics

When evaluating the performance of a model, we use accuracy, sensitivity, specificity, and AUC (area under the receiver operating characteristic curve) metrics:

1. Accuracy represents the ratio of correct predictions. The sum of a true positive and a false negative is divided by the total number of events. Accuracy is useful when the classes in the dataset are nearly balanced. One of the primary advantages of accuracy is its simplicity and interpretability. Accuracy directly represents the proportion of correct predictions out of the total number of instances. It's an intuitive metric that's easy to understand and can be used to explain the results to a general audience.

2. Sensitivity, or recall, measures the ratio of predicted positive classes. The number of true positive events is divided by the sum of true positive and false negative events.

3. Specificity or True Negative Rate measures the rate of actual negatives correctly identified.

4. AUC stands for "area under the curve" and captures the area under the ROC (Receiver Operating Characteristic) curve and compares the relationship between the true positive rate with the false positive rate across different cut-off thresholds. The metric is useful when datasets are imbalanced, because it is not affected by the proportion of samples at each class. Even for balanced binary datasets, there are still advantages of AUC over accuracy. First, the AUC considers predictions across all thresholds, not just the default threshold of 0.5 used by accuracy. This can be useful if we want to specify a threshold in case when false positives and false negatives have different costs. Second, AUC

score is not affected if distribution of classes in the test set is different from the distribution in the training set, unlike accuracy. The model trained on a balanced dataset, might be used to make predictions on unbalanced datasets in the future, in that case, AUC-ROC could be a better measure of performance.

It is important to consider the application of the ASD detection method when choosing the evaluation metric. In the context of autism spectrum disorder prediction, sensitivity, and specificity are important metrics for evaluating the accuracy of a diagnostic test. A high sensitivity value is important for avoiding false negatives, which could lead to missed diagnoses and delayed treatment. On the other hand, a high specificity value is necessary for avoiding false positives, which could lead to unnecessary treatments.

3.6 Results and Discussion

In this section, we evaluate the performance of our SGC framework and graph-based ensemble model in comparison with the baselines.

3.6.1 Results for SGC Model

Comparison with baseline models. To verify the proposed graph construction process and efficiency of the SGC-based model, we compare our method with several baselines. The first two baselines are our PyTorch implementations of the GCN-based population model and the competitive GCN-based model, proposed by Parisot et al., 2018. These baselines were introduced to test the performance of the SGC model over other graph convolutional models. Additionally, we introduce three baselines, *Graph_random*, *Graph_identity*, and *Graph_no_features* in order to evaluate graph construction and the importance of subjects' imaging and non-imaging features in the classification task. Underlying these three baselines is the same SGC model running on the same

TABLE 3.3: Test accuracy(%) and training time per one fold (seconds) of baselines and the proposed model. Numbers are averaged over 10 runs. GCN denotes our PyTorch implementation of Parisot et al., 2018. The best results are highlighted in bold.

Model	Test accuracy	Time
GCN (Parisot et al., 2018)	67.53±4.55	58.22±0.56
Graph_no_features	50.50±3.79	1.15±0.02
Graph_random	67.84±4.80	1.29±0.05
Graph_identity	65.34±5.30	1.28±0.04
Ours	68.56±4.33	1.28±0.04

population set. More specifically, *Graph_random* and *Graph_no_features* baselines are designed to test the meaningfulness of the proposed edge construction method. In *Graph_random*, edges are randomly rewired so that the edge density of the original graph is preserved. *Graph_identity* baseline is introduced to test the meaningfulness of fMRI features regardless of non-imaging phenotypic data. This means that each node is still associated with the fMRI feature vector, however, the identity matrix is used in place of the adjacency matrix to represent a completely disconnected graph, i.e. no two nodes are connected. On the other hand, *Graph_no_features* baseline is implemented to test the performance of the model with respect to graph structure acquired using phenotypic features.

We chose 10-fold cross-validation to make a fair comparison with the baseline method Parisot et al., 2018. For GCN implementation on ABIDE, the model parameters were chosen according to Parisot et al., 2018. For other cases, hyperparameters for GCN, and SGC were tuned using grid search. For the ABIDE dataset, we trained the model on 2,000 epochs with Adam optimizer, learning rate = 0.1, and propagation step $K = 2$. All models were trained on NVIDIA GeForce GTX TITAN X GPU with 12GB memory size. The accuracy values of all baselines across three fMRI datasets using stratified cross-validation are presented in Table 3.3.

Evaluation. Based on the results presented in Table 3.3, we conclude that our model is competitive with other models and baselines. For example, the

classification accuracy of our linear model on the ABIDE dataset is comparable to the performance of GCN based method proposed by Parisot et al., 2018 (68.56% vs. 67.53%). Moreover, the linear model achieves the best performance when compared to *Graph_random*, *Graph_identity*, and *Graph_no_features* baselines, which justifies the usefulness of the constructed population graphs based on phenotypic features. The consistently higher accuracy of our model over other baselines across three datasets shows its robustness in capturing the population graph structure and extracting useful features.

In addition to accuracy improvement, by using a simplified linear graph convolutional model we were also able to achieve a better performance in terms of training time. In the last column of Table 3.3, we present the training time for each of the models. Our model shows up to 50 times a speedup on the ABIDE dataset compared to the GCN model. We attribute the performance boost of our SGC-based model to a significantly lower number of parameters the model has to learn and the linearity of the learning function. This makes our population-based model lightweight and allows it to run on the CPU. In addition, the model is able to scale up to larger datasets, which is especially important given the increasing volume of the fMRI imaging data.

3.6.2 Results for Multi-model Ensemble

Comparison with baseline models. For fair evaluation, we compare the proposed method using the same evaluation procedure (10-fold stratified cross validation) and the same subset of subjects (871). First, we compare our approach with single-model competitive methods presented at the top of the Table 3.4. These models convolutional neural networks CNN (Sherkatghanad et al., 2019), ASD-DiagNet (Eslami et al., 2019) which use only RSFC features as an input and graph-based GCN model (Parisot et al., 2018), which uses both RSFC and phenotypic features. In addition to single-model approaches, we compare our method with recent multi-model ensembles on ABIDE dataset, which are Ensemble_bootstrap (Anirudh and Thiagarajan, 2019) and majority voting-based ensemble method Ensemble_mv used in

(Kam, Suk, and Lee, 2017). To provide a fair comparison, we ran all the baseline models on the same subset of subjects used in (Parisot et al., 2018), consisting of 403 patients with ASD and 468 healthy individuals.

Additionally, we introduce three baselines: *No graph*, *Random graph*, and fully connected *FC graph* to evaluate the meaningfulness of the proposed edge construction method. *No graph* baseline evaluates the contribution of fMRI features alone, regardless of the underlying population graph. This means that the underlying graph becomes completely disconnected, i.e., no two nodes are connected. In *Graph random*, edges are randomly rewired. *FC graph* baseline is implemented so that all the connections in a population graph are equal to one.

In terms of accuracy results, we observe that our ensemble model achieves highest average accuracy of 73%. Unlike other methods, our method's accuracy (0.73 ± 0.04) is significantly higher than the accuracy by *No graph* baseline (0.68 ± 0.04) (using t-test with confidence level 95%, $p<0.05$). The higher accuracy of our model over *No graph* baselines shows its robustness in capturing the population graph structure and extracting useful features. In terms of AUC, our model also achieves the highest score, performing as well as CNN method by Sherkatghanad et al., 2019.

TABLE 3.4: Classification results of baselines and ablation studies. Top: non-graph methods, middle: graph-based methods, bottom: our proposed ensemble. The results for Kernel Regression (He et al., 2018), GCN (Parisot et al., 2018), and Ensemble_bootstrap (Anirudh and Thiagarajan, 2019) were calculated using our implementation. The results are using the same procedure for evaluation and the same subset of subjects. * results of methods were calculated using our implementation.

	Accuracy	AUC	Sensitivity	Specificity
CNN(Sherkatghanad et al., 2019)	0.70±0.09	0.75±0.01	0.77±0.01	0.61±0.01
ASD-DiagNet (Eslami et al., 2019)*	0.70±0.05	0.67±0.23	0.68±0.06	0.72±0.07
GCN (Parisot et al., 2018)*	0.69±0.04	0.73±0.05	0.75±0.06	0.57±0.05
Ensemble_bootsrap (Anirudh and Thiagarajan, 2019)*	0.68±0.04	0.73±0.04	0.75±0.05	0.59±0.05
Ensemble_mv (Kam, Suk, and Lee, 2017)*	0.66±0.04	0.72±0.04	0.68±0.07	0.63±0.09
No graph	0.68±0.04	0.74±0.05	0.76±0.03	0.31±0.04
Random graph	0.66±0.05	0.74±0.05	0.86±0.10	0.25±0.17
FC graph	0.62±0.06	0.66±0.04	0.86±0.07	0.34±0.10
Ensemble_gsp (ours)	0.73±0.04	0.75±0.04	0.76±0.07	0.69±0.05

Optimizing balance between sensitivity and specificity. Since the class labels are relatively balanced in the ABIDE dataset, the common approach in the field is to use the probability output of a classification model thresholded at 0.5 for computing accuracy. That means that the predicted values above 0.5 are classified as positive and predicted probabilities below 0.5 are classified as negative. It's important to note that the choice of the threshold will depend on the specific context and goals of the analysis, and different thresholds may be appropriate for different applications. To find a balance between sensitivity and specificity, we adjust the classification threshold to prioritize high sensitivity and specificity.

To select the optimal threshold for a binary classifier based on the receiver operating characteristic (ROC) curve we use the J statistic as a metric (Youden, 1950). It is defined as $J = sensitivity + specificity - 1$ and ranges from 0 to 1, with higher values indicating better performance. The J statistic can be used to identify the threshold that maximizes the sum of sensitivity and specificity, which is often considered to be the optimal operating point for a binary classifier.

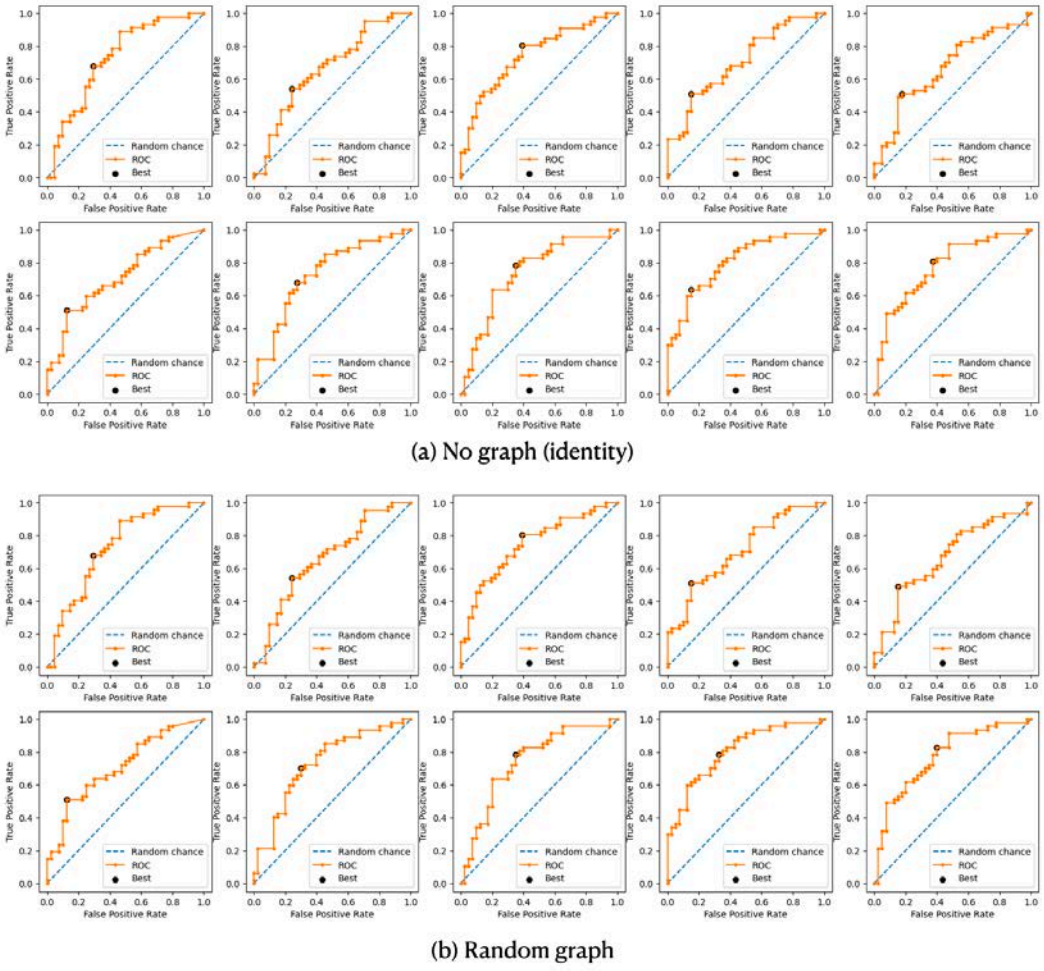


FIGURE 3.10: ROC line plot for No graph (identity) and Random graph ablation methods. Best denotes the optimal threshold value.

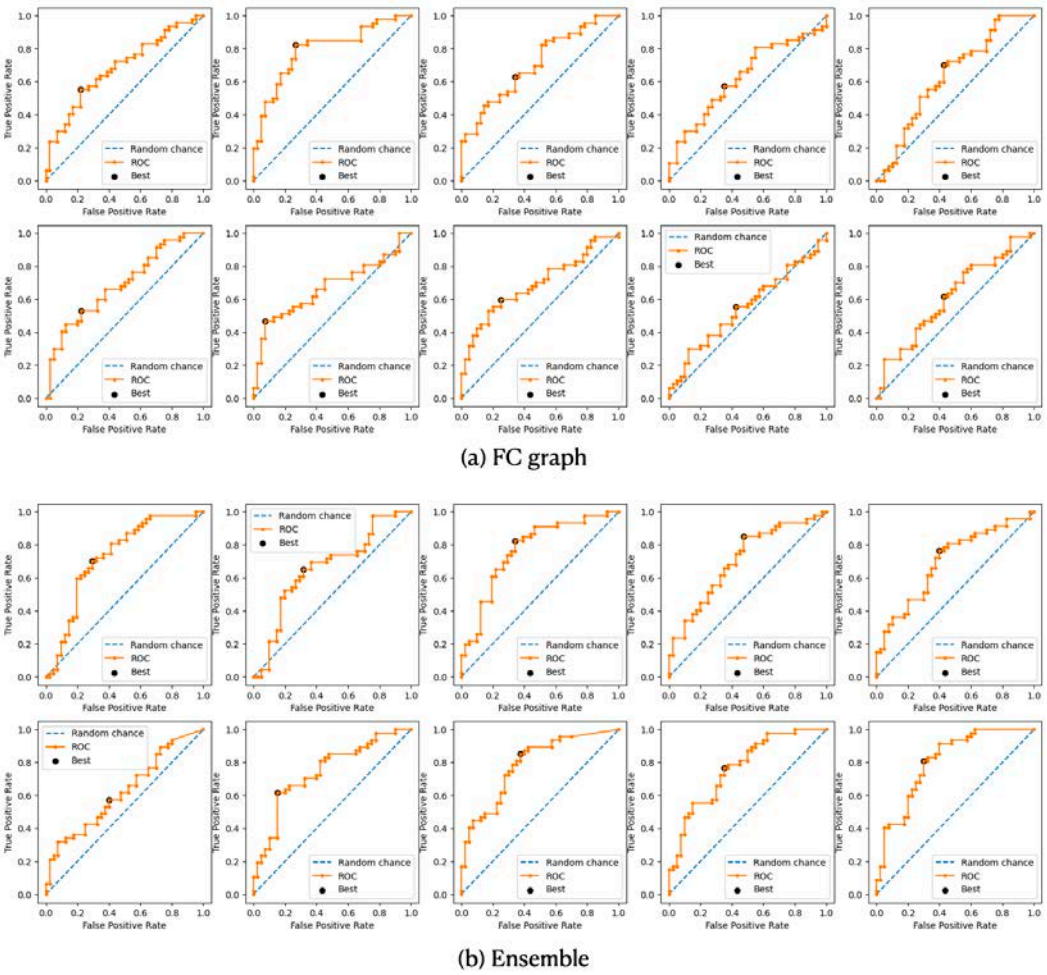


FIGURE 3.11: ROC line plot for FC graph and Ensemble_gsp (ours) methods. Best denotes the optimal threshold value.

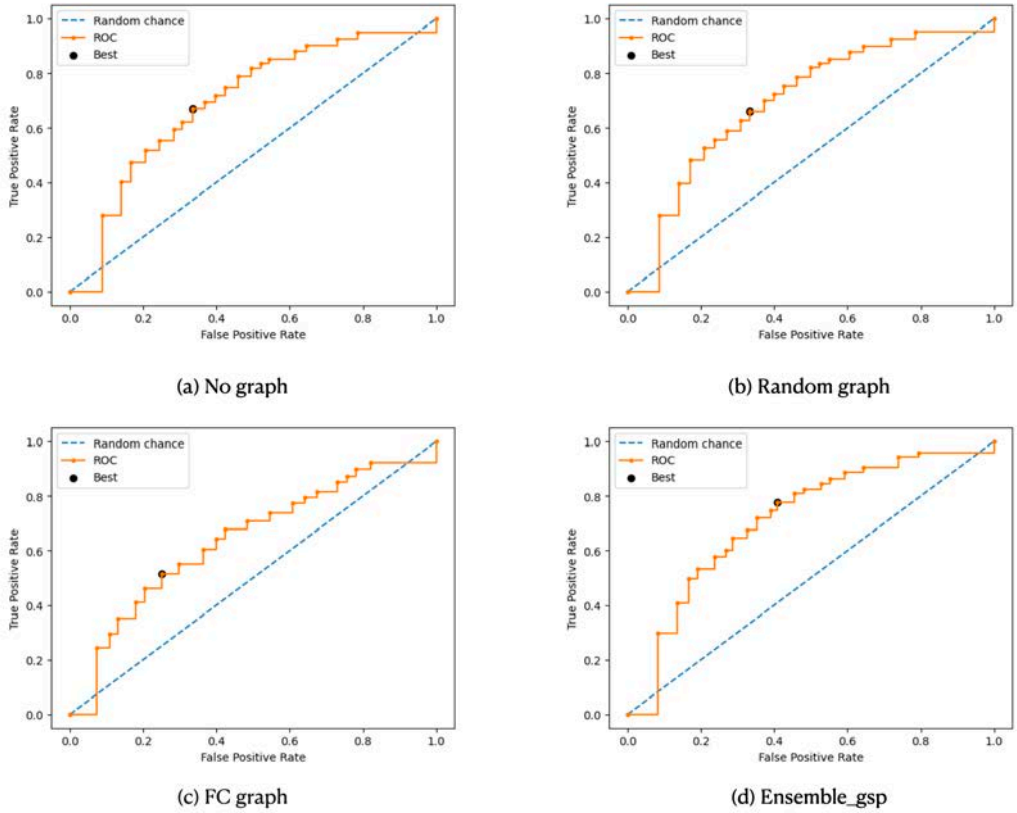


FIGURE 3.12: ROC line plot for optimized threshold values. Best denotes the optimal threshold value.

TABLE 3.5: Classification results with optimized balanced between sensitivity and specificity values in ablation study.

	Accuracy	AUC	Sensitivity	Specificity
Random graph	0.66±0.03	0.74	0.66±0.05	0.67±0.00
FC graph	0.64±0.05	0.66	0.51±0.09	0.75±0.00
No graph	0.68±0.03	0.74	0.66±0.05	0.67±0.00
Ensemble_gsp (ours)	0.73±0.04	0.75	0.75±0.07	0.69±0.00

As can be seen from Figures 3.10 and 3.11, the optimum threshold value changes when computed on 10 different splits used in the stratified cross-validation evaluation. To optimize the sensitivity and specificity values we first define a range of thresholds between 0 and 1 and then compute 10-fold cross-validation for fixed threshold values. In Table 3.5 and Figure 3.12, we present the optimal result based on balancing sensitivity and specificity. As a

result of threshold optimization, we achieved balanced sensitivity and specificity values, which are not too high and not too low. However, balancing did not affect the accuracy results because the dataset is balanced.

Discussion. By definition of the edge assigning function, the population graph is being constructed in such a way that neighboring nodes have similar features (based on age, sex, site, and RSFC). If we look from eigendecomposition, then eigenvectors will be the smoothest (those that correspond to low eigenvalues) in the case when similar features lie on the neighboring nodes. Thus we are interested in the case when the graph definition reinforces the smoothness property. In a graph with a smooth signal, the unknown subject will have a higher chance of being labeled healthy if connected to many healthy subjects. When we incorporate other non-imaging features to define edges in the graph, we want to reinforce the signal smoothness. High-frequency components are more likely to increase the contribution of signal noise that arises from edges connecting nodes with dissimilar features. Therefore, by analyzing the performance of the model on low-frequency components, we can distinguish between features that contribute towards the signal smoothness and improve the performance, and features that reinforce edge weights between dissimilar nodes and, thus, introduce the noise.

The ensemble approach we proposed in this work improves the performance by setting a frequency filtering threshold and efficiently combining weighted contributions from each graph. Our choice of integrating multiple graph-based models into an ensemble was motivated by comparison with several ensembling schemes. Since simple averaging (Anirudh and Thiagarajan, 2019) of the results from individual models and majority voting algorithm (Kam, Suk, and Lee, 2017) did not yield superior performance, we proposed an end-to-end deep neural network-based ensemble model that uses weighting mechanism to assign different levels of importance for each model's prediction.

Sensitivity to frequency filtering. The analysis of the underlying graph property allowed us to choose the best-performing low-frequency graphs

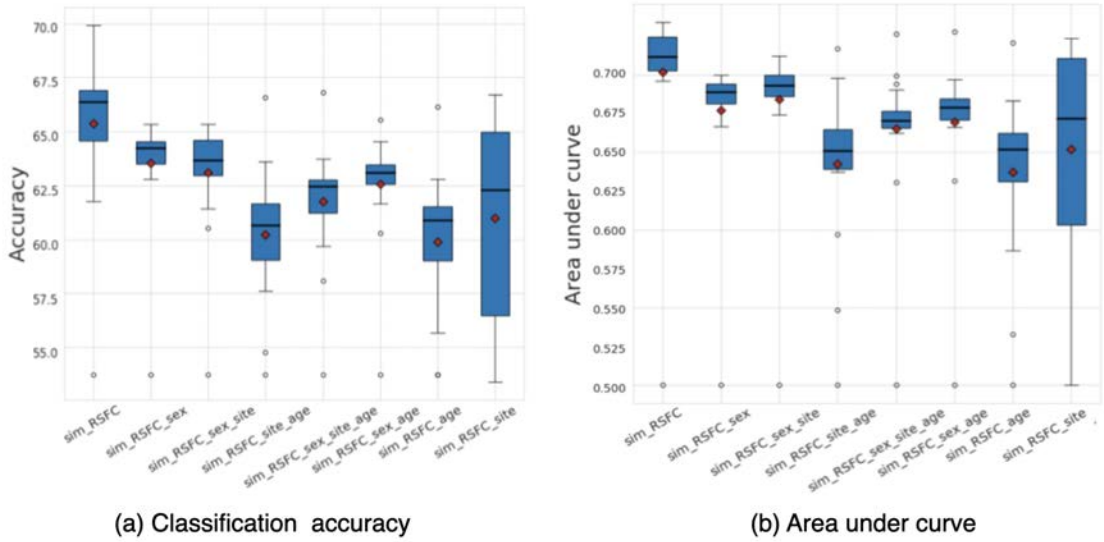


FIGURE 3.13: (a) Classification accuracy and (b) area under curve (AUC) results for each of the eight input graphs run separately for the ABIDE dataset. The mean values are shown as red diamonds.

for building a multi-modal framework. To explore the effect of graph frequency filtering, we present the accuracy and AUC results calculated for different frequency thresholds for each of the eight input graphs in Figure 3.13. The results demonstrate that the most informative graphs are those constructed using RSFC data, which, due to graph construction design, are low-frequency by nature. By analyzing the performance of each of the eight input graphs separately, we can see that the *sim_RSFC* graph alone achieves relatively higher performance than other graphs, with a the highest accuracy of 69.8% and AUC of 71.0%.

We explored the influence of frequency filtering threshold k, which we varied from 100 to 900. We found that lower frequency filtering thresholds resulted in improved performance. Figure 3.14 illustrates the box plot for average accuracy and area under curve (AUC) calculated across ten folds on the ABIDE dataset. The model's highest performance is achieved within the range of frequency filtering threshold between 200 and 400. The model's best performance with an accuracy of 73.13% and AUC 0.75 is achieved when the frequency filtering threshold equals 200.

Based on the analysis, we build a multi-model classification ensemble

that focuses on low-frequency population graphs. The results show that the proposed ensemble improves prediction performance compared to the single-model case by 3.33%, and presents robustness to the choice of the input graphs.

In this work, the signal filtering parameter k is fixed to be the same for all of the selected graphs. Based on our model's performance on different k-frequency components, we specify k to be 200. Determining the optimal k for each graph configuration can lead to improved performance of the ensemble overall and is yet to be explored in future work. For example, an attention mechanism can be explored to learn the best-performing k for each graph configuration. Another potential future work can focus on learning the population graph structure so that the input data forms graph signals with smooth variations (Dong et al., 2016; Dong et al., 2019; Kalofolias, 2016).

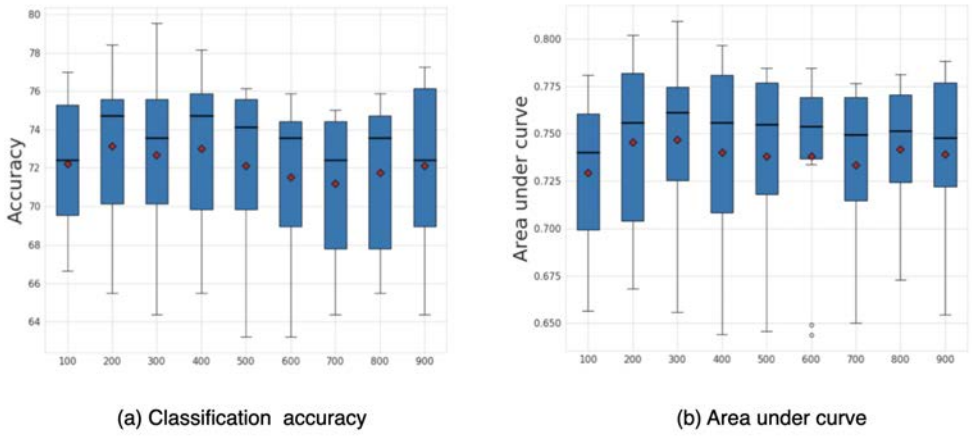


FIGURE 3.14: Sensitivity analysis for k. Average accuracy and AUC results with respect to frequency filtering threshold k. The mean values are shown as red diamonds.

3.6.3 Subsequent Research Trend

Subsequent research on the topic of ASD prediction using fMRI data is summarized by Ahmedt-Aristizabal et al., 2021. Here we outline the most relevant to the scope of this work.

Kazi et al., 2019 proposed Inception-GCN (Figure 3.15), a spectral domain architecture for deep learning on graphs. Unlike traditional GCNs which

force the features of every node to be learned using neighbors at a fixed number of hops away via constant filter size, Inception-GCN uses filters with different kernel sizes. According to the authors, different kernel size help capture intra- and inter-graph structural heterogeneity in the graph. The best-reported accuracy is $70.26 \pm 04.58\%$.

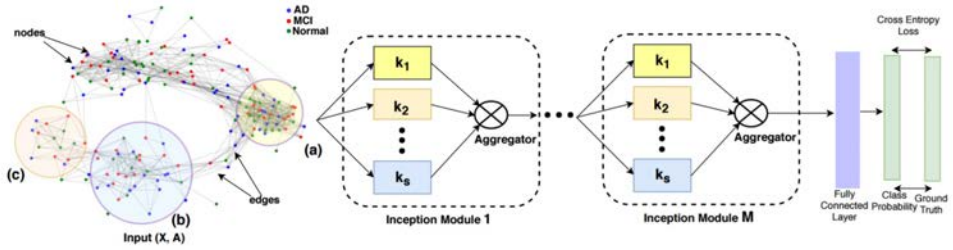


FIGURE 3.15: Different cluster sizes are depicted at points (a), (b), and (c). InceptionGCN framework takes feature matrix X as input and processed it using several GC-layers with specified neighborhood kernels $k_1, \ldots, k_S$. Image is taken from Kazi et al., 2019.

In Jiang et al., 2020, the authors build upon a graph neural network to learn useful representations for graph classification in an end-to-end manner. The authors propose a hierarchical GCN framework, called hi-GCN, to learn the graph embedding which captures both network topology information and the subject's interconnectivity. The authors achieved higher classification accuracy of 73.1% and AUC score of 0.82 for the subset of ABIDE dataset containing 402 ASD patients and 466 control subjects.

Finally, Li et al., 2020a proposed a GCN algorithm to discover biomarkers of ASD from task-fMRI. Different from the spectral GCN algorithm Kipf and Welling, 2017, the GCN classifier is based on graph isomorphism and can interpret graphs with different nodes and edges. Specifically, the training of the GCN model is performed on the whole graph while testing - on sub-graphs. In Li et al., 2020b, the authors also investigated the brain region related to a neurological disorder from task-fMRI data and modeled the whole brain fMRI as a graph. This helped to preserve the temporal and structural information and learn a better graph embedding. Their results indicated a more robust classification of ASD.

The value of this research. Our discovery of the simplified and well-performing graph neural network contrasts with the existing literature, where the general trend is toward more complex neural network architecture. We suppose that more complex deep learning architectures are more suited to datasets with significantly higher amounts of training data with a high signal-to-noise ratio. Therefore, the main contribution of this research is presenting a community with a simple interpretable linear graph-based model and explaining why graph structure provides a boost in performance through the analysis of graph spectral features.

3.7 Conclusion

In this chapter, we first propose a linear model for the problem of population-based disease prediction that utilizes the novel concept of simplified graph convolutions. We improved upon previous works by designing simple linear end-to-end graph convolutional model. The experiments we conducted demonstrate the comparable performance of the proposed linear graph-based model to complex multilayer graph-based models, confirming our initial assumption of unnecessary complication of graph convolutional models. Due to the improved computational efficiency, we encourage a simple linear model like SGC to be used as a baseline for comparison with other graph-based models involving deep neural network architecture.

Next, since different combinations of RSFC and phenotypic features imply different graph structures, this affects the prediction results. Therefore, choosing a particular graph definition over the other has remained a problem. Next, we address this issue by integrating multiple graph-based prediction models into an ensemble. First, using graph filtering, we selected the best-performing graphs. Second, we combine multiple graph-based models to construct a more powerful ensemble for detecting Autism Spectrum Disorder. The performance on the ABIDE dataset achieved 73% in terms of accuracy and 0.75 in terms of AUC score.

Chapter 4

Task-based fMRI for Classifying Autism Spectrum Disorder

4.1 Introduction

In Chapter 3, we introduced two graph-based neural networks to discriminate control subjects from patients with ASD. In short, the model uses brain imaging information in the form of resting-state fMRI and uses relationships between subjects' phenotypical information (such as age and cite information) to build a population graph. By utilizing both imaging and non-imaging data, the model outperforms other related methods using resting-state fMRI data based on the number of evaluation metrics.

In this chapter, we introduce another type of brain imaging data – task-based visual fMRI – which finds an application in a range of tasks, from stimuli classification to the study of brain disorders. The first question we ask is: can we generalize the proposed graph-based neural network to new visual task-fMRI data for ASD prediction? We approach the above question by applying the linear graph-based prediction framework introduced in Chapter 3 to task-fMRI data for the ASD study. Specifically, we extract brain fMRI signals from highly informative visual areas, which activate when a visual stimulus is presented and capture visual information of different complexity. In addition to the proposed graph-based framework, we conduct experiments using different neural network types and ablations with different model configurations. **Our contributions can be summarized as follows:**

- For the first time, we explore a graph-based framework for detecting ASD from fMRI collected in visual stimuli repetition task.
- We run experiments on multiple graph configurations to demonstrate the contribution of each of the components, i.e., graph neural network and the task fMRI, we show the results as part of ablation study.
- We experimentally show the benefit of using task fMRI data in increasing the training sample size, which subsequently improves the model's performance.
- Finally, we explore modern DNN architectures that can efficiently learn from visual task-based fMRI data.

The rest of the chapter is organized as follows. Section 4.2 focuses on the prediction of ASD using visual task-based fMRI data. Section 4.2.1 describes the application of visual task fMRI for ASD detection and the related works on this topic. In Section 4.3, we explore modern DNN architectures that can efficiently learn from visual task-based fMRI data. Finally, Section 4.4 summarizes the chapter.

4.2 Prediction of ASD using Visual Task-based fMRI Data

4.2.1 Application of Task-based fMRI for ASD Study

Although atypical social behavior remains a core feature of ASD, unusual characteristics of perceptual processing have recently been given more focus. These characteristics are usually defined by the tendency of people with ASD to focus on local details rather than a global picture. We refer to Figure 4.1 for a metaphoric example of how social interaction could be misprocessed in ASD. Visual information processing in people with ASD is different from typical perceptual processing, where the processing stages are organized in time so that they move from global structuring to the analysis of finer visual

FIGURE 4.1: Metaphoric example of how social interaction could be "misprocessed" in ASD from Kessler, Seymour, and Rippon, 2016. The image on the left shows how individuals with typical neurological development systems process a sequence of movements into meaningful chunks. Kessler, Seymour, and Rippon, 2016 hypothesize that patients with ASD system experience a "blended" sequence when elements cannot be effectively separated into meaningful chunks.

details. Recent studies have also suggested that patients with autism perform better than healthy controls on visual search tasks and have superior visual discrimination ability (O'riordan, 2004). This attention to local detail in ASD is often accompanied by a pronounced deficit in global processing (Kessler, Seymour, and Rippon, 2016).

Several research works have empirically demonstrated that the visual information processing of individuals with ASD is significantly different from those of normally developing individuals. This difference is revealed externally in visual eye fixation to the stimuli in fixation time, fixation latency (Gong et al., 2021). In addition, compared to other typically developing (TD) groups, individuals with ASD prefer to attend to highly repetitive images, such as repeating geometric shapes. The differences in responses to visual stimuli have also been captured by studies using task visual fMRI. In particular, Utzerath et al., 2018 demonstrated different neural responses to stimulus expectation and repetition. The authors observed a tendency for brain

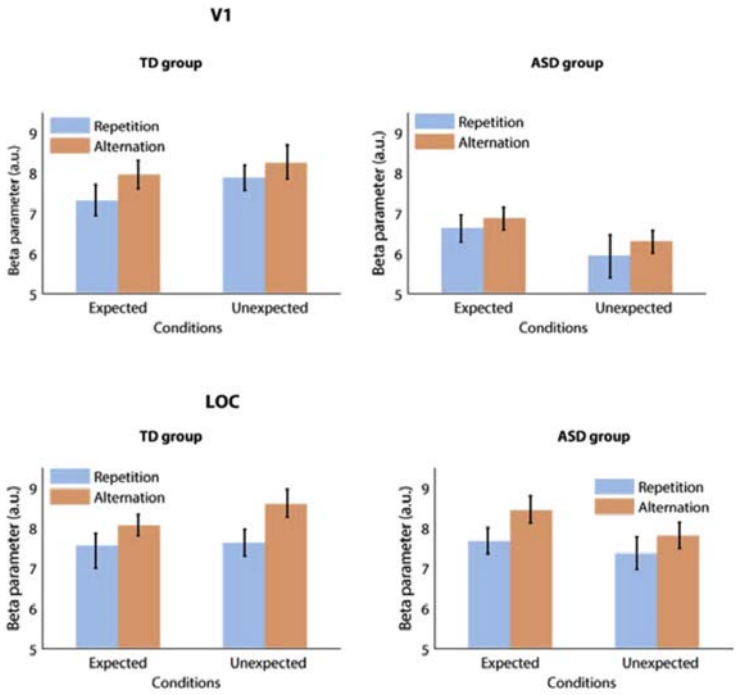


FIGURE 4.2: Beta parameters per group and regions of interest (V1 and LOC). The first row shows responses in V1. The second row shows responses in LOC. The first and second columns represent the TD and ASD groups from the present experiment, respectively. Error bars reflect SEM. The figure is adapted from Utzerath et al., 2018.

activity in visual brain regions (LOC and V1) to increase in response to unexpected stimuli in the TD group, but to decrease in the ASD group (see Figure 4.2).

Visual preference patterns have been widely studied until now, and many frameworks have been proposed to explain differences between expectations and outcomes in ASD that may contribute to the symptoms of the disorder. Therefore our goal in using task data is to explore the potential of using visual task-based fMRI for ASD detection. The present study aimed to utilize the features of the responses to repetitive and alternating visual stimuli for studying ASD. We combine our graph-based classification framework with visual task-based fMRI features. To the author's best knowledge, no study has considered a graph-based framework for classifying ASD patients using visual fMRI data.

4.2.2 Problem Formulation

Since we generalize the method proposed in Chapter 3, we formulate our problem in the same manner as Section 3.2.2. We define the population graph, where nodes represent the subjects from the dataset $V = \{v_1, \ldots, v_n\}$, and edges connecting the nodes represent the similarity between subjects' imaging and non-imaging features. The resulting graph is an undirected weighted graph $G = (V, E, \mathcal{W})$, Each node v_i is associated with a d-dimensional feature vector $\mathbf{x}_i$ extracted from fMRI imaging data. The feature matrix $\mathbf{X} \in \mathbb{R}^{n \times d}$ consists of stacked feature vectors of n nodes in the graph. The set of edges $E \subseteq V \times V$ corresponds to links between the nodes, and $\mathcal{W} : E \mapsto \mathbb{R}$ is a function which assigns weight w_{ij} to each edge as follows:

$$\mathcal{W}(i,j) = sim\left(\mathbf{x}_i, \mathbf{x}_j\right) \sum_{h=1}^{H} \mathbf{1}[M_h(v_i) = M_h(v_j)], \tag{4.1}$$

where $sim\left(\mathbf{x}_i, \mathbf{x}_j\right)$ is defined based on correlation distance between fMRI features, and $M_h(v_i)$ is categorical phenotypic feature value[1] corresponding to node v_i.

4.2.3 Dataset

The visual task-based fMRI dataset we use is quite different from the resting-state ABIDE dataset from Chapter 3 in terms of data acquisition paradigm and the number of participants. The dataset is part of the project for studying hypo-priors in autistic visual perception by Utzerath et al., 2018 and consists of fMRI data from 44 participants: 22 with an ASD diagnosis and 22 without ASD.

We extract brain fMRI signals from highly informative visual areas that activate when a visual stimulus is presented and capture visual information of different complexity. While in a resting-state fMRI analysis, we used a flattened connectivity matrix (resting state connectivity), in the case of task-fMRI, we use the features that represent normalized signals from multiple

[1]Categorical phenotypic features are gender (male/female) and age (categorical values corresponding to age groups).

TABLE 4.1: The statistics of the visual task-based fMRI dataset used in this work.

	Visual fMRI dataset
Total subjects	44
Patients	22
Healthy controls	22
Female/Male	3/19
Age range	12 - 18

brain regions. For this task, the data from the following brain regions were used: 1) first of the cortical regions called V1 responding to lower-level stimuli; and 2) the high-level Lateral Occipital Cortex (LOC) responsible for processing the object's semantic information.

The dataset samples were group-matched on gender, age, and intellectual ability. The inclusion criteria were the following: 1) age between 12 and 18 years old, 2) normal or corrected to normal vision, and 3) an IQ above 85. To be considered for the ASD group, participants had to have a clinical diagnosis of Autism Spectrum Disorder. The details of the dataset are presented in Table 4.1.

The perceptual expectation paradigm. The participants were presented with visual stimuli, namely, four images of a *lion, a turtle, a bike,* and *a car*. In each trial, a pair of stimuli were shown to participants. These pairs consisted of either the repetition of a single image or an alternation between two images. This resulted in four possible outcomes: expected and unexpected repetitions, as well as expected and unexpected alternations (see Figure 4.3 and Figure 4.4). During a pre-imaging practice task, participants were given time to learn which stimuli are most likely to repeat and which to alter. Which stimuli would repeat was counter-balanced between participants. In the fMRI task, two image stimuli had a 75% probability of repeating, and two had a 75% probability of alternating. Each stimulus was shown for 250 msec, separated by an inter-stimulus interval of 500 msec. The dataset is publicly available online at the Donders Research Data Repository.[2]

[2] http://hdl.handle.net/11633/di.dccn.DSC_3018026.01_044

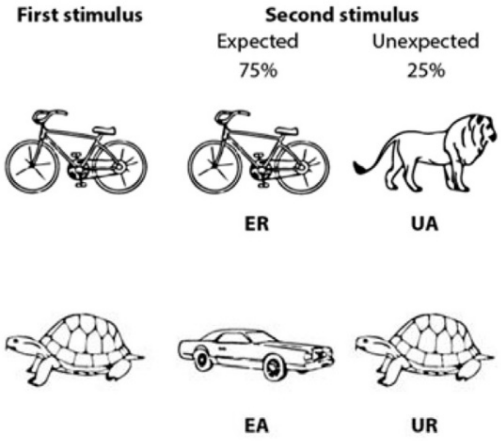


FIGURE 4.3: Examples of fixed stimulus pairings. The prior knowledge of which stimuli would repeat created four experimental settings: expected repetitions (ER) and expected alternations (EA): unexpected repetitions (UR) and unexpected alternations (UA). The figure is adapted from Utzerath et al., 2018.

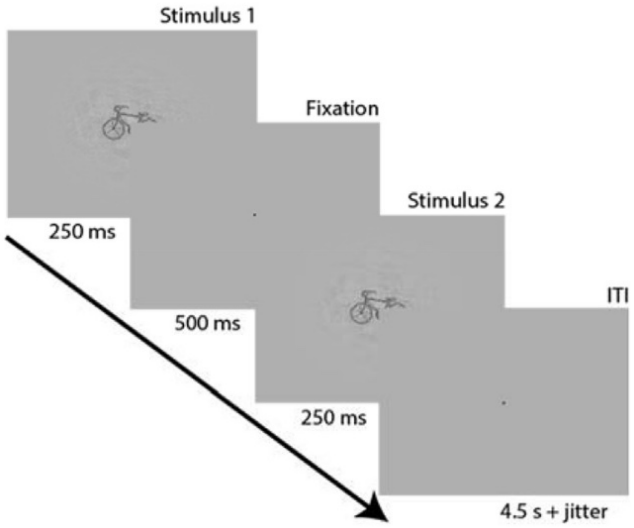


FIGURE 4.4: Display of visual stimuli for expected repetition (ER) trial. The figure is adapted from Utzerath et al., 2018.

4.2.4 Experimental Settings

Model Design

Analogous to the methodology for analyzing resting-state fMRI data with population graphs described in Section 3.4, we perform classification using linear graph architecture, SGC. The simple graph architecture was shown to perform on par or even better than more sophisticated graph architectures (Wu et al., 2019). This result was also supported in our analysis of resting-state fMRI data in Section 3.6.

Comparison with Baselines and Ablation Studies

We compare our method with traditional machine learning methods and deep learning methods to evaluate classification performance. The machine learning methods baseline methods include Multi-layer perceptron (MLP), Random Forest and 2-layer GCN Parisot et al., 2018, which is one of the current best performing deep learning methods with the available implementation.

Multilayer Perceptron (MLP) is well-suited as a baseline model due to its ability to learn complex patterns and relationships in data. It can handle both linear and non-linear relationships. MLP's simple and flexible architecture makes it suitable for various tasks. Random Forest is commonly used for tasks such as classification and regression, particularly when dealing with high-dimensional data. It combines the knowledge of multiple decision trees to give more reliable results and can show which features are most important in making prediction.

Graph Convolutional Network (GCN) is a baseline for tasks involving graph-structured data, such as node classification. It is used when the objective is to analyze relationships between entities in a graph.

For non-graph methods (Random Forest and MLP) we use flattened fMRI feature vectors as input. In the case of graph-based methods (GCN and ablations) we use flattened brain feature vectors as node features of the model,

and graph adjacency, which measures the correlation between fMRI and non-fMRI features.

Training setup

Due to the limited size of the new task fMRI dataset (44 samples) and for the sake of fair comparison with the baselines, we used five-fold stratified cross-validation, with the stratification criterion being an equal allocation of the class labels: patients with ASD and healthy subjects. The hyperparameters of the methods, such as learning rate, number of epochs, number of layers, hidden layers size, degree for graph convolutions, were tuned using grid search. For our method we use 100 training epochs, the degree parameter 2, and learning rate 0.1. Multi-layer perceptron (MLP) was also trained using Adam optimizer on 100 epochs, with learning rate 0.01, and two layers of size 64 and 16. To train Random forest classifier we used 8 tree with maximum depth of 8. Random forest classifier is trained with 8 trees and maximum depth of 16. Finally, GCN was trained on 100 epochs with degree 2, and learning rate 0.01.

The input features for graph models are the adjacency matrix $n \times n$ and the fMRI feature vector $n \times d$, where n is the number of samples and d is the dimension of the fMRI feature vector (voxels). Non-graph methods use only the fMRI feature vector.

Ablation study

We perform an ablation study to compare the contribution of task fMRI data and the graph-based framework. Our framework, a linear SGC model on task-fMRI, is denoted as a Graph with fMRI features. In addition, we introduce four ablations, where we use: 1) a graph without fMRI features; 2) a random graph structure with fMRI features; 3) a fully connected graph with fMRI features. The first two ablations are intended to evaluate the importance of subjects' imaging features, while the last two baselines evaluate the

importance of graph structure in the classification task. Underlying these baselines is the same GNN model running on the same population set.

The first baseline is introduced to test the meaningfulness of fMRI features regardless of non-imaging phenotypic data. We implement it by using an identity graph at the core, which implies no contribution from the graph, thus computation-wise the model is reduced to the MLP model. This means that each node is still associated with the fMRI feature vector, however, the identity matrix is used in place of the adjacency matrix to represent a completely disconnected graph. A graph without features baseline is implemented to test the performance of the model with respect to graph structure acquired using phenotypic features. In a random graph and a fully connected graph, we modify the edges. The edges are randomly rewired so that the edge density of the original graph is preserved. On the other hand, a fully connected graph assumes that all the nodes are connected.

4.2.5 Results and Discussion

In this section, we evaluate the performance of our linear graph-based model when applied to visual task-fMRI data. We emphasize the advantage of utilizing task-based fMRI data as it enables us to leverage larger training sample size. Due to multiple visual tasks which are performed in visual task fMRI, the dataset can be increased several folds, by using fMRI for each task as a separate data sample. Since there are four tasks being performed by each person (ER, UR, EA, UA), this allowed us to train the model with four times the number of original samples.

Integrating information from multiple brain regions is a common practice in neuroimaging and can help to capture the complex patterns of brain activity. Therefore, utilizing the domain knowledge we combine fMRI voxels from two visual brain regions: V1 (primary visual cortex) and LOC (lateral occipital complex). These regions process visual information at different levels. V1 is involved in early visual processing, such as detecting edges and

basic visual features. Whereas the LOC, is involved in higher-level processing and is responsible for recognizing complex visual stimuli.

In Table 4.2 we present prediction results using a combination of brain signal data from V1 and LOC. Table 4.2 shows the results on 5-fold stratified cross validation of the proposed method, ablation study and the baselines methods. Using the domain knowledge, the baselines GCN, MLP, Random forest, and SVM achieved over 0.61 in terms of mean AUC score and over 58% in terms of mean accuracy. Our model achieves the highest accuracy (69%) and AUC score (0.7). The baselines from the ablation study verify the contribution of the graph structure. The lowest scores (random prediction) correspond to Graph w/o fMRI, a baseline using only graph structure without important fMRI feature vectors.

TABLE 4.2: Classification results (mean and standard error) using combined V1 and LOC brain regions. The experiment performed with increased number of samples using five-fold cross validation. Abbreviations: 'w/' denotes 'with', 'w/o' denotes 'with out', and 'FC' denotes 'fully connected'.

	AUC	Accuracy	Sensitivity	Specificity
Graph w/ fMRI (ours)	0.70 ± 0.04	0.69 ± 0.05	0.69 ± 0.10	0.58 ± 0.11
Graph w/o fMRI	0.53 ± 0.12	0.52 ± 0.09	0.51 ± 0.14	0.53 ± 0.15
Graph random w/ fMRI	0.58 ± 0.08	0.57 ± 0.09	0.50 ± 0.28	0.74 ± 0.22
Graph FC w/ fMRI	0.59 ± 0.15	0.57 ± 0.16	0.60 ± 0.09	0.52 ± 0.25
GCN	0.66 ± 0.06	0.65 ± 0.06	0.63 ± 0.11	0.51 ± 0.15
MLP	0.63 ± 0.05	0.58 ± 0.07	0.65 ± 0.13	0.51 ± 0.20
Random forest	0.63 ± 0.04	0.63 ± 0.05	0.60 ± 0.08	0.58 ± 0.10
SVM	0.61 ± 0.06	0.61 ± 0.06	0.77 ± 0.04	0.32 ± 0.11

Discussion. The data described in this chapter focused on the perceptual differences in response to expected and unexpected visual patterns. The existing studies suggest that these differences in visual information processing are key to understanding ASD. Moreover, one of the recent studies by Gong et al., 2021 emphasizes that people with ASD exhibit a visual preference for non-social objects than typically developing individuals. This implies that visual preference for social and nonsocial stimuli in patients with ASD can

aid our understanding of the key signs of social interaction disorder. Visual stimuli presentation in task-fMRI may provide the possibility to screen ASD by discovering new disease biomarkers. Therefore we suggest complex visual stimuli with social and nonsocial components might improve the quality of the signal from visual task fMRI. The potential drawback in using the task fMRI data is the limited size of datasets. In the case of resting-state fMRI, the absence of the task allows the combining of the imaging datasets from different laboratories, as was done in ABIDE dataset. Combining task-based fMRI data requires a coordinated effort in devising a common visual experiment and performing post-imaging data processing.

4.3 Other DNN Methods for Task-based fMRI

In Section 4.2, we generalized our proposed graph-based framework for detecting Autism Spectrum Disorder to visual fMRI data, an alternative form of brain imaging data collected during a task, which is also used for medical diagnosis of neurodevelopmental disorders including ASD (Ahmedt-Aristizabal et al., 2021). Our interest in visual task fMRI is motivated by the findings showing an association of alterations in visual perception with ASD. Moreover, the task-fMRI data finds increased applicability in deep-learning-based research in neuroscience.

In this section, we continue exploring task-based fMRI data. However, the main focus will be shifted from ASD prediction towards developing an improved DNN architecture for analyzing visual task-fMRI. Visual task-based fMRI is accompanied by the corresponding stimuli. The presence of two different complex inputs requires designing new DNNs which can learn from multi-modal input, i.e., fMRI data and stimuli images.

We, therefore, explore how well neural network models can extract the meaningful signal from the task-based fMRI data by focusing on two tasks: 1) visual stimuli reconstruction, and 2) visual stimuli classification.

Despite the focus being set on only healthy subjects in this section, we consider the utility of visual task-fMRI data analysis for improving ASD detection in future research. As discussed in Section 2.1.2, visual information processing is shown to have a link to ASD, and more studies advocate for the use of task fMRI in neurodevelopmental diagnosis (Finn, 2021; Kolodny et al., 2020).

4.3.1 Motivation and Objective

The goal of this section can be summarized as follows:

- To evaluate the effectiveness of computational methods (DNNs) on visual task-based fMRI. Specifically, we survey visual stimuli reconstruction and visual stimuli classification. We are interested in qualitatively and quantitatively evaluating the representation that DNNs can learn from task-based fMRI signals.
- To identify the advantages of using task fMRI with deep learning models.
- To propose a multi-modal framework that can utilize both visual tasks-based fMRI and stimuli for prediction. Additionally, we introduce an experimental setting, through which we can assess the contribution of fMRI features toward improved classification accuracy. This experimental setup gradually reduces the quality of non-imaging data, i.e. the stimuli image, and thus requires the model to learn from the brain imaging data, proving its practicality to the task.

4.3.2 Modern Applications of Visual Task-fMRI Data

Visual Brain Decoding

Many brain imaging studies focus on decoding how the human brain represents information about the outer world. The majority of external sensory

information is processed by the human visual system (Logothetis and Sheinberg, 1996). In recent years, the study of visual information processing in the human brain has gained increasing attention and is commonly referred to as *human visual decoding*. The goal of neural decoding models is to learn a mapping function $f : \mathcal{V} \rightarrow \mathcal{X}$, where $\mathcal{X}$ and $\mathcal{V}$ denote two sets corresponding to stimulus images and fMRI activity patterns extracted from the visual cortex. A framework diagram for visual reconstruction is shown in Figure 4.5.

A great advancement in recent neuroscience research has been achieved through visual task fMRI (Poldrack and Farah, 2015; Nestor et al., 2020). The recent progress in human visual decoding has shown that beyond merely encoding the information about visual stimuli (Poldrack and Farah, 2015), brain activity captured by fMRI can be used to reconstruct visual stimuli information (Roelfsema, Denys, and Klink, 2018; Kay et al., 2008).

Based on the target objective, human visual decoding can be categorized into stimuli category classification and reconstruction (Naselaris et al., 2011). In classification, brain activity is used to predict discrete object categories of the presented stimuli (Haxby et al., 2001; Horikawa and Kamitani, 2017). In stimuli reconstruction, the aim is to recover image-specific details, such as object position, size, and angle. Compared to stimuli classification, reconstruction is a more challenging task, in which a replica of the stimulus image needs to be generated for a given fMRI signal (see Figure 4.5). Furthermore, stimulus-related information encoded in the fMRI activity, which allows high-accuracy identification, may only partially characterize stimuli images and thus be insufficient for proper image reconstruction (Kay et al., 2008; St-Yves and Naselaris, 2018). With the development of sophisticated image reconstruction methods and the increasing amount of brain imaging data, more attention has been directed toward visual stimuli reconstruction from fMRI activity in the visual cortex (Miyawaki et al., 2008; Naselaris et al., 2009; Gerven, Lange, and Heskes, 2010). fMRI-based visual reconstruction can improve our understanding of the brain's visual processing mechanisms, and researchers can incorporate this knowledge into other domains, including neurodevelopmental disorder prediction.

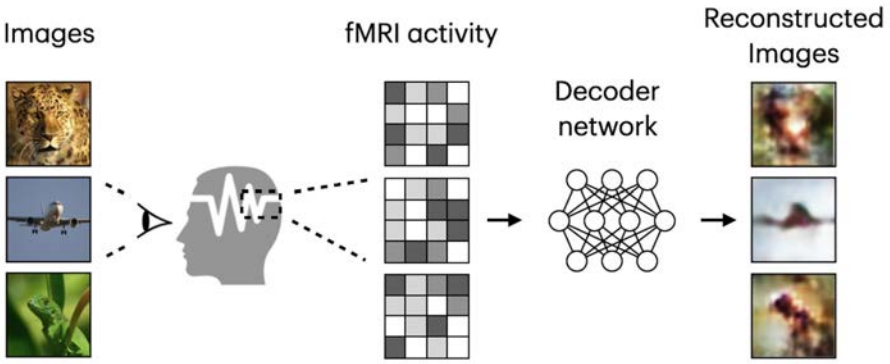


FIGURE 4.5: Diagram for natural image reconstruction task.

Visual Stimuli Reconstruction

The variety of visual stimuli used in visual reconstruction tasks can range from simple low-level detail images, such as Gabor wavelets and domino patterns (Thirion et al., 2006), to more elaborate images depicting alphabetical characters, digits (Gerven, Lange, and Heskes, 2010; Schoenmakers et al., 2013), natural objects, and scenes (Haxby et al., 2001; Horikawa and Kamitani, 2017). Among the variety of visual stimuli, natural images are considered to be the most challenging, as they require accurate reconstruction of color, shape, and high-level perceptual features of an object. We focus on the natural image stimuli since they require higher information processing in the brain and thus activate a higher number of visual areas, both low-level and high-level.

In recent years, deep neural networks (DNNs) have significantly advanced computer vision research, replacing models based on hand-crafted features. In particular, DNN models have achieved improved performance in various computer vision tasks, including image classification (Krizhevsky, Sutskever, and Hinton, 2012), image segmentation (Chen et al., 2015), and image restoration (Zhang et al., 2017). In visual decoding tasks using brain imaging data, deep learning approaches have been applied to image classification (Haxby et al., 2001; Horikawa and Kamitani, 2017), object segmentation (Kamnitsas et al., 2017), and natural image reconstruction (Shen et al., 2019a; Shen et al., 2019b). DNNs were shown to be more powerful than traditional methods

(Zhang et al., 2021a; Kriegeskorte, 2015) primarily due to the multilayer architecture allowing to learn nonlinear mappings from stimulus images (Beliy et al., 2019; Shen et al., 2019b).

Motivated by the success of deep learning in image generation, many recent studies have widely used DNN models in natural image reconstruction for several reasons. First, the deep learning framework conforms to some degree to the visual encoding–decoding process occurring in the hierarchical regions of the human visual system (Pinto et al., 2009; Krizhevsky, Sutskever, and Hinton, 2012; Schrimpf et al., 2018). Second, the application of deep generative models allows the synthesis of high-quality natural-looking images, which is achieved by learning the underlying data distribution (Goodfellow et al., 2014). Additionally, the training process can be aided by models pretrained on large-scale image datasets (Shen et al., 2019a; Shen et al., 2019b).

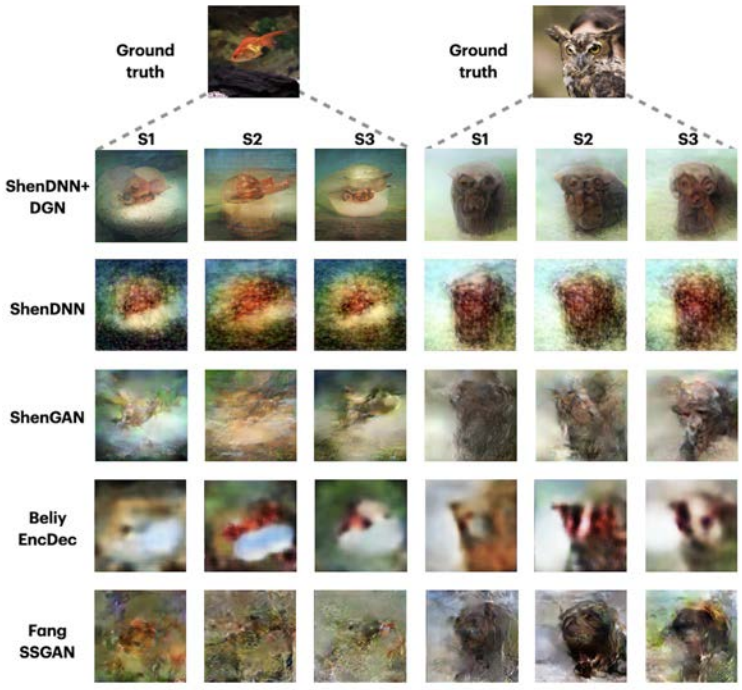


FIGURE 4.6: Reconstructions for two images across three subjects (s1, s2, and s3) from DIR dataset. The reconstructions are shown for the two natural image stimuli: *golden fish* and *owl*.

In our survey (Rakhimberdina et al., 2021), we study different DNN methods for natural stimuli reconstruction. We present qualitative and quantitative comparisons of the state-of-the-art deep learning methods, and demonstrate the power of learning representations from visual task fMRI. In Figure 4.6, we show the reconstructions across the methods corresponding to three different subjects. Depending on the subject, the reconstructions by different methods show varying degrees of resemblance to the original stimuli but preserve the global object information such as shape, position, and color. None of the traditional machine learning methods are able to achieve such quality of reconstruction.

4.3.3 Proposed Method

In this section, we propose a framework that integrates both fMRI data and stimuli information for improving classification performance. With this framework, we aim to show the benefits of aggregating imaging and non-imaging stimuli information and demonstrate the robustness of fMRI features.

Dataset Description

We utilize the publicly available Imagenet-fMRI dataset from Shen et al., 2019b. The dataset consists of pairs of high-quality 500×500 pixels stimuli images from the ImageNet dataset Deng et al., 2009 and the corresponding fMRI recordings collected from five healthy subjects. There are 1,200 distinct Imagenet images spanning 150 object categories in the training set, and each stimuli image is presented five times. As a result, we use 6,000 train-fMRI samples consisting of five repeated recordings per stimulus. For each fMRI, we considered approximately 10,000 voxels from the visual cortex area corresponding to visual brain areas V1, V2, V3, V4, LOC, FFA, and PPA. The results in the current study correspond to subject 1.

For training and evaluating our proposed method, we used MNIST and CIFAR-10 datasets. MNIST Lecun et al., 1998 consists of 28×28 grayscale images of ten handwritten digits. CIFAR-10 dataset Krizhevsky, 2009 contains

32×32 color images from ten different object categories. Note that the lack of complex, large-scale Imagenet-fMRI datasets did not allow us to test the method on more complex datasets. With the development of brain imaging techniques, we expect that more large-scale ImageNet-based fMRI datasets will become available.

Multi-modal Stimuli Prediction

In our survey Rakhimberdina et al., 2021, we showed that using task fMRI data one can obtain high-quality representations, the quality of which can be verified both qualitatively and quantitatively. In this section, we present a potential application of these high-quality brain visual representations in the stimuli classification task. In this scenario, we want to use the rich information captured in the form of fMRI data to recover the original visual stimuli. To understand how much information is learned from the fMRI, as opposed to images, we introduce different levels of noise to the image input. We propose a novel solution by combining the neural network representations from task fMRI and the original stimuli as multi-modal input. The term *modality* refers to a different type of information about the same scene obtained using multiple measurement devices, or acquisition techniques (Lahat, Adali, and Jutten, 2015). In a multi-modal setup, one can integrate data obtained from multiple modalities to obtain rich representations of the scene and enhance system robustness.

Our framework, proposed in (Rakhimberdina, Liu, and Murata, 2022), consists of the following components: voxel-wise (image-to-fMRI) encoder E and multi-modal image classification framework F . Individual components are illustrated in Figure 4.7. Next, we describe each component in more detail.

Encoder

To overcome the limitation of the small size of the fMRI dataset, we train a voxel-wise encoder on complex natural stimuli dataset Imagenet-fMRI from

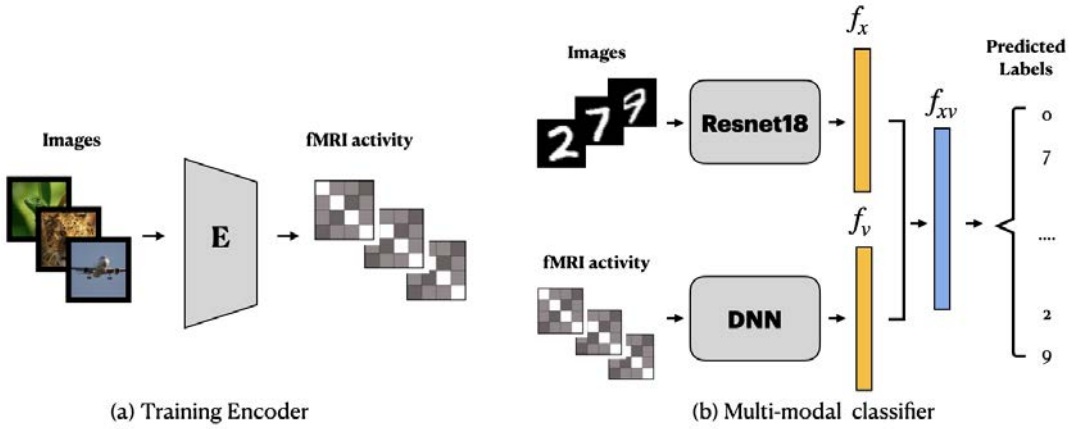


FIGURE 4.7: Our proposed multi-modal framework consists of the following components: a) image-to-fMRI encoder E and b) multi-modal classification network that integrates images and the corresponding brain codes to predict the output class of the stimuli.

(Shen et al., 2019b). The encoder E is used to learn the mapping from stimulus images to the corresponding fMRI activity pattern. The trained encoder further acts as a data augmentation tool for generating fMRI signals corresponding to images from less complex MNIST and CIFAR-10 datasets (see Figure 4.7a).

The architecture of the encoder is described in Table 4.3 and consists of four blocks of 3×3 convolution layers with ReLU activation. We utilized a fully connected layer to map the features to voxel space. Using the training image-fMRI samples from the Imagenet-fMRI dataset, we train encoder $E(x)$ in a supervised manner to predict the fMRI responses of input image x. The training loss is defined as the mean square error between the encoded fMRI representation $E(x)$ and the ground truth fMRI signal v.

Multi-modal Framework

Let $x \in \mathbb{R}^{w \times h \times c}$ be an image with width w, height h and c color channels. Let $v \in \mathbb{R}^d$ be the corresponding preprocessed fMRI response, with d denoting the number of voxels in the visual cortex. To integrate information from both modalities, we introduce a multi-modal framework as shown in Figure 4.7b. The framework consists of an image classifier and an fMRI classifier. Each of

TABLE 4.3: Encoder and Decoder architectures. The numbers in brackets for (de-)convolutional layers represent the number of channels and the kernel size.

Encoder	Decoder
Conv(32, 3×3)	FC()
ReLU()	DeConv(64, 2×2)
Conv(32, 3×3)	ReLU()
ReLU()	DeConv(64, 2×2)
Conv(32, 3×3)	ReLU()
ReLU()	DeConv(3, 2×2)
Conv(32, 3×3)	Sigmoid()
ReLU()	
Dropout(0.5)	
FC()	

the classifiers viewed separately can be viewed as an independent network adapted for image classification. In our multi-modal framework, for each image $x_i \in X$ we have an fMRI pair $v_i \in V$, where V denotes fMRI activity patterns extracted from the visual cortex. The image-fMRI dataset comprises N paired samples $S = \{(x_i, v_i) | x_i \in X, v_i \in V\}_{i=1}^{N}$. The goal of a multi-modal classifier is to learn a mapping function $f : X, V \rightarrow Y$.

In terms of architecture, the image classifier is based on ResNet18 architecture (He et al., 2016). fMRI classifier adopts a 2-layer deep neural network (DNN) with two fully connected layers separated by a non-linear activation function. While the convolutional network networks are considered effective for extracting information from images, we find simple feed-forward DNNs sufficient for performing fMRI classification.

The visual information from the image and brain codes from the fMRI is integrated by extracting the visual feature vector f_x from the image and fMRI feature vector f_v. Specifically, these extracted features f_x and f_v correspond to penultimate or pre-classification layer activations for image and fMRI classifiers. Then, the multi-modal method integrates the visual and fMRI feature vectors into a single vector via concatenation $f_{xv} = f_x \parallel f_v$. Finally, using the cross-entropy loss function, a fully-connected layer with a softmax predicts the output class probability from f_{xv}.

Experimental Settings

We train our networks using NVIDIA TITAN RTX GPUs. The encoder was the Adam optimization for 100 epochs with the learning rate of 1e-4 and 1e-3, and batch size 32 and 50, respectively. The multi-modal framework was trained for 50 epochs using learning rate 1e-2 and batch size 64.

4.3.4 Results and Discussion

Robustness under Gaussian Noise We first investigate the robustness of our multi-modal ensemble using image Gaussian noise perturbations. Gaussian noise corruption occurs commonly due to the limitations of visual sensors under poor lighting conditions. The Gaussian noise perturbation of input image x is $x + \delta$, where δ is the noise level $\delta \sim \mathcal{N}(0, \sigma^2_{noise} I)$. Figure 4.8 a,b shows classification accuracy under single-modality Gaussian noise perturbations for MNIST and CIFAR-10, respectively. In our setting, perturbations are added to images $x + \delta_x$ as a single-modality corruption.

On unperturbed data, all methods show the highest classification accuracy. With more perturbation added, the accuracy of uni-modal and multi-modal methods gradually drops. fMRI classifier is the uni-modal classifier using only human brain neural activity. The fMRI baseline is indicated by a black dashed line. Note that for the relatively simple MNIST dataset, the baseline performance of the fMRI classifier is comparable to the unperturbed image classifier (98.89%). The multi-modal classifier is more robust to image perturbation. That is, our multi-modal classifier outperforms the image classifier across different perturbation levels. For example, on MNIST, the multi-modal classifier can improve the accuracy of the corrupted image from 12.65% to 96.37% on the maximum perturbation level[3].

The robustness results obtained on the CIFAR-10 dataset are significantly lower than on the MNIST dataset for all methods. We attributed the higher robustness results on the MNIST dataset to the ability of the image-to-fMRI

[3]To determine the amount of perturbation σ_x we used signal-to-noise variance ratio (SNVR), the maximum σ_x for both MNIST and CIFAR datasets corresponds to SNVR = 0.1.

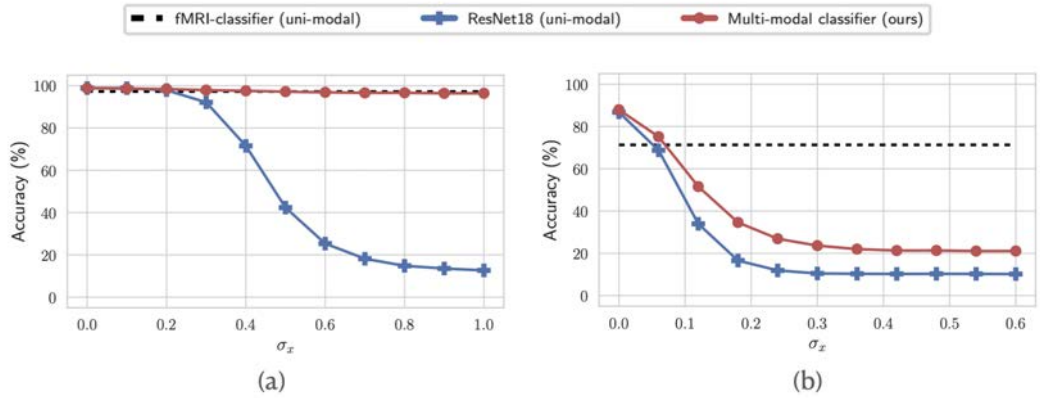


FIGURE 4.8: Robustness to Gaussian image noise perturbations on a) MNIST and b) CIFAR-10 datasets. The horizontal axis represents the strength of corruption in terms of the standard deviation σ. $\sigma \in [0, 1]$ for images.

encoder to generalize better to simpler images of hand-written MNIST digits rather than to more complex CIFAR-10 images. Overall, for both datasets, an additional classifier strengthens the perception of the multi-modal classifiers by utilizing the information from the visual task fMRI.

4.3.5 Discussion

This work serves as proof for the idea that brain codes in the form of visual task-based fMRI encode sufficiently to perform complex tasks such as reconstruction and classification. Unlike other multi-modal fusion methods employing various visual-audio or visual-text input integration (Tian and Xu, 2021), our method shows that fMRI signal can be used as a reliable input source resulting in high performance on visual classification.

Applications. There are multiple applications of our proposed method. The first application concerns the use of the developed multi-modal framework for classification tasks relevant to brain imaging. This application is more relevant to the topic of the thesis, as the integration of task-based fMRI and visual image input can be potentially used for disorder detection. Perturbing a stimuli image makes the fMRI-stimuli correlation weaker. We assume this setting can be adapted to ASD classification, where it is expected

that healthy people have strong fMRI-stimuli correlation (and thus prediction accuracy of the stimuli class), while ASD patients have a weaker correlation.

Second, the idea of combining visual input and human brain signals in a multi-modal framework adds to a growing literature on human-AI complementarity benefiting from the combined human and machine predictions (Steyvers et al., 2022; Jang, McCormack, and Tong, 2021). Nowadays, much effort is devoted to developing AI systems that complement human creativity and innovation in arts and music, augment human physical abilities in sports and medicine, and enhance human cognitive abilities such as memory and attention. Since the model was shown to be robust to various noise inputs, it can be used as an alternative to existing defense methods in cases when combining visual and human brain signals is feasible, such as in brain-computer interfaces. Incorporating reliable fMRI input acts as human guidance to the classification process when the visual input is corrupted. We believe future research in this area could have significant implications for healthcare, industry, and society as a whole.

Contribution of multi-modal stimuli classification to study of ASD. Previous studies demonstrated the usefulness of task fMRI in investigating autism spectrum disorder (ASD), as individuals with ASD often experience visual impairments. Therefore, information derived from both visual fMRI and the stimuli source is essential in these studies. Therefore, in our multi-modal framework, we proposed combining information from fMRI and stimuli to make the model learn from both. Though we tested our framework on label prediction, it can be extended to labeling patients and healthy subjects.

In the case of ASD, stimuli information can be used to determine the correlation between image and fMRI pair in individuals with ASD. Any misprocessing of visual information in such individuals can impact the strength of the signal between the image and fMRI. A weaker correlation between the two can impact the accuracy of the prediction. Our experiments have shown that introducing Gaussian noise can decrease the model's performance.

In the case of ASD, stimuli information can be useful to learn the strength

of the correlation between image and fMRI pair. If there is a misprocessing of visual info in an ASD patient, we can see the effect on the strength of the image-fMRI signal. The weaker correlation between image-fMRI affects the prediction. It has been shown in the experiment that the accuracy of the model goes down when Gaussian noise is introduced in the experiment.

We introduce an experimental setting, through which we assess the contribution of fMRI features toward model's performance. This experimental setup gradually reduces the quality of non-imaging data, i.e. the stimuli image, and thus requires the model to learn from the brain imaging data rather than stimuli. For that purpose, we introduce a Gaussian noise of different severity commonly used in the literature.

Perturbing a stimuli image makes the fMRI-stimuli correlation weaker. We assume this setting can be adapted to ASD classification, where it is expected that healthy people have strong fMRI-stimuli correlation (and thus prediction accuracy of the stimuli class), while ASD patients have a weaker correlation.

By introducing noise to the imaging modality, we weaken the correlation between fMRI and stimuli. We expect that if information distortion happens in ASD, the correlation of the fMRI-stimuli pair will also decrease. Then the proposed multi-modal framework can be adapted for training on fMRI-image samples for healthy and ASD patients.

Availability of large-scale imaging data. The primary challenge in using deep learning methods is the limited size of fMRI data. Compared to the existing resting-state fMRI and task-based fMRI datasets used in brain diagnosis, the visual task fMRI dataset contains more than 10 times of data sample. Moreover, the lack of training data is compensated by pretraining DNN components on external image data (Shen et al., 2019b; Shen et al., 2019a; Beliy et al., 2019), self-supervision on additional image-fMRI pairs (Beliy et al., 2019; Gaziv et al., 2020) and generation of new surrogate fMRI via pretrained encoding models (St-Yves and Naselaris, 2018). Using efficient models trained on visual decoding task, we can enhance the current research on ASD prediction through transfer learning or use the pretraining models on the image

domain. This, in turn, may lead to broader adoption of the proposed frameworks for real-world purposes.

4.4 Conclusion

More researchers nowadays emphasize the increasing prospect of using deep learning models to extract neural activity responses to naturalistic stimuli (i.e., visual task) (Zhang et al., 2021b). Visual task-induced brain imaging data can potentially improve the prediction of individual markers better.

To verify this, in this chapter, we generalized our proposed graph-based framework for detecting Autism Spectrum Disorder to visual fMRI data. The results demonstrated the improvement in classification accuracy due to the introduction of the population graph structure. Both the graph-based analysis approach and visual signal extraction from this study can be utilized to make predictions about ASD in the presence of a visual fMRI dataset for ASD, in a similar manner as for the resting-state fMRI data.

As a follow-up work, we focus on improving of visual task-based fMRI processing by introducing a multi-modal framework based on images and fMRI encodings.

The presence of visual stimuli allowed us to adapt the existing image classifiers for improving accuracy. Such analysis would not be possible with resting-state fMRI data, which is recorded in the absence of a task.

The contributions of the presented analysis are two-fold. First, we quantify the quality of the task fMRI feature, as it was shown to improve the accuracy despite the corrupted stimuli information. Second, we present a stimuli-fMRI integration framework, which can be applied to a prediction of ASD and variety of related downstream tasks in brain-computer interfaces. The last point may be of interest to scientists and engineers working with task fMRI data.

Chapter 5

Discussion

In this chapter, we first explain the significance and implications of research results in relation to research questions. We discuss the design considerations and limitations of our population graph-based models proposed in Chapter 3. Next, we integrate our findings from the study of visual fMRI data in Chapters 4 and show the usefulness of combined task-fMRI features and visual stimuli-related features as input. Furthermore, we discuss the implications and limitations of using resting-state fMRI and task-based fMRI, and DNN-based analysis techniques that can be shared with these two fMRI acquisition paradigms.

5.1 Significance and Implications of Current Research

To recap, the aim of this study is to *propose a general framework for ASD classification that can be applied to both resting-state and task fMRI data*. Below we summarise the key findings that relate to each of the research questions:

- **R1. What kind of neural networks can efficiently model the population of ASD subjects described with fMRI and non-fMRI features?** The performance analysis of deep learning methods using resting-state brain fMRI data suggests that graph neural networks can effectively represent the group of healthy individuals and patients. This is supported by high accuracy and AUC results of the proposed SGC method

in predicting ASD. Moreover, the graph-based architecture is capable of integrating different types of input data: brain imaging data and non-imaging phenotypic data, which non of the rival methods can accomplish without significant changes in model architecture. While brain imaging signals are used as features of the individual nodes in the graph, the connection between nodes is reinforced by additional information important in predicting ASD disorder. In addition, we would like to note that the simplest GNN architecture was able to outperform the more elaborate CNN-based models. Our discovery of the simplified and well-performing graph neural network contrasts with the existing literature, where the general trend is toward more complex neural network architecture. We suppose that more complex deep learning architectures are more suited to datasets with significantly higher amounts of training data and a high signal-to-noise ratio. Therefore, we believe that the main contribution of this research is presenting a community with a simple graph-based model and explaining why it provides a performance improvement through the analysis of graph spectral features.

- **R2. Can a graph-based classification framework improve ASD prediction in task-based fMRI?** We generalized our proposed graph-based framework for diagnosing Autism Spectrum Disorder to visual fMRI data in Chapter 4. The results demonstrated the improvement in classification accuracy and AUC scores when using the graph-based neural network. Additionally, the performance improvement is due to the increased number of training samples corresponding to each fMRI-stimuli pair. Both the graph-based analysis approach and visual signal extraction from this study can be adopted to make predictions for ASD using visual fMRI dataset.

- **R3. How effective are neural networks with task-based fMRI data compared to resting-state data?** Considering the unique properties of task-fMRI data, we focused on different neural network architectures developed for task-fMRI. We found that the analysis of task-based fMRI

data, in general, can be seen as advantageous compared to the analysis of resting-state data in two ways. First, the task-based fMRI is less noisy because of smaller regions of interest and precise time frames corresponding to stimuli presentation. The presence of the task allows combining data from different individuals by reducing the variance of the signal corresponding to a particular frame. Second, the presence of the target stimuli allows deep learning methods to learn from fMRI by encoding and decoding the stimuli to the brain signal. We demonstrated this both on visual stimuli reconstruction and stimuli classification. In visual stimuli reconstruction, we showed that the brain signal can be reconstructed with a high accuracy. In visual stimuli classification, we showed that despite any perturbations to visual stimuli, the DNN can efficiently use visual task-based fMRI information to correctly classify the label of the stimuli.

5.1.1 Evaluating ASD Detection Performance

The TP rate, or true positive rate, reflects the portion of individuals accurately identified with ASD by a diagnostic test. While a perfect TP rate of 100% is ideal, achieving this is challenging due to the absence of a flawless ASD diagnostic test. Even a test with a high TP rate may miss some cases of ASD.

It's crucial to note that a high true positive rate (TPR) coupled with a low true negative rate (TNR) isn't favorable in diagnostics or screening. A low TNR implies more false positives, wrongly classifying non-ASD individuals as positive, potentially leading to unnecessary anxiety, additional tests, or treatment for the wrong reasons.

In ASD screening, achieving both high TPR and TNR is vital. Striking this balance ensures accurate identification of ASD individuals (high TPR) while minimizing false positives, accurately labeling non-ASD individuals as negative (high TNR). The aim is a balanced approach maximizing sensitivity (TPR), specificity (TNR), and ultimately, classification accuracy. To represent

the overall correctness of prediction model on the balanced dataset we calculate accuracy. i.e, the proportion of correct classifications (both true positives and true negatives) among all cases. It provides a general overview of the model's performance. Along with AUC score, sensitivity, specificity, and accuracy help assess the reliability and performance of classification models. A balanced approach is desired, where high both false positives and false negatives are minimized, achieving an overall accuracy of classification

5.1.2 Exploring Applications of the Proposed Method

We present a scenario where our ASD prediction model can be used. Traditional diagnostic methods for ASD rely on behavioral observations, which can be time-consuming and subjective. In contrast, our method uses machine learning and neuroimaging fMRI data to explore a more objective approach to ASD prediction. Considering the high cost of fMRI, our method can be an effective alternative when the behavioral test can not be used, that is when patients do not exhibit the expected behavior. For example, the model might serve as a screening tool for identifying ASD in small children before symptoms become apparent.

The increasing cases of autism spectrum disorder (ASD) require expanding ASD diagnostic capacity in terms of number of professionals trained in this field. Evidence suggests that a significant portion of children remains not diagnosed until they reach 4 or 5 years, partially due to long waiting period between referral and receiving a team-based ASD behavioural evaluation, which can frequently exceed a year (Brian, Zwaigenbaum, and Ip, 2019).

The potential of fMRI in predicting ASD in infants has already been shown by Emerson et al., 2017 on a small fMRI dataset of 59 6-month-old infants. The authors achieved relatively high results in predicting which infants would later be diagnosed with ASD. Given the absence of efficient alternative non-behavioral methods, the expense of fMRI screening may be justifiable in such scenarios.

For a newly developed diagnostic method in Autism Spectrum Disorder (ASD) to be used in a clinical setting, generally, a high sensitivity is desirable to ensure that individuals with ASD are not missed or falsely classified as negative. As a guideline we suggest using the proposed model at a threshold where sensitivity reaches 70%. This is in agreement with studies and guidelines suggest aiming for higher sensitivity (Charman and Gotham, 2013). This means that the diagnostic method correctly identifies at least 70% of individuals who truly have ASD. Further inspections can be made to reduce false positives.

In cases when both high sensitivity and specificity are important, our proposed method could be integrated within a multi-model framework, where the initial model prioritizes high sensitivity and the second model emphasizes high specificity. This arrangement can help reduce instances of unclear predictions, prioritizing cases that both models agree upon and flagging cases with conflicting predictions for further examination.

While in our work we presented the balanced sensitivity and specificity scores, if the diagnostic method is intended for early detection and intervention, a higher sensitivity may be desirable to identify ASD cases as accurately as possible, even at the cost of a slightly higher false positive rate. In this case, a high sensitivity level should consider the availability of follow-up evaluations and confirmatory assessments. A higher sensitivity may result in more referrals for further evaluations, requiring appropriate resources to accommodate the increased caseload.

5.2 Simplified Graph Neural Network

5.2.1 Design Considerations

In Chapter 3, we study the problem of classifying ASD using resting-state fMRI data. We design a population graph-based approach, where we incorporate resting-state fMRI features and non-imaging features for improving

the model's performance. Our model includes three design considerations as follows:

1. Due to the introduction of the graph structure, the model is capable of integrating two types of different imaging and non-imaging features, which have different sources, dimensionalities, categorical/numerical representations, etc. Brain imaging features are used as nodes' features, while non-imaging are incorporated through the edge assigning function.

2. The model can be generalized to different types of fMRI, both resting-state and task-based. In the case of resting-state fMRI, the brain signal is represented as a flattened connectivity matrix (resting-state connectivity), while in the case of task fMRI, the feature vector represents the normalized signal from multiple ROIs.

3. Simplistic architecture is proposed with the intention to emphasize the role of features rather than complex architecture on classification performance. The simplistic graph architecture (without non-linearity layers) also achieves comparable performance compared to results from more complex architectures, such as Graph Convolutional Networks or Graph Attention Networks. The simple graph architecture was shown to perform on par or even better than more sophisticated graph architectures on a variety of large-scale benchmark datasets (Wu et al., 2019). This result was also supported in our analysis of resting-state fMRI data in Section 3.6.

5.2.2 Limitations of the SGC Model

Our SGC model for classifying ASD from a population of subjects can only consider a single graph and requires prior knowledge about which features are most important for classification. Another limitation concerns the classification of new unseen data samples. In that case, the graph-based model

needs to be retrained from scratch to account for changes in population samples.

5.3 Multi-modal Ensemble for Classifying ASD

5.3.1 Implications of Model Design

Most of the population-based models proposed in the literature do not consider the variety of graphs that result from the choice of different features. Our multi-modal ensemble model differs from the previous methods by two different design considerations: graph preprocessing and graph aggregation. First, we investigate the importance of different graph configurations. For that purpose, spectral filtering acts as noise filtering and favors the information flow along the cluster of similar nodes which helps distinguish healthy controls from patients. Next, after we identified the best-performing graph structures (and the corresponding best-performing features), we combine each of these graphs in a multi-modal ensemble. Both preprocessing and ensembling techniques help to improve the performance of ASD detection.

5.3.2 Limitations of the Multi-modal Ensemble Model

In our multi-modal ensemble model, we aim to use information from multiple best-performing graphs to improve performance. Each of the input graphs passes through a spectral filtering operation, which requires the computation of the graph's eigenvector matrices. In the case of an ASD dataset like ABIDE, each of the population graphs contains only $n = 871$ nodes. However, doubling the number of nodes will increase the complexity 8 times since the eigenvalue computations scale to $\mathcal{O}(n^3)$.

5.4 Comparison of Resting-state and Task-fMRI Data

In this section, we compare the advantages and disadvantages of resting-state and task-based fMRI acquisition paradigms when used with DNNs.

TABLE 5.1: Comparison of resting-state fMRI and visual task-based fMRI.

	Resting-state fMRI	Task-based fMRI
Number of fMRI samples per subject	1	multiple
Ease of fMRI acquisition	✓	✗
Quality of the signal - Signal-to-noise ratio - Reconstructable signal	 lower ✗	 higher ✓
Deep Learning - Variety of DNNs	DNN, 2D-CNN, 3D-CNN	DNN, 2D-CNN, 3D-CNN, AE, GAN, etc.
- Pretraining	✗	✓

We compare the two types of datasets using the following characteristics:

- Number of fMRI samples per subject
- Ease of fMRI acquisition
- Signal-to-noise ratio
- Variety of DNNs that can be applied

The summary of the comparison metrics is given in Table 5.1. Next, we discuss each point in detail.

5.4.1 Number of fMRI Samples

The absence of stimuli in a resting-state paradigm usually requires a longer acquisition time resulting in single several minutes for recording. For example, for ABIDE dataset, resting-state fMRI acquisition duration varied from five to eight minutes per individual (Di Martino et al., 2017). In contrast, the task paradigm allows the acquisition of more data samples per subject. Depending on the number of visual stimuli or a number of visual tasks presented to the subject, the task-based fMRI contains multiple stimuli-fMRI

pairs which can be used as training samples. For example, in the Generic Object Decoding dataset by Horikawa and Kamitani, 2017 each stimulus requires only a few seconds for brain data acquisition, thus thousands of fMRI-stimuli pairs can be acquired during a single session. For the ASD dataset by Utzerath et al., 2018, the number of visual task presentation experiments is equal to four, which allows for increasing the training data 4 times. In addition, the presence of the task allows more control over the data processing such as averaging the samples across different runs or scanning sessions and across subjects. Thus, task paradigms enable researchers to parse signals into meaningful and non-meaningful segments and look for generalization patterns in the behavior (Finn, 2021).

Increased data samples. Due to the presence of multiple tasks, the dataset can be increased by several folds, which allows training the model with a higher number of samples. For example in the case of visual task-fMRI data for ASD, each person participated in 4 visual tasks. Subsequently, we were able to increase the size of the data samples 4 times, from 44 to 176. We investigated the role of the increased data size on prediction accuracy and AUC score in Section 4.2.5. Our observation is that a higher number of samples resulted in an improvement in model's performance.

5.4.2 Ease of fMRI Acquisition

In terms of the fMRI acquisition, resting-state fMRI is greatly simplified compared to the task fMRI. First, the overall data acquisition time is usually shorter, especially when samples need to be collected from a large group of people (Albuquerque et al., 2021; O'Connor and Zeffiro, 2019). Second, due to the absence of the specific task, the data preprocessing and analysis is significantly simplified (Liu et al., 2022). This makes it easy to share the data across the laboratories in the pursuit of creating a large-scale database. Finally, rsfMRI is more practical when individuals, such as children, paralyzed, or cognitively impaired patients, are unable to perform an fMRI task (O'Connor and Zeffiro, 2019).

5.4.3 Signal-to-Noise Ratio

The brain activity patterns in the resting state can potentially take a large number of possible configurations. The lack of a stimulus or an experimental task causes great variability in activity patterns. A big variation between training and test data samples across a single individual subject makes the task quite challenging for computational models designed for extracting common patterns from data (Guidotti et al., 2015). In addition to the variability within single-subject scans, the variability of fMRI signals across different subjects can be a source of concern. The evidence shows that it has been particularly challenging to obtain consistent fMRI activations across different brains and populations due to the huge variability between individuals (Zhang et al., 2016). All of these decrease the signal-to-noise ratio of resting-state fMRI, making it challenging to compare the data from different individuals.

5.4.4 Deep Learning

Variety of DNNs. More researchers nowadays emphasize the increasing role of deep learning models in learning neural activity responses caused by visual task stimuli (Zhang et al., 2021b). Task-induced brain imaging data improves the prediction of individual markers better than rest-based. DNN, 2D-CNN, 3D-CNN, and DNNs models for processing sequential data are used for the analysis of the rest-state fMRI. However, the presence of the image stimuli in the visual task fMRI allows using a greater variety of DNNs models.

Both the graph-based analysis approach and visual signal extraction from this study can be utilized to make predictions about ASD in the presence of a naturalistic visual fMRI dataset for ASD. For example, we can use the proposed encoder from Section 4.3.3 for extracting the brain representations corresponding to different stimuli and use this representation as characterizing features of subjects-nodes in the graph neural network architecture.

Pretaining. Visual task-fMRI can use the abundance of the pretrained visual models for such large-scale datasets as Imagenet or MNIST for extracting high-quality features and mapping them into the domain of fMRI signal. Existing trained DNN models and pretraining techniques may foster ASD detection methods in task-based paradigms, which is not possible in the resting-state paradigm.

5.4.5 Limitations of fMRI Data

Due to the absence of the task, the variability of resting-state fMRI data from subject to subject is high. Both limited dataset size and the low-signal-to-noise ratio of resting-state fMRI result in high experimental standard deviation. A limiting factor for using task fMRI in the clinical application is task difficulty: not all of the participants, including patients, will be able to perform a given task. Hence, participants are limited to those whose cognition is least affected by the disease (Greicius et al., 2003).

Although resting-state fMRI is relatively straightforward to use, their practical interpretation is often questioned due to the lack of functional context in a task-free resting paradigm. On the other hand, the presence of diverse, naturalistic stimuli provides rich spatial and temporal contexts for a wide range of neural processes (Zhang et al., 2021b). In Chapter 4, we show that these naturalistic stimuli induce highly reproducible brain responses, and many computational models, such as neural networks, are increasingly available nowadays to map stimuli in the image domain to brain activity. Therefore, in this work, we advocate naturalistic stimuli, such as visual stimuli, as a paradigm complementary to traditional resting-state paradigms.

New computational techniques are unlocked via computer vision research, and an increasing amount of research suggests that computer vision methods have a strong impact on autism research (Belen et al., 2020). Prior to applying task-based visual fMRI data for ASD prediction, we draw inspiration from the research on brain visual decoding. In particular, we investigate the

recent computation methods used to extract the information from the task-based dataset. To assess the quality of visual fMRI, we studied visual stimuli reconstruction and compared how well signals can be recovered from task-based visual fMRI. Our findings suggest the significance of natural stimuli for DNN-based frameworks. However, it would be ideal if we had both resting and task-based fMRI from the same dataset, to compare both, but up to our knowledge, no such datasets currently exist for ASD studies.

5.5 Conclusion

In this chapter, we discussed the design considerations and limitations of our proposed graph-based models. Also, we integrated our findings from the study of visual fMRI data and demonstrated the usefulness of combined task-fMRI features and visual stimuli-related features as input. Based on our observations and the additional experiments, we provided a comparative review of resting-state fMRI and task-based fMRI. We can conclude that our proposed model has the potential to be used for both resting-state and task-fMRI. The task-based data remains underrepresented in the research of neurodevelopmental pathologies such as ASD and further exploration of task-fMRI using deep learning is left for future work.

Chapter 6

Conclusion

Autism spectrum disorder is a neurodevelopmental disorder characterized by difficulties in social communication, including speech and nonverbal interaction, and repetitive behavior patterns. The screening of ASD is difficult due to uncertainties in symptoms and subsequently biases of human experts' examination. Along with resting-state fMRI, task fMRI is used to identify the population containing ASD patients. The objective of this thesis is the proposal of a general framework for ASD classification that can be applied to both resting-state and task fMRI data. Specifically, given the differences between each type of imaging data, we aim to have a model that incorporates the unique features of each dataset, which are expected to be helpful in improving classification performance. For concreteness, we summarize the key findings with respect to the research goals. In addition to that, we state the value and contributions as well as the limitations of our study. Finally, we will propose opportunities for future research.

6.1 Key Contributions

In this thesis, we aim to investigate the problem of ASD detection using fMRI data. The thesis begins in Chapter 1 with the introduction of the ASD detection problem, fMRI datasets used for the study of ASD, and the existing deep learning frameworks. In Chapter 2, a literature review was provided. Specifically, a large body of recent work and the advantages and disadvantages of resting-state and task fMRI data were explained and summarized. Since

we are interested in classifying ASD patients in a population containing both patients and healthy controls, we rely on a graph-based approach for modeling the community (Chapter 3). We are interested in identifying the role of graphs in aggregating imaging and non-imaging features. We also generalize the opposed graph neural network framework for the task fMRI data in Chapter 4. For the first time, we apply a population graph-based framework in the task paradigm paving a new way to solve the ASD detection problem.

While the resting paradigm still dominates in many sub-fields of neuroscience, data from tasks have empirically shown benefits. These empirical benefits include enhanced interpretability and sensitivity to brain-behavior relationships. Based on natural image reconstruction and classification tasks, our findings suggest that task-based fMRI preserves high-quality, interpretable information that can be extracted by neural networks.

6.2 Limitations

The limitations of the current study can be summarized as follows:

- Nowadays, using fMRI as screening tool for ASD is quite costly, and thus traditional behavioral and developmental observation tests may be preferable in most cases. However, despite high cost, the fMRI-based screening can be advantageous in cases when patients do not express any symptoms, for example small children as in the study by Emerson et al., 2017.

  Despite its limitations, fMRI shows promise for ASD screening. As technology advances and becomes more affordable, it is possible that fMRI will become more widely utilized in ASD screening. Unlike behavioral observations, fMRI provides a more objective measure of ASD.

  With the continuous development of technology and the availability of more data, there is an anticipation for the use of Deep Neural Networks (DNNs) in ASD prediction. Using large-scale data, DNNs have

the potential to effectively learn and generalize patterns related to ASD, further enhancing the utility of fMRI in this domain.

- It's important to note that there is ongoing debate and research around the classification and diagnosis of ASD, and some experts argue that it should not be treated as a binary condition. The multi-label classification of ASD, where the labels represent the severity of the diagnosis, is one of the potential improvements of this research. However, such a type of multi-class prediction is restricted to the availability of the data and will require considerably more data samples for analysis. Such a type of multi-class prediction would require considerably more data for analysis containing a sufficient amount of samples representing each group of the severity of the diagnosis.

- The lack of an established task paradigm for the study of ASD, and as a result, the lack of a large-scale dataset prevents us from providing an empirical comparison of these two fMRI types. Whereas it is a common practice to establish collaborative datasets from different laboratories for the study of ASD, such large-scale open data still does not exist for task-based paradigms.

- In addition, there is a great variety in visual stimuli presentation for identifying ASD, ranging from the presentation of simple stimuli with varying luminance (i.e., gratings) and shapes, to more complex perceptual tasks, such as stimuli repetitions. The lack of consensus on the established visual paradigm makes it challenging for establishing a large-scale visual fMRI dataset.

6.3 Future research

Despite a common assumption that resting-state data is traditionally used in clinical diagnosis, while task-based fMRI is used to study a specific function, this study highlights the role of task-based fMRI in detecting brain disorders.

With more evidence suggesting the involvement of sensory perception, such as vision or touch, in the existence of brain pathology, more work can be done toward designing task-based paradigms for studying the disease.

Therefore, as a future direction, we suggest applying visual task-based fMRI in ASD detection. Future studies can build on this study by using a task-based paradigm specifically designed for ASD diagnosis. In addition, the experiments with the combined use of resting-state and task-based fMRI are of great interest.

Our multi-modal image classification framework, which we presented in Chapter 4 can be used as a foundational framework for the study of visual information processing using visual task-fMRI for studying ASD.

Nowadays, DNNs offer techniques for processing natural real-world information. If we want to make good use of deep learning methods in the analysis of brain imaging data, we need to design experimental paradigms which use complex naturalistic stimuli, that is real-world input, such as images, video, and language. Since recent findings suggest a link between ASD and abnormal visual processing, more research needs to be done to design valid task-based fMRI protocols for studying ASD and assess the differences between two existing fMRI acquisition paradigms.

Appendix A

Additional Information on Male/Female Imbalance

The effect of male/female imbalance on classification accuracy.

To test the effect of male/female imbalance, we performed the experiment where we changed the male/female ratios in a dataset as follows:

- a fully imbalanced subset consisting of only males subjects
- a fully imbalanced subset consisting of only females subjects
- male/female balanced subset (144 males/144 females)
- the original dataset (727 males/144 females).

Note that both fully imbalanced datasets and male/female balanced subsets are imbalanced with respect to the output class labels (healthy or patient with Autism spectrum disorder). Therefore, we use evaluation metrics suitable for imbalanced datasets, i.e., AUC, Sensitivity, and Specificity (please see Table 1 below). For a fully imbalanced subset consisting of only female subjects and a male/female balanced subset, there is a significant increase in sensitivity score (the ability to identify subjects with Autism spectrum disorder correctly). However, it comes at the expense of reduced specificity (more false positives). The inferior performance might also be a result of the significantly smaller dataset used for training/testing. We did not find significant differences in performance between models trained only on male data and the original dataset. The classification using the original dataset showed the

best balance between Sensitivity and Specificity scores and the highest AUC score.

TABLE A.1: The classification results using different subsets of data evaluated using AUC, Sensitivity, and Specificity scores.

Subset	**Number of samples**	**AUC**	**Sensitivity**	**Specificity**
Males	727	0.74	0.74±0.12	0.72±0.11
Female	144	0.68	0.73±0.08	0.47±0.29
Male/female	288	0.69	0.84±0.15	0.53±0.25
Original dataset	871	0.75	0.75±0.07	0.69±0.00

Appendix B

Additional Information on Using Different Brain Regions

The effect using different brain regions on classification performance. To see the contributions of using combined information from V1 and LOC, we present the results when using the data from V1 and LOC separately in Tables B.1 and B.2.

Tables B.1 and B.2 show the results on 5-fold stratified cross validation of the proposed method, ablation study and the baselines methods. Using information from only a single brain region, the performance of all methods including the baselines GCN, MLP, Random forest, and SVM is close to random. Abbreviations: 'w/' denotes 'with', 'w/o' denotes 'with out', and 'FC' denotes 'fully connected'.

TABLE B.1: Classification results (mean and standard error) using only V1 brain region.

	AUC	Accuracy	Sensitivity	Specificity
Graph w/ fMRI (ours)	0.58 ± 0.15	0.59 ± 0.15	0.68 ± 0.22	0.68 ± 0.24
Graph w/o fMRI	0.44 ± 0.11	0.47 ± 0.10	0.53 ± 0.36	0.69 ± 0.29
Graph random w/ fMRI	0.53 ± 0.10	0.52 ± 0.10	0.67 ± 0.24	0.62 ± 0.28
Graph FC w/ fMRI	0.53 ± 0.11	0.53 ± 0.10	0.70 ± 0.37	0.61 ± 0.28
GCN	0.51 ± 0.02	0.50 ± 0.02	0.60 ± 0.33	0.68 ± 0.23
MLP	0.52 ± 0.05	0.50 ± 0.02	0.61 ± 0.32	0.73 ± 0.25
Random forest	0.58 ± 0.13	0.57 ± 0.14	0.49 ± 0.37	0.67 ± 0.30
SVM	0.47 ± 0.04	0.45 ± 0.03	0.65 ± 0.22	0.65 ± 0.25

TABLE B.2: Classification results (mean and standard error) using only LOC brain region.

	AUC	Accuracy	Sensitivity	Specificity
Graph w/ fMRI (ours)	0.57 ± 0.12	0.57 ± 0.13	0.61 ± 0.27	0.60 ± 0.28
Graph w/o fMRI	0.44 ± 0.11	0.47 ± 0.10	0.53 ± 0.36	0.69 ± 0.29
Graph random w/ fMRI	0.57 ± 0.14	0.57 ± 0.13	0.49 ± 0.29	0.75 ± 0.26
Graph FC w/ fMRI	0.57 ± 0.08	0.55 ± 0.08	0.51 ± 0.37	0.70 ± 0.32
GCN	0.52 ± 0.12	0.51 ± 0.04	0.51 ± 0.40	0.66 ± 0.36
MLP	0.56 ± 0.06	0.56 ± 0.06	0.66 ± 0.27	0.51 ± 0.26
Random forest	0.57 ± 0.09	0.55 ± 0.08	0.47 ± 0.39	0.69 ± 0.31
SVM	0.54 ± 0.12	0.51 ± 0.10	0.51 ± 0.29	0.74 ± 0.22

Bibliography

Abraham, Alexandre et al. (Feb. 2017). "Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example". en. In: *NeuroImage* 147, pp. 736–745. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2016.10.045. URL: https://www.sciencedirect.com/science/article/pii/S1053811916305924 (visited on 03/21/2023).

Ahmedt-Aristizabal, David et al. (Jan. 2021). "Graph-Based Deep Learning for Medical Diagnosis and Analysis: Past, Present and Future". en. In: *Sensors* 21.14. Number: 14 Publisher: Multidisciplinary Digital Publishing Institute, p. 4758. ISSN: 1424-8220. DOI: 10.3390/s21144758. URL: https://www.mdpi.com/1424-8220/21/14/4758 (visited on 01/31/2023).

Akiba, Takuya et al. (July 2019). "Optuna: A Next-generation Hyperparameter Optimization Framework". In: *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. KDD '19. New York, NY, USA: Association for Computing Machinery, pp. 2623–2631. ISBN: 978-1-4503-6201-6. DOI: 10.1145/3292500.3330701. URL: https://doi.org/10.1145/3292500.3330701 (visited on 11/18/2022).

Albuquerque, Daniela de et al. (2021). "Deep Generative Analysis for Task-Based Functional MRI Experiments". In: *Proceedings of machine learning research* 149, pp. 146–175. ISSN: 2640-3498. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8871581/ (visited on 11/06/2022).

Anirudh, Rushil and Jayaraman J. Thiagarajan (May 2019). "Bootstrapping Graph Convolutional Neural Networks for Autism Spectrum Disorder Classification". In: *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. ISSN: 2379-190X, pp. 3197–3201. DOI: 10.1109/ICASSP.2019.8683547.

Armstrong, Richard A. (Nov. 2018). "Visual problems associated with traumatic brain injury". eng. In: *Clinical & Experimental Optometry* 101.6, pp. 716–726. ISSN: 1444-0938. DOI: 10.1111/cxo.12670.

Bassett, Danielle S and Edward T Bullmore (2009). "Human brain networks in health and disease". In: *Current Opinion in Neurology* 22.4. Publisher: NIH Public Access, p. 340.

Belen, Ryan Anthony J. de et al. (Dec. 2020). "Computer vision in autism spectrum disorder research: a systematic review of published studies from 2009 to 2019". en. In: *Translational Psychiatry* 10.1, p. 333. ISSN: 2158-3188. DOI: 10.1038/s41398-020-01015-w. URL: https://www.nature.com/articles/s41398-020-01015-w (visited on 11/13/2022).

Beliy, Roman et al. (2019). "From voxels to pixels and back: Self-supervision in natural-image reconstruction from fMRI". In: *Advances in Neural Information Processing Systems*. Ed. by H. Wallach et al. Vol. 32. Vancouver.

Biswal, B. et al. (Oct. 1995). "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI". eng. In: *Magnetic Resonance in Medicine* 34.4, pp. 537–541. ISSN: 0740-3194. DOI: 10.1002/mrm.1910340409.

Brian, Jessica A, Lonnie Zwaigenbaum, and Angie Ip (Nov. 2019). "Standards of diagnostic assessment for autism spectrum disorder". In: *Paediatrics & Child Health* 24.7, pp. 444–451. ISSN: 1205-7088. DOI: 10.1093/pch/pxz117. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6812299/ (visited on 07/02/2023).

Bullmore, E. and O. Sporns (2009). "Complex brain networks: graph theoretical analysis of structural and functional systems". In: *Nature Reviews Neuroscience* 10.3. Publisher: Nature Publishing Group, pp. 186–198.

Bullmore, Ed and Olaf Sporns (May 2012). "The economy of brain network organization". en. In: *Nature Reviews Neuroscience* 13.5. Number: 5 Publisher: Nature Publishing Group, pp. 336–349. ISSN: 1471-0048. DOI: 10.1038/nrn3214. URL: https://www.nature.com/articles/nrn3214 (visited on 11/06/2022).

Charman, Tony and Katherine Gotham (Feb. 2013). "Measurement Issues: Screening and diagnostic instruments for autism spectrum disorders – lessons from research and practice". In: *Child and adolescent mental health* 18.1, pp. 52–63. ISSN: 1475-357X. DOI: 10.1111/j.1475-3588.2012.00664.x. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3607539/ (visited on 07/28/2023).

Chen, Liang-Chieh et al. (2015). "Semantic Image Segmentation with Deep Convolutional Nets and Fully Connected CRFs". In: *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*. Ed. by Yoshua Bengio and Yann LeCun. URL: http://arxiv.org/abs/1412.7062.

Cho, Kyunghyun et al. (Oct. 2014). "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation". In: *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Doha, Qatar: Association for Computational Linguistics, pp. 1724–1734. DOI: 10.3115/v1/D14-1179. URL: https://www.aclweb.org/anthology/D14-1179 (visited on 04/08/2021).

Chung, Fan (Dec. 1996). *Spectral Graph Theory*. en. Vol. 92. CBMS Regional Conference Series in Mathematics. ISSN: 0160-7642, 2380-5668. American Mathematical Society. ISBN: 978-0-8218-0315-8 978-0-8218-8936-7 978-1-4704-2452-7. DOI: 10.1090/cbms/092. URL: https://www.ams.org/cbms/092 (visited on 11/19/2022).

Cohen, Ori et al. (Oct. 2014). "Fmri-based robotic embodiment: Controlling a humanoid robot by thought using real-time fmri". In: *Presence: Teleoperators and Virtual Environments* 23.3, pp. 229–241. ISSN: 1054-7460. DOI: 10.1162/PRES_a_00191. URL: https://doi.org/10.1162/PRES_a_00191 (visited on 05/17/2022).

Defferrard, Michaël, Xavier Bresson, and Pierre Vandergheynst (Dec. 2016). "Convolutional neural networks on graphs with fast localized spectral filtering". In: *Proceedings of the 30th International Conference on Neural Information Processing Systems*. NIPS'16. Red Hook, NY, USA: Curran Associates Inc., pp. 3844–3852. ISBN: 978-1-5108-3881-9. (Visited on 11/18/2022).

Deng, Jia et al. (June 2009). "ImageNet: A large-scale hierarchical image database". In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. ISSN: 1063-6919. Miami, pp. 248–255. DOI: 10.1109/CVPR.2009.5206848.

Di Martino, A. et al. (June 2014). "The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism". en. In: *Molecular Psychiatry* 19.6. Number: 6 Publisher: Nature Publishing Group, pp. 659–667. ISSN: 1476-5578. DOI: 10.1038/mp.2013.78. URL: https://www.nature.com/articles/mp201378 (visited on 07/10/2023).

Di Martino, Adriana et al. (Mar. 2017). "Enhancing studies of the connectome in autism using the autism brain imaging data exchange II". en. In: *Scientific Data* 4.1. Number: 1 Publisher: Nature Publishing Group, p. 170010. ISSN: 2052-4463. DOI: 10.1038/sdata.2017.10. URL: https://www.nature.com/articles/sdata201710 (visited on 02/02/2023).

Dong, Xiaowen et al. (Dec. 2016). "Learning Laplacian Matrix in Smooth Graph Signal Representations". In: *IEEE Transactions on Signal Processing* 64.23. Conference Name: IEEE Transactions on Signal Processing, pp. 6160–6173. ISSN: 1941-0476. DOI: 10.1109/TSP.2016.2602809.

Dong, Xiaowen et al. (May 2019). "Learning Graphs From Data: A Signal Representation Perspective". In: *IEEE Signal Processing Magazine* 36.3. Conference Name: IEEE Signal Processing Magazine, pp. 44–63. ISSN: 1558-0792. DOI: 10.1109/MSP.2018.2887284.

Eickenberg, Michael et al. (May 2017). "Seeing it all: Convolutional network layers map the function of the human visual system". In: *NeuroImage* 152, pp. 184–194. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2016.10.001. URL: http://www.sciencedirect.com/science/article/pii/S1053811916305481 (visited on 12/16/2020).

Emerson, Robert W. et al. (June 2017). "Functional neuroimaging of high-risk 6-month-old infants predicts a diagnosis of autism at 24 months of age". eng. In: *Science Translational Medicine* 9.393, eaag2882. ISSN: 1946-6242. DOI: 10.1126/scitranslmed.aag2882.

Escala-Garcia, Maria et al. (Jan. 2020). "A network analysis to identify mediators of germline-driven differences in breast cancer prognosis". en.

In: *Nature Communications* 11.1. Number: 1 Publisher: Nature Publishing Group, p. 312. ISSN: 2041-1723. DOI: 10.1038/s41467-019-14100-6. URL: https://www.nature.com/articles/s41467-019-14100-6 (visited on 11/19/2022).

Eslami, Taban et al. (Nov. 2019). "ASD-DiagNet: A Hybrid Learning Approach for Detection of Autism Spectrum Disorder Using fMRI Data". In: *Frontiers in Neuroinformatics* 13, p. 70. ISSN: 1662-5196. DOI: 10.3389/fninf.2019.00070. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6890833/ (visited on 11/19/2022).

Fang, Tao, Yu Qi, and Gang Pan (2020). "Reconstructing Perceptive Images from Brain Activity by Shape-Semantic GAN". In: *Advances in Neural Information Processing Systems*. Ed. by H. Larochelle et al. Vol. 33. Curran Associates, Inc., pp. 13038–13048. (Visited on 12/01/2020).

Finn, Emily S. (Dec. 2021). "Is it time to put rest to rest?" en. In: *Trends in Cognitive Sciences* 25.12, pp. 1021–1032. ISSN: 1364-6613. DOI: 10.1016/j.tics.2021.09.005. URL: https://www.sciencedirect.com/science/article/pii/S1364661321002345 (visited on 11/08/2022).

Fox, Michael D. and Marcus E. Raichle (Sept. 2007). "Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging". en. In: *Nature Reviews Neuroscience* 8.9. Number: 9 Publisher: Nature Publishing Group, pp. 700–711. ISSN: 1471-0048. DOI: 10.1038/nrn2201. URL: https://www.nature.com/articles/nrn2201 (visited on 11/22/2022).

Gandal, Michael J. et al. (Nov. 2022). "Broad transcriptomic dysregulation occurs across the cerebral cortex in ASD". en. In: *Nature*. Publisher: Nature Publishing Group, pp. 1–8. ISSN: 1476-4687. DOI: 10.1038/s41586-022-05377-7. URL: https://www.nature.com/articles/s41586-022-05377-7 (visited on 11/10/2022).

Gaziv, Guy et al. (Sept. 2020). *Self-Supervised Natural Image Reconstruction and Rich Semantic Classification from Brain Activity*. preprint. Neuroscience. DOI: 10.1101/2020.09.06.284794. URL: http://biorxiv.org/lookup/doi/10.1101/2020.09.06.284794 (visited on 03/19/2021).

Gerven, Marcel A. J. van, Floris P. de Lange, and Tom Heskes (Sept. 2010). "Neural Decoding with Hierarchical Generative Models". In: *Neural Computation* 22.12. Number: 12 Publisher: MIT Press, pp. 3127–3142. ISSN: 0899-7667. DOI: 10.1162/NECO_a_00047. (Visited on 11/02/2020).

Gong, Xiaoyun et al. (Aug. 2021). "Comparing visual preferences between autism spectrum disorder (ASD) and normal children to explore the characteristics of visual preference of ASD children by improved visual preference paradigm: a case-control study". In: *Translational Pediatrics* 10.8, pp. 2006–2015. ISSN: 2224-4336. DOI: 10.21037/tp-21-294. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8429864/ (visited on 11/22/2022).

Goodfellow, Ian J. et al. (Dec. 2014). "Generative adversarial nets". In: *Proceedings of the 27th International Conference on Neural Information Processing Systems - Volume 2*. NIPS'14. Cambridge, MA, USA: MIT Press, pp. 2672–2680. (Visited on 04/17/2021).

Greene, Abigail S. et al. (Aug. 2020). "How Tasks Change Whole-Brain Functional Organization to Reveal Brain-Phenotype Relationships". eng. In: *Cell Reports* 32.8, p. 108066. ISSN: 2211-1247. DOI: 10.1016/j.celrep.2020.108066.

Greene, Rachel K. et al. (2019). "Social and nonsocial visual prediction errors in autism spectrum disorder". en. In: *Autism Research* 12.6, pp. 878–883. ISSN: 1939-3806. DOI: 10.1002/aur.2090. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/aur.2090 (visited on 11/21/2022).

Greicius, Michael D. et al. (Jan. 2003). "Functional connectivity in the resting brain: A network analysis of the default mode hypothesis". In: *Proceedings of the National Academy of Sciences* 100.1. Publisher: Proceedings of the National Academy of Sciences, pp. 253–258. DOI: 10.1073/pnas.0135058100. URL: https://www.pnas.org/doi/10.1073/pnas.0135058100 (visited on 02/02/2023).

Guidotti, Roberto et al. (July 2015). "Visual Learning Induces Changes in Resting-State fMRI Multivariate Pattern of Information". en. In: *Journal*

of Neuroscience 35.27. Publisher: Society for Neuroscience Section: Articles, pp. 9786–9798. ISSN: 0270-6474, 1529-2401. DOI: 10.1523/JNEUROSCI.3920-14.2015. URL: https://www.jneurosci.org/content/35/27/9786 (visited on 06/01/2022).

Hagen, Susan (Apr. 2012). "The Mind's Eye". In: *Rochester Review* Vol. 74,No. 4. URL: https://www.rochester.edu/pr/Review/V74N4/0402_brainscience.html (visited on 11/22/2022).

Haxby, James V. et al. (Sept. 2001). "Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex". In: *Science* 293.5539. Number: 5539 Publisher: American Association for the Advancement of Science Section: Research Article, pp. 2425–2430. ISSN: 0036-8075, 1095-9203. DOI: 10.1126/science.1063736. URL: https://science.sciencemag.org/content/293/5539/2425 (visited on 12/24/2020).

He, K. et al. (June 2016). "Deep Residual Learning for Image Recognition". In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. ISSN: 1063-6919, pp. 770–778. DOI: 10.1109/CVPR.2016.90.

He, Tong et al. (June 2018). "Is deep learning better than kernel regression for functional connectivity prediction of fluid intelligence?" In: *2018 International Workshop on Pattern Recognition in Neuroimaging (PRNI)*. Singapore: IEEE, pp. 1–4. ISBN: 978-1-5386-6859-7. DOI: 10.1109/PRNI.2018.8423958. URL: https://ieeexplore.ieee.org/document/8423958/ (visited on 08/11/2019).

Heinsfeld, Anibal Sólon et al. (Aug. 2017). "Identification of autism spectrum disorder using deep learning and the ABIDE dataset". In: *NeuroImage : Clinical* 17, pp. 16–23. ISSN: 2213-1582. DOI: 10.1016/j.nicl.2017.08.017. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5635344/ (visited on 11/19/2022).

Horikawa, Tomoyasu and Yukiyasu Kamitani (May 2017). "Generic decoding of seen and imagined objects using hierarchical visual features". In: *Nature Communications* 8.1. Number: 1 Publisher: Nature Publishing Group, p. 15037. ISSN: 2041-1723. DOI: 10.1038/ncomms15037. URL: https://www.nature.com/articles/ncomms15037 (visited on 09/29/2020).

Hsieh, Tsung-Hao, Ming-Jian Sun, and Sheng-Fu Liang (2014). "Diagnosis of Schizophrenia Patients Based on Brain Network Complexity Analysis of Resting-State fMRI". en. In: *The 15th International Conference on Biomedical Engineering*. Ed. by James Goh. IFMBE Proceedings. Cham: Springer International Publishing, pp. 203–206. ISBN: 978-3-319-02913-9. DOI: 10.1007/978-3-319-02913-9_52.

Hulvershorn, Leslie A. et al. (Sept. 2014). "Developmental Resting State Functional Connectivity for Clinicians". eng. In: *Current Behavioral Neuroscience Reports* 1.3, pp. 161–169. ISSN: 2196-2979. DOI: 10.1007/s40473-014-0020-3.

Isola, Phillip et al. (2017). "Image-To-Image Translation With Conditional Adversarial Networks". In: pp. 1125–1134. URL: https://openaccess.thecvf.com/content_cvpr_2017/html/Isola_Image-To-Image_Translation_With_CVPR_2017_paper.html (visited on 04/08/2021).

Jaiswal, Ayush et al. (May 2018). *Large-Scale Unsupervised Deep Representation Learning for Brain Structure*. arXiv:1805.01049 [cs, stat]. DOI: 10.48550/arXiv.1805.01049. URL: http://arxiv.org/abs/1805.01049 (visited on 11/22/2022).

Jang, Hojin, Devin McCormack, and Frank Tong (Dec. 2021). "Noise-trained deep neural networks effectively predict human vision and its neural responses to challenging images". en. In: *PLOS Biology* 19.12. Publisher: Public Library of Science, e3001418. ISSN: 1545-7885. DOI: 10.1371/journal.pbio.3001418. (Visited on 07/10/2022).

Ji, Chengtao, Natasha M. Maurits, and Jos B. T. M. Roerdink (2019). "Comparison of Brain Connectivity Networks Using Local Structure Analysis". In: *Complex Networks and Their Applications VII*. Ed. by Luca Maria Aiello et al. Vol. 813. Cham: Springer International Publishing, pp. 639–651. ISBN: 978-3-030-05413-7 978-3-030-05414-4. DOI: 10.1007/978-3-030-05414-4_51. URL: http://link.springer.com/10.1007/978-3-030-05414-4_51 (visited on 08/16/2019).

Jiang, Hao et al. (Dec. 2020). "Hi-GCN: A hierarchical graph convolution network for graph embedding learning of brain network and brain disorders prediction". en. In: *Computers in Biology and Medicine* 127, p. 104096. ISSN: 0010-4825. DOI: 10.1016/j.compbiomed.2020.104096. URL: https://www.sciencedirect.com/science/article/pii/S0010482520304273 (visited on 02/01/2023).

Jordan, M. I. and T. M. Mitchell (July 2015). "Machine learning: Trends, perspectives, and prospects". In: *Science* 349.6245. Publisher: American Association for the Advancement of Science, pp. 255–260. DOI: 10.1126/science.aaa8415. URL: https://www.science.org/doi/abs/10.1126/science.aaa8415 (visited on 11/12/2022).

Kalofolias, Vassilis (May 2016). "How to Learn a Graph from Smooth Signals". en. In: *Proceedings of the 19th International Conference on Artificial Intelligence and Statistics*. ISSN: 1938-7228. PMLR, pp. 920–929. URL: https://proceedings.mlr.press/v51/kalofolias16.html (visited on 11/19/2022).

Kam, Tae-Eui, Heung-Il Suk, and Seong-Whan Lee (2017). "Multiple functional networks modeling for autism spectrum disorder diagnosis". en. In: *Human Brain Mapping* 38.11, pp. 5804–5821. ISSN: 1097-0193. DOI: 10.1002/hbm.23769. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/hbm.23769 (visited on 11/19/2022).

Kamitani, Yukiyasu and Frank Tong (May 2005). "Decoding the visual and subjective contents of the human brain". In: *Nature Neuroscience* 8.5. Number: 5, pp. 679–685. ISSN: 1097-6256. DOI: 10.1038/nn1444.

Kamnitsas, Konstantinos et al. (Feb. 2017). "Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation". In: *Medical Image Analysis* 36, pp. 61–78. ISSN: 1361-8415. DOI: 10.1016/j.media.2016.10.004. URL: https://www.sciencedirect.com/science/article/pii/S1361841516301839 (visited on 03/31/2021).

Kay, Kendrick N. et al. (Mar. 2008). "Identifying natural images from human brain activity". In: *Nature* 452.7185. Number: 7185, pp. 352–355. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/nature06713. URL: http://www.nature.com/articles/nature06713 (visited on 11/27/2020).

Kazi, Anees et al. (2019). "InceptionGCN: Receptive Field Aware Graph Convolutional Network for Disease Prediction". en. In: *Information Processing in Medical Imaging*. Ed. by Albert C. S. Chung et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 73–85. ISBN: 978-3-030-20351-1. DOI: 10.1007/978-3-030-20351-1_6.

Kessler, K., R. A. Seymour, and G. Rippon (Dec. 2016). "Brain oscillations and connectivity in autism spectrum disorders (ASD): new approaches to methodology, measurement and modelling". en. In: *Neuroscience & Biobehavioral Reviews* 71, pp. 601–620. ISSN: 0149-7634. DOI: 10.1016/j.neubiorev.2016.10.002. URL: https://www.sciencedirect.com/science/article/pii/S0149763416303530 (visited on 05/27/2022).

Khosla, Meenakshi et al. (2018). "3D Convolutional Neural Networks for Classification of Functional Connectomes". en. In: *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Ed. by Danail Stoyanov et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 137–145. ISBN: 978-3-030-00889-5. DOI: 10.1007/978-3-030-00889-5_16.

Kingma, Diederik P. and Jimmy Ba (2015). "Adam: A Method for Stochastic Optimization". In: *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*. Ed. by Yoshua Bengio and Yann LeCun. URL: http://arxiv.org/abs/1412.6980.

Kingma, Diederik P. and Max Welling (2014). "Auto-Encoding Variational Bayes". In: *2nd International Conference on Learning Representations*. Ed. by Yoshua Bengio and Yann LeCun. Banff. URL: http://arxiv.org/abs/1312.6114.

Kipf, Thomas N and Max Welling (2017). "SEMI-SUPERVISED CLASSIFICATION WITH GRAPH CONVOLUTIONAL NETWORKS". en. In: p. 14.

Kolodny, Tamar et al. (Mar. 2020). "Response Dissociation in Hierarchical Cortical Circuits: a Unique Feature of Autism Spectrum Disorder". In: *The Journal of Neuroscience* 40.11, pp. 2269–2281. ISSN: 0270-6474. DOI: 10.

1523/JNEUROSCI.2376-19.2020. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7083290/ (visited on 05/20/2022).

Kowalczyk, Olivia S. et al. (Oct. 2021). "Task-Based Functional Connectivity in Attention-Deficit/Hyperactivity Disorder: A Systematic Review". en. In: *Biological Psychiatry Global Open Science*. ISSN: 2667-1743. DOI: 10.1016/j.bpsgos.2021.10.006. URL: https://www.sciencedirect.com/science/article/pii/S2667174321001233 (visited on 06/14/2022).

Kriegeskorte, Nikolaus (Nov. 2015). "Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing". In: *Annual Review of Vision Science* 1, pp. 417–446. ISSN: 2374-4650. DOI: 10.1146/annurev-vision-082114-035447.

Krizhevsky, Alex (2009). *Learning multiple layers of features from tiny images*. Tech. rep. University of Toronto.

Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton (2012). "ImageNet Classification with Deep Convolutional Neural Networks". In: *Advances in Neural Information Processing Systems*. Vol. 25. Lake Tahoe, pp. 1097–1105. (Visited on 12/14/2020).

Ktena, Sofia Ira et al. (Apr. 2018). "Metric learning with spectral graph convolutions on brain connectivity networks". en. In: *NeuroImage* 169, pp. 431–442. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2017.12.052. URL: https://www.sciencedirect.com/science/article/pii/S1053811917310765 (visited on 03/21/2023).

Lahat, Dana, Tülay Adali, and Christian Jutten (Sept. 2015). "Multimodal Data Fusion: An Overview of Methods, Challenges, and Prospects". In: *Proceedings of the IEEE* 103.9. Conference Name: Proceedings of the IEEE, pp. 1449–1477. ISSN: 1558-2256. DOI: 10.1109/JPROC.2015.2460697.

LeCun, Y. et al. (Dec. 1989). "Backpropagation Applied to Handwritten Zip Code Recognition". In: *Neural Computation* 1.4. Number: 4 Publisher: MIT Press, pp. 541–551. ISSN: 0899-7667. DOI: 10.1162/neco.1989.1.4.541. URL: https://doi.org/10.1162/neco.1989.1.4.541 (visited on 03/03/2021).

Lecun, Y. et al. (Nov. 1998). "Gradient-based learning applied to document recognition". In: *Proceedings of the IEEE* 86.11, pp. 2278–2324. ISSN: 1558-2256. DOI: 10.1109/5.726791.

LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton (May 2015). "Deep learning". en. In: *Nature* 521.7553. Number: 7553 Publisher: Nature Publishing Group, pp. 436–444. ISSN: 1476-4687. DOI: 10.1038/nature14539. URL: https://www.nature.com/articles/nature14539 (visited on 11/12/2022).

Leskovec, Jure, Daniel Huttenlocher, and Jon Kleinberg (Apr. 2010). "Predicting positive and negative links in online social networks". In: *Proceedings of the 19th international conference on World wide web*. WWW '10. New York, NY, USA: Association for Computing Machinery, pp. 641–650. ISBN: 978-1-60558-799-8. DOI: 10.1145/1772690.1772756. URL: https://doi.org/10.1145/1772690.1772756 (visited on 11/18/2022).

Li, Xiaoxiao et al. (2018). "Brain Biomarker Interpretation in ASD Using Deep Learning and fMRI". en. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*. Ed. by Alejandro F. Frangi et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 206–214. ISBN: 978-3-030-00931-1. DOI: 10.1007/978-3-030-00931-1_24.

Li, Xiaoxiao et al. (Feb. 2020a). "Graph Embedding Using Infomax for ASD Classification and Brain Functional Difference Detection". In: *Proceedings of SPIE–the International Society for Optical Engineering* 11317, p. 1131702. ISSN: 0277-786X. DOI: 10.1117/12.2549451. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7569478/ (visited on 01/31/2023).

Li, Xiaoxiao et al. (Oct. 2020b). "Pooling Regularized Graph Neural Network for fMRI Biomarker Analysis". In: *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention* 12267, pp. 625–635. DOI: 10.1007/978-3-030-59728-3_61. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7544244/ (visited on 01/31/2023).

Liu, Yuan et al. (2022). "Functional Reorganization After Four-Week Brain–Computer Interface-Controlled Supernumerary Robotic Finger Training: A Pilot Study of Longitudinal Resting-State fMRI". In: *Frontiers in Neuroscience* 15. ISSN: 1662-453X. URL: https://www.frontiersin.org/articles/10.3389/fnins.2021.766648 (visited on 11/06/2022).

Logothetis, Nikos K. and David L. Sheinberg (1996). "Visual Object Recognition". In: *Annual Review of Neuroscience* 19.1, pp. 577–621. DOI: 10.1146/annurev.ne.19.030196.003045. URL: https://doi.org/10.1146/annurev.ne.19.030196.003045 (visited on 02/23/2021).

Mahendran, A. and A. Vedaldi (June 2015). "Understanding deep image representations by inverting them". In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. ISSN: 1063-6919. Boston, pp. 5188–5196. DOI: 10.1109/CVPR.2015.7299155.

Minaee, S. et al. (2021). "Image Segmentation Using Deep Learning: A Survey". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 1–1. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2021.3059968.

Miyawaki, Yoichi et al. (Dec. 2008). "Visual Image Reconstruction from Human Brain Activity using a Combination of Multiscale Local Image Decoders". In: *Neuron* 60.5. Number: 5, pp. 915–929. ISSN: 08966273. DOI: 10.1016/j.neuron.2008.11.004. URL: https://linkinghub.elsevier.com/retrieve/pii/S0896627308009586 (visited on 12/05/2020).

Naselaris, Thomas et al. (Sept. 2009). "Bayesian Reconstruction of Natural Images from Human Brain Activity". In: *Neuron* 63.6. Number: 6, pp. 902–915. ISSN: 0896-6273. DOI: 10.1016/j.neuron.2009.09.006. URL: http://www.sciencedirect.com/science/article/pii/S0896627309006850 (visited on 12/02/2020).

Naselaris, Thomas et al. (May 2011). "Encoding and decoding in fMRI". In: *NeuroImage*. Multivariate Decoding and Brain Reading 56.2. Number: 2, pp. 400–410. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2010.07.073. URL: http://www.sciencedirect.com/science/article/pii/S1053811910010657 (visited on 12/17/2020).

Nestor, Adrian et al. (Sept. 2020). "The Face of Image Reconstruction: Progress, Pitfalls, Prospects". In: *Trends in Cognitive Sciences* 24.9. Publisher: Elsevier, pp. 747–759. ISSN: 1364-6613, 1879-307X. DOI: 10.1016/j.tics.2020.06.006. URL: https://www.cell.com/trends/cognitive-sciences/abstract/S1364-6613(20)30147-9 (visited on 08/23/2021).

Ng, Andrew, Michael Jordan, and Yair Weiss (2001). "On Spectral Clustering: Analysis and an algorithm". In: *Advances in Neural Information Processing Systems*. Vol. 14. MIT Press. (Visited on 11/19/2022).

Nielsen, Jared et al. (2013). "Multisite functional connectivity MRI classification of autism: ABIDE results". In: *Frontiers in Human Neuroscience* 7. ISSN: 1662-5161. URL: https://www.frontiersin.org/articles/10.3389/fnhum.2013.00599 (visited on 07/10/2023).

Niepert, Mathias, Mohamed Ahmed, and Konstantin Kutzkov (June 2016). "Learning convolutional neural networks for graphs". In: *Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48*. ICML'16. New York, NY, USA: JMLR.org, pp. 2014–2023. (Visited on 01/31/2023).

O'Connor, Erin E. and Thomas A. Zeffiro (2019). "Why is Clinical fMRI in a Resting State?" In: *Frontiers in Neurology* 10. ISSN: 1664-2295. URL: https://www.frontiersin.org/article/10.3389/fneur.2019.00420 (visited on 06/01/2022).

Ogawa, S et al. (Dec. 1990). "Brain magnetic resonance imaging with contrast dependent on blood oxygenation." In: *Proceedings of the National Academy of Sciences of the United States of America* 87.24. Number: 24, pp. 9868–9872. ISSN: 0027-8424. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC55275/ (visited on 12/24/2020).

O'riordan, Michelle A. (Sept. 2004). "Superior visual search in adults with autism". eng. In: *Autism: The International Journal of Research and Practice* 8.3, pp. 229–248. ISSN: 1362-3613. DOI: 10.1177/1362361304045219.

Pagliaccio, David et al. (Oct. 2019). "Task-based fMRI predicts response and remission to exposure therapy in obsessive-compulsive disorder". In: *Proceedings of the National Academy of Sciences* 116.41. Publisher: Proceedings

of the National Academy of Sciences, pp. 20346–20353. DOI: 10.1073/pnas.1909199116. URL: https://www.pnas.org/doi/10.1073/pnas.1909199116 (visited on 11/07/2022).

Parisot, Sarah et al. (2017). "Spectral Graph Convolutions for Population-Based Disease Prediction". en. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2017*. Ed. by Maxime Descoteaux et al. Lecture Notes in Computer Science. Cham: Springer International Publishing, pp. 177–185. ISBN: 978-3-319-66179-7. DOI: 10.1007/978-3-319-66179-7_21.

Parisot, Sarah et al. (Aug. 2018). "Disease prediction using graph convolutional networks: Application to Autism Spectrum Disorder and Alzheimer's disease". en. In: *Medical Image Analysis* 48, pp. 117–130. ISSN: 1361-8415. DOI: 10.1016/j.media.2018.06.001. URL: https://www.sciencedirect.com/science/article/pii/S1361841518303554 (visited on 08/18/2021).

Pinto, Nicolas et al. (Nov. 2009). "A High-Throughput Screening Approach to Discovering Good Forms of Biologically Inspired Visual Representation". In: *PLOS Computational Biology* 5.11. Publisher: Public Library of Science, e1000579. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1000579. URL: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1000579 (visited on 03/31/2021).

Poldrack, Russell A. and Martha J. Farah (Oct. 2015). "Progress and challenges in probing the human brain". In: *Nature* 526.7573. Number: 7573 Publisher: Nature Publishing Group, pp. 371–379. ISSN: 1476-4687. DOI: 10.1038/nature15692. URL: https://www.nature.com/articles/nature15692 (visited on 12/24/2020).

Rakhimberdina, Zarina, Xin Liu, and Tsuyoshi Murata (Jan. 2020). "Population Graph-Based Multi-Model Ensemble Method for Diagnosing Autism Spectrum Disorder". In: *Sensors* 20.21. Number: 21 Publisher: Multidisciplinary Digital Publishing Institute, p. 6001. DOI: 10.3390/s20216001. URL: https://www.mdpi.com/1424-8220/20/21/6001 (visited on 12/01/2020).

Rakhimberdina, Zarina, Xin Liu, and Tsuyoshi Murata (2022). "Strengthening Robustness Under Adversarial Attacks Using Brain Visual Codes". In: *IEEE Access* 10. Conference Name: IEEE Access, pp. 96149–96158. ISSN: 2169-3536. DOI: 10.1109/ACCESS.2022.3204995.

Rakhimberdina, Zarina and Tsuyoshi Murata (2020). "Linear Graph Convolutional Model for Diagnosing Brain Disorders". en. In: *Complex Networks and Their Applications VIII*. Ed. by Hocine Cherifi et al. Studies in Computational Intelligence. Cham: Springer International Publishing, pp. 815–826. ISBN: 978-3-030-36683-4. DOI: 10.1007/978-3-030-36683-4_65.

Rakhimberdina, Zarina et al. (Dec. 2021). "Natural Image Reconstruction From fMRI Using Deep Learning: A Survey". en. In: *Frontiers in Neuroscience* 15, p. 795488. ISSN: 1662-453X. DOI: 10.3389/fnins.2021.795488. (Visited on 03/09/2022).

Ripley, David L. and Thomas Politzer (2010). "Vision disturbance after TBI". eng. In: *NeuroRehabilitation* 27.3, pp. 215–216. ISSN: 1878-6448. DOI: 10.3233/NRE-2010-0599.

Roelfsema, Pieter R., Damiaan Denys, and P. Christiaan Klink (July 2018). "Mind Reading and Writing: The Future of Neurotechnology". In: *Trends in Cognitive Sciences* 22.7. Number: 7, pp. 598–610. ISSN: 1364-6613. DOI: 10.1016/j.tics.2018.04.001. (Visited on 03/11/2021).

Schoenmakers, Sanne et al. (Dec. 2013). "Linear reconstruction of perceived images from human brain activity". In: *NeuroImage* 83, pp. 951–961. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2013.07.043. URL: http://www.sciencedirect.com/science/article/pii/S1053811913007994 (visited on 11/16/2020).

Schrimpf, Martin et al. (Sept. 2018). "Brain-Score: Which Artificial Neural Network for Object Recognition is most Brain-Like?" In: *bioRxiv*. Publisher: Cold Spring Harbor Laboratory Section: New Results, p. 407007. DOI: 10.1101/407007. URL: https://www.biorxiv.org/content/10.1101/407007v1 (visited on 11/05/2020).

Shen, Guohua et al. (Jan. 2019a). "Deep image reconstruction from human brain activity". In: *PLOS Computational Biology* 15.1, e1006633. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1006633. (Visited on 11/02/2020).

Shen, Guohua et al. (2019b). "End-to-End Deep Image Reconstruction From Human Brain Activity". In: *Frontiers in Computational Neuroscience* 13. ISSN: 1662-5188. DOI: 10.3389/fncom.2019.00021. (Visited on 09/23/2020).

Sherkatghanad, Zeinab et al. (2019). "Automated Detection of Autism Spectrum Disorder Using a Convolutional Neural Network". eng. In: *Frontiers in Neuroscience* 13, p. 1325. ISSN: 1662-4548. DOI: 10.3389/fnins.2019.01325.

Shuman, David I, Benjamin Ricaud, and Pierre Vandergheynst (Mar. 2016). "Vertex-frequency analysis on graphs". en. In: *Applied and Computational Harmonic Analysis* 40.2, pp. 260–291. ISSN: 1063-5203. DOI: 10.1016/j.acha.2015.02.005. URL: https://www.sciencedirect.com/science/article/pii/S1063520315000214 (visited on 11/19/2022).

Shuman, David I et al. (May 2013). "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains". In: *IEEE Signal Processing Magazine* 30.3. Conference Name: IEEE Signal Processing Magazine, pp. 83–98. ISSN: 1558-0792. DOI: 10.1109/MSP.2012.2235192.

St-Yves, G. and T. Naselaris (Oct. 2018). "Generative Adversarial Networks Conditioned on Brain Activity Reconstruct Seen Images". In: *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. ISSN: 2577-1655. Miyazaki, pp. 1054–1061. DOI: 10.1109/SMC.2018.00187.

Steyvers, Mark et al. (Mar. 2022). "Bayesian modeling of human–AI complementarity". In: *Proceedings of the National Academy of Sciences* 119.11. Publisher: Proceedings of the National Academy of Sciences, e2111547119. DOI: 10.1073/pnas.2111547119. URL: https://www.pnas.org/doi/10.1073/pnas.2111547119 (visited on 06/29/2022).

Straaten, Elisabeth C. W. van and Cornelis J. Stam (Jan. 2013). "Structure out of chaos: Functional brain network analysis with EEG, MEG, and functional MRI". en. In: *European Neuropsychopharmacology*. Neural Networks

in Psychiatry 23.1, pp. 7–18. ISSN: 0924-977X. DOI: 10.1016/j.euroneuro.2012.10.010. URL: https://www.sciencedirect.com/science/article/pii/S0924977X12002891 (visited on 11/22/2022).

Thirion, Bertrand et al. (Dec. 2006). "Inverse retinotopy: inferring the visual content of images from brain activation patterns". In: *NeuroImage* 33.4, pp. 1104–1116. ISSN: 1053-8119. DOI: 10.1016/j.neuroimage.2006.06.062.

Tian, Yapeng and Chenliang Xu (2021). "Can Audio-Visual Integration Strengthen Robustness Under Multimodal Attacks?" en. In: *Proceedings of the IEEE/CVF Conference on CVPR*, pp. 5601–5611. (Visited on 07/14/2021).

Utzerath, Christian et al. (Dec. 2018). "Adolescents with autism show typical fMRI repetition suppression, but atypical surprise response". en. In: *Cortex* 109, pp. 25–34. ISSN: 0010-9452. DOI: 10.1016/j.cortex.2018.08.019. URL: https://www.sciencedirect.com/science/article/pii/S0010945218302715 (visited on 11/21/2022).

Ventresca, Mario (2019). "Using Algorithmic Complexity to Differentiate Cognitive States in fMRI". en. In: *Complex Networks and Their Applications VII*. Ed. by Luca Maria Aiello et al. Studies in Computational Intelligence. Cham: Springer International Publishing, pp. 663–674. ISBN: 978-3-030-05414-4. DOI: 10.1007/978-3-030-05414-4_53.

Werling, Donna M. and Daniel H. Geschwind (Apr. 2013). "Sex differences in autism spectrum disorders". eng. In: *Current Opinion in Neurology* 26.2, pp. 146–153. ISSN: 1473-6551. DOI: 10.1097/WCO.0b013e32835ee548.

Wu, Felix et al. (May 2019). "Simplifying Graph Convolutional Networks". en. In: *Proceedings of the 36th International Conference on Machine Learning*. ISSN: 2640-3498. PMLR, pp. 6861–6871. URL: https://proceedings.mlr.press/v97/wu19e.html (visited on 11/19/2022).

Youden, W. J. (1950). "Index for rating diagnostic tests". en. In: *Cancer* 3.1, pp. 32–35. ISSN: 1097-0142. DOI: 10.1002/1097-0142(1950)3:1<32::AID-CNCR2820030106>3.0.CO;2-3. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/1097-0142%281950%293%3A1%3C32%3A%3AAID-CNCR2820030106%3E3.0.CO%3B2-3 (visited on 04/24/2023).

Zhang, K. et al. (July 2017). "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising". In: *IEEE Transactions on Image Processing* 26.7. Conference Name: IEEE Transactions on Image Processing, pp. 3142–3155. ISSN: 1941-0042. DOI: 10.1109/TIP.2017.2662206.

Zhang, Shu et al. (Mar. 2016). "Characterizing and differentiating task-based and resting state fMRI signals via two-stage sparse representations". en. In: *Brain Imaging and Behavior* 10.1, pp. 21–32. ISSN: 1931-7557, 1931-7565. DOI: 10.1007/s11682-015-9359-7. URL: http://link.springer.com/10.1007/s11682-015-9359-7 (visited on 11/07/2022).

Zhang, Xiang et al. (Mar. 2021a). "A survey on deep learning-based non-invasive brain signals: recent advances and new frontiers". eng. In: *Journal of Neural Engineering* 18.3. ISSN: 1741-2552. DOI: 10.1088/1741-2552/abc902.

Zhang, Yizhen et al. (Sept. 2021b). "Naturalistic Stimuli: A Paradigm for Multi-Scale Functional Characterization of the Human Brain". eng. In: *Current Opinion in Biomedical Engineering* 19, p. 100298. ISSN: 2468-4511. DOI: 10.1016/j.cobme.2021.100298.

Zhu, Xiaofan and Michael Rabbat (Mar. 2012). "Approximating signals supported on graphs". In: *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. ISSN: 2379-190X, pp. 3921–3924. DOI: 10.1109/ICASSP.2012.6288775.