Static and Dynamic Functional Connectivity analysis in individuals with Autism Spectrum Disorder

Thesis submitted in partial fulfillment
of the requirements for the degree of

Master of Science
in
Computer Science and Engineering by Research

by

Pindi Krishna Chandra Prasad
2020701035
krishna.chandra@research.iiit.ac.in

International Institute of Information Technology, Hyderabad
(Deemed to be University)
Hyderabad - 500 032, INDIA
July 2023

International Institute of Information Technology
Hyderabad, India

CERTIFICATE

It is certified that the work contained in this thesis, titled " **Static and Dynamic Function Connectivity analysis in individuals with Autism Spectrum Disorder** " by **Pindi Krishna Chandra Prasad**, has been carried out under my supervision and is not submitted elsewhere for a degree.

Date

Adviser: Dr. Bapi Raju S

To *my Mother*

Acknowledgements

My adviser Dr. Bapi Raju S, my parents, and my friends deserve my gratitude for their unwavering support and direction during my time at the Cognitive Science Lab at IIIT Hyderabad.

If this thesis and my graduate school experience are a tree, it originated from a seed laid by a year-long internship at the Health care and Artificial Intelligence unit. The knowledge and strength I gained during this internship gave me the foundation to pursue my research during my MS program. I am immensely grateful to Dr. Bapi Raju, Dr. Vinod P K, Dr. C V Jawahar, Dr. Deva Priyakumar, my project coordinator Dr. Manasa K, my mentor Ruchi Chauhan, and my teammates Yash Khare, Anirudh Reddy, Viraj Bagal, and Rishal Agarwal for their unwavering support throughout the internship. Without their guidance and assistance, I would not have been able to lay the groundwork for my MS by research programme under the supervision of my adviser.

I would also like to thank IIIT Hyderabad and especially the faculty of Cognitive Science for the classroom courses that helped me understand the length and breadth of quality research work in my field. I would especially like to thank Dr. Vineeth Gandhi for my knowledge of Machine Learning and Dr. Vemuri for my knowledge of Neuroscience through their courses.

Finally, I sincerely and humbly thank Dr. Kamalaker Dadi and my adviser Dr. Bapi Raju for the technical discussions and contributions that lead to this thesis. I would also like to thank iHub-Data and IIIT Hyderabad for their support.

Abstract

Autism Spectrum Disorder (ASD) is a lifelong heterogeneous developmental disorder that is characterized by abnormal development of the brain. Deficits in social skills, incapability to articulate language, abnormal sensory-motor movements, and stereotyped behaviors are mainly observed in children with ASD. The underlying biological markers for the diagnosis and treatment of ASD are not known yet. The current behavior-based diagnosis of ASD is arduous and requires expertise. It has been demonstrated that the non-invasive method such as resting-state functional magnetic resonance imaging (rs-fMRI) provides a means to examine functional connectivity patterns in the brain, which can help diagnose various neurodegenerative and psychiatric disorders, including ASD. Most previous studies associated ASD with atypical functional connectivity (FC) between different pairs of regions. However, brain connectivity is dynamic and varies extensively among brain states. In this thesis, we explored both static and dynamic functional connectivity based features to understand the fundamental group differences between ASD patients and typically developing (TD) subjects.

Firstly, we proposed a Multilayer Perceptron (MLP) based classification model with autoencoder pretraining for classifying ASD from TD based on static functional connectivity (sFNC) extracted from rs-fMRI scans of the ABIDE-1 (a publicly available dataset from Autism Brain Imaging Data Exchange consortium). Our model achieves new state-of-the-art performance on the ABIDE-1 dataset with a 10-fold cross-validation accuracy of 74.82%. Further, we use the Integrated Gradients (IG) and DeepLIFT techniques to identify the correlations between brain regions that contribute most to the classification task. Our analysis identifies the following regions associated with ASD: Left Lingual Gyrus, Right Insula, Right Cuneus, Right Middle Frontal Gyrus, and Left Superior Temporal Gyrus. Interestingly, these regions in the brain are primarily responsible for social cognition, language, attention, decision-making, and visual processing, which are known to be altered in ASD.

Secondly, we investigated the dynamic functional connectivity (dFNC) between 53 independent components among 188 ASD and 195 TD subjects sampled from the ABIDE-I consortium. We estimated dFNC using sliding window-based approaches and identified four distinct dynamic states through hard-clustering analysis. Hyper-connectivity within the cognitive control domain, between cognitive control and default mode network have been identified among ASD subjects. Hyper-connectivity within the default mode network has been found among TD individuals. Further, we estimated the dynamic temporal properties such as fractional time spent, mean dwell time per state and observed significant differences between ASD and TD groups. ASD subjects are found to have significantly longer dwell

time in one of the states (4) when compared to TD individuals. We also found a significantly increased occurrence of the same state (4) in ASD subjects, whereas other states (1 and 3) are more frequent in TD subjects.

Overall, both static and dynamic-based methods have been explored to find out the ASD biomarkers. While there is broad consensus in the brain network profiles between sFNC and dFNC, the temporal profile of brain state dynamics is additionally available with dFNC analysis and may potentially contribute to disease biomarkers.

In addition, we have done extensive replication studies of both sFNC and dFNC based models for ASD classification or characterization in the literature. We observe great discrepancy between what is reported and what could be replicated or reproduced. Based on these analyses, we propose a set of recommendations for future studies that encompass factors such as dataset selection criteria, preprocessing pipepline, proper reporting of selected samples, atlas selection, and hyperparameter choices and reporting for the proposed models.

Contents

[1] Prasad, P.K.C., Khare, Y., Dadi, K., Vinod, P.K. and Surampudi, B.R. (2022). Deep Learning Approach for Classification and Interpretation of Autism Spectrum Disorder. In International Joint Conference on Neural Networks (IJCNN-2022) (pp. 1-8). https://ieeexplore.ieee.org/document/9892350/

[2] Prasad, P.K.C., Dadi, K., and Surampudi, B.R. (2023). Dynamic functional connectivity analysis in individuals with Autism Spectrum Disorder. In International Joint Conference on Neural Networks (IJCNN-2023).

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Autism Spectrum Disorder

Introduction Autism Spectrum Disorder, commonly known as ASD, is a developmental disorder that affects an individuals social communication and interaction, as well as their behavior and interests. It is a complex and multifaceted disorder that can manifest itself in different ways and can range from mild to severe. ASD, as the very name suggests, is a spectrum disorder with a broad range of types, severities, and symptoms [49].

Symptoms The symptoms of ASD typically appear in early childhood, although they may not be noticeable until later in life. Some of the most common signs of ASD include difficulty with social interactions, repetitive behaviors, difficulty with communication, and a narrow range of interests or activities. Children with ASD may have trouble making eye contact, understanding social cues, or engaging in back-and-forth communication. They may also have specific routines or rituals that they follow, and may become very upset if these routines are disrupted.

Causes and Prevalence There are many different factors that can contribute to the development of ASD, including genetic, environmental, and neurological factors. While the exact cause of ASD is not yet fully understood, it is believed that a combination of genetic and environmental factors may play a role. Centers for disease control and prevention reported that 1 out of 54 children in the United States and 1 out of 160 children worldwide are diagnosed with ASD. All socioeconomic and ethnic groups have been affected by ASD. The prevalence of ASD has been increasing worldwide, and the underlying cause is unclear [50].

Diagnosis Diagnosing ASD can be challenging, as there is no specific medical test or imaging study that can definitively diagnose the disorder. Instead, diagnosis is typically based on a combination of behavioral assessments, developmental screening, and observation of the childs behavior and social interactions. Early diagnosis is important, as it can lead to earlier intervention and support for the child

and their family. Early detection of ASD is crucial, and it can significantly improve the quality of life of individuals with ASD, their careers, and families, as mentioned in the clinical study by Elder et al. [29].

Treatment There are a number of different treatments and interventions that can be helpful for children with ASD, depending on their individual needs and symptoms. Some of the most common treatments for ASD include behavioral therapy, speech and language therapy, and medication. These treatments can help to improve communication and social skills, reduce anxiety and repetitive behaviors, and promote independence and self-sufficiency. Despite the challenges of living with ASD, many individuals with the disorder are able to lead happy and fulfilling lives. With appropriate support and intervention, individuals with ASD can learn to communicate effectively, develop meaningful relationships, and pursue their interests and passions.

Summary In recent years, there has been an increased focus on raising awareness about ASD and improving support for individuals and families affected by the disorder. This has led to greater understanding and acceptance of individuals with ASD, as well as improved access to services and support. Overall, ASD is a complex and challenging disorder that can have a significant impact on individuals and their families. However, with early diagnosis, appropriate support, and effective treatments, individuals with ASD can lead happy and fulfilling lives, and contribute to their communities in meaningful ways.

1.2 Diagnosis of ASD

The DSM-IV (Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition) is a manual published by the American Psychiatric Association (APA) that provides criteria for the diagnosis of mental disorders [6]. ADOS and ADI-R are two standardized diagnostic tools widely used in the diagnosis of autism spectrum disorder (ASD). Both of these assessments involve a structured interview and observation of the individual being assessed, as well as information gathered from caregivers or family members.

ADOS Scores The ADOS (Autism Diagnostic Observation Schedule) is a standardized observational assessment that is designed to measure social interaction, communication, play, and imaginative use of materials. The assessment is designed to be used with individuals of all ages who are suspected of having ASD. During the assessment, the clinician will engage the individual in various activities that are designed to elicit social and communication behaviors. The clinician will then score the individual's responses on a scale that ranges from 0 to 3, with higher scores indicating more severe deficits in social and communication behaviors [5].

ADI-R scores The ADI-R (Autism Diagnostic Interview-Revised) is a structured interview that is designed to gather information from caregivers or family members about the individual's behavior, communication, and social interactions. The interview covers a range of topics including early development, language, social interaction, and repetitive behaviors. The clinician will then score the responses on a scale that ranges from 0 to 3, with higher scores indicating more severe deficits in social and communication behaviors [61].

Summary Both the ADOS and ADI-R are considered to be gold standard assessments for the diagnosis of ASD. The scores obtained from these assessments can help clinicians to determine whether an individual meets the diagnostic criteria for ASD. Specifically, a diagnosis of ASD requires persistent deficits in social communication and interaction, and the presence of restricted, repetitive patterns of behavior, interests, or activities.

1.3 Problems faced by Clinicians in diagnosing ASD

Diagnosing ASD can be challenging for clinicians as it is a complex disorder with a wide range of symptoms and severity [41]. Some of the common problems faced by clinicians in diagnosing ASD are:

- Overlap with other disorders: The symptoms of ASD often overlap with those of other disorders, such as attention-deficit/hyperactivity disorder (ADHD), anxiety, and depression. This can make it difficult for clinicians to differentiate between ASD and other conditions.
- Heterogeneity of symptoms: The symptoms of ASD vary widely from person to person and can also change over time. This heterogeneity makes it challenging to identify a set of criteria that can be used consistently to diagnose ASD.
- Limited access to specialized diagnostic tools: Diagnosing ASD often requires specialized diagnostic tools, such as standardized tests and questionnaires. However, not all clinicians have access to these tools, which can make it difficult to accurately diagnose ASD.
- Difficulty in assessing social communication skills: ASD is characterized by deficits in social communication skills, which can be difficult to assess accurately. Clinicians may have to rely on parent and teacher reports, which may not always provide a complete picture of a child's social communication skills.
- Cultural and language barriers: ASD symptoms may be perceived differently in different cultures, and language barriers can make it difficult for clinicians to accurately assess and diagnose ASD in individuals from non-English-speaking backgrounds.

Overall, diagnosing ASD requires a comprehensive evaluation that takes into account a wide range of clinical and behavioral factors. Clinicians need to be aware of the challenges and limitations of the

diagnostic process and work closely with families and other professionals to ensure accurate diagnosis and appropriate interventions.

1.4 Motivation

Diagnosis using the above mentioned procedures suffers from subjectivity, and lack of accuracy and also requires experts, which often results in a significant economic burden on the families of ASD patients [33]. With these various layers of complexity, there has been significant interest in the discovery of biological markers (or biomarkers) for ASD, with the hopes of identifying more homogeneous groups for biological study, assisting in diagnosis (including early detection before the onset or worsening of symptoms), and generating more reliable and sensitive markers for individualizing treatment and gauging treatment response. Therefore, developing a computer-aided tool for accurately diagnosing ASD will make the diagnosis objective, swift, and economical.

Machine learning (ML) offers a potential solution to these challenges by analyzing large amounts of data and identifying patterns that may be difficult for human clinicians to detect. The use of ML on large datasets of clinical and neuroimaging data can identify patterns and relationships that are indicative of ASD and improve the accuracy and consistency of diagnoses. By using standardized and validated algorithms, clinicians can improve the reliability of ASD diagnoses, leading to earlier interventions and better outcomes for individuals with ASD. The goal is to improve the efficiency, objectivity, and accuracy of ASD diagnoses using ML.

1.5 Resting State Magnetic Resonance Imaging (rs-fMRI)

Resting-state functional magnetic resonance imaging (rs-fMRI) is a non-invasive neuroimaging technique that measures brain activity at rest. During an rs-fMRI scan, subjects lie still in the scanner while the machine takes images of their brain every few seconds. Blood Oxygenation Level Dependent (BOLD) signals, which reflect changes in the deoxyhemoglobin and are a proxy measure of brain activation, are captured during the functional magnetic resonance imaging (fMRI) scan. By measuring BOLD signals in different areas of the brain, rs-fMRI can identify brain activity patterns associated with different neurological disorders. The underlying principle behind this method is that neuronal activity is directly proportional to the oxygen consumption of neurons induced by the alterations in blood flow. A series of brain images are acquired during fMRI while the subject performs a set of tasks.

One way that rs-fMRI is used in the diagnosis of neurological disorders is by measuring functional connectivity in the brain. Functional connectivity refers to the degree to which different regions of the brain are synchronized in their activity [24]. In individuals with neurological disorders, functional connectivity patterns may be disrupted or abnormal, which can provide important diagnostic information of various neurodegenerative and psychiatric disorders like ASD, Mild cognitive impairment, Bipolar

disorder, Attention deficit hyperactivity disorders, and Schizophrenia [12]. Thus, rs-fMRI is a powerful tool for studying and diagnosing neurological disorders.

1.6 ABIDE-I Dataset

Introduction The Autism Brain Imaging Data Exchange I (ABIDE-I) [1] dataset is a large, publicly available dataset of neuroimaging data from individuals with Autism Spectrum Disorder (ASD) and typically developing individuals (TD). It was released in 2012 and consists of data collected from 17 international sites, making it one of the largest datasets of its kind. The ABIDE-I dataset includes data from $1,112$ individuals, including 539 individuals with ASD and 573 TD individuals. The individuals range in age from 6 to 64 years old and include both males and females. The dataset includes structural MRI and rs-fMRI data, behavioral and clinical measures, such as Full IQ, Performance IQ, Verbal IQ scores and diagnostic measures such as ADOS and ADI-R scores. These measures provide important information about the severity and nature of ASD in the individuals included in the dataset and allow researchers to investigate the relationship between brain function and behavior in ASD.

Quality Measures The ABIDE-I dataset provides various quality measures to ensure the reliability and validity of the data. Some of these quality measures include:

- Motion measures: The dataset provides motion measures during the scanning process, such as the mean frame-wise displacement (FD) and the percentage of volumes exceeding a certain threshold of FD. These measures help assess the amount of motion artifacts in the data.
- Signal-to-noise ratio (SNR): The dataset provides measures of the SNR of the functional data, which reflects the quality of the data and can affect the accuracy of the analyses.
- Temporal signal-to-noise ratio (tSNR): This measure quantifies the stability of the signal over time and provides information on the quality of the functional data.
- Quality control (QC) ratings: The dataset includes QC ratings of the data by expert reviewers, who assess the quality of the imaging data, the preprocessing, and the registration of the data.
- Structural measures: The dataset provides measures of the quality of the structural imaging data, such as the signal-to-noise ratio (SNR) and the contrast-to-noise ratio (CNR).

These quality measures are important for assessing the reliability and validity of the data and can help ensure that the results obtained from the data are accurate and reproducible. They can also be used to identify data that may need to be excluded from the analyses or to adjust for potential confounding factors.

[1] http://preprocessed-connectomes-project.org/abide/download.html

Table 1.1 Site-wise Phenotypic information of the ABIDE-1 dataset. ASD and TD count represent the number of autism and typically developing samples in each site, respectively. Average Mean Frame-wise Displacement (MFD) represents the average mean frame-wise displacement of all scans in the same site measured in millimeters (mm). Timeseries represent the number of frames in each scan along the temporal dimension. Duration represents the time taken for each scan in seconds(s). Repetition Time (TR) represents the amount of time between successive pulse sequences applied to the same slice in seconds (s)

.

SITE	ASD count	TD count	Average Age (years)	Male count	Female count	Average MFD (mm)	Time-series	Duration (s)	TR (s)
CALTECH	19	18	27.721	29	8	0.0699	146	292.0	2
CMU	14	13	26.59	21	6	0.2944	236/316	632.0	2
KKI	20	28	10.013	36	12	0.1430	152	380	2.5
LEUVEN	29	34	18	55	8	0.090	246	410	1.6
MAX MUN	24	28	25.30	48	4	0.1381	116/196	588	3
NYU	75	100	15.26	139	36	0.069	176	352	2
OHSU	12	14	10.7	26	0	0.095	78	195	2.5
OLIN	19	15	16.5	29	5	0.187	206	309	1.5
PITT	29	27	18.9	48	8	0.152	196	294	1.5
SBL	15	15	34.3	30	0	0.161	196	431	2.2
SDSU	14	22	14.4	29	7	0.0944	176	352	2
STANFORD	19	20	9.97	31	8	0.1066	176	352	2
TRINITY	22	25	16.95	47	0	0.1085	146	292	2
UCLA	54	44	13	86	12	0.184	116	348	3
UM	66	74	14.03	113	27	0.1576	296	592	2
USM	46	25	22.69	71	0	0.1453	236	472	2
YALE	28	28	12.71	40	16	0.1093	196	392	2

Site-wise distribution of ABIDE-I dataset ABIDE-I dataset includes rs-fMRI scans from 17 international sites. Distribution of participants, their average age, repetition time, duration of scan, average mean frame wise displacement and number of volumes per scan is shown in 1.1.

Preprocessing Pipelines There is no agreement on the most effective approach to preprocessing rs-fMRI data. Instead of promoting a single processing method, the data has been preprocessed using four different pipelines with the developer's chosen parameters and settings. The preprocessing steps[2] are quite similar across the different pipelines, but what differs are the specific algorithms used for each step, their software implementations, and the parameters utilized, as shown in 1.2.

Table 1.2 Basic Preprocessing parameters. Table from here [2]

Step	CCS	C-PAC	DPARSF	NIAK
Drop first "N" volumes	4	0	4	0
Slice timing correction	Yes	Yes	Yes	No
Motion realignment	Yes	Yes	Yes	Yes
Intensity normalization	4D Global mean = 1000	4D Global mean = 1000	No	Non-uniformity correction using median volume

Table 1.3 Nuisance Signal Removal parameters. Table from here [2]

Regressor	CCS	C-PAC	DPARSF	NIAK
Motion	24-param	24-param	24-param	scrubbing and 1st principal component of 6 motion parameters and their squares
Tissue signals	mean WM and CSF signals	CompCor (5 PCs)	mean WM and CSF signals	mean WM and CSF signals
Motion realignment	Yes	Yes	Yes	Yes
Low-frequency drifts	linear and quadratic trends	linear and quadratic trends	linear and quadratic trends	discrete cosine basis with a 0.01 Hz high-pass cutoff

[2] http://preprocessed-connectomes-project.org/abide/Pipelines.html

To remove confounding variance from the fMRI data caused by physiological processes (heartbeat and respiration), head motion, and low frequency scanner drifts, each pipeline incorporated some type of nuisance variable regression as shown in table 1.3. Each pipeline was used to calculate four different preprocessing strategies as shown in table 1.4.

Table 1.4 Preprocessing Strtegies. Table from here [2]

Strategy	Band-Pass Filtering	Global Signal Regression
filt_global	Yes	Yes
filt_noglobal	Yes	No
nofilt_global	No	Yes
nofilt_noglobal	No	No

Strengths & Limitations One of the key strengths of the ABIDE-I dataset is its large size and international scope. The dataset includes data from individuals with ASD and TD individuals from various countries, making it a valuable resource for investigating the cross-cultural variability of ASD. Additionally, the large size of the dataset provides ample statistical power for identifying subtle differences in brain function between individuals with ASD and TD individuals. However, the ABIDE-I dataset also has some limitations. For example, the dataset includes data from individuals with a wide range of ages and ASD severity, which can make it challenging to draw clear conclusions about the neural mechanisms underlying ASD. Additionally, the dataset is limited to individuals who could undergo MRI scanning, which may not represent the broader population of individuals with ASD.

Summary The ABIDE-I dataset is a valuable resource for investigating the neural mechanisms underlying ASD and for developing and validating neuroimaging-based biomarkers for ASD diagnosis. Its large size and international scope makes it a unique resource for investigating the cross-cultural variability of ASD and for identifying subtle differences in brain function between ASD subjects and TD individuals.

1.7 Static and Dynamic Functional Connectivity

Static Functional Connectivity (sFNC) is estimated from the temporal correlation of spontaneous BOLD signals among two or more anatomically distinct brain regions. If the signals from two or more brain regions show synchronized fluctuations, then the regions are said to be functionally correlated (connected). Previous studies modeled sFNC for classifying ASD from typically developing (TD) and achieved promising results. However, these approaches based on sFNC lack model interpretability, i.e., identifying the brain regions that contribute most to the ASD classification task. Therefore, we proposed

a classification model that achieves new state-of-the-art performance on the ABIDE-1 dataset in 10x less training time than the previous best method [77] followed by model interpretation.

Dynamic functional connectivity (dFNC) is a neuroimaging technique that examines how connectivity between different brain regions changes over time, as opposed to sFNC, which assumes that connectivity is stable over a fixed period. dFNC enables researchers to explore how the functional connections between brain regions change in response to different stimuli or cognitive states. Therefore, we explored clustering, statistical and meta-state analysis based on dynamic functional connectivity to understand the fundamental characteristics of ASD. Contrasting with most extant studies, we compared dynamic brain states to sFNC. Specifically, the dynamic characteristics of the brain states were estimated while applying clustering analysis on the time series segments extracted from 7 brain networks consisting of 53 regions of interest on a large subset of 188 ASD and 195 TD subjects from the ABIDE-I consortium.

1.8 Contributions

In this thesis, we utilized the ABIDE-I dataset and examined multiple techniques to diagnose ASD and gain insights into the core features of ASD. This thesis makes the following significant contributions:

Classification and Interpretation

The main contributions of this study are as follows:

1. Proposed a two-hidden layer feed-forward neural network with autoencoder pretraining.
2. Applied Integrated gradients and DeepLIFT, prominent feature attribution methods, to interpret brain biomarkers in ASD subjects.
3. Revealed the impact of heterogeneity in the dataset and compared the effect of different preprocessing pipelines, brain parcellation schemes on the classification performance.

Dynamic Functional Connectivity analysis

This study aimed to

1. Compare connectograms of ASD and TD subjects estimated with dFNC and sFNC (i.e., can dFNC delineate any other meaningful information in terms of brain network differences and information that can be evident from sFNC?)
2. Understand the dFNC based network group differences in each state and reveal the dFNC patterns that showed increased or decreased connectivity in ASD subjects.

3. Explore the temporal properties such as fractional windows in each state, the mean dwell time of each state, and the number of transitions between each pair of dynamic states using statistical two-sample *T-test*.

1.9 Thesis overview

In this thesis, we investigated both sFNC and dFNC differences between the groups ASD subjects and TD individuals and identified the bio-markers for ASD. Chapter 2 provides an overview of the extant literature on ASD classification and interpretation based on rs-fMRI scans and dFNC analysis. Chapter 3 discusses the reproducibility challenges encountered in replicating the results of papers published in this domain, along with potential causes, instances of papers we tried to implement, and potential best practices to overcome the challenges. Chapter 4 presents the classification and interpretation model proposed based on sFNC in detail. Chapter 5 discusses the clustering, statistical and meta-state analysis based on dFNC and the bio-markers identified. Finally, chapter 6 discusses conclusions, limitations, and future scope.

Chapter 2

Background and Literature

In the previous chapter, we discussed ASD, the challenges faced by clinicians in diagnosing it, and how computer-aided tools can help with the diagnosis and identification of ASD biomarkers. The current chapter focuses on the existing literature on this topic. Previously many studies have focused on using ML and deep learning (DL) based methods to classify children with ASD from the TD. Different rs-fMRI representations (as shown in 2.1) can be passed as an input to the ML/DL models for the ASD classification from TD individuals.

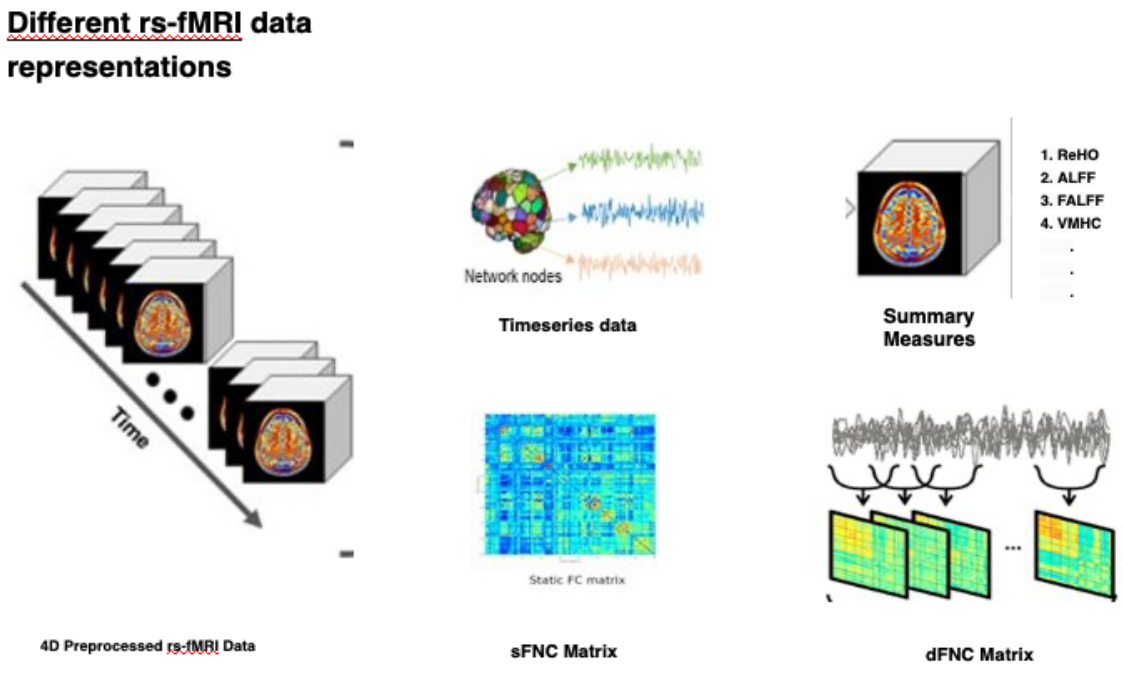

Figure 2.1 Different rs-fMRI representations. This figure has been generated from the papers [73, 59]
.

2.1 4D scan

The pre-processed rs-fMRI scan is a four dimensional data (3-spatial and 1 temporal). Each scan would consist of approximately 1 million voxels and take up to $200 - 300$ MB space. There exist spatial measures (as discussed in section 2.2) which summarize the temporal information and various brain parcellation schemes (as discussed in section 2.3) which summarize the spatial information to reduce the dimensions of 4D data and remove noise. Such methods would lead to the loss of either spatial or temporal information. Hence, modelling the 4D data using the deep learning models is one solution proposed for classifying ASD from TD. Typically, we can model the flattened 4D data using traditional machine learning models or use the 3D convolutional network (CNN) to extract pertinent features from the 4D data and then pass it to fully connected layers to classify ASD from TD.

Machine learning-based methods for the classification of ASD

El-Gazzar et al. [28] proposed an end-to-end pipeline that uses 3D CNNs and 3D Convolutional LSTMs to extract spatiotemporal features from the complete 4D data. To extract lower-level coherent spatial feature maps required for the subsequent stage of learning spatio-temporal features, they used a 3D CNN with coupled weights at all input time steps. Each CNN layer has a kernel size of $3 \times 3 \times 3$ and a stride of $2 \times 2 \times 2$ to down-sample the input feature vector. To reduce the overfitting, they added dropout with a rate of 0.2 to each convolution's output. The spatial feature maps extracted by 3D CNNs were flattened and passed to the fully connected layers for classification. However, flattening the spatial feature maps would lead to losing spatial patterns. Using LSTM, without flattening the spatial dimensions, allows modeling the temporal information in a memory-efficient manner. The 3D CNN model yielded 54% classification accuracy, F1 score of 50%, whereas 3D CNN with LSTM model yielded 58% classification accuracy, F1 score of 53% on the whole ABIDE-I consortium dataset. They also tried modeling the flattened 4D data using a support vector machine (SVM) model, yielding 58% classification accuracy.

2.2 Summary measures

The summary measures are used in functional connectivity analysis of neuroimaging data. Numerous instances of lowering the dimensionality of brain volumes while preserving the data's three-dimensional spatial dimensions may be found in the literature on neuroimaging. These various techniques extract various elements of the time series. The following are some examples of commonly used measures:

1. Regional homogeneity (ReHo): ReHo [85] is a measure of functional homogeneity or similarity of the time series within a region of interest (ROI) in fMRI data. It quantifies the consistency of the temporal behavior of the voxels within an ROI by computing the local coherence of the time series of each voxel with its neighboring voxels. The measure is calculated by comparing the

time series of a central voxel with its surrounding neighboring voxels within a certain radius. The homogeneity score is higher if the time series of the central voxel is more similar to its neighbors, indicating a higher degree of functional homogeneity within the ROI. ReHo is one of the summary measures commonly used in functional connectivity studies.

2. Amplitude of low-frequency fluctuations (ALFF): ALFF [87] is a neuroimaging metric used to measure the magnitude of low-frequency spontaneous fluctuations in BOLD signal in fMRI data. The ALFF metric is computed by first bandpass filtering the fMRI data to retain low-frequency oscillations (typically $0.01 - 0.08$ Hz) and then calculating the root mean square of the amplitude within this frequency band for each voxel in the brain.

3. Fractional amplitude of low-frequency fluctuations (fALFF): fALFF [88] is a neuroimaging measure that characterizes the relative power of low-frequency ($0.01 - 0.1$ Hz) fluctuations within the brain. It is calculated as the ratio of the amplitude of low-frequency oscillations (ALFF) to the total power within the frequency range of interest. Unlike ALFF, fALFF is a relative measure that provides information about the fraction of total power within the low-frequency range.

4. Degree centrality (DC): DC [88] is a graph theoretical measure that quantifies the importance of a node (ROI) in a network based on the number of connections (or edges) it has with other nodes in the network. In other words, DC measures how well connected a node is to other nodes in the network. A node with a high DC indicates that it is highly connected and plays an important role in information transfer within the network.

5. Eigenvector centrality (EC): EC [89] is a measure of node importance in a network based on its connections with other important nodes. It calculates the centrality of a node by considering not only the number of connections it has but also the centrality of its neighboring nodes. It is often used to identify nodes critical for the spread of information or influence in a network, as these nodes can have a strong impact on the rest of the network due to their high centrality score.

6. Local functional connectivity density (LFCD): LFCD [78] used to assess the local connectivity properties of a brain region or a voxel in functional brain networks. It is defined as the number of functional connections a given voxel has with its neighboring voxels, as measured by rs-fMRI. The neighboring voxels are defined based on a pre-specified distance threshold, typically in the range of 10-20 mm. The LFCD metric is calculated by first constructing a functional brain network, where each voxel is represented as a node, and the edges between nodes represent the functional connectivity strength between the corresponding voxels. Then, the LFCD value of a given voxel is calculated as the number of functional connections it has with its neighbors. A connection is defined as a significant correlation between the BOLD time series of two voxels. The LFCD metric thus provides a measure of the local connectivity strength of a voxel, taking into account both the number of connections and the strength of those connections.

7. Entropy: Entropy [85] is a measure of the randomness or disorderliness of a system. In fMRI, it refers to the variability of the BOLD signal over time within a given brain region or voxel. Entropy can be used to assess the complexity and variability of brain activity, which may reflect the brain's functional organization or changes in brain states or conditions. High entropy indicates high variability and complexity, while low entropy indicates low variability and regularity in the BOLD signal.

8. Voxel-mirrored homotopic connectivity (VMHC): VMHC [88, 89] is a measure of functional connectivity that quantifies the temporal synchronization of neural activity between homotopic brain regions in each hemisphere. Homotopic brain regions are anatomically and functionally equivalent and located in opposite brain hemispheres. VMHC is calculated by computing the functional connectivity between each voxel in one hemisphere and its mirrored counterpart in the other hemisphere. VMHC can be used to investigate the alterations in interhemispheric connectivity in various neurological and psychiatric disorders and to identify potential biomarkers for diagnosis and treatment.

9. Auto-correlation lag (ACL): ACL [43] refers to the time delay between a given fMRI signal and its copy that has been shifted in time.

These measures can be flattened, feature engineered, and then fed to a neural network or passed directly to a CNN for the classification of ASD.

Machine learning-based methods for the classification of ASD

Thomas et al. [73] investigated different transformations that preserve the full spatial resolution, making it possible to train a full three-dimensional convolutional neural network (3D-CNN). They determined ReHo, ALFF, fALFF, DC, EC, LFCD, entropy, VMHC, and ACL for each brain volume. The 3D CNN models has been trained on all these summary measures independently as well as all nine measures combined for the ASD classification. All the CNN models were trained for 50 epochs with a batch size of 32. They used binary cross entropy as a loss function and stochastic gradient descent (SGD) as an optimizer with a learning rate of 0.01. This model yielded 66% accuracy on both ABIDE-I and ABIDE-II datasets. The performance of CNN and SVM models trained separately on these and all measures combined can be seen in table 2.1.

Even after combining all nine summary metrics, the outcome was still inferior to that of the best single indicator, ReHo. The model's performance was very bad when all the measurements were concatenated and provided as input. The integration of uninformative summary measures with informative ones may have introduced additional noise to the input data, potentially posing challenges for the model in accurately classifying the data. The ensemble model yielded much better results when compared to the model trained on all measures combined input. Furthermore, they also used the support vector machine (SVM) technique on the same dataset and obtained equivalent findings, indicating that 3D-CNNs

Table 2.1 Performance of CNN and SVM models on summary measures. Table from here [73]

	CNN		SVM	
Input	Accuracy	F1-score	Accuracy	F1-score
ReHo	0.64	0.65	0.66	0.66
ALFF	0.59	0.60	0.57	0.58
fALFF	0.62	0.63	0.57	0.58
DC	0.61	0.62	0.63	0.64
EC	0.60	0.61	0.61	0.63
VMHC	0.61	0.62	0.62	0.62
LFCD	0.59	0.60	0.63	0.65
Entropy	0.54	0.49	0.56	0.57
ACL	0.59	0.61	0.57	0.59
All measures ensemble	0.64	0.66	0.66	0.67
All measures combined	0.59	0.60	0.61	0.62

could not learn any new information from these temporal transformations that would be more useful to distinguish ASD from TD.

2.3 Timeseries data

Brain parcellation schemes summarize the spatial information and extract the time-series data. Parcellation refers to the process of dividing the brain into discrete regions, or parcels, to analyze the functional connectivity between these regions [26]. rs-fMRI data can be parcellated using a variety of strategies.

1. Anatomical parcellation: This approach involves dividing the brain into regions based on its anatomy. Anatomical parcellation is commonly used as a reference space for defining functional parcellation.

2. Functional parcellation: This approach involves using the rs-fMRI signal itself to define functional parcels of the brain. This can be done using clustering techniques, such as k-means or hierarchical clustering, or by using independent component analysis (ICA) to separate the brain into different networks.

3. Hybrid parcellation: This approach combines both anatomical and functional information to define the parcels. For example, one could use anatomical information to guide the selection of voxels for functional parcellation or vice versa.

4. Data-driven parcellation: This approach involves using machine learning techniques to identify patterns in the data that can be used to define the parcels. These techniques include SVM, deep learning, and random forests (RF).

Each parcellation strategy has its advantages and disadvantages, and the choice of strategy often depends on the research question being addressed and the specific rs-fMRI dataset being analyzed. The extracted time-series data would be a $2D$ matrix representing the time-series signal for every ROI. Different brain atlas/parcellations would yield different numbers of ROIs and details can be found in table 2.2. Timeseries matrix is directly passed to $1D$ CNNs or the upper triangular part (as it is symmetric) is extracted and passed to feature engineering pipelines followed by ML/DL models.

Table 2.2 Different brain atlas with the corresponding number of ROIs

Parcellation Type	Atlas	Number of ROIs
Anatomical	Harvard-Oxford (HO)	111
	Dosenbach	160
	Automated Anatomical Labelling (AAL)	116
Functional	Cameroon Craddock 200 (CC200)	200
	Cameroon Craddock 400 (CC400)	392
	Brainnettome	246
	ICA NeuroMark (ICANM)	53

Machine learning-based methods for the classification of ASD

Supekar et al. [72] proposed a spatiotemporal DNN (stDNN) model trained on the timeseries data extracted using the Brainnettome atlas. The stDNN model consists of two blocks of 1D convolutional layers and max-pooling layers, a temporal average layer, a dropout layer, and a fully connected layer with a sigmoid activation function. The number of inputs to the fully connected layer equals the output channels of the second convolution block layer because a temporal averaging operation averages the temporal information for each filter. They also took sex and site into account by creating a one-hot feature encoding scheme combined with age. They trained the model using binary cross entropy cost function and Adam optimizer with a learning rate of 0.00005 for 300 epochs. They trained the model on ABIDE-I and II datasets in a 5-fold cross-validation strategy. They achieved an F1 score of 0.79, average precision of 0.76, recall of 0.82, and classification accuracy of 78.2% across the 5 folds. They evaluated this trained model on the GENDAAR and independent Stanford cohorts, and they achieved promising results suggesting that the model can generalize well. They used Integrated gradients (a state-of-the-art attribution algorithm) to interpret the features contributing to the classification results. The posterior cingulate cortex (PCC) and precuneus, dorsolateral and ventrolateral prefrontal cortex, and superior temporal sulcus, which anchor the default mode network (DMN), cognitive control, and

human voice processing systems, respectively, were found to be the brain features that most clearly distinguished ASD from TD in the three cohorts.

2.4 Static Functional Connectivity

Static Functional Connectivity (sFNC)can be considered a measure of the "strength" of the functional connections between different brain regions. sFNC has been used in various neuroimaging studies to investigate brain function and dysfunction. For example, studies using sFNC have identified disruptions in connectivity patterns associated with different tasks or cognitive states, such as memory encoding, attention, and emotion regulation in individuals with neurological and psychiatric disorders. Functional Connectivity in the resting human brain was first identified by Biswal et al. [10]. Methods based on sFNC involve passing the upper triangular part of the sFNC matrix as input to the linear ML models or non-linear deep neural networks (DNN) for the classification of ASD. Feature engineering on the sFNC vector is also performed before training the ML/DL models.

Machine learning-based methods for the classification of ASD

Sherkatghanad et al. [65] proposed a novel 1D CNN architecture to classify ASD from TD based on sFNC matrix. Data is preprocessed using CPAC preprocessing pipeline and parcellated using the CC400 atlas. Each subject sFNC matrix would be of size 392×392 and fed directly to the CNN model. The CNN model was evaluated on the ABIDE-I dataset using 10-fold cross validation (CV) strategy and leave one site out cross validation strategy (LOOCV) in which the model is tested on data from one site and trained on subjects from other sites. The model achieved a classification accuracy of 70.22%, sensitivity of 77.46%, and specificity of 61.82% using a 10-fold CV strategy. The model yielded the highest classification accuracy of 77% when tested on USM site and trained on other sites. They also trained SVM, K-Nearrest Neighbors (KNN), and RF algorithms, achieving accuracies of 69%, 62%, and 59%, respectively. They used the saliency technique to visualize the brain regions that contributed the most to the classification task. They identified that the right supramarginal gyrus, fusiform gyrus, cerebellar vermis and anterior-posterior are associated with ASD.

Heinsfeld et al. [38] used two stacked denoising autoencoders (SDA) at the pretraining stage followed by multilayer perceptron (MLP) to classify ASD from TD. The upper triangular part of the sFNC matrix was extracted from the data preprocessed using CPAC pipeline and parcellated using CC400 parcellation scheme. The input feature vector would be of size $200C_2 = 19900$ and used stacked autoencoders to reduce dimensions. They reported an overall 10-fold cross-validation accuracy of 70%, a sensitivity of 74%, and a specificity of 63% on the ABIDE-1 dataset. They also trained SVM and RF on the same data and achieved a 10-fold CV accuracy of 65% and 63%, respectively. This suggests that DNN models yield better performance when compared to traditional ML models.

Eslami et al. [30] proposed a joint learning procedure using an autoencoder and a single layer perceptron for the ASD classification task, called *ASD-Diagnet*. Preprocessed the data using CPAC preprocessing pipeline and extracted the timeseries data using CC200 parcellation scheme. sFNC matrix was estimated using Pearson correlation coefficients and considered only the upper triangular part of the matrix. They calculated the average correlation across all subjects in the training set. They considered only the first and last $\frac{1}{4}$ of them for training the model (i.e.) choosing the indices with the highest negative and positive values from the averaged correlation array. To augment the data, they employed the Synthetic Minority Over-sampling Technique (SMOTE) [16]. The model yielded classification accuracy of 70.3%, sensitivity of 68.3%, and specificity of 72.2% on the ABIDE-I dataset. With feature selection and data augmentation techniques, the model's performance was boosted by 3%.

Wang et al. [77] proposed a multi-atlas deep feature representation and ensemble learning method based on SDA and MLP for the ASD identification task. They trained three SDA to reduce the dimensions of input followed by MLP classifier. This model was trained independently on sFNC matrices extracted from the CC200, Dosenbach, and AAL atlases, and the final prediction was based on a majority vote. Models trained independently on the sFNC matrices extracted from CC200, Dosenbach and AAL parcellation schemes yielded 73.39%, 69.74%, and 71.42%, respectively. The ensemble model achieved a classification accuracy of 74.52%, a sensitivity of 80.69%, and a specificity of 66.71% on the ABIDE-1 dataset. Though the proposed model yielded cutting-edge accuracy, training the three SDA prior to MLP classifier training requires a significant amount of computation and time.

2.5 Dynamic functional connectivity

Dynamic functional connectivity (dFNC) is a type of functional connectivity analysis in neuroimaging that examines how the strength and pattern of connectivity between different brain regions change over time. Unlike sFNC, which assumes that the strength of connectivity between regions is stable over a fixed period of time, dFNC allows researchers to investigate how the functional connections between different regions of the brain change in response to different stimuli or cognitive states [62]. dFNC analyses can be conducted using a variety of different approaches, such as sliding-window correlation, time-varying graphical modeling, and independent component analysis (ICA). These approaches allow researchers to identify patterns of connectivity that are specific to certain time periods or tasks, and to examine how these patterns are related to cognitive or behavioral outcomes [8, 64]. Presently dFNC analysis has shown to be a successful procedure for examining the complex and dynamic nature of functional connections in the brain, and for understanding how these connections are altered in various neurological disorders, including ASD. A sliding window-based dFNC matrix is determined and different measures that capture the brain dynamics are extracted. Previous studies conducted hard clustering and statistical analysis to identify the fundamental characteristics of ASD. Sliding windows-based dFNC matrix can be passed as input to ML/DL models for diagnosing ASD.

Machine learning-based solutions based on dFNC for diagnosing ASD

Zou et al. [39], proposed a dynamic weighted functional connectivity matrix to capture the brain dynamics and passed as an input to the SVM classifier. Preprocessed the ABIDE-I dataset using the DPARSF pipeline and parcellated using AAL atlas. In this dynamic weighted sliding-based approach, different time instances within a window are assigned different weights followed by computing Pearson correlation coefficients. The authors state that information from the recent time instances would contain more information than the later ones, thus they multiplied weights to time instances in a linearly decreasing fashion. They experimented with various window sizes in range of 15 to 100 and found the optimal window length to be 30 and 35. Computed the variance of all the dynamic windows to capture the temporal variability followed by feature selection using the LASSO operator [74]. The final selected features are passed to the SVM classifier for training. The classifier yielded accuracy of 85.25% on the NYU and 80.61% on UCLA datasets sampled from ABIDE-I consortium.

These traditional sliding window based dFNC suffer from temporal mismatch issues i.e., subnetworks do not exhibit temporal connection between various subjects within the same temporal window. To study the higher level and more intricate interaction linkages among many ROIs, Zhao et al. [86] built a novel high-order dFNC based on the "correlation's correlation" principle. High-order dFNC is estimated by computing PCC for each window of the dFNC matrix (computing the correlation of the correlation). To reduce the heterogeneity in the dataset, they considered only 47 TD and 45 ASD NYU subjects from the ABIDE-I dataset. They first computed sFNC, dFNC, and high-order dFNC, followed by applying the central moment technique to extract the temporal invariance characteristics present in either dFNC or high-order dFNC. Feature selection techniques such as two-sample t-test and LASSO regression have been applied to reduce the dimensions and select the relevant features. These dynamic features extracted from dFNC, high-order dFNC and static features from sFNC are passed independently to SVM classifiers for training. SVM classifiers trained independently on sFNC, dFNC and high-order dFNC yielded classification accuracy of 74%, 79% and 78%, respectively.

Clustering and Statistical analysis based on dFNC

Fu et al. [35] investigated the dFNC between 51 intrinsic connectivity networks in 170 ASD subjects and 195 age-matched typically developing (TD) individuals. Hard clustering analysis on the dFNC windows has yielded five dynamic brain states. They found hyper-connectivity between sub-cortical network (hypothalamus/subthalamus) and some sensory-motor networks (lingual gyrus, bi-paracentral lobule, and right postcentral gyrus) in certain functional states.

Ma et al. [51] examined the dFNC between 29 regions among 88 ASD and 87 TD subjects. Significantly stronger dFNC between the left fusiform / lingual gyrus and right dorsolateral prefrontal cortex (DLPFC), between the the superior temporal gyrus, and DLPFC, between the left middle frontal gyrus and right frontal eye field (FEF), between the right angular gyrus and the FEF, and between the right

middle temporal gyrus and left intra parietal sulcus has been observed among ASD subjects when compared to TD.

Harlalka et al. [37] explored the dynamic variability in the connection strength to characterize the differences between 100 ASD and 75 TD subjects. They observed significantly higher dFNC between the attention network and default mode network (DM), in patients with ASDs when compared to TD individuals. Li et al. [48] conducted dFNC and clustering analyses using the sliding-window and K-means clustering methods on 62 ASD and 63 TD subjects. They identified that individuals with ASD have increased dFNC variability between the middle temporal pole and the posterior cingulate gyrus. The study [34] incorporated the use of cosine similarity to evaluate alterations in dFNC and introduced stepwise functional network reconfiguration (sFNR) as a unified dynamic measure to effectively capture the temporal characteristics of diverse functional brain organizations in ASD individuals. The analysis was conducted on an independent dataset comprising 314 TD and 255 ASD subjects. The findings of the study revealed notable increases in whole-brain sFNR among individuals with ASD. Additionally, they found the sFNR between sensorimotor and cerebellar domains correlated with Autism Diagnostic Observation Schedule (ADOS) among ASD subjects.

Using the sliding window, hard clustering-based approaches, many studies investigated altered dFNC patterns in other disorders such as Parkinson's disease (PD), Subjective Cognitive Decline (SCD) and healthy controls (HCs). Chen et al. [17] investigated the dFNC and static parameters obtained from graph theory among 33 HCs and 32 SCD subjects. They found 4 optimal dynamic brain states through hard clustering analysis on dFNC windows and identified hyperconnectivity within and between auditory domain, visual domain, and, somatomotor networks among SCD subjects.

Fiorenzato et al. [31] analyzed the dFNC between 118 patients suffering from PD and 35 HCs. The sliding window approach and hard clustering analysis have found two dynamic brain states. They hypothesized that the underlying cognitive deficits in PD subjects are characterized by strong connectivity in default mode and cognitive executive networks.

2.6 Research Gaps and Opportunities

Direct classification using complex deep learning models on the raw 4D rs-fMRI data is resource-intensive and yields poor results. To address this, downsampling techniques such as summary measures and brain atlases have been employed. Summary measures involve downsampling the data temporally, but they assume stationary brain activity and may be influenced by various factors, leading to unreliable results. Brain parcellation schemes reduce the spatial dimensionality by assigning regions of interest (ROIs), but choosing parcellation schemes and resolution impacts classification metrics. sFNC has been used as an input feature for ASD classification, but it fails to capture temporal changes in connectivity and the complexity of brain networks. dFNC analysis examines how connectivity changes over time and shows promise in discriminating ASD from TD individuals. However, dFNC analysis requires advanced algorithms, may lack consistency across individuals or scans, and interpreting results can be complex.

Data sharing, standardized preprocessing, feature extraction, and model selection are essential to improve rs-fMRI-based classification. Cross-validation and replication studies can test robustness while addressing limitations such as parcellation scheme and data quality is crucial. Further research is needed to enhance classification accuracy and understand the neurobiological basis of rs-fMRI measures.

2.7 Summary

While different approaches have shown promise for classifying ASD from TD using rs-fMRI data, further research is needed to determine their effectiveness and generalizability. According to the literature, modeling the temporal aspects for diagnosing ASD produced better outcomes than other approaches. Future research could focus on capturing this temporal information to increase the classifier's performance and discover the inherent traits of ASD. Numerous research papers are being published on this topic every year, but many lack reliability and replicability. To address this issue, the upcoming chapter outlines the challenges encountered when reproducing research papers on this topic and suggests good practices to enhance replicability and reliability.

Chapter 3

ASD Classification Literature: Reproducibility Challenges

This chapter focuses on the difficulties encountered when attempting to replicate the results obtained in various research papers. We will discuss the discrepancies between the reported and replicated results, as well as potential reasons for these differences and good practices to improve reproducibility.

3.1 Introduction

The reproducibility crisis in neuroimaging refers to the difficulty researchers face in reproducing the results of studies in the field of neuroimaging. There are several reasons why reproducibility is a challenge in neuroimaging research. One reason is that the field is highly interdisciplinary and involves complex data analysis techniques that require specialized expertise. Additionally, neuroimaging studies often involve small sample sizes either cherry picked or randomly sampled from a dataset, which can make it difficult to draw general conclusions from the results [44, 57]. Another factor that contributes to the reproducibility crisis in neuroimaging is the use of flexible analysis pipelines that allow researchers to selectively report only the results that support their hypotheses. This can lead to a publication bias, where studies that fail to find significant results are less likely to be published, skewing the scientific literature towards positive results [13]. To address these issues, researchers have called for greater transparency and openness in neuroimaging research. This includes sharing data, code, and analytical methods to ensure that other researchers can reproduce the results of a study. Additionally, the use of standardized analysis pipelines and larger sample sizes can help improve the reliability of neuroimaging research [36].

3.1.1 Does reliability of a study improve with sample size?

Yes, the sample size in neuroimaging studies can have a significant impact on the reliability and generalizability of the results. In general, larger sample sizes can increase the statistical power of a study, which means that researchers are more likely to detect true effects and less likely to mistakenly interpret random fluctuations as significant findings [11]. For example, if a neuroimaging study only

includes a small number of participants, it may not be possible to reliably detect subtle differences in brain activity between groups or conditions, or to generalize the findings to other populations. This is because the variability in the data may be too high relative to the size of the effect being studied, which can lead to low statistical power and a higher likelihood of false positives. On the other hand, studies with larger sample sizes can provide a more accurate representation of the population being studied, enhancing the generalizability of the findings. Larger samples can also help to reduce the impact of outliers or other sources of noise in the data, which can improve the reliability of the results [57]. However, it's worth noting that increasing the sample size alone is not a solution to all the challenges in neuroimaging research. Other factors such as data quality, experimental design, and analytical methods also play important roles in determining reliability and validity of the results.

3.1.2 Machine Learning Solutions in neuroimaging research

When a research paper proposes a novel neural network architecture, there are several parameters that should be reported for reproducibility. These parameters include:

1. Network architecture: This includes the number and type of layers in the network, the number of neurons in each layer, and the activation functions used in each layer.

2. Initialization: This refers to how the network weights and biases were initialized, as this can affect the network's performance.

3. Optimization algorithm: This refers to the algorithm used to update the weights and biases of the network during training, such as stochastic gradient descent (SGD) or Adam.

4. Learning rate: This refers to the step size used to update the weights and biases during training, and can significantly impact the network's convergence.

5. Regularization: This refers to techniques used to prevent overfitting, such as L1 or L2 regularization, dropout, or early stopping.

6. Batch size: This refers to the number of training examples used in each iteration of the optimization algorithm, and can affect the speed of training and the convergence of the network.

7. Number of epochs: This refers to the number of times the entire training dataset is passed through the network during training.

By reporting these parameters, researchers can enable others to reproduce and build upon their work, as well as facilitate comparisons between different approaches and identify areas for improvement. Most papers published in the domain of ASD classification involving neural networks failed to report the batch size, regularization techniques, and weight initialization strategies employed, which could impact the outcomes significantly.

3.2 Experiments on ASD literature

All the papers presented below have not made their code publicly available. The reported results in the paper and reproduced results when we implemented the paper are significantly different.

3.2.1 Classification of ASD using a Convolutional Neural Network [65]

Introduction & Methods Sherkatghanad et al [65] proposed a 1D CNN architecture to classify ASD subjects based on rs-fMRI data. They passed the 2D symmetric sFNC matrix as input to the CNN architecture for the classification of ASD from TD. The sFNC matrix was estimated on timeseries data extracted from CC400 parcellation and CPAC preprocessing pipeline. The input was passed to seven 1D convolutional layers with filters of different dimensions from 1×392 to 7×392 simultaneously followed by corresponding seven max-pooling layers (as shown in figure 3.1). The Max-Pooling layer's outputs were concatenated and passed to the fully connected dense layer for classification. Dropout regularization after max pooling layers was used to reduce overfitting.

Figure 3.1 Sherkatghanad et al [65] proposed 1D CNN architecture

Implementation & Results The model was trained in a batches of batch size 32 for 300 epochs, with a learning rate of 0.005. The proposed model have achieved a 10-fold cross validation accuracy of 70.22%, sensitivity of 77% and specificity of 61%.

Reproduced results & Observations The weight initialization strategy and optimizer used are not specified in this paper. We attempted to replicate the results using various combinations of weight initialization schemes and optimizers, but best accuracy we could achieve was 68% with Adam optimizer and Xavier normal weight initialization strategy. However, during the training, we observed significant fluctuation in the validation loss. A smooth decreasing trend is not observed in the validation loss even after experimenting with other parameters such as learning rate, using batch normalization etc. Therefore, our reproduced result doesn't seem to be a reliable one.

3.2.2 Diagnosing ASD using Neural Networks [84]

Introduction & Methods Yin et al. [84], proposed a Auto-Encoder (AE) based Deep Neural Network (DNN) for diagnosing ASD based on brain networks created using rs-fMRI data. The proposed method is evaluated on 871 subjects downloaded from the ABIDE-1 dataset. CPAC pre-processing pipeline was used to preprocess the data and power atlas to parcellate the brain into 264 ROIs. sFNC matrix of shape 264×264 is computed using Pearson correlation coefficient and thresholded to a value T.

$$\begin{aligned} a_{ij} &= 1, \quad \text{if} \quad cm_{ij} \geq T \\ &= -1, \quad \text{if} \quad cm_{ij} \leq -T \\ &= 0, \quad \text{if} \quad i = j \\ &= 0, \quad \text{otherwise} \end{aligned} \tag{3.1}$$

where $cm_{i.j}$ is the sFNC matrix before thresholding and a_{ij} is after thresholding.

Eigenvalues of the Laplacian matrix, which is the difference between the adjacency matrix and degree matrix are computed and normalized. These 264 eigenvalues were concatenated with graph theoretical measures [37] such as assortativity, clustering coefficient, and the average degree to create an input vector of length 267 for every subject. This input vector is passed to the AE to extract the advanced features. Then, decoder part is being removed and two more hidden layers were added to the encoder for classification (as shown in figure 3.3)). The classifier was trained on these advanced features.

Implementation & Results To prevent data split bias from affecting the results, the 10-fold cross validation was carried out using random shuffling. The proposed method DNN model without a pre-trained AE achieved a receiving operating characteristic curve (AUC) of 79.7% and a classification accuracy of 76.2% using 10-fold cross-validation strategy for threshold $T = 0.2$. The classification accuracy boosted to 79.2% and the AUC to 82.4% when the DNN and the pre-trained AE are combined and trained on the raw features. They also compared the classification results of DNN architecture with those of traditional machine learning algorithms such as SVM and KNN using the same advanced features.

Reproduced results & Observations We implemented the above AE architecture and classifier using Pytorch. Although the method is fairly easy to implement, the paper skips crucial information like the

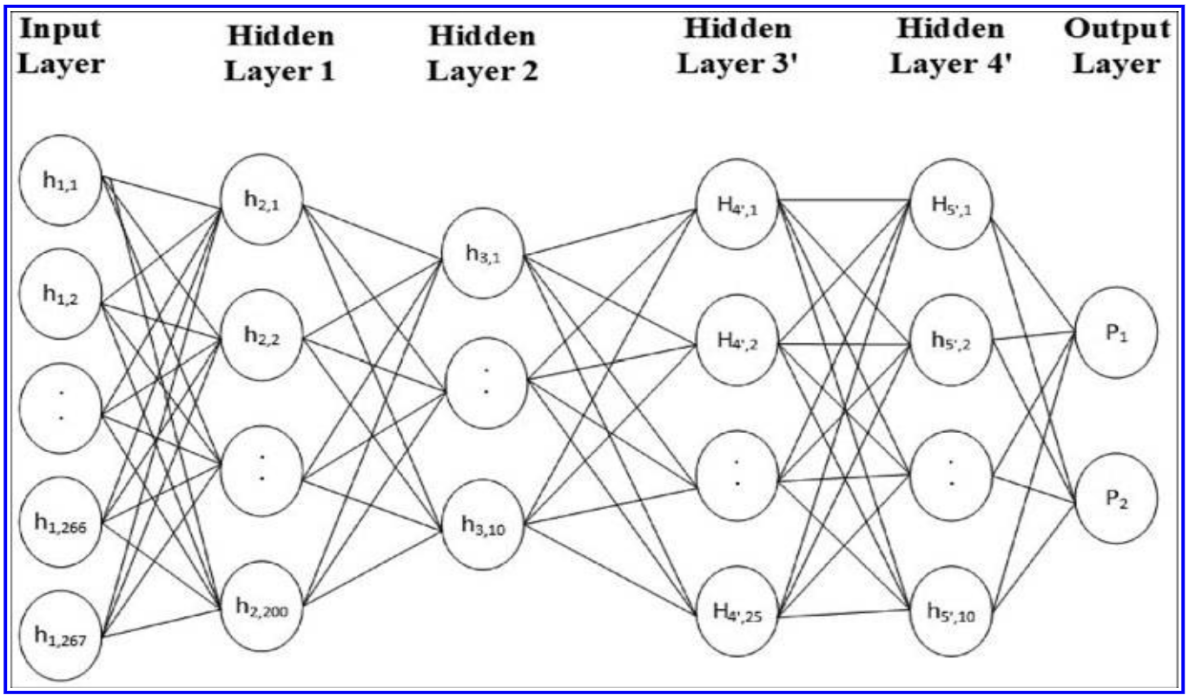

Figure 3.2 Yin et al. [84] autoencoder-based classifier for classification of ASD

weight initialization strategy, activation functions, optimizer, batch size, number of epochs, and learning rate in the case of autoencoder and classifier. We attempted a variety of combinations of these hyper parameters, but the best we could achieve was 56.48% classification accuracy, a sensitivity of 27.34% and a specificity of 81.63% with the default neural network parameters. The classification accuracy reproduced was almost close to the random chance (50%). Neural network hyperparameters play a vital role in training the model and have to report in the paper.

3.2.3 Classifying ASD using Machine Learning based on temporal properties [21]

Introduction & Methods Dammu et al. [21], proposed approach captures brain dynamics by extracting graph-theoretical methods such as occupancy rates based on dynamic functional connectivity. Data is pre-processed using CPAC pipeline and parcellated using AAL atlas. Participants with repetition time (TR) of 2 seconds were included in this study and dFNC matrix was estimated based on sliding window approach with window size 22 and step size 1. Clustered all the windows extracted from all the subjects using K-Means algorithm to identify the dFNC patterns or brain states. The number of brain states could be different in different age groups. To create temporal attributes and enable the classifier to learn the best features, they implemented group clustering for K values ranging from 6 to 31 on all participant data. An input vector of size 481 has been created for each subject i.e. the subject-wise mean lifetimes for each state are computed for all K values and stacked horizontally. This input vector describing

the mean-lifetimes of several brain states is passed as input to DNN and ML algorithms such as SVM, Random Forest (RF) and XG Boost algorithm.

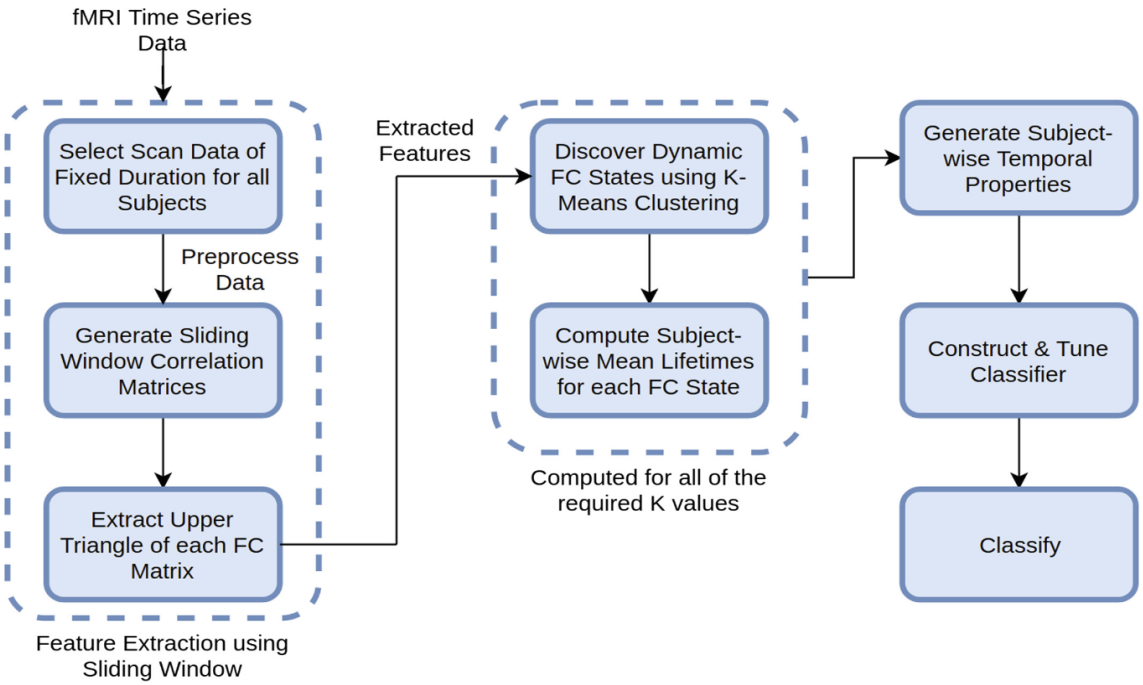

Figure 3.3 Dammu et al. [21] proposed pipeline

Implementation & Results The implementation details and hyper parameters are not mentioned in the paper. Table 3.1, summarizes the results reported in the paper.

Table 3.1 Different brain atlas with corresponding number of ROIs. This table adapted from [21]

SVM	0.735	0.764	0.704	0.748
RF	0.717	0.759	0.673	0.733
XGB	0.736	0.750	0.720	0.744
DNN	0.710	0.713	0.706	0.714

Reproduced results & Observations Hyper-parameters are crucial for training the DNN. The hyper-parameters which are essential to training SVM are regularization constant (λ), kernel and, to train RF, XG Boost are number of estimators, criterion, and max depth. The paper fails to mention these hyper-parameters of the classifiers and other implementation details. Despite our best efforts and several combinations of various hyper-parameters, we could only reach a classification accuracy of 65.93%.

3.2.4 Dynamic functional connectivity analysis among ASD subjects [35]

Introduction & Methods Fu et al. [35] investigated the dFNC between 51 intrinsic connectivity networks in 170 ASD subjects and 195 age-matched TD individuals. Using group-independent component analysis (GICA), 51 intrinsic connectivity networks (ICNs) were extracted based on the spatial mappings of the components. dFNC was estimated using a sliding window approach with window size of 20 and a step size of 1 based on the time courses (TCs) of ICNs. To look into regional and global dynamic patterns of functional connectivity, time-varying correlation coefficients (dFNC estimations) were subjected to hard clustering state analysis and fuzzy meta-state analysis. To examine the frequency of different functional states, they computed the percentage of occurrence of windows per state per subject.

Implementation & Results All the above analysis was implemented using GIFT toolbox [60]. From the hard-clustering analysis, they found 5 optimal brain states or dFNC patterns. The occurrence of state 1 is significantly higher in TD individuals, and state 4 is significantly greater in ASD subjects $(p < 0.05)$. Hyper-connectivity between the subthalamus/hypothalamus and right postcentral gyrus, bi-paracentral lobule in state 2, and lingual gyrus in state 4 has been observed among ASD subjects.

Reproduced results & Observations First, the paper lacks low-level details needed to pre-process the data, extract ICNs using the GIFT toolbox, and parameters to perform KMeans clustering. Second, we obtained only 189 TD subjects using the data selection criteria proposed in the paper while they reported 195 TD subjects. We found four optimal states while performing the clustering analysis on the dFNC windowed matrices, as opposed to the paper's report of five. Therefore, as a result, the remaining analysis and final conclusions are entirely different from what we reproduced and what was reported in the publication.

Neccessary details which would have helped in reproducing the reported results :

1. The subject IDs of the final subjects selected for the analysis should have been provided in the supplementary material.

2. The pre-processed data should have been made publicly available because it is a crucial and intricate step.

3. They should have provided the parameter files generated by GIFT toolbox which contains all the required details to extract the ICNs, estimate dFNC and cluster the dFNC windows.

3.3 Good Practices to improve reproducibility

Neuroimaging studies involving ML and neural networks can be particularly challenging to reproduce. However, there are several best practices that can be employed to improve reproducibility in these studies [76]. Here are some of them:

- Data sharing: Sharing the data and code publicly is important in order to facilitate replication of the study by other researchers. This can also promote transparency and trustworthiness of the findings.
- Pre-processing: Document all pre-processing steps, including image normalization, smoothing, and registration. Ensure that the pre-processing is done in a consistent manner across all participants. Making the pre-processed data available saves a lot of time and efforts of other researchers.
- Feature extraction: Clearly define the features that are being extracted from the neuroimaging data. Use established feature extraction methods that are widely accepted in the field.
- Model selection: Clearly document the ML or neural network models and its parameters that are being used. Use models that have been previously validated and showed consistent performance across different datasets.
- Cross-validation: Use cross-validation to test the robustness of the model. This involves splitting the data into training and testing sets and repeating the analysis multiple times with different splits. Make sure to avoid data leakage i.e. training and testing the model on same or overlapping datasets. This could lead to false and unreliable results.
- Replication: Conduct replication studies to confirm the robustness of the findings. Replicating the study using different samples or datasets can help to establish the validity of the findings.

By following these best practices, researchers can improve the reproducibility of their neuroimaging studies involving machine learning and neural networks, and help to ensure that the findings are reliable and robust.

3.4 Summary

The reproducibility crisis in neuroimaging is caused by the interdisciplinary nature of the field, small sample sizes, and flexible analysis pipelines that allow researchers to report only results that support their hypotheses. There is urgent need for greater transparency, openness, and the use of standardized analysis pipelines and larger sample sizes to improve the reliability of neuroimaging research. The use of ML and neural networks in neuroimaging studies can be challenging to reproduce. However, employing best practices as described in section 3.3 can improve reproducibility. In the next chapter we will discuss our study where we introduced a novel neural network with two hidden layers based on sFNC for ASD classification, producing better results than the current state-of-the-art methods. Additionally, we applied attribution methods to explain and interpret the model.

Chapter 4

Diagnosis of ASD using Static Functional Connectivity [1]

The current diagnosis of Autism Spectrum Disorder (ASD) is subjective, lacks accuracy, and requires experts, resulting in an economic burden on families. sFNC (as explained in subsection 2.4) measures the functional connections between different brain regions and can provide important diagnostic information of various neurodegenerative and psychiatric disorders like ASD. ML algorithms can potentially analyze and identify functional connectivity patterns that are difficult for human clinicians to detect, improving the accuracy and consistency of diagnoses. In this chapter, we will present our method for classifying ASD from TD using sFNC and interpreting the model to determine which brain regions correlations are most relevant for the classification task. The material in this chapter has been taken from [58].

4.1 Dataset

We utilized the dataset from the preprocessed ABIDE-I [22]. 86 ASD scans have been discarded due to their missing time series [77]. There is significant noise due to patient-specific difficulties such as repetitive head motion, body trembling, keeping eyes open/closed during the scan, and heterogeneity due to high variation in demographics, scan parameters, and age groups of participants collected from different sites. We experimented with four different pipelines such as CCS [80], C-PAC [19], DPARSF [82], NIAK [7] and four brain parcellation schemes such as HO, AAL [75], Dosenbach 160 [23], CC200 [20].

[1] **Prasad, P.K.C., Khare, Y., Dadi, K., Vinod, P.K. and Surampudi, B.R. (2022). Deep Learning Approach for Classification and Interpretation of Autism Spectrum Disorder. In International Joint Conference on Neural Networks (IJCNN-2022) (pp. 1-8). https://ieeexplore.ieee.org/document/9892350/**

4.2 Methods

4.2.1 Feature Extraction

sFNC is measured by calculating the pair-wise correlations between every pair of brain regions. Mostly these correlations are linear, which are captured using the Pearson Correlation Coefficient (PCC). The PCC, ρ_{xy} for two signals, x and y each of length T and mean $\hat{x}$ and $\hat{y}$ respectively, can be computed using the following equation.

$$\rho_{xy} = \frac{\sum_{t=1}^{T}(x_t - \hat{x})(y_t - \hat{y})}{\sqrt{\sum_{t=1}^{T}(x_t - \hat{x})^2}\sqrt{\sum_{t=1}^{T}(y_t - \hat{y})^2}} \tag{4.1}$$

Given n brain regions, we obtain an $n \times n$, sFNC symmetric matrix where each $(i, j)^{th}$ entry represents the PCC between i^{th} and j^{th} regions. We extract the upper triangular values of the sFNC matrix and use them as an input to our model.

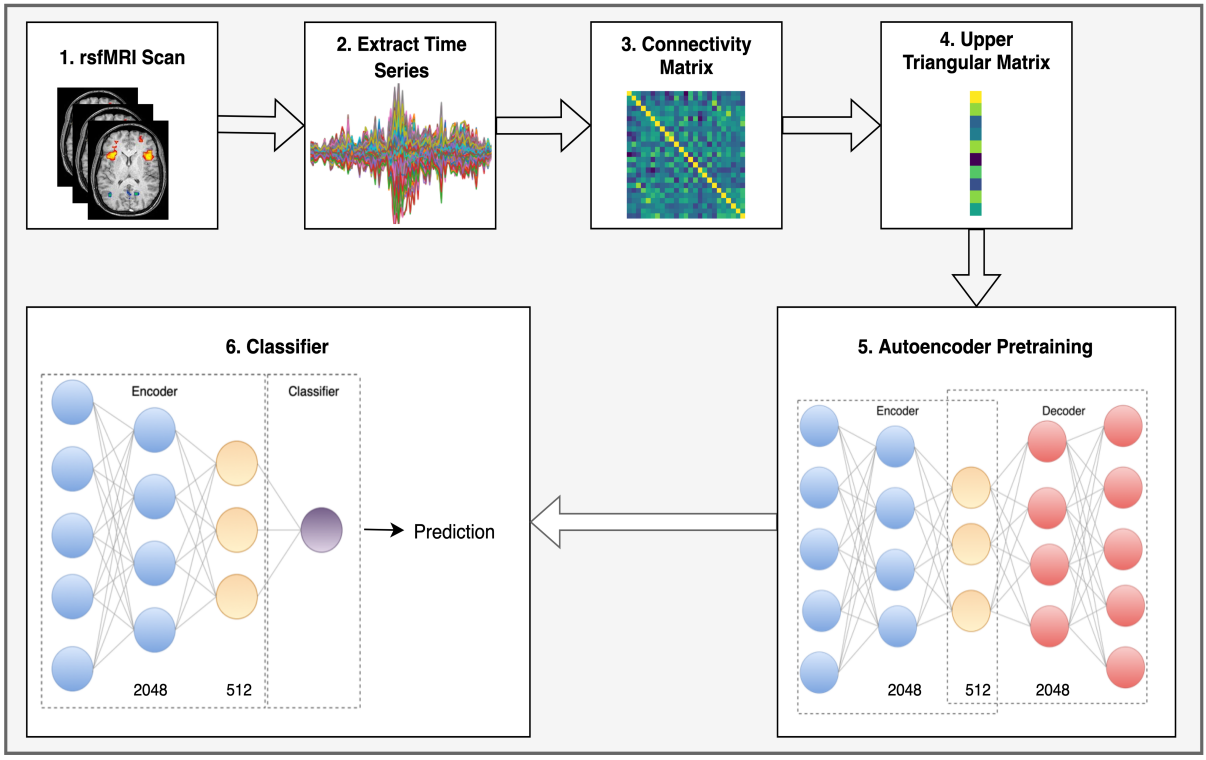

Figure 4.1 Classification Pipeline. 1. Pre-processing the rs-fMRI scan. 2. Extracting the time-series data using pre-defined anatomical or functional brain parcellations. 3. Extracting the static functional connectivity matrix. 4. Flattening the upper triangular part of functional connectivity matrix and passing as input to the auto-encoder. 5. Training the auto-encoder. 6. Fine-tuning the classifier with pre-trained auto encoder weights.

4.2.2 Autoencoder and Classification Method

Autoencoder (AE) [4] is a type of neural network used for unsupervised learning, particularly for data compression, denoising, and feature extraction tasks. AE has been applied in various fields, including image processing, natural language processing, and recommendation systems. AE consists of an encoder network that maps input data to a lower-dimensional representation, called the "latent space", and a decoder network that reconstructs the input data from the latent space. The encoder takes the input vector (x) and maps it to a lower-dimensional representation in the latent vector (z), while the decoder takes the latent representation and reconstructs the input vector ($\hat{x}$). The loss function measures the difference between the original input data and the reconstructed output, and the AE aims to minimize this loss during training. Usually, the reconstruction loss is computed using the Mean Squared Error (MSE).

$$Loss(x, \hat{x}) = \sum_{i=1}^{n} (x_i - \hat{x}_i)^2 \tag{4.2}$$

One of the key features of autoencoders is that they learn to represent data in a compact and meaningful way in the latent space. The encoder learns to extract relevant features or representations from the input data, which are then used by the decoder to reconstruct the input data. By doing so, autoencoders can learn to capture important patterns, structures, and features in the data, even if the data is high-dimensional or noisy.

In this study, we use the autoencoder in the pretraining stage. The encoder architecture consists of two hidden layers of size 2048 and 512 respectively, as shown in Fig.4.1. We use $tanh$ as the nonlinear activation function for both the layers. The decoder architecture mirrors the encoder's, with two hidden layers of size 512 and 2048, respectively. For each fold, we first train the autoencoder, remove the decoder part, and add the classifier, which has a single output layer as shown in Fig.4.1. The model comprising the encoder and classifier is then fine-tuned for the classification task by minimizing the binary cross-entropy loss H between the original class label, y_i and the model prediction, $\hat{y_i}$ for all the N subjects.

$$H = \frac{-1}{N} \sum_{i=1}^{N} (y_i * \log \hat{y}_i + (1 - y_i) * \log(1 - \hat{y}_i)) \tag{4.3}$$

4.2.3 Model Interpretation

Motivation Many studies have developed classification pipelines to detect ASD from TD with high accuracy. However, identifying the brain bio-markers in ASD subjects is less explored. Hence, it is important to identify which features or connectivities are contributing to the prediction of Autism. Therefore, in this study, we have attempted to find out the regions associated with ASD using Integrated Gradients (IG) [70] and DeepLIFT [66]. To the best of our knowledge, this is the first study in our field to use such cutting-edge attribution techniques.

Introduction to Attribution Methods Attribution methods used to understand and interpret the contributions of different input features or variables towards the prediction of a model. Attribution methods are divided into two categories: Perturbation-based approaches and Backpropagation-based approaches. The perturbation-based approaches compute the feature importance by making perturbations to individual inputs or neurons and observing how this affects the network. As each perturbation requires a separate forward propagation through the network, such methods can be computationally expensive. Backpropagation-based approaches compute the feature importance by propagating the signal from the output back to the input in one pass, making them efficient. Attribution methods such as Deconvolutional networks, Guided backpropagation, Layer-wise relevance propagation compute feature importance based on the gradient of input with respect to models prediction. These gradient-based methods suffer from saturation and thresholding problems as explained in [70]. IG and DeepLIFT attempt to tackle these limitations. Therefore, this study uses IG to determine which ROIs in the brain are associated with ASD and are further validated by DeepLIFT.

Integrated Gradients IG is a gradient-based method aggregates the gradients along the inputs that fall on the straight line between the baseline and the input. The baseline is typically a reference point, such as the mean or zero value of the input features, and the integral is computed along a linear path connecting the baseline and the input data point. The integral of gradients measures the change in the prediction caused by the change in the input features along the path, and it helps to quantify the importance of each feature in the prediction. Computationally, the IG algorithm involves the following steps:

- Define a baseline: A baseline is chosen as a reference point, typically the mean or zero value of the input features. This baseline serves as a starting point for computing the integral of gradients.

- Compute gradients: Gradients of the output of the neural network with respect to the input features are computed. These gradients represent the sensitivity of the prediction to changes in the input features.

- Compute path: A linear path is defined from the baseline to the input data point of interest. This path is typically parameterized by a scalar value, and intermediate points along the path are computed.

- Compute integral: The integral of the gradients along the path is computed using numerical integration techniques, such as the trapezoidal rule or Simpson's rule. This integral quantifies the change in the prediction caused by the change in the input features along the path.

- Attribute prediction: The integrated gradients are used to attribute the prediction to the input features. The integrated gradients represent the contributions of each input feature to the prediction, and they can be visualized or analyzed to interpret the neural network's behavior.

The integrated gradient along the i^{th} dimension for an input x and baseline x' is defined as follows. Here, $\partial F(x)/\partial x_i$ is the gradient of $F(x)$ along the i^{th} dimension.

$$IG_i(x) = (x_i - x_i') \times \int_{\alpha=0}^{1} \frac{\partial F(x' + \alpha \times (x - x'))}{\partial x_i} d\alpha \tag{4.4}$$

DeepLIFT DeepLIFT method computes feature importance based on explicating the difference in output from some reference output in terms of the difference of the input from some reference input. Though the gradient is zero, information can be propagated using this approach based on differences from the reference values [66]. The contribution of a feature is computed as the difference between the activation of the feature at the input layer and the reference activation, multiplied by the gradient of the prediction with respect to the feature. This attribution method helps to quantify the importance of each feature in the prediction.

Summary IG and DeepLIFT are both methods for interpreting neural networks, but they differ in their approaches, baseline selection, gradient computation methods, interpretation results, and computational complexity. Feature importance for any input is computed against baseline input. A baseline represents the beginning point to compute the feature's significance and should be unbiased in terms of being independent of the final attribution method chosen. Defining a baseline is challenging and depends on the dataset. Passing the input with null feature values reveals how the model performs when no information is provided and is neutral to the models prediction as recommended in [69]. Therefore, initializing the baseline as input of zeros would be appropriate in this study. IG and DeepLIFT has several advantages as an interpretation method for neural networks. Firstly, it provides a way to quantify the contributions of input features to the predictions, allowing for a better understanding of which features are important for the neural network's decision-making process. Secondly, these methods are model-agnostic, meaning it can be applied to any neural network architecture, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and transformer models, among others. Thirdly, these methods can be used for both binary and multi-class classification problems and regression problems, making them versatile in interpreting different types of neural networks.

4.3 Experiments and Results

4.3.1 Experimental Setup

This study evaluates the proposed method using 10-fold cross-validation [56] on the ABIDE-1 dataset with 949 samples. In each fold, 20% of the training set is used as a validation set for hyperparameter tuning. The upper triangular part of the sFNC matrix is flattened and passed as input to train the autoencoder part (pretraining stage) for 50 epochs and the classifier part for 50 epochs using

Adam optimizer with a learning rate of 10^{-4} and weight decay of 0.1. The encoder part of the classifier is initialized with pretrained auto-encoder weights and the whole classifier has been trained. All the hyper-parameters were tuned using cross-validation The code is implemented in pytorch [54] and made available here[2]. Computations are performed on a computing system with 16GB RAM and NVidia K80 GPUs for the entire training stage.

4.3.2 Classification

We report the accuracy, sensitivity, and specificity of our model. Our model achieves a new state-of-the-art performance with an accuracy of 74.82%, sensitivity of 67.33%, specificity of 80.75% on the ABIDE-1 dataset. Table 4.1 reports the performance of the proposed method in comparison with previous studies for the ASD classification task.

4.3.3 Impact of different Preprocessing Pipelines

In this section, we discuss the impact of using different preprocessing pipelines on the performance of our proposed method. We train our model on data obtained using four preprocessing pipelines, CPAC, NIAK, DPARSF, and CCS. Table 4.2 shows the classification results using these pipelines. We get the best accuracy of 74.82% for the CPAC pipeline followed by 71.21% for DPARSF, 71.04% for CCS, and 64.55% for NIAK. As we can see, there is a 13% difference in the accuracy when using the CPAC pipeline compared to the NIAK pipeline, which suggests that choosing the correct preprocessing pipeline for the classification task plays a significant role in the overall results.

Table 4.1 Our method outperforms all the previous methods in overall accuracy.

Method	Accuracy(%)	Sensitivity(%)	Specificity(%)
Heinsfeld[38]	70.0	74.0	63.0
Parisot[53]	70.4	-	-
ASD-Diagnet[30]	70.3	68.3	72.20
CNN[65]	70.22	77.46	61.82
AIMAFE[77]	74.52	**80.69**	66.71
Proposed Method	**74.82**	67.33	**80.75**

4.3.4 Impact of different parcellation schemes

This section presents the results of a comparative study using four different brain parcellations – CC200, Dosenbach, AAL, and HO – which parcellate the brain into 200, 160, 116, and 111 regions, respectively. As the number of regions of interest (ROIs) is different for each atlas, the input vector size

[2]https://github.com/pindi-krishna/Classification-and-Interpretation-of-ASD.git

Table 4.2 Results using different preprocessing pipelines on ABIDE-1 dataset using CC200 brain parcellation. CPAC pipeline has yielded the best performance

Pipeline	Accuracy(%)	Sensitivity(%)	Specificity(%)
CPAC	**74.82**	**67.33**	**80.75**
NIAK	64.55	46.78	78.6
DPARSF	71.21	64.46	76.57
CCS	71.04	66.02	75.02

Table 4.3 Results using different atlases on ABIDE-1 dataset preprocessed using CPAC pipeline

Atlas	Accuracy(%)	Sensitivity(%)	Specificity(%)
CC200	**74.82**	**67.33**	**80.75**
HO	72.29	66.06	77.21
Dosenbach	70.92	64.69	75.85
AAL	69.95	60.90	77.09

Table 4.4 Leave-one-site-out results on ABIDE-1 dataset preprocessed using CPAC pipeline and CC200 brain parcellation scheme

Site	Accuracy(%)	Sensitivity(%)	Specificity(%)
CALTECH	71.25	64.29	76.67
CMU	71.54	58.46	84.62
KKI	73.91	65.56	79.28
LEUVEN	74.07	33.0	98.24
MAXMUN	57.34	48.24	62.86
NYU	78.75	72.33	82.60
OHSU	60	44.44	70
OLIN	74.9	66.67	82.22
PITT	73.73	63.33	82.96
SBL	54.82	18.34	84.0
SDSU	76.13	40.0	90.91
STANFORD	58.38	78.82	41.0
TRINITY	65.0	65.26	64.8
UCLA	76.0	71.31	80.91
UM	75.76	75.86	75.67
USM	**85.08**	**84.29**	**86.4**
YALE	84.31	89.56	80.0
Mean	71.23	61.16	77.83

also changes accordingly, affecting the models overall performance. Table 4.3 shows the classification results using the above mentioned atlases. We get the best accuracy of 74.82% using the CC200 atlas.

4.3.5 Leave-One-Site-Out Results

One major challenge in the ASD classification task using the ABIDE-1 dataset is the inter-site variability as it contains rs-fMRI scans from 17 different sites. To evaluate the performance of our proposed method on new sites, we train our model using the leave-one-site-out approach. In this approach, we keep the data from each site separately for testing, and we train our model on the data from the remaining sites. This approach allows us to estimate the models generalizability to new, unseen sites. Table 4.4 shows the leave-one-site-out results. Our model achieves the best accuracy of 85.08% on scans collected from the University of Southern Mississippi (USM) site, followed by 84.31% for Yale and 78.75% for NYU.

4.3.6 Interpretation

Attribution methods have been applied on the model trained using the data obtained from CPAC preprocessing pipeline and CC200 brain parcellation scheme as this combination yielded the best result. The steps for finding the associated regions using IG and DeepLIFT are as follows:

1. Group all the correctly predicted autism samples from the test set as we are interested in regions that led to the prediction of autism
2. Apply attribution method on these autism samples to find the attributions for each feature passed as input to the model with a zero embedded vector as the baseline.
3. Replace the attributions in the top one percentile with 1 and the remaining values with 0.
4. Construct a 2D matrix of size 200×200 with attribution vector contributing to the upper triangular matrix.
5. Calculate the row-wise sum and pick those ROIs with the maximum value.
6. Repeat steps $1 - 5$ on each fold and find out the most repeated, common ROIs in all the folds.

The results are qualitatively similar when the attributions in top k percentile ($k = 0.5, 2$, and 3) are replaced with 1. We also employed the same attribution methods on the TD subjects to ensure the consistency of the identified regions. It was observed that the regions influencing the prediction of ASD were also found to have an impact on the prediction of TD, considering the binary nature of the classification task. These attribution methods have been implemented using *Captum*, a model interpretability library for PyTorch [45]. IG and DeepLIFT analysis identifies the following regions: Left Lingual Gyrus (LLG), Right Insula Lobe (RIL), Right Cuneus (RC), Right Middle Frontal Gyrus (RMFG), Left

Superior Temporal Gyrus (LSTG) to be associated with ASD classification (see Fig.4.2). The primary function and coordinates of these regions based on CC200 parcellation are shown in Table 4.5. We conclude that feature attribution methods such as IG and DeepLIFT enable accurate identification of brain regions whose activation seems to be altered in ASD compared to TD.

Table 4.5 ROIs that contributed most to the ASD classification task based on CC200 brain parcellation

ROI name	ROI number	Center of mass	Primary Function
LLG	177	(-14.3;-74.2;-10.1)	Face/Object Recognition
RIL	59	(36.7;17.2;3.6)	Decision making
RC	142	(17.8;-89.5;22.7)	Visual Processing
RMFG	106	(28.6;34.6;42.0)	Attention
LSTG	200	(-41.9;-31.5;15.2)	Language Comprehension

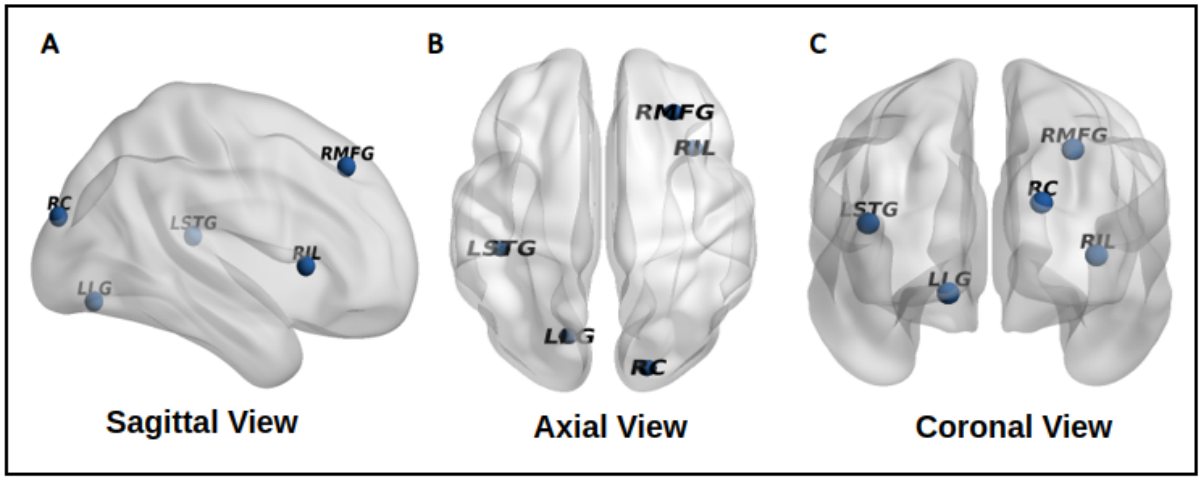

Figure 4.2 Visualization of regions associated with ASD in 3 different views using Brain Net Viewer [79]. A. Sagittal view B. Axial View C. Coronal View. LLG : Left Lingual Gyrus, RIL : Right Insula Lobe, RC : Right Cuneus, RMFG : Right Middle Frontal Gyrus, LSTG : Left Superior Temporal Gyrus

4.4 Discussion

This work shows that the proposed method demonstrates effective results on the ASD classification task. Among all the studies reported in Table 4.1, our method achieved the best classification accuracy of 74.82%. The entire training time of our model is less than an hour, approximately 10x faster than the previous state-of-the-art method. This suggests that the autoencoder pretraining weights acted as an excellent initialization point for training the classifier, which helped in the faster convergence of the model.

The data have significant heterogeneity as the scans are collected from various sites, and each site uses a different set of scan parameters and protocols. The variations in the leave-one-site-out results as shown in Table 4.4 highlight the impact of the heterogeneity on the classification task. Although there is no consensus and conceptual clarity on how different preprocessing pipelines and brain parcellations affect the ASD classification results, we observe a significant variation in the results as shown in Table 4.2 and 4.3, respectively. These variations may occur due to the implementation of different algorithms and parameters used in preprocessing pipelines and different ROIs extracted from each parcellation scheme. The empirical results obtained in this study reveal that the combination of CPAC preprocessing pipeline and CC200 parcellation yield the best accuracy for the ASD classification task. Going forward, these results might be useful in defining a common benchmark dataset and specifications to enable a fair and viable comparison of methods being proposed for ASD classification.

This study identifies LLG, RIL, RC, RMFG, LSTG associated with ASD. A lingual gyrus plays a crucial role in vision processing, primarily related to letters. Chandran et al. [15] found a significant association between the left lingual gyrus cortical thickness and the right lateral occipital cortex surface area among autism subjects. The greater volume and gyrification of the lingual gyrus and lateral occipital cortex may cause abnormal visual processing in individuals with higher autistic symptoms. Yamada et al. study [81] on 36 ASD subjects and 38 TD subjects found a significant change in the anterior sector of the left insula and the middle ventral sub-region of the right insula in the ASD brain. They noticed a notable volumetric increase in the ASD brain compared with the TD brain in the middle ventral sub-region, in the right insula. The right cuneus is responsible for visual processing. Stock et al. study [27] on 66 adults with high-functioning autism and 66 TD investigated gray matter abnormalities in the two groups. They found increased gray-matter volume in frontal brain regions, including the medial prefrontal cortex, superior and inferior frontal gyri, and middle temporal gyrus, and reduced gray-matter volume in posterior brain regions, including the posterior hippocampus, cuneus, in individuals with ASD in relative to TD. Middle Frontal Gyrus and Face processing seem quite relevant as ASD patients have difficulty with face processing and facial emotion identification. In ASD subjects, significantly reduced activity in the middle frontal gyrus when involving face processing tasks and middle temporal gyrus during nonface social tasks is found when compared to TD subjects [55]. The superior temporal gyrus (STG) is responsible for language comprehension. Bigler et al. study [9] on 30 autistic children and 39 controls of similar age, education, and head circumference, investigated the link between autism and intellectual-language-based abilities. Clinical Evaluation of Language FundamentalsThird Edition (CELF3) [63], divided into three index scores: expressive, receptive, and total, has been used to measure language ability. This study observed a positive correlation between receptive language scores and STG volume in control subjects and zero correlation in autistic subjects.

4.5 Summary

We proposed a model with two hidden layers and an autoencoder (AE) pertaining that achieves state-of-the-art performance. In addition, for model interpretation we used integrated gradients (IG) and DeepLIFT to identify ROIs associated with ASD. We demonstrated the effects of different preprocessing pipelines and brain parcellation schemes on classification performance. Although the BOLD time series signal is a time-varying signal throughout the rs-fMRI scan, in the proposed approach, we calculated the grand mean FC value over the entire time series, thus yielding sFNC. Features from the sFNC matrix were used in the proposed approach to obtain the best results. Thus, there is still scope for improvement using various measures to capture the brain dynamics using the dFNC. In the next chapter, we will discuss clustering, statistical and meta-state analysis conducted based on dFNC to understand the fundamental characteristics of ASD.

Chapter 5

Characterizing ASD using Dynamic Functional Connectivity[1]

Dynamic functional connectivity (dFNC) is a neuroimaging technique that analyzes how connectivity between different brain regions changes over time. Currently, the dFNC analysis has proven to be effective in studying the neurological basis of various psychiatric disorders, including ASD. The functional correlation matrix for each overlapping interval was created using the BOLD signals in the rs-fMRI time series. The changes in functional connectivity between brain regions can be captured continuously over time and analyzed to understand the ASD characterstics.

5.1 Dataset & Preprocessing

We used the pre-processed dataset from the ABIDE-I initiative [22]. For the analysis, we chose 383 subjects (188 ASD and 195 TD). The criteria (figure 7.1) to include subjects are as follows:

1. Subjects with DSM-IV diagnosis,
2. Male subjects to remove the gender bias.
3. Subjects with mean frame-wise displacement (mFD) corresponding to two standard deviations above the sample mean, i.e., smaller than 0.4432
4. Subjects with atleast 175, volumes in fMRI acquisition
5. Subjects with repetition time $(TR) = 2$ sec while scanning
6. Subjects with Full IQ score
7. Subject mask with a spatial correlation greater than 0.8 with the group mask computed using the subjects qualified for the above five conditions. Group mask computation is as follows:

[1] **Prasad, P.K.C., Dadi, K., and Surampudi, B.R. (2023). Dynamic functional connectivity analysis in individuals with Autism Spectrum Disorder. In International Joint Conference on Neural Networks (IJCNN-2023).**

(a) Firstly, we determined the individual mask for each subject by setting the value of voxels larger than 70 percent of the entire brain mean value to 1.

(b) Next, we set the voxels present in more than 70% of the individual masks to 1 to compute a group mask.

We used publicly available preprocessed four-dimensional rs-fMRI scans using Connectome Computation System (CCS) [80]. Dataset can be downloaded from here [2] by setting the pipeline as 'ccs' and strategy as 'nofilt_noglobal'.

5.2 Methods

5.2.1 Extraction of Independent Component Networks

We used the ICA NeuroMark template [25] to extract 53 Independent Components (ICs) belonging to 7 resting state networks. The ICA NeuroMark Template is a functional parcellation atlas developed using independent component analysis (ICA) applied to rs-fMRI data. The atlas is based on data from a large sample of healthy individuals, and its regions of interest (ROIs) are defined by functional networks rather than anatomical boundaries. This approach allows for a more precise mapping of functional connectivity networks in the brain and has the potential to reveal new insights into brain organization and function. The atlas is based on a standardized template space, which allows for easy comparison and integration with other neuroimaging datasets. The ICA NeuroMark Template has shown promise in diagnosing neurological disorders, such as Alzheimer's disease, Parkinson's disease, and multiple sclerosis. Therefore, we used it for the dFNC analysis of ASD to identify the reliable bio-markers.

The number of components is denoted by C. These networks were categorized into the following seven domains: Subcortical Domain (SC; ICs: $1-5$), Auditory Domain (AUD; ICs: $6, 7$), Visual Domain (VIS; ICs: $8-16$), Somatomotor Domain (SM; ICs: $17-25$), Cognitive Control domain (CC; ICs: $26-42$), Default-Mode Domain (DM; ICs: $43-49$), and Cerebellar Domain (CB; ICs: $50-53$). Each of these networks has specific functional roles and is involved in different cognitive and motor processes described as follows [1]:

1. Sub-cortical Network (SC): SC is a functional network that includes the thalamus, caudate nucleus, and putamen. This network is involved in several functions, such as motor control, sensory processing, and attention. The thalamus is responsible for relaying sensory information to the cerebral cortex, while the caudate nucleus and putamen are involved in motor control.

2. Visual Network (VIS): VIS is a functional network that includes the primary visual cortex, secondary visual cortex, and higher-order visual areas. This network is involved in processing visual information and is responsible for visual perception, object recognition, and spatial attention.

[2] http://preprocessed-connectomes-project.org/abide/download.html

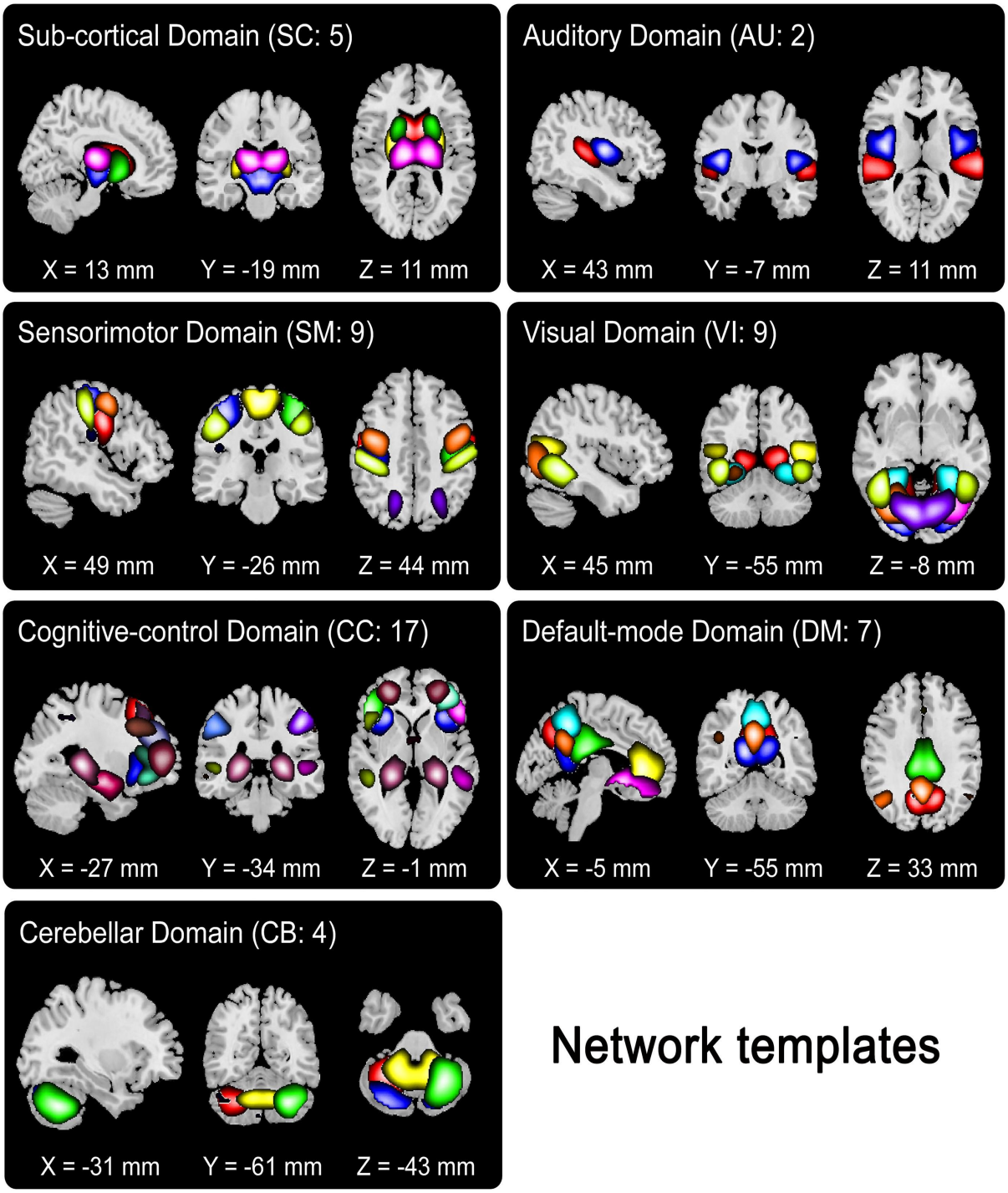

Figure 5.1 Visual representation of the discovered network templates, separated into seven functional areas based on anatomical and functional characteristics. One of the composite maps' colours in each subfigure corresponds to an ICN.

3. Auditory Domain (AUD): AUD is a functional network that includes the primary auditory cortex, secondary auditory cortex, and higher-order auditory areas. This network is responsible for processing auditory information and is involved in auditory perception, speech processing, and sound localization.

4. Default Mode Network (DMN): DMN is a functional network that includes the medial prefrontal cortex, posterior cingulate cortex, and inferior parietal lobule. This network is active during rest and is deactivated during task performance. The DMN is involved in several functions, such as self-referential processing, episodic memory, and social cognition.

5. Cerebellum (CB): The cerebellum is a functional network that includes the cerebellar hemispheres and the vermis. This network is involved in motor control, motor learning, and sensory processing. The cerebellum plays an important role in fine-tuning motor movements and maintaining postural stability.

6. Cognitive Execution Network (CEN): CEN is a functional network that includes the dorsolateral prefrontal cortex, anterior cingulate cortex, and posterior parietal cortex. This network is involved in several cognitive functions, such as working memory, attention, and decision-making.

7. Somatomotor Domain (SM): SM is a functional network that includes the primary motor cortex, supplementary motor area, and somatosensory cortex. This network is involved in motor control and somatosensory processing. The somato-motor domain plays an important role in the execution of voluntary movements and the perception of touch and pain.

ICA Neuromark template can be downloaded here [3] and its corresponding labels can be found in appendix 7.1. Visual representation of these ICNs can be seen in figure 5.1. Time courses (TCs) are extracted from these ICNs followed by temporal pre-processing strategies that includes: detrending linear, quadratic, and cubic trends, de-spiking detected outliers, and low-pass filtering with cut-off frequency of 0.15 Hz.

Table 5.1 Dynamic functional connectivity parameters

Parameters	Values
Number of components (C)	53
Timepoints per subject (T)	175
Window size (w)	20
Stride (s)	1
Windows extraction per subjects (N)	155
Windows extracted from all subjects (W)	59365

[3] https://trendscenter.org/trends/data/neuromark/Neuromark_fMRI_1.0.nii

5.2.2 Dynamic Functional Connectivity estimation

A sliding window-based method was used to estimate dFNC for each individual. We only chose the first T time points of each subject's component TCs for the dFNC estimation because the subject's data from different sites could have varying scan lengths. This was done to minimize the effects of different scan lengths. To localize the dataset at each time point, we created a tapered window by convolving a rectangle with a Gaussian ($\sigma = 3$). We chose a window size w TRs corresponding to t seconds. We slid the window in s TR increments, producing a total of N windows per subject. We estimated the regularised precision matrix using the graphical LASSO approach (with the $L1$ norm to encourage sparsity) and then deduced the covariance matrix from the precision matrix [32]. Let the total number of windows extracted from all the subjects be W. All the dFNC parameters can be seen in table 5.1

5.2.3 Hard Clustering Analysis

The fundamental presumption behind the hard clustering state analysis is that functional brain networks will enter different states with unique dFNC patterns. All the windows extracted from all the subjects (W) are divided into distinct states (discrete FC patterns) utilizing the K-Means clustering technique and the $L1$ norm as the distance function. To reduce the redundancy between windows and computational demands, we utilized the subset of windows (consisting of local maxima in functional connectivity variance) as subject exemplars. As K-means algorithm is sensitive to initialization, it was repeated for 100 times (with random initialization of centroid position) with a maximum of 250 iterations to obtain the group cluster centroids. To find the optimal number of states, we used the elbow criterion which is defined as the ratio of within clustering distance to between clusters distance. The optimal number of dynamic states was determined as $K = 4$. The optimal dynamic states can be seen in figure 5.2.

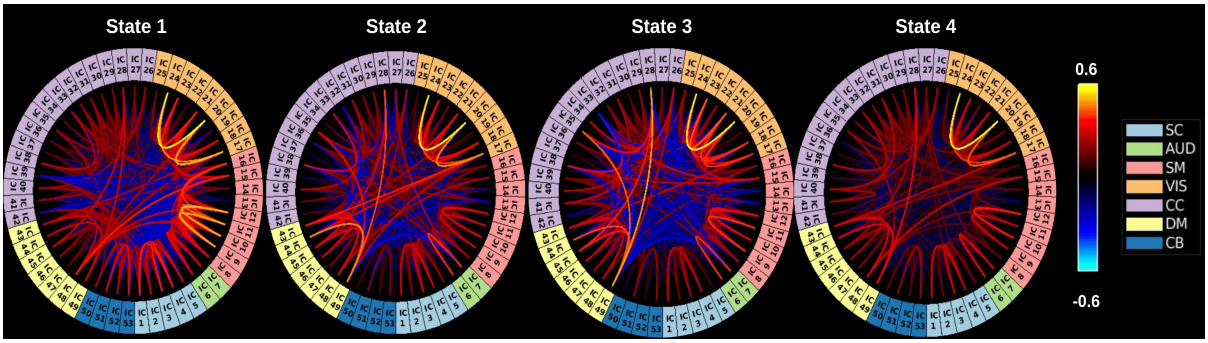

Figure 5.2 Discrete Functional Connectivity (FC) Patterns (optimal number of brain states 4). These are the cluster centroids obtained through hard clustering analysis on the FC patterns from all subjects, both ASD and TD. The Connectograms depict the connectivity among the regions that comprise seven resting state networks indicated in different colors. The strengths of the inter-regional connectivity is depicted using a color map varying from -0.6 to 0.6.

5.2.4 Group differences in static and dynamic-connectivity strength

Two-sample *T-test* analysis has been done to compare static and dynamic connectivity strength between ASD and TD groups. For dFNC, we repeated the below steps (A, B, C) for all four dynamic brain states. A) Consider those subjects in which at least a window belonging to that particular state exists. B) In all those subjects, compute the median of windows per state. C) There exist 1378 distinct connectivities for every pair of ICNs. For every distinct connectivity, performed the two-sample *T-test* between groups ASD and TD per state to find significantly different connectivities ($p < 0.05$, false discovery rate (FDR) correction). In case of sFNC, we can conduct the two-sample *T-test* directly on connectivity between pairs of ICNs to compare the groups.

5.2.5 Temporal properties

We also investigated the temporal properties of four dynamic brain states by computing the fractional windows (the number of total windows belonging to a given state), mean dwell time (the number of consecutive windows belonging to a given state) per state and the number of transitions between each pair of states. Statistical two-sample *T-test* are used to identify the differences between groups ASD and TD per state ($p < 0.05$, FDR correction).

5.2.6 Meta-state Analysis

Meta state analysis [52], has been conducted to investigate abnormal dFNC patterns. While dFNC clustering and statistical analysis have yielded some noteworthy findings, reducing the connectivity space to a single dimension inevitably hides and alters important aspects of dynamic network-coupling behavior that may be distinctive features of clinically or demographically defined groups. This method represents the time-varying functional network connectivity as a combination of independent dFNC patterns with varying weights. The discretized vector of dFNC patterns weights is called a meta-state. This approach better comprehends the underlying dynamics of network connectivity in terms of distinct patterns of correlated network pairs that appear and disappear independently in the observed functional network connectivity.

A four-dimensional weight vector was obtained for each time window by fitting the windowed dFNC estimates to the group cluster centroids. This weight vector represents how much each basic dFNC pattern contributes to the windowed dFNC. To make these real-value weight vectors discrete, the vector weights were replaced with values of $\pm(1, 2, 3, 4)$, depending on the quartile of weights the vector falls into. These different combinations of $\pm(1, 2, 3, 4)$ are then defined as meta-states. There exist 8^4 meta-states (all permutations of $\pm$ {1,2,3,4}).

Four dynamic measures were derived based on the meta-states vector to assess the overall dynamic features of the functional brain network.

1. Meta-states number: The count of unique meta-states traversed by an individual.

2. Meta-states switching times: The measure of how often individuals transition or switch between different meta-states.

3. State span: Measures the degree of dissimilarity among the meta-states that have been occupied.

4. Overall traveled distance: Quantifies the total distance traveled by an individual when switching between different meta-states.

Table 5.2 Distribution of number of subjects and windows per state for each group: ASD and TD. The 'Subjects' represents the number of subjects occurred in each state, and the 'Windows' represents the number of windows assigned to each state within ASD and TD.

State	ASD		TD	
	Subjects	Windows	Subjects	Windows
State1	59	2241	79	3653
State2	145	5333	165	6575
State3	131	4849	158	6583
State4	185	16717	187	13414

Table 5.3 Distribution of strength in connections characterized from pairs of IC networks for both dynamic and static functional connectivity. For each state, the number indicates the dominant connections that are estimated from statistical differences while comparing both ASD and TD groups. For comparison, we also show strength in connections that emerged from static functional connectivity.

State	Total		CC-CC		CC-DM		DM-DM		SM-CC		VIS-DM		VIS-CC		SC-CC	
	ASD	TD	ASD	TD	ASD	TD	ASD	TD	ASD	TD	ASD	TD	ASD	TD	ASD	TD
State1	56	40	8	2	8	5	0	4	10	2	1	0	8	7	4	7
State2	100	87	17	9	20	10	0	10	4	7	6	3	6	7	9	4
State3	103	94	21	13	21	14	0	13	1	5	17	2	6	18	2	6
State4	70	68	14	7	12	6	0	8	3	6	10	4	6	14	3	5
sFNC	111	114	19	16	24	18	0	12	4	6	9	5	7	9	2	11

5.3 Implementation

ICNs extraction using ICA neuromark template, dFNC clustering analysis followed by hard clustering analysis and two sample *T-test* analysis has been done using Group ICA of fMRI Toolbox (GIFT) [60]. The toolbox scripts were implemented in matlab and can be downloaded from here [4]. All the sub-

[4]https://github.com/trendscenter/gift

ject IDs of the ABIDE-I samples used in this paper, and parameters required to extract ICNs, estimate dFNC, and perform the clustering analysis is made publicly available here[5] for reproducibility.

5.4 Results

5.4.1 Characteristics of Different Brain States

Four optimal dynamic states have been found using K-Means clustering on all windows (W) extracted from all subjects in both ASD and TD. As shown in Table 5.2, we can observe that 10% of the total windows were assigned to state 1 and characterized by strong positive intra-connectivity within VIS and SM networks and sparse connectivity within DM and CC networks. 20% of the windows were assigned to state 2 and characterized by strong positive intra-connectivity within VIS, DM networks and inter-connectivity between CC-DM networks. 19% of the windows were assigned to state 3 and characterized by strong positive intra-connectivity within VIS, SM, and inter-connectivity between CC-DM networks and negative inter-connectivity between AUD, SM, and CC networks. 51% of the windows were assigned to state 4 and characterized by positive intra-connectivity within SM, VIS, DM, and inter-connectivity between CC and DM. All the state characteristics mentioned above can be observed in figure 5.2. The group and state-specific distribution of windows and subjects can be seen in Table 5.2. It is to be noted that not all the states occurred in all the subjects. From Table 5.2, states 1 and 3 occurred less frequently among ASD subjects than in TD (state 1: ASD - 3.77%, TD - 6.15%; state 3: ASD - 8.16%, TD - 11.08%). State 4 occurred more frequently among ASD subjects (ASD - 28.15%, TD - 22.59%).

5.4.2 Dynamic Functional Connectivity Differences with two-sample T-test

For each state, Table 5.3, shows the number of stronger connections in ASD and TD groups among pairs of within or between IC networks that are statistically significant ($p < 0.05$, FDR correction). Overall, for state 1, we observed 56 stronger within- and between-network connections in ASD compared to TD subjects. Similarly, for other states, $2, 3$ and $\&4$, the number of connections emerged to be higher in ASD compared to TD. Looking at the number of stronger connections among pairs of IC networks, we observed that within the CC network and between CC-DM networks, most of them were stronger in ASD subjects, whereas, within DM network, connections were stronger among TD individuals in all four states. Figure 5.3 highlights the qualitative visualization of regional differences in the CC-DM network. Interestingly, the patterns were quite similarly observed for static functional connectivity in ASD and TD groups, as evident from figure 5.3 focusing on the CC-DM network. However, the statistical comparisons on the temporal characterization of dynamic states (clusters) between ASD and TD brought additional regional differences between IC networks that may be evident in static functional

[5]https://github.com/BCCL-IIITH/Dynamic-Functional-Connectivity-Analysis-of-ASD

connectivity or vice-versa. Overall, we observe 329 stronger intra- and inter-connectivity differences in ASD while using dynamic functional connectivity compared to 111 significant connections in ASD while using static functional connectivity.

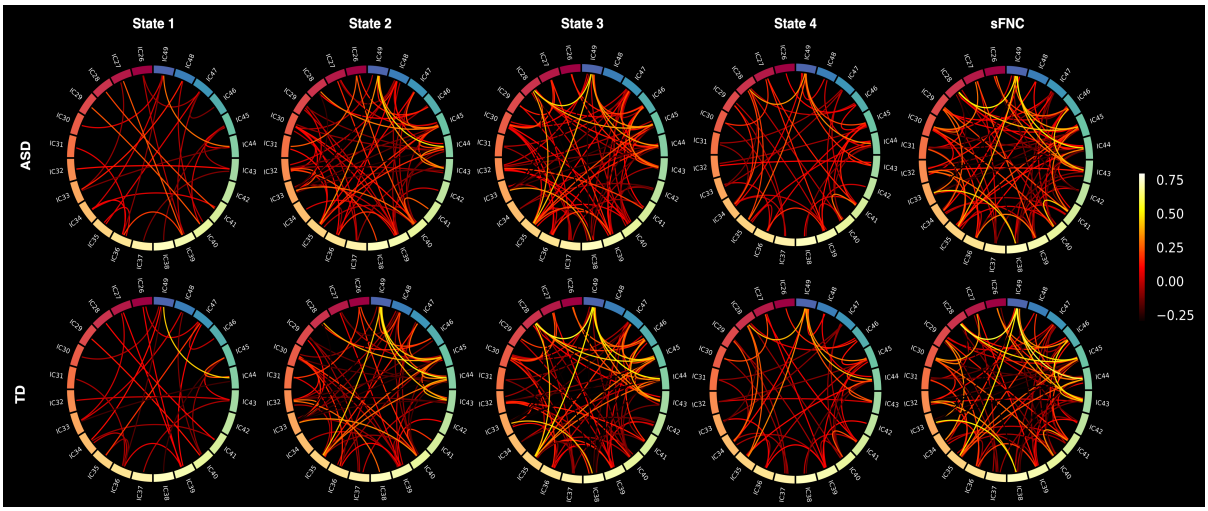

Figure 5.3 Connectograms of ASD vs TD and dFNC states vs sFNC. Activations within and between CC and DM networks that are statistically significant in both groups ASD and TD. The first row represents the significantly stronger activations among ASD subjects and the second row represents significantly stronger activations among TD subjects in all four dynamic brain states. The last column represents the significantly stronger activations found in both groups found using sFNC analysis. The names of each IC and its corresponding brain networks are part of Neuromark template release.

5.4.3 Temporal Properties Differences

From figure 5.4, it can be observed that fractional time spent among ASD subjects is significantly greater than TD subjects in state 4 ($p = 0.0005$) and significantly lower in state 1 ($p = 0.04$) and state 3 ($p = 0.019$). Significant differences were found in mean dwell time of state 4. ASD subjects exhibited significantly longer mean dwell time in state 4 compared to TD subjects ($p = 0.00023$), while there was a trend for longer mean dwell time in state 3 among the TD subjects ($p = 0.071$). No significant differences between groups have been found based on the number of transitions between states (ASD: $6.27 \pm 3.68; TD : 6.90 \pm 3.17; p = 0.071$, FDR correction).

5.4.4 Meta-state Properties Differences

All the meta-state properties are found to be significantly different in the groups ASD and TD ($p < 0.05$, FDR corrected). Table 5.4 displays the mean values of both groups, p-values, and t-values. From the results, we can clearly infer that dynamic fluidity among ASD subjects is less because the occurrence of meta-states and transition between states is significantly less in ASD patients compared

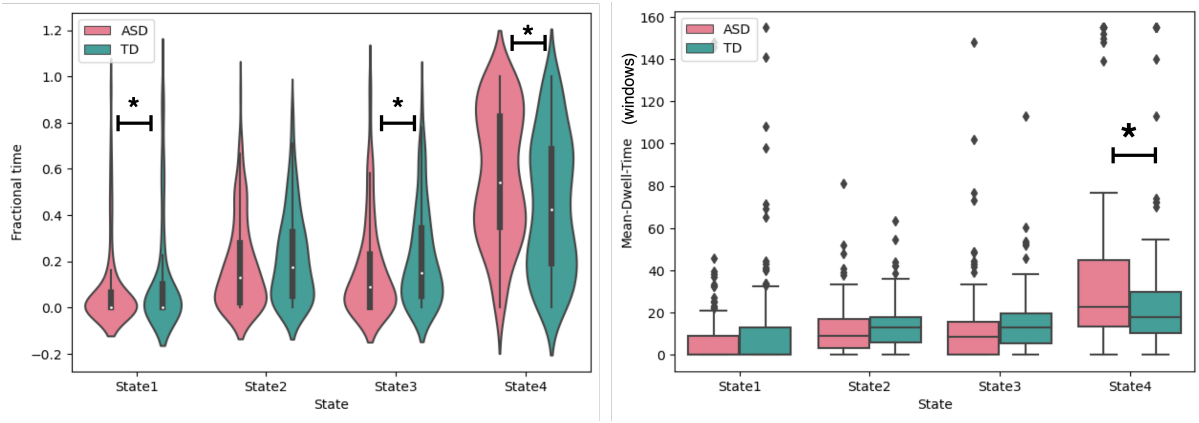

Figure 5.4 Fractional Time spent and Mean dwell time of ASD and TD groups per state. The left plot represents the fractional time spent in each state and the right plot represents the mean dwell time of each state specific to the group. * indicate that the distributions are significantly different ($p < 0.05$, FDR correction)

to TD individuals. It can be observed that individuals with ASD tend to stay within a smaller region of the state space compared to TD as determined by measuring the maximum $L1$ distance between the occupied meta-states. Total $L1$ distance traveled by ASD subjects between consecutive meta-states through state space is significantly less than TD individuals. Finally, we can conclude that there exists significant differences between ASD and TD in terms of meta-state dynamic properties.

Table 5.4 Meta state properties differences between ASD and TD

	Num States	Change Between States	State Span	Total distance
T-value	-3.5761	-2.7889	-2.5348	-3.0081
P-value	0.0004	0.0056	0.0117	0.0028
ASD Group Mean	16.8777	31.4043	6.7074	34.9628
TD Group Mean	19.7590	34.0513	7.2667	38.6667

5.5 Discussion

We studied dFNC patterns for ASD and TD on a large number of subjects sampled from the ABIDE-I consortium. We extracted dFNC patterns and applied hard clustering analysis on those patterns to find the optimal brain states or discrete FC patterns, which helps us to understand the network-level differences in the fundamental brain organization of ASD with respect to TD subjects. We performed a two-sample *T-test* and found many activations between ICNs within and between networks that are

significantly different among the groups ASD and TD per state ($p < 0.05$, FDR correction). The results of the two-sample t-test conducted on the temporal and meta-state properties suggest that these features can serve as reliable indicators for diagnosing Autism Spectrum Disorder (ASD). We observed hyper-connectivity among ASD subjects within the CC network and between CC-DM networks in all four states. Among TD subjects, hyper-connectivity has been found within DM network in all four states.

Regions that contribute to the hyper-connectivity in the CC network among ASD subjects in most of the dynamic states are found to be the Hippocampus, Middle cingulate cortex, Left inferior parietal lobule, Inferior frontal gyrus, and hypo-connectivity in the DM network are Precuneus, Posterior cingulate cortex, and Anterior cingulate cortex. The region hippocampus controls memory encoding, learning, and memory consolidation [2], the inferior frontal gyrus controls language comprehension and production [40], and the middle cingulate cortex is in charge of cognitive processing, particularly decision-making [42]. There are numerous neurological processes that the inferior parietal lobe (IPL) is assumed to be engaged in, such as spatial attention, multimodal sensory integration, and oculomotor control [18]. The precuneus is a part of the default mode network that performs a wide range of complicated tasks, such as information integration related to the perception of the environment, cue response, mental imagery techniques, and retrieval of episodic memory [14]. The anterior and posterior cingulate cortex are responsible for social cognition [68, 46]. All of these functions are found to be identified in patients with ASD [67, 47].

Earlier research [3, 83] has investigated the role of altered FC of DM sub-networks among high-functioning ASD subjects based on rs-fMRI scans. These studies hypothesized that the FC between the precuneus and the medial prefrontal cortex/anterior cingulate cortex, and DM core areas were weaker in ASD patients. The findings are consistent with the concept that the fundamental deficiencies in ASD are a result of the under-connectivity of DM sub-networks [72]. Harlalka et al. [37] observed that patients with ASD demonstrated higher dFNC between the attentiveness network (AN) and default mode network (DM) than TD individuals.

5.6 Summary

We investigated significant differences between ASD and TD groups in terms of the strength of features estimated from sFNC and dFNC between distinct pairs of regions, dynamic temporal properties such as fractional windows in each state, the mean dwell time of each state and the number of transitions between each pair of dynamic states using statistical two-sample *T-test* followed by meta-state analysis. The upcoming chapter will summarize and conclude the two studies presented in this thesis and suggest potential scope for future research.

Chapter 6

Conclusions and Future Scope

6.1 Summary & Conclusions

In this thesis, we proposed a two-hidden layer neural network, receiving autoencoder pre-training based features for the classification of ASD from TD individuals followed by applying attribution methods such as Integrated Gradients (IG) and DeepLIFT to find out the regions that contributed most to the classification task. This study identifies Left Lingual Gyrus, Right Insula, Right Cuneus, Right Middle Frontal Gyrus, Left Superior Temporal Gyrus associated with ASD. We experimented with various brain parcellation schemes and preprocessing pipelines to reveal the impact of model performance on the classifiers. This study found that the combination of CPAC preprocessing pipeline and CC200 parcellation yielded the best outcomes. The results might be useful in defining a common benchmark dataset and specifications for a fair and viable comparison of methods proposed for ASD classification in the future. The classification of ASD using ABIDE-1 dataset faces a significant obstacle due to the presence of rs-fMRI scans from 17 geographically distinct sites. A leave-one-site-out cross-validation strategy is employed to assess the proposed method's effectiveness on new sites. This strategy involves reserving the data from each site exclusively for testing and training the model on data from the other sites. By doing so, the strategy provides insight into the model's ability to perform on new, unseen sites. However, this proposed method relies on sFNC features for the classification of ASD, that are not capable of capturing the dynamic properties of ASD.

Therefore, in the next part of the thesis, we studied the dFNC patterns from seven different brain networks between ASD and TD for characterization of ASD. The hard clustering analysis revealed four optimal brain states or dFNC patterns. Our systematic comparisons across four optimal brain states using a statistical two-sample t-test revealed hyper-connectivity intra- and inter-level Cognitive control and Default mode networks in ASD. Examining the temporal properties such as fractional time spent and mean dwell time found that there were significant differences between ASD and TD subjects in terms of the amount of time spent in different states. ASD subjects spent more time in state 4, but less time in states 1 and 3, compared to TD subjects. The mean dwell time in state 4 was also significantly longer in ASD subjects. All the meta-state properties were significantly different in the two groups,

indicating that dynamic fluidity is less in ASD patients. Individuals with ASD tend to stay within a smaller region of the state space. Although the derived dFNC networks are consistent with sFNC networks, our study provides additional insights into the brain network-level differences. This can provide a greater understanding of disease biomarkers.

6.2 Limitations & Future scope

In future, efforts can focus on integrating phenotypic and demographic information with dFNC to attempt superior classification performance and possible biomarkers for the diagnosis of ASD. The shortcoming of the dFNC study presented in this thesis includes considering only male participants. This is to avoid potentially biased conclusions arising from our sample selection criteria. As part of future work, it would be valuable to assess the robustness of the proposed methods using the ABIDE-II dataset. Additionally, exploring different values of dFNC parameters could provide insights into their impact on the analysis and further enhance our understanding of the findings. In future, we can extend this selection criterion to both males and females as gender influences in ASD subjects are proven to be different in terms of functional organization [71]. From a machine learning perspective, the connectivity strengths that are shown to be statistically significant (could act as a prior feature selection criteria) to classify ASD and TD. This may reveal the usefulness of dFNC over sFNC – sFNC considers the information from multiple brain states but the representation corresponds to time-averaged BOLD signal recorded over the resting state scanning period. Within this space, another possible extension could be correlating these hyper-connectivity findings in brain networks to clinical scores (such as IQ) and severity scores (such as ADOS and ADI-R) to better understand sub-types such as Autistic, Aspergers, and Pervasive developmental disorder. The major challenge encountered in classifying ASD subtypes using ML methods is the limited availability of labeled data for each subtype. This makes it difficult to train accurate models for each subtype separately. Another challenge is the data heterogeneity as individuals with ASD can present with a wide range of symptoms and severity levels. Additionally, the symptoms may overlap between different subtypes, making it difficult to distinguish them based on the available features. Future studies could focus on developing techniques to address problems such as limited data and class imbalance, potentially leading to better classification outcomes and more useful clinical biomarkers for distinguishing ASD from TD.

Related Publications

Conference Publication

Prasad, P.K.C., Khare, Y., Dadi, K., Vinod, P.K. and Surampudi, B.R., 2022, July. Deep Learning Approach for Classification and Interpretation of Autism Spectrum Disorder. In 2022 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.

Conference Publication Accepted

Prasad, P.K.C., Dadi, K., and Surampudi, B.R., 2023, July. Dynamic functional connectivity analysis in individuals with Autism Spectrum Disorder. In 2023 International Joint Conference on Neural Networks (IJCNN). IEEE

Bibliography

[1] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun. Tracking whole-brain connectivity dynamics in the resting state. *Cerebral cortex*, 24(3):663–676, 2014.

[2] K. S. Anand and V. Dhikav. Hippocampus in health and disease: An overview. *Annals of Indian Academy of Neurology*, 15(4):239, 2012.

[3] M. Assaf, K. Jagannathan, V. D. Calhoun, L. Miller, M. C. Stevens, R. Sahl, J. G. O'Boyle, R. T. Schultz, and G. D. Pearlson. Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients. *Neuroimage*, 53(1):247–256, 2010.

[4] P. Baldi. Autoencoders, unsupervised learning, and deep architectures. In *Proceedings of ICML workshop on unsupervised and transfer learning*, pages 37–49. JMLR Workshop and Conference Proceedings, 2012.

[5] J. A. Bastiaansen, H. Meffert, S. Hein, et al. Diagnosing autism spectrum disorders in adults: the use of autism diagnostic observation schedule (ados) module 4. *Journal of autism and developmental disorders*, 41:1256–1266, 2011.

[6] C. C. Bell. Dsm-iv: diagnostic and statistical manual of mental disorders. *Jama*, 272(10):828–829, 1994.

[7] P. Bellec, F. Carbonnell, V. Perlbarg, C. Lepage, et al. A neuroimaging analyses kit for matlab and octave. In *Human Brain Mapping (HBM), Quebec City, Canada, June 26-30, 2011*, pages 1–5. Organization on Human Brain Mapping, 2011.

[8] R. F. Betzel, M. Fukushima, Y. He, X.-N. Zuo, and O. Sporns. Dynamic fluctuations coincide with periods of high and low modularity in resting-state functional brain networks. *NeuroImage*, 127:287–297, 2016.

[9] E. D. Bigler, S. Mortensen, E. S. Neeley, S. Ozonoff, L. Krasny, M. Johnson, J. Lu, S. L. Provencal, W. McMahon, and J. E. Lainhart. Superior temporal gyrus, language function, and autism. *Developmental neuropsychology*, 31(2):217–238, 2007.

[10] B. Biswal, F. Zerrin Yetkin, V. M. Haughton, and J. S. Hyde. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic resonance in medicine*, 34(4):537–541, 1995.

[11] K. S. Button, J. P. Ioannidis, C. Mokrysz, B. A. Nosek, J. Flint, E. S. Robinson, and M. R. Munafò. Power failure: why small sample size undermines the reliability of neuroscience. *Nature reviews neuroscience*, 14(5):365–376, 2013.

[12] E. Canario, D. Chen, and B. Biswal. A review of resting-state fmri and its use to examine psychiatric disorders. *Psychoradiology*, 1(1):42–53, 2021.

[13] J. Carp. The secret lives of experiments: methods reporting in the fmri literature. *Neuroimage*, 63(1):289–300, 2012.

[14] A. E. Cavanna and M. R. Trimble. The precuneus: a review of its functional anatomy and behavioural correlates. *Brain*, 129(3):564–583, 2006.

[15] V. A. Chandran, C. Pliatsikas, J. Neufeld, G. O'Connell, A. Haffey, V. DeLuca, and B. Chakrabarti. Brain structural correlates of autistic traits across the diagnostic divide: A grey matter and white matter microstructure study. *NeuroImage: Clinical*, 32:102897, 2021.

[16] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer. Smote: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, 16:321–357, 2002.

[17] Q. Chen, J. Lu, X. Zhang, Y. Sun, W. Chen, X. Li, W. Zhang, Z. Qing, and B. Zhang. Alterations in dynamic functional connectivity in individuals with subjective cognitive decline. *Frontiers in aging neuroscience*, 13:646017, 2021.

[18] D. M. Clower, R. A. West, J. C. Lynch, and P. L. Strick. The inferior parietal lobule is the target of output from the superior colliculus, hippocampus, and cerebellum. *Journal of Neuroscience*, 21(16):6283–6291, 2001.

[19] C. Craddock, S. Sikka, B. Cheung, R. Khanuja, S. S. Ghosh, C. Yan, Q. Li, D. Lurie, J. Vogelstein, R. Burns, et al. Towards automated analysis of connectomes: The configurable pipeline for the analysis of connectomes (c-pac). *Front Neuroinform*, 42:10–3389, 2013.

[20] R. C. Craddock, G. A. James, P. E. Holtzheimer III, X. P. Hu, and H. S. Mayberg. A whole brain fmri atlas generated via spatially constrained spectral clustering. *Human brain mapping*, 33(8):1914–1928, 2012.

[21] P. S. Dammu and R. S. Bapi. Employing temporal properties of brain activity for classifying autism using machine learning. In *Pattern Recognition and Machine Intelligence: 8th International Conference, PReMI 2019, Tezpur, India, December 17-20, 2019, Proceedings, Part II*, pages 193–200. Springer, 2019.

[22] A. Di Martino, C.-G. Yan, Q. Li, E. Denio, F. X. Castellanos, K. Alaerts, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Molecular psychiatry*, 19(6):659–667, 2014.

[23] N. U. Dosenbach, B. Nardos, A. L. Cohen, D. A. Fair, J. D. Power, J. A. Church, S. M. Nelson, G. S. Wig, A. C. Vogel, C. N. Lessov-Schlaggar, et al. Prediction of individual brain maturity using fmri. *Science*, 329(5997):1358–1361, 2010.

[24] Y. Du, Z. Fu, and V. D. Calhoun. Classification and prediction of brain disorders using functional connectivity: promising but challenging. *Frontiers in neuroscience*, 12:525, 2018.

[25] Y. Du, Z. Fu, J. Sui, S. Gao, Y. Xing, D. Lin, M. Salman, A. Abrol, M. A. Rahaman, J. Chen, et al. Neuromark: An automated and adaptive ica based pipeline to identify reproducible fmri markers of brain disorders. *NeuroImage: Clinical*, 28:102375, 2020.

[26] S. B. Eickhoff, B. T. Yeo, and S. Genon. Imaging-based parcellations of the human brain. *Nature Reviews Neuroscience*, 19(11):672–686, 2018.

[27] T. Eilam-Stock, T. Wu, A. Spagna, L. J. Egan, and J. Fan. Neuroanatomical alterations in high-functioning adults with autism spectrum disorder. *Frontiers in neuroscience*, 10:237, 2016.

[28] A. El-Gazzar, M. Quaak, et al. A hybrid 3dcnn and 3dc-lstm based model for 4d spatio-temporal fmri data: an abide autism classification study. In *Machine Learning in Clinical Neuroimaging: Second International Workshop, MLCN 2019 China, October 13 and 17, 2019, Proceedings 2*, pages 95–102. Springer, 2019.

[29] J. H. Elder, C. M. Kreider, S. N. Brasher, and M. Ansell. Clinical impact of early diagnosis of autism on the prognosis and parent–child relationships. *Psychology research and behavior management*, 2017.

[30] T. Eslami, V. Mirjalili, A. Fong, A. R. Laird, and F. Saeed. Asd-diagnet: a hybrid learning approach for detection of autism spectrum disorder using fmri data. *Frontiers in neuroinformatics*, 13:70, 2019.

[31] E. Fiorenzato, A. P. Strafella, J. Kim, R. Schifano, L. Weis, A. Antonini, and R. Biundo. Dynamic functional connectivity changes associated with dementia in parkinsons disease. *Brain*, 142(9):2860–2872, 2019.

[32] J. Friedman, T. Hastie, and R. Tibshirani. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*, 9(3):432–441, 2008.

[33] A. Frigaux, R. Evrard, and J. Lighezzolo-Alnot. Adi-r and ados and the differential diagnosis of autism spectrum disorders: Interests, limits and openings. *L'encephale*, 45(5):441–448, 2019.

[34] Z. Fu, J. Sui, J. A. Turner, Y. Du, M. Assaf, G. D. Pearlson, and V. D. Calhoun. Dynamic functional network reconfiguration underlying the pathophysiology of schizophrenia and autism spectrum disorder. *Human brain mapping*, 42(1):80–94, 2021.

[35] Z. Fu, Y. Tu, X. Di, Y. Du, J. Sui, B. B. Biswal, Z. Zhang, N. de Lacy, and V. D. Calhoun. Transient increased thalamic-sensory connectivity and decreased whole-brain dynamism in autism. *Neuroimage*, 190:191–204, 2019.

[36] R. O. Gilmore, M. T. Diaz, B. A. Wyble, and T. Yarkoni. Progress toward openness, transparency, and reproducibility in cognitive neuroscience. *Annals of the New York Academy of Sciences*, 1396(1):5–18, 2017.

[37] V. Harlalka, R. S. Bapi, P. K. Vinod, and D. Roy. Age, disease, and their interaction effects on intrinsic connectivity of children and adolescents in autism spectrum disorder using functional connectomics. *Brain Connectivity*, 8(7):407–419, 2018.

[38] A. S. Heinsfeld, A. R. Franco, R. C. Craddock, A. Buchweitz, and F. Meneguzzi. Identification of autism spectrum disorder using deep learning and the abide dataset. *NeuroImage: Clinical*, 17:16–23, 2018.

[39] Z. Hongliang and J. Yang. Dynamic weighted functional connectivity analysis for autism spectrum disorders classification. *Dual Diagn*, 2019.

[40] B. Ishkhanyan, V. Michel Lange, K. Boye, J. Mogensen, et al. Anterior and posterior left inferior frontal gyrus contribute to the implementation of grammatical determiners during language production. *Frontiers in Psychology*, 11:685, 2020.

[41] R. Jellett and J. Muggleton. Implications of applying clinically significant impairment to autism assessment: Commentary on six problems encountered in clinical practice. *Journal of autism and developmental disorders*, 52(3):1412–1421, 2022.

[42] F. R. Jumah and R. H. Dossani. Neuroanatomy, cingulate cortex. 2019.

[43] Y. Kaneoke, T. Donishi, J. Iwatani, S. Ukai, K. Shinosaki, and M. Terada. Variance and autocorrelation of the spontaneous slow brain activity. *PLoS One*, 7(5):e38131, 2012.

[44] R. E. Kelly Jr and M. J. Hoptman. Replicability in brain imaging, 2022.

[45] N. Kokhlikyan, V. Miglani, M. Martin, E. Wang, B. Alsallakh, J. Reynolds, A. Melnikov, N. Kliushkina, C. Araya, S. Yan, et al. Captum: A unified and generic model interpretability library for pytorch. *arXiv preprint arXiv:2009.07896*, 2020.

[46] R. Leech and D. J. Sharp. The role of the posterior cingulate cortex in cognition and disease. *Brain*, 137(1):12–32, 2014.

[47] I. P. Levin, G. J. Gaeth, M. Foley-Nicpon, V. Yegorova, C. Cederberg, and H. Yan. Extending decision making competence to special populations: a pilot study of persons on the autism spectrum. *Frontiers in Psychology*, 6:539, 2015.

[48] Y. Li, Y. Zhu, B. A. Nguchu, Y. Wang, H. Wang, B. Qiu, and X. Wang. Dynamic functional connectivity reveals abnormal variability and hyper-connected pattern in autism spectrum disorder. *Autism Research*, 13(2):230–243, 2020.

[49] C. Lord, T. S. Brugha, T. Charman, J. Cusack, G. Dumas, T. Frazier, et al. Autism spectrum disorder. *Nature Reviews Disease Primers*, 6(1):1–23, 2020.

[50] C. Lord, M. Elsabbagh, G. Baird, and J. Veenstra-Vanderweele. Autism spectrum disorder. *The Lancet*, 392(10146):508–520, 2018.

[51] L. Ma, T. Yuan, W. Li, L. Guo, D. Zhu, Z. Wang, Z. Liu, K. Xue, Y. Wang, J. Liu, et al. Dynamic functional connectivity alterations and their associated gene expression pattern in autism spectrum disorders. *Frontiers in Neuroscience*, 15:1837, 2022.

[52] R. L. Miller, M. Yaesoubi, J. A. Turner, D. Mathalon, A. Preda, G. Pearlson, T. Adali, and V. D. Calhoun. Higher dimensional meta-state analysis reveals reduced resting fmri connectivity dynamism in schizophrenia patients. *PloS one*, 11(3):e0149849, 2016.

[53] S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker, and D. Rueckert. Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimers disease. *Medical image analysis*, 48:117–130, 2018.

[54] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32:8026–8037, 2019.

[55] M. A. Patriquin, T. DeRamus, L. E. Libero, A. Laird, and R. K. Kana. Neuroanatomical and neurofunctional markers of social cognition in autism spectrum disorder. *Human brain mapping*, 37(11):3957–3978, 2016.

[56] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, et al. Scikit-learn: Machine learning in python. *Journal of machine Learning research*, 12:2825–2830, 2011.

[57] R. A. Poldrack, C. I. Baker, J. Durnez, K. J. Gorgolewski, P. M. Matthews, M. R. Munafò, T. E. Nichols, J.-B. Poline, E. Vul, and T. Yarkoni. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nature reviews neuroscience*, 18(2):115–126, 2017.

[58] P. K. C. Prasad, Y. Khare, K. Dadi, P. Vinod, and B. R. Surampudi. Deep learning approach for classification and interpretation of autism spectrum disorder. In *2022 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2022.

[59] L. Rabany, S. Brocke, V. D. Calhoun, B. Pittman, S. Corbera, B. E. Wexler, M. D. Bell, K. Pelphrey, G. D. Pearlson, and M. Assaf. Dynamic functional connectivity in schizophrenia and autism spectrum disorder: Convergence, divergence and classification. *NeuroImage: Clinical*, 24:101966, 2019.

[60] S. Rachakonda, E. Egolf, N. Correa, and V. Calhoun. Group ica of fmri toolbox (gift) manual. *Dostupnez [cit 2011-11-5]*, 2007.

[61] M. Rutter, A. Le Couteur, C. Lord, et al. Autism diagnostic interview-revised. *Los Angeles, CA: Western Psychological Services*, 29(2003):30, 2003.

[62] J. Schumacher, L. R. Peraza, M. Firbank, A. J. Thomas, M. Kaiser, P. Gallagher, J. T. O'Brien, A. M. Blamire, and J.-P. Taylor. Dynamic functional connectivity changes in dementia with lewy bodies and alzheimer's disease. *NeuroImage: Clinical*, 22:101812, 2019.

[63] E. Semel, E. Wiig, and W. Secord. Clinical evaluation of language fundamentalsthird edition (celf–3) psychological corporation. *San Antonio, TX*, 1995.

[64] S. Shakil, C.-H. Lee, and S. D. Keilholz. Evaluation of sliding window correlation performance for characterizing dynamic functional connectivity and brain states. *Neuroimage*, 133:111–128, 2016.

[65] Z. Sherkatghanad, M. Akhondzadeh, S. Salari, M. Zomorodi-Moghadam, M. Abdar, U. R. Acharya, et al. Automated detection of autism spectrum disorder using a convolutional neural network. *Frontiers in neuroscience*, 13:1325, 2020.

[66] A. Shrikumar, P. Greenside, and A. Kundaje. Learning important features through propagating activation differences. In *International Conference on Machine Learning*, pages 3145–3153. PMLR, 2017.

[67] M. Solomon, J. B. McCauley, A.-M. Iosif, C. S. Carter, and J. D. Ragland. Cognitive control and episodic memory in adolescents with autism spectrum disorders. *Neuropsychologia*, 89:31–41, 2016.

[68] F. L. Stevens, R. A. Hurley, and K. H. Taber. Anterior cingulate cortex: unique role in cognition and emotion. *The Journal of neuropsychiatry and clinical neurosciences*, 23(2):121–125, 2011.

[69] P. Sturmfels, S. Lundberg, and S.-I. Lee. Visualizing the impact of feature attribution baselines. *Distill*, 5(1):e22, 2020.

[70] M. Sundararajan, A. Taly, and Q. Yan. Axiomatic attribution for deep networks. In *International Conference on Machine Learning*, pages 3319–3328. PMLR, 2017.

[71] K. Supekar, C. de Los Angeles, S. Ryali, K. Cao, T. Ma, and V. Menon. Deep learning identifies robust gender differences in functional brain organization and their dissociable links to clinical symptoms in autism. *The British Journal of Psychiatry*, 220(4):202–209, 2022.

[72] K. Supekar, S. Ryali, R. Yuan, D. Kumar, C. de Los Angeles, and V. Menon. Robust, generalizable, and interpretable artificial intelligence–derived brain fingerprints of autism and social communication symptom severity. *Biological Psychiatry*, 92(8):643–653, 2022.

[73] R. M. Thomas, S. Gallo, L. Cerliani, P. Zhutovsky, A. El-Gazzar, and G. Van Wingen. Classifying autism spectrum disorder using the temporal statistics of resting-state functional mri data with 3d convolutional neural networks. *Frontiers in psychiatry*, 11:440, 2020.

[74] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1):267–288, 1996.

[75] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 15(1):273–289, 2002.

[76] G. Varoquaux and V. Cheplygina. Machine learning for medical imaging: methodological failures and recommendations for the future. *NPJ digital medicine*, 5(1):48, 2022.

[77] Y. Wang, J. Wang, F.-X. Wu, R. Hayrat, and J. Liu. Aimafe: autism spectrum disorder identification with multi-atlas deep feature representation and ensemble learning. *Journal of Neuroscience Methods*, page 108840, 2020.

[78] Z. Wang, Y. Li, A. R. Childress, and J. A. Detre. Brain entropy mapping using fmri. *PloS one*, 9(3):e89948, 2014.

[79] M. Xia, J. Wang, and Y. He. Brainnet viewer: a network visualization tool for human brain connectomics. *PloS one*, 8(7):e68910, 2013.

[80] T. Xu, Z. Yang, L. Jiang, X.-X. Xing, and X.-N. Zuo. A connectome computation system for discovery science of brain. *Science Bulletin*, 60(1):86–95, 2015.

[81] T. Yamada, T. Itahashi, M. Nakamura, H. Watanabe, M. Kuroda, H. Ohta, C. Kanai, N. Kato, and R.-i. Hashimoto. Altered functional organization within the insular cortex in adult males with high-functioning autism spectrum disorder: evidence from connectivity-based parcellation. *Molecular autism*, 7(1):1–15, 2016.

[82] C. Yan and Y. Zang. Dparsf: a matlab toolbox for" pipeline" data analysis of resting-state fmri. *Frontiers in systems neuroscience*, 4:13, 2010.

[83] B. Yang, M. Wang, W. Zhou, X. Wang, S. Chen, M. N. Potenza, L.-x. Yuan, and G.-H. Dong. Disrupted network integration and segregation involving the default mode network in autism spectrum disorder. *Journal of Affective Disorders*, 323:309–319, 2023.

[84] W. Yin, S. Mostafa, and F.-X. Wu. Diagnosis of autism spectrum disorder based on functional brain networks with deep learning. *Journal of Computational Biology*, 28(2):146–165, 2021.

[85] Y. Zang, T. Jiang, Y. Lu, Y. He, and L. Tian. Regional homogeneity approach to fmri data analysis. *Neuroimage*, 22(1):394–400, 2004.

[86] F. Zhao, Z. Chen, I. Rekik, S.-W. Lee, and D. Shen. Diagnosis of autism spectrum disorder using central-moment features from low-and high-order dynamic resting-state functional connectivity networks. *Frontiers in neuroscience*, 14:258, 2020.

[87] Q.-H. Zou, C.-Z. Zhu, Y. Yang, X.-N. Zuo, X.-Y. Long, Q.-J. Cao, Y.-F. Wang, and Y.-F. Zang. An improved approach to detection of amplitude of low-frequency fluctuation (alff) for resting-state fmri: fractional alff. *Journal of neuroscience methods*, 172(1):137–141, 2008.

[88] X.-N. Zuo, A. Di Martino, C. Kelly, Z. E. Shehzad, D. G. Gee, D. F. Klein, F. X. Castellanos, B. B. Biswal, and M. P. Milham. The oscillating brain: complex and reliable. *Neuroimage*, 49(2):1432–1445, 2010.

[89] X.-N. Zuo, R. Ehmke, M. Mennes, D. Imperati, F. X. Castellanos, O. Sporns, and M. P. Milham. Network centrality in the human functional connectome. *Cerebral cortex*, 22(8):1862–1875, 2012.

Chapter 7

Appendix

Table 7.1: NeuroMark Template ICN labels

Network	ICN	Index	X	Y	Z
Subcortical network (SC)	Caudate	IC 1	6.5	10.5	5.5
	Subthalamus/hypothalamus	IC 2	-2.5	-13.5	-1.5
	Putamen	IC 3	-26.5	1.5	-0.5
	Caudate	IC 4	21.5	10.5	-3.5
	Thalamus	IC 5	-12.5	-18.5	11.5
Auditory network (AUD)	Superior temporal gyrus ([STG])	IC 6	62.5	-22.5	7.5
	Middle temporal gyrus ([MTG])	IC 7	-42.5	-6.5	10.5
Sensorimotor network (SM)	Postcentral gyrus [PoCG]	IC 8	56.5	-4.5	28.5
	Left postcentral gyrus [L PoCG]	IC 9	-38.5	-22.5	56.5
	Paracentral lobule [ParaCL]	IC 10	0.5	-22.5	65.5
	Right postcentral gyrus [R PoCG]	IC 11	38.5	-19.5	55.5
	Superior parietal lobule [SPL]	IC 12	-18.5	-43.5	65.5
	Paracentral lobule [ParaCL]	IC 13	-18.5	-9.5	56.5
	Precentral gyrus [PreCG]	IC 14	-42.5	-7.5	46.5
	Superior parietal lobule [SPL]	IC 15	20.5	-63.5	58.5
	Postcentral gyrus [PoCG]	IC 16	-47.5	-27.5	43.5
Visual network (VIS)	Calcarine gyrus [CalcarineG]	IC 17	-12.5	-66.5	8.5
	Middle occipital gyrus [MOG]	IC 18	-23.5	-93.5	-0.5
	Middle temporal gyrus [MTG]	IC 19	48.5	-60.5	10.5
	Cuneus	IC 20	15.5	-91.5	22.5
	Right middle occipital gyrus [R MOG]	IC 21	38.5	-73.5	6.5
	Fusiform gyrus	IC 22	29.5	-42.5	-12.5
	Inferior occipital gyrus [IOG]	IC 23	-36.5	-76.5	-4.5
	Lingual gyrus [LingualG]	IC 24	-8.5	-81.5	-4.5

	Middle temporal gyrus [MTG]	IC 25	-44.5	-57.5	-7.5
Cognitive-control network (CC)	Inferior parietal lobule [IPL]	IC 26	45.5	-61.5	43.5
	Insula	IC 27	-30.5	22.5	-3.5
	Superior medial frontal gyrus [SMFG]	IC 28	-0.5	50.5	29.5
	Inferior frontal gyrus [IFG]	IC 29	-48.5	34.5	-0.5
	Right inferior frontal gyrus [R IFG]	IC 30	53.5	22.5	13.5
	Middle frontal gyrus [MiFG]	IC 31	-41.5	19.5	26.5
	Inferior parietal lobule [IPL]	IC 32	-53.5	-49.5	43.5
	Left inferior parietal lobue [R IPL]	IC 33	44.5	-34.5	46.5
	Supplementary motor area [SMA]	IC 34	-6.5	13.5	64.5
	Superior frontal gyrus [SFG]	IC 35	-24.5	26.5	49.5
	Middle frontal gyrus [MiFG]	IC 36	30.5	41.5	28.5
	Hippocampus [HiPP]	IC 37	23.5	-9.5	-16.5
	Left inferior parietal lobule [L IPL]	IC 38	-47.5	5.5	22.5
	Middle cingulate cortex [MCC]	IC 39	-15.5	20.5	37.5
	Inferior frontal gyrus [IFG]	IC 40	39.5	44.5	-0.5
	Middle frontal gyrus [MiFG]	IC 41	-26.5	47.5	5.5
	Hippocampus [HiPP]	IC 42	-24.5	-36.5	1.5
Default-mode network (DM)	Precuneus	IC 43	-8.5	-66.5	35.5
	Precuneus	IC 44	-12.5	-54.5	14.5
	Anterior cingulate cortex [ACC]	IC 45	-2.5	35.5	2.5
	Posterior cingulate cortex [PCC]	IC 46	-5.5	-28.5	26.5
	Anterior cingulate cortex [ACC]	IC 47	-9.5	46.5	-10.5
	Precuneus	IC 48	-0.5	-48.5	49.5
	Posterior cingulate cortex [PCC]	IC 49	-2.5	54.5	31.5
Cerebellar network [CB]	Cerebellum [CB]	IC 50	-30.5	-54.5	-42.5
	Cerebellum [CB]	IC 51	-32.5	-79.5	-37.5
	Cerebellum [CB]	IC 52	20.5	-48.5	-40.5
	Cerebellum [CB]	IC 53	30.5	-63.5	-40.5

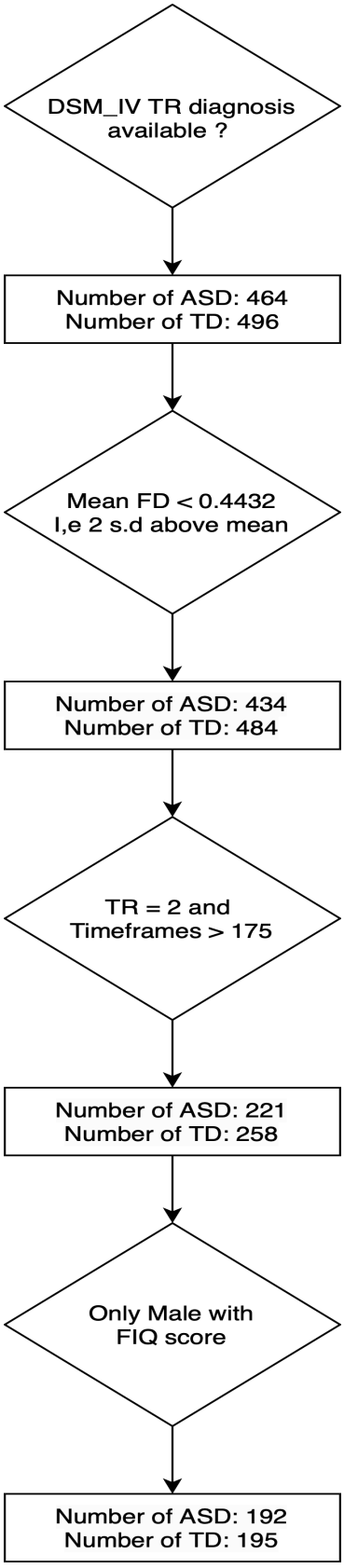

Figure 7.1 Participant Selection procedure for the dFNC analysis.