Studying functional connectome of the human brain

Gupta, Sukrit

2020

Gupta, S. (2020). Studying functional connectome of the human brain. Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/144674>

<https://doi.org/10.32657/10356/144674>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0).

Downloaded on 26 Apr 2023 16:30:07 SGT

STUDYING FUNCTIONAL CONNECTOME OF THE HUMAN BRAIN

Sukrit Gupta

School of Computer Science and Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy (PhD)

2020

Statement of Originality

I hereby certify that the work embodied in this thesis is the result of original research, is free of plagiarised materials, and has not been submitted for a higher degree to any other University or Institution.

6 July, 2020
.....
Date

.....
Sukrit Gupta

Supervisor Declaration Statement

I have reviewed the content and presentation style of this thesis and declare it is free of plagiarism and of sufficient grammatical clarity to be examined. To the best of my knowledge, the research and writing are those of the candidate except as acknowledged in the Author Attribution Statement. I confirm that the investigations were conducted in accord with the ethics policies and integrity standards of Nanyang Technological University and that the research data are presented honestly and without prejudice.

10 July 2020
.....
Date

.....
Jagath C. Rajapakse

Authorship Attribution Statement

This thesis contains material from seven papers published in the following peer-reviewed journals in which I am listed as an author.

1. Parts of chapter 2 are published as Sukrit Gupta, and Jagath C. Rajapakse. "Iterative consensus spectral clustering improves detection of subject and group level brain functional modules." *Scientific Reports* 10.1 (2020): 1-15.
2. Parts of chapter 3 are published as Sukrit Gupta, Jagath C. Rajapakse, and Roy E. Welsch. "Ambivert degree identifies crucial brain functional hubs and improves detection of Alzheimer's disease and Autism spectrum disorder." *NeuroImage: Clinical* 25 (2020): 102186, Jagath C. Rajapakse, Sukrit Gupta, and Xiuchao Sui. "Fitting networks models for functional brain connectivity." *IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)* and Sukrit Gupta, and Jagath C. Rajapakse. "Nodal degree distributions of resting-state functional brain modules." *IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*.
3. Parts of chapter 4 are submitted for publication to the journal *Human Brain Mapping* as Sukrit Gupta, Marcus Lim, and Jagath C. Rajapakse. "Decoding task-specific and task-general functional architectures of the brain".
4. Parts of chapter 5 are published as Sukrit Gupta, Yi Hao Chan, and Jagath C. Rajapakse "Decoding brain functional connectivity implicated in AD and MCI." *International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2019* and Sukrit Gupta, Yi Hao Chan, and Jagath Chandana Rajapakse. "Obtaining leaner deep neural networks for decoding brain functional connectome in a single shot.", *Neurocomputing* (Accepted, In Press).

The contributions of the co-authors are as follows:

- For 1, Prof. Jagath C. Rajapakse provided the initial research direction, and reviewed and revised the manuscript.

- For 2, Prof. Jagath C. Rajapakse provided the initial research direction and suggested the set of experiments. Both Prof. Jagath C. Rajapakse and Prof. Roy E. Welsch reviewed and revised the manuscript.
- For 3, Marcus Lim worked on writing the initial codes. I wrote the manuscript and rewrote all the associated codes. The experiments were performed solely by me. Prof. Jagath C. Rajapakse suggested the set of experiments to be performed, and reviewed and revised the manuscript.
- For 4, The codes were written together with Mr. Yi Hao Chan. We wrote the manuscript together and performed the associated experiments. I derived all the associated mathematical equations for the analytical frameworks proposed in the manuscripts. Prof. Jagath C. Rajapakse provided the initial research direction, and reviewed and revised the manuscript.
- I analyzed and interpreted the simulation and numerical results for all manuscripts. Assistance for manuscripts in 4) was provided by Yi Hao Chan.
- I prepared all the manuscript drafts. The manuscripts were revised together with Prof. Jagath C. Rajapakse and Yi Hao Chan and Prof. Roy E. Welsch.

6 July, 2020
 Date

 Sukrit Gupta

Abstract

Advances in neuroimaging techniques have made it possible to access intricate details on brain function and structure. Functional magnetic resonance imaging (MRI) gives us access to brain's functional activations and aids in understanding its functional architecture during rest and varied states. Functional MRI data can be modeled as a network, known as the brain functional connectome, such that the connectome's nodes represent brain regions and the edges between these nodes represent functional relationships between the regions. Studying the functional connectome has led to an improved understanding of cognitive and diseased brain states. In this thesis, we propose techniques to investigate the functional modules and hubs in the brain, decode task-related modulations, and encode diseased brain states.

We propose a novel Iterative Consensus Spectral Clustering (ICSC) algorithm that detects brain functional modules while considering the inter-individual variability in brain functional architecture and complete functional connectivity information. The ICSC algorithm can detect group level modularizations by considering functional MRI scans from multiple subjects and subject level modularizations by considering multiple functional MRI scans from the same subject. Using multiple real functional MRI scan datasets, we show that the ICSC algorithm gives biologically relevant modularizations. We also compare the performance of the ICSC algorithm with previous methods and show its application to detect diseases characterized by modular disruption.

We propose a novel measure, the ambivert degree, to derive a set of crucial brain regions that act as focal points of information processing. These regions are known as brain hubs and we show that ambivert degree identifies hubs that are more crucial for brain's information processing than the previously used measures. We use the brain hubs detected by ambivert degree as disease biomarkers and use them for diagnosing Alzheimer's disease and Autism spectrum disorder.

Besides studying the brain during resting state, it is important to study the brain’s functional architecture during varied brain states such as task performance or neurological diseases. We propose a novel approach to decode alterations in the brain’s functional architecture during task performance using model interpretability techniques. The proposed approach considers subject level data and models non-linear relationships between the functional connectivity features. Using the proposed approach we delineated the task-general architecture, that is common across all tasks, and task-specific architecture that captures changes particular to a task. We show that although there is pronounced inter-subject variability across subjects, a small common set of modulations can help us distinguish between rest and task states.

Besides studying the brain functional connectome during rest and task states, we also study computer aided detection of neurological disorders with deep learning models. We propose the Layerwise Elimination of Accessory Nodes (LEAN) and Correlation-based eLImination of InPuts (CLIP) algorithms to perform pruning of deep neural networks to obtain leaner and more generalizable classification models for neurological disease classification. We perform experiments with datasets for neurodegenerative, developmental and psychiatric disorders to show that LEAN and CLIP give models that have similar classification performance (as the base network) with just a small fraction of the trainable weights. We show that the disease biomarkers derived from LEAN are consistent with previous studies in the domain.

Acknowledgement

I would like to thank my supervisor, Prof. Jagath C. Rajapakse, for his time and support. Prof. Jagath's patience and efforts have helped me in orientating my attention in the right research direction. His expectation of high standard and rigor have been instrumental in improving the quality of my work.

I would like to express my gratitude to my colleagues in our research group at Biomedical Informatics Laboratory. I would like to thank Dr. Xiuchao Sui, Dr. Satya Prakash Singh, Dr. Rama Kaalia, Dr. L. L. Chamara Kasun, Dr. Parvin Kumar, Dr. Katharina Baum, Mr. Yi Hao Chan and Dr. L.C. Tranchevent for their comments and help along the way. My sincere gratitude to Dr. Rami Puzis and Mr. Avtar Mahajan for guiding me through my PhD years. I would also like to thank Mr. Tan Sing Yau and Ms. Wan Chee Heng for their technical assistance.

My gratitude for various collaborators that have helped me along the way. I would like to thank the teams at the Human Connectome Project, Alzheimer's Disease Neuroimaging Initiative, Autism Brain Imaging Data Exchange, ADHD-200 Consortium, and Dr. Jiang Qiu and Dr. Dongtao Wei from Southwest University for providing us with datasets used in this thesis.

Sincere thanks for my parents, Ashwani and Taruna, and my brother, Sushant, for their perpetual and kind interest in my work. A special thank you to my friends (and fellow candidates) Ravikiran, Garima and Saumya for standing with me and supporting me through these years. Finally, I would like to thank the School of Computer Science & Engineering, Nanyang Technological University for providing me with world-class research facilities.

For the suffering and fight put up by Aishwarya, Naani and Naanu.

Contents

Abstract	v
Acknowledgement	vii
List of Figures	xiv
List of Tables	xvii
Acronyms	xx
1 Introduction	1
1.1 From functional MRI to the brain functional connectome	1
1.2 Motivation	4
1.3 Major contributions	9
1.4 Organisation of the thesis	11
2 Detecting functional modules of the brain	14
2.1 Brain functional modules and related work	14
2.2 Previous methods to detect brain functional modules	17
2.2.1 Module detection at individual-level	17
2.2.1.1 Modularity maximization	18
2.2.1.2 Infomap	19
2.2.1.3 Normalized Cut Cost	19
2.2.2 Group-level consensus modules	20
2.3 Iterative consensus spectral clustering (ICSC)	21
2.3.1 Inter-subject variability of functional modules	23
2.3.2 Detecting subject-level functional modules using ICSC	25

2.3.3	Using ICSC modularizations to detect major depressive disorder	26
2.4	Results	27
2.4.1	Dataset	27
2.4.2	Brain Functional ROI	28
2.4.3	Group-level modules from multiple subject data	28
2.4.4	Group-level modules from a validation dataset	32
2.4.5	Inter-subject variability of functional modules	32
2.4.6	Subject-level modules from multiple scans	35
2.4.7	Comparison with previous modularization methods	35
2.4.8	Experiments with synthetic data	37
2.4.9	Using ICSC modularizations for disease detection	42
2.5	Discussion	43
2.5.1	Detecting brain functional modules by the ICSC	43
2.5.2	Inter-subject variability of functional modules	44
2.5.3	Using ICSC algorithm for deriving subject-level modularization	45
2.5.4	Comparison of ICSC algorithm with previous studies and methods . . .	46
2.6	Chapter Summary	48
3	Ambivert degree for detecting functional hubs of the brain	49
3.1	Existing work on brain's functional hubs	50
3.2	Detecting brain's functional hubs	52
3.2.1	Detecting brain functional modules	52
3.2.2	Thresholding functional connectivity	53
3.2.3	Degree distribution of the modules	53
3.2.4	Modular hubs	55
3.2.4.1	Intra-modular degree	55
3.2.4.2	Ambivert degree	55
3.2.4.3	Betweenness centrality	56
3.2.4.4	Detection of modular hubs	56
3.2.5	Connector hubs	57
3.2.5.1	Participation coefficient	57

3.2.5.2	Gateway coefficient	57
3.2.6	Network hubs	58
3.2.6.1	Inter-subject variability of hubs	58
3.2.6.2	Artificial lesioning: efficacy of hubs	59
3.3	Analyzing disruption of brain's functional hubs in AD and ASD and using hubs as disease biomarkers	59
3.3.1	AD dataset	60
3.3.2	ASD dataset	60
3.4	Results	61
3.4.1	HCP dataset	61
3.4.2	Brain functional modules	61
3.4.3	Best fit for nodal degree distributions	61
3.4.4	Artificial lesioning to identify the most crucial modular and network hubs	64
3.4.5	Disruption of network hubs in AD and ASD	66
3.4.6	Classification of patients and healthy controls using nodal hub scores	71
3.5	Discussion	74
3.5.1	Considering brain functional modules while detecting network hubs	74
3.5.2	Using ambivert degree to detect modular and network hubs	75
3.5.3	Disruption of network hubs in AD and ASD	76
3.5.4	Brain hubs as diagnostic features for AD and ASD	77
3.6	Chapter Summary	78
4	Decoding task-general and task-specific brain architectures	79
4.1	Task related brain alterations and related work	80
4.2	Proposed methods to study task related modulations in the brain	83
4.2.1	Feedforward DNN	83
4.2.2	Salience of input features	84
4.2.3	Decoding functional brain connectivity	85
4.2.4	Task-general brain architecture	86
4.2.5	Task-specific brain architecture	86
4.2.6	Salient interactions between functional modules	87

4.2.7	Inter-subject variability in the task-related functional architectures . . .	87
4.3	Experiments and results	88
4.3.1	Dataset	88
4.3.2	Building the classifiers	88
4.3.3	Decoding the task-general brain architecture	91
4.3.4	Decoding motor task-specific brain architecture	95
4.3.5	Decoding language task-specific brain architecture	97
4.3.6	Inter-subject variability in task-related architectures	101
4.4	Discussion	102
4.4.1	Classification of rest and task brain functional scans	102
4.4.2	Task-general architecture	104
4.4.3	Task-specific architecture	105
4.5	Chapter Summary	106
5	Optimizing brain encoders for detecting diseased brain states	107
5.1	Disease detection using functional connectivity	108
5.2	Methods to optimize brain encoding and decoding	110
5.2.1	Feedforward DNN encoder	110
5.2.2	Computing the saliency of nodes in network layers	110
5.2.3	Layerwise elimination of accessory nodes (LEAN)	111
5.2.4	Correlation-based elimination of inputs (CLIP)	112
5.2.4.1	Finding clusters of correlated features	113
5.2.4.2	Selecting subsets of correlated features	115
5.2.5	Combining LEAN and CLIP	116
5.2.6	Decoding the brain functional connectome associated with brain disease	116
5.3	Results	117
5.3.1	Datasets	117
5.3.1.1	ABIDE	117
5.3.1.2	ADNI	117
5.3.1.3	ADHD	118
5.3.1.4	MDD	119

5.3.2	Features for classification	119
5.3.3	Classification for full feature set	120
5.3.4	Classification with feature subsets from CLIP and LEAN	120
5.3.4.1	Reduction in overfitting	122
5.3.4.2	Comparison with other redundancy removal methods	123
5.3.4.3	Comparison of algorithm runtime	124
5.3.5	Salient brain features	125
5.4	Discussion	126
5.4.1	DNN models for full feature set	126
5.4.2	LEAN + CLIP: high accuracy with few parameters	132
5.4.3	Identified salient features are clinically relevant	133
5.5	Chapter Summary	134
6	Conclusion and Future Work	135
6.1	Future Work	137
6.1.1	Hierarchical modularization	137
6.1.2	Studying modular disruption during diseases	138
6.1.3	Mixing topological features for disease classification	138
6.1.4	Sub-task detection from task fMRI data	139
	Appendices	141
A	Appendix for Chapter 2	141
B	Appendix for Chapter 3	148
C	Appendix for Chapter 4	154
D	Appendix for Chapter 5	164
	References	167

List of Figures

1.1	From functional MRI to the brain functional connectome	5
1.2	Outline of topics covered in this thesis	6
2.1	Similarity between the group-level modularizations computed for 3 folds of the HCP dataset and the whole HCP dataset.	30
2.2	Group-level modules detected by the ICSC algorithm on the resting-state fMRI scans from the HCP dataset.	31
2.3	Inter-subject variability of functional modules	33
2.4	Comparison of inter-subject and intra-subject variability of brain ROI of different functions modules	34
2.5	The group consensus matrices obtained for different modular detection approaches.	37
2.6	Comparison of performance of different module detection approaches.	38
2.7	Performances for the ICSC algorithm on synthetic data with varying purity for nodes.	40
2.8	MCC of the ICSC algorithm on synthetic data with varying amounts of noise in the scans.	40
2.9	Performances for the ICSC algorithm on synthetic data with different number of individual-level scans.	41
3.1	The intuition behind ambivert degree.	56
3.2	Increase in modular path length after artificial lesioning of modular hubs detected by different hub measures.	65
3.3	Increase in network path length after artificial lesioning of network hubs detected by different hub measures.	66

3.4	The inter-subject variability of hubs detected by different hub measures and the participation coefficient at different percentiles.	67
3.5	The anatomical locations of ambivert brain hubs.	67
3.6	Different brain regions that are identified as top 10% network hubs, using ambivert degree and participation coefficient in cognitively normal (CN) and diseased subjects.	70
4.1	Accuracy and time taken for convergence for classification of rest vs all tasks, rest vs motor task and rest vs language task.	92
4.2	Decoded task-general brain architecture.	94
4.4	Decoded motor task-specific architecture.	97
4.5	Decoded language task-specific architecture.	99
4.6	The salient anatomical connections for the task-specific language architecture in the brain.	100
4.7	Inter-subject variability in task-related architectures	103
5.1	The intuition behind the CLIP algorithm.	113
5.2	Schematic of how the features are distributed in different clusters throughout the distribution of salience scores and how salience scores consider only features from a few clusters.	114
5.3	Selecting subsets of correlated features	115
5.4	Salient brain functional connectivity and functional regions for classification of Alzheimer's Disease	127
5.5	Salient brain functional connectivity and functional regions for classification of major depressive disorder	128
5.6	Salient brain functional connectivity and functional regions for classification of mild cognitive impairment	129
5.7	Salient brain functional connectivity and functional regions for classification of autism spectrum disorder	130
5.8	Salient brain functional connectivity and functional regions for classification of attention deficit hyperactivity disorder	131

A1	Modules detected by the ICSC algorithm on the resting-state fMRI scans from the validation dataset.	142
A2	Choosing L_{max} for ICSC	143
A3	The intra-subject purity of brain ROIs.	144
A4	Modular structure obtained with different module detection algorithms for different connectivity thresholds.	146
A5	Group-level modules detected by the Louvain algorithm.	147
B1	Increase in the path length after artificial lesioning for different network hubs. .	150
B2	The different hubs identified using ambivert degree and participation coefficient, and intra-modular degree and participation coefficient.	151
B3	The different hubs identified using ambivert degree and participation coefficient in cognitively normal (CN) and AD subjects for different percentiles.	152
B4	The different hubs identified using ambivert degree and participation coefficient in cognitively normal (CN) and AD subjects for different percentiles.	153
C2	Accuracy and time taken for convergence for rest vs task classification.	156
C3	Accuracy and time taken for convergence for rest vs task classification.	157
D1	Accuracy plots of SVM, FFN and CNN models applied for different diseases. .	166

List of Tables

2.1	MDD subject details, Hamilton Depression Rating Scale (HAM-D)	41
2.2	Classification accuracies for MDD and CN subjects using deep neural networks and module features as features. * refers to a statistically significant result in comparison to all the others.	43
3.1	p – value for paired t-test on Bayesian Information Criterion for fits of power law, exponential, log-normal and stretched exponential degree distributions for different modules.	62
3.2	Fitting parameters of stretched exponential for brain functional network and modules.	63
3.3	Disrupted brain regions in AD. MNI coordinates of the most disrupted regions (R: Right, L: Left) detected by ambivert degree. Hub measures are given as the sum of ambivert degree and participation coefficient.	69
3.4	Disrupted brain regions in ASD. MNI coordinates of the most disrupted regions (Right, L: Left) detected by ambivert degree. Hub measures are given as the sum of ambivert degree and participation coefficient.	69
3.5	Accuracies of classification of AD and NC using deep neural networks and nodal hub scores as features. * refers to a statistically significant result in comparison to all the others.	72
3.6	Classification of ASD and CN with deep neural networks with nodal hub measures and connectivity features. * refers to a statistically significant result in comparison to all the others.	73
4.1	Summary of HCP task and resting state data	89

4.2	Performance of different classification models on the HCP data with 5-fold cross-validation.	91
4.3	Parameters of the 4 layer FFN used for classification of rest vs all tasks and rest vs language task.	93
4.4	Performance of the FFN and the SVM (on the test set) for rest vs all task classification using different fractions of features. MAE denotes mean absolute error.	93
4.5	Significant task-general modular interactions.	95
4.6	Parameters of the 5 layer FFN used for classification of rest vs motor task . . .	96
4.7	Performance of the FFN and the SVM for rest vs motor task classification using different fraction of features. MAE is mean absolute error.	96
4.8	Significant modular interactions specific to the motor task.	98
4.9	Performance of the FFN and the SVM for rest vs language task classification using different fraction of features. MAE is mean absolute error.	100
4.10	Significant modular interactions specific to the language task.	101
5.1	ASD subject details, FIQ Score	118
5.2	ADNI subject details, including Mini-Mental State Examination (MMSE) and Clinical Dementia Rating (CDR) scores	118
5.3	ADHD subject details, Conners' Parent Rating Scale-Revised, Long version (CPRS-LV)	119
5.4	MDD subject details, Hamilton Depression Rating Scale (HAM-D)	119
5.5	Performances of the best models of different models on the neuroimaging datasets.	121
5.6	Comparison of functional connectivity based classifiers. Params represent the number of parameters left in the reduced model, relative to the model that uses the full feature set.	122
5.7	The differences of test and train losses computed using cross-entropy.	122
5.8	Accuracies of models produced by keeping X% most important features from SVM and logistic regression models. C + L = CLIP + LEAN.	123
5.9	Accuracies of baseline models before and after CLIP was applied.	124
5.10	Average time taken per epoch (in seconds) for each variant of the proposed algorithms. mil = million.	124

C1	Best architectures of SVM Models on the neuroimaging datasets	157
C2	Results of recursive feature elimination for the Emotion, Gambling, Relational, Social and Working Memory tasks	160
C3	Significant task specific modular interactions corresponding to the Emotion, Gambling, Relational, and Social tasks.	162
C4	Significant task specific modular interactions corresponding to the social and working memory tasks.	163
D1	Best architectures of SVM Models on the neuroimaging datasets	165

Acronyms

2D	Two Dimensional
ABIDE	Autism Brain Imaging Data Exchange
AD	Alzheimer's Disease
ADHD	Attention Deficit Hyperactivity Disorder
ADNI	Alzheimer's Disease Neuroimaging Initiative
AFNI	Analysis of Functional Neuro Images
AMI	Adjusted Mutual Information
ASD	Autism Spectrum Disorder
BIC	Bayesian Information Criterion
BOLD	Blood Oxygen Level Dependent
CDR	Clinical Dementia Rating
CN	Cognitively Normal
CNN	Convolutional Neural Network
CLIP	Correlation-based eLImination of InPuts
DNN	Deep Neural Network
DL	Deep Learning
EEG	Electroencephalography
EPI	Echo Planar Imaging
FFN	Feed Forward Network
fMRI	functional Magnetic Resonance Imaging
HCP	Human Connectome Project
ICSC	Iterative Consensus Spectral Clustering
LEAN	Layerwise Elimination of Accessory Nodes
LFR	Lancichinetti-Fortunato-Radicchi
MCC	Matthew Correlation Coefficient
MCI	Mild Cognitive Impairment
MDD	Major Depressive Disorder
MELODIC	Multivariate Exploratory Linear Optimized Decomposi- tion Into Independent Components
MEG	Magnetoencephalography
ML	Machine Learning
MMSE	Mini Mental State Examination

MRI	Magnetic Resonance Imaging
NMI	Normalized Mutual Information
PET	Positron Emission Tomography
ROI	Region of Interest
rs-fMRI	resting state fMRI
SNR	Signal to Noise Ratio
SPECT	Single Photon Emission Computed Tomography
SPM	Statistical Parametric Mapping
SVM	Support Vector Machine
TE	Echo Time
TR	Repetition Time

Author's Publications

Journal Publications

1. **Sukrit Gupta**, Jagath C. Rajapakse, and Roy E. Welsch. Ambivert degree identifies crucial brain functional hubs and improves detection of Alzheimer's Disease and Autism Spectrum Disorder. *NeuroImage: Clinical* (2020): 102186.
2. **Sukrit Gupta**, and Jagath C. Rajapakse. Iterative consensus spectral clustering improves detection of subject and group level brain functional modules. *Scientific reports* 10.1 (2020): 1-15
3. **Sukrit Gupta***, Yi Hao Chan*, Jagath C. Rajapakse. Obtaining leaner DNN for decoding brain functional connectome in a single shot, *Neurocomputing* (Accepted, In Press), *equal contribution
4. **Sukrit Gupta**, Marcus Lim, and Jagath C. Rajapakse. Decoding task-specific and task-general functional architectures of the brain. Submitted to *Human Brain Mapping* for possible publication.

Conference Publications

1. Yi Hao Chan, **Sukrit Gupta**, L.L. Chamara Kasun, and Jagath C. Rajapakse. Decoding task states by spotting salient patterns at time points and brain regions. *Proceedings of International Workshop on Machine Learning in Clinical Neuroimaging*.
2. **Sukrit Gupta**, Yi Hao Chan , and Jagath C. Rajapakse. Decoding brain functional connectivity implicated in AD and MCI. *Proceedings of the 22nd International Conference on Medical Image Computing and Computer Assisted Intervention*, 2019.

3. **Sukrit Gupta**, and Jagath C. Rajapakse. Nodal degree distributions of resting-state functional brain modules. In 2018 *IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (pp. 1256-1259).
4. Jagath C. Rajapakse, **Sukrit Gupta**, and Xiuchao Sui. Fitting network models for functional brain connectivity. In 2017 *IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)* (pp. 515-519).

Chapter 1

Introduction

1.1 From functional MRI to the brain functional connectome

Functional Magnetic Resonance Imaging (fMRI) is a non-invasive neuroimaging modality that measures *in vivo* changes in the functional activity of brain regions over a period of time. These changes occur due to the changes in brain metabolism that are caused by both task performance and intrinsic processing in the resting state brain. Since its inception, almost thirty years ago, fMRI has been widely used in studies in cognitive neuroscience, clinical psychiatry, detecting disease biomarkers and studying the efficacy of the administered treatment. FMRI has become a method of choice for studies on brain mapping because of its higher spatial resolution than other functional neuroimaging modalities such as positron emission tomography (PET) and the commonly used electroencephalography (EEG). Also, unlike PET or single photon emission computed tomography (SPECT), fMRI does not need the patient to be injected with any radioactive substance before image acquisition. Deciphering fMRI data with computer science models has played an important role in furthering our understanding of the brain's functional organization during rest and its alterations during task performance and neurological diseases.

Neurons do not possess any intrinsic energy reserves, either in the form of glucose or oxygen. Upon activation, the neurons are supplied with energy by neighboring capillaries. This process of supply of energy to neurons by blood capillaries, known as the hemodynamic response, leads to an alteration in the relative levels of oxyhemoglobin and deoxyhemoglobin in the blood flowing near the activated neurons. Functional MRI detects this alteration on the basis of the differential magnetic susceptibilities of blood oxyhemoglobin and deoxyhemoglobin.

The biological significance of the fluctuations in neural activity were first elucidated by Biswal et al [1]. Biswal and colleagues demonstrated that the left and right hemispheric regions of the primary motor cortex show a high correlation between their fMRI blood oxygen level dependent (BOLD) time-series even during rest [1,2]. This suggests the presence of intrinsic information processing in these regions even in the absence of tasks or during rest [2–4]. With the advent of high resolution functional MRI, sub-millimetre spatial resolutions can be obtained [5]. However, voxels with sizes ranging from 2-3 mm are most commonly used. Functional MRI has been widely used to analyze functional characteristics for both healthy and diseased subjects that are at rest or performing some predefined tasks. While studying the resting state functional brain architecture of healthy subjects informs about the cognitive mechanisms behind cognition and decision making, studying alterations in the brain during task and disease states gives information about underlying mechanisms and regions that regulate task performance and neurological diseases. Although traditional applications of deep learning were limited to analyzing anatomical medical images of different human organs to identify diseases like breast and lung cancer or lesions in the brain, multiple recent studies have demonstrated the application of deep learning on fMRI data to automate diagnosis of neurological diseases.

fMRI scan acquisition can be done during rest or when the subject is asked to perform a predefined sequence of tasks. During resting state fMRI (rs-fMRI) acquisition, the participant is not exposed to any stimulus or task, i.e. the participant is at rest. The participants are usually asked to relax, lie still with their eyes closed or open, and not concentrate on anything particular during the rs-fMRI scan acquisition. Absence of a set of task based instructions renders rs-fMRI attractive for patients who may have difficulty in performing tasks. These include patients with neurological, neurosurgical, psychiatric and pediatric conditions. The brain activations during the resting state represent intrinsic brain activity, depicting spontaneous fluctuations in BOLD signal.

Analyzing rs-fMRI data aids in understanding the functional organization of the brain. Two fundamental organization principles in the brain are functional specialization and functional integration. While functional specialization focuses on mapping functions to specific brain regions or groups of regions, functional integration techniques study the functional relationships between different brain regions. However, the analysis of rs-fMRI data is challenging due to the large amount of generated data, which requires sophisticated analysis. Multiple preprocessing

steps need to be applied to the rs-fMRI data before the actual analysis is performed. These include (but are not limited to) realignment and removal of confounding artifacts (head motion, cerebral spinal fluid signal, et cetera), removal of the initial time points, slice timing correction, data normalization, smoothing, and band-pass filtering. Commonly used software packages for rs-fMRI data analysis include Statistical Parametric Mapping (SPM) [6], Analysis of Functional Neuro Images (AFNI) [7], the CONN toolbox [8], Multivariate Exploratory Linear Optimized Decomposition into Independent Components (MELODIC) [9], and fMRIPrep [10].

There are multiple methods to analyze the functional connectivity from fMRI scan data. One method uses *a priori* knowledge to select subsets of brain regions/voxels and performs connectivity analysis on these selected regions/voxels [11]. The other approach considers all the brain voxels and forms different subnetworks by performing decomposition or clustering from the time-series of the voxels, such that voxels assigned to the same component/cluster are highly correlated. For the latter method, two basic approaches are (1) using predefined seeds and performing correlation analysis with those seeds [12], and (2) independent component analysis [13]. For seed-based approaches, multiple seed regions are selected and pairwise relationships between these seed regions and the brain voxels are computed. Conversely, independent component analysis identifies clusters of brain regions on the basis of the statistical patterns in their BOLD time series signal. While both these approaches have enjoyed considerable success for analyzing fMRI data, none of these are able to characterize the brain's functional architecture in its entirety. In case of the seed-based method, only the first-degree neighbors of nodes are considered and higher degree neighbors are excluded from the analysis. Similarly, using independent component analysis, built on the strict assumption of the independence of subnetworks, we can uncover clusters of functionally homogeneous regions, but we do not get information about the patterns of connectivity within these clusters.

In this thesis, we use predefined brain regions of interest (ROI) for modelling functional architecture of the brain. This involves selecting the locations of brain's ROIs, computing the representative time series for each ROI, and finally computing the functional connectivity among different ROIs. Selecting a subset of ROIs that represent anatomically and functionally diverse regions in the brain is a much researched area [14–16]. Often the ROIs encompass multiple voxels, and the representative time-series for each ROI is obtained by computing the mean time series or the first principal component (using principal component analysis) of the

time series of all voxels in the ROI. Different measures can be used to estimate the functional connectivity between computed time series for ROIs, depending upon the context. Pearson's correlation coefficient is commonly used to measure linear relationships between the ROIs, mutual information can identify non-linear relationships [17], co-activations between regions can be measured using regional homogeneity (ReHo) [18], and coherence can be used to estimate linear relationships in the frequency domain [19]. In this thesis, we assume an undirected, symmetric relationship between brain regions that is computed using Pearson's correlation coefficient. Conversely, there also exist measures for directed (and asymmetric) functional connectivity that can be computed using Granger causality [20, 21], directed coherence [22], dynamic causal modeling [23], linear non-Gaussian [24], and Bayes network methods [25]. Since we discount causal effects (captured by directed or asymmetric connectivity), the functional connectivity measure used in this thesis serves as a useful summary of distributed activity and can be used to classify or distinguish subjects given some measurement of distributed brain activity.

We model the brain functional connectome as a network or graph by defining the network nodes as the anatomically defined ROI and the edges between nodes given by the correlations of inter-regional BOLD time series signal. The pairwise information between all nodes is compiled to form an adjacency matrix that can be further analyzed using graph theoretical methods and measures. Multiple properties of the brain's functional connectome such as the network path length or efficiency [26], high degree nodes [27], modularizations or clustering [14] have been widely studied. Studies have also been conducted to measure the difference in brain functional connectome's network measures due to neurological diseases [28–31] and aging [32, 33].

1.2 Motivation

The human brain is composed of approximately 100 billion neurons that interact with each other to form a complex and massive neural network that regulates a range of essential functions such as perception, cognition, and decision making. The human brain can be further partitioned into a set of regions that are spatially distant but functionally linked. The neuronal populations in these regions coordinate with each other to give rise to the functional capabilities of the brain. Advances in functional neuroimaging, especially fMRI techniques, have made it possible to

Fig. 1.1: Illustration of how the change in the blood oxygenation is used to derive the BOLD time series signal for all the brain voxels and further processed to derive the functional connectivity matrix and finally the functional connectome. Part of the figure is obtained from [34].

Fig. 1.2: The figure shows the outline of topics covered in this thesis. We study the brain functional connectome during rest, task and diseased states. Part of the figure is obtained from [35].

study the human brain's functioning during different states noninvasively. By studying the functional connectome built from fMRI data, brain's functional architecture can be studied at different scales during rest, task, and diseased states. Thus the functional connectome has become a powerful tool for studying the organization of the normal human brain and alterations related to varied tasks and neurological diseases. In this thesis, we develop methods to study the brain's functional connectome during different states, viz. the rest, task and diseased states. Applications of the proposed methods in studying the varied brain states are also demonstrated in the respective chapters.

Detecting functional modules of the brain

The human brain is composed of several independent sub-systems that are involved in performing specialized tasks. These sub-systems, or modules, have been derived by applying module detection algorithms on the brain functional connectome (see [36] for review), such that each detected module performs subject specific intrinsic functional tasks [37]. Brain modules are characterized by dense functional connectivity with the nodes in the same module (intra-modular) and sparse functional connectivity with nodes in different modules (inter-modular).

Modular networks are known to engage in specific information processing [38] and are generally more robust and immune to local perturbations [39]. As a result, module detection has been researched extensively in graph theory and network science, and different module detection algorithms have been proposed (see [40] for review). Although, modules associated with the brain functional connectome have been uncovered by multiple studies [14, 41–44], there are multiple challenges associated with the area. In most scenarios, we have fMRI data for multiple subjects or multiple scan data for a single subject. The goal is to find brain functional modularizations at the group-level for the multiple subject cohort or at the subject-level for the multiple scan cohort. This introduces complexities in the modularization process due to variability in brain functional architecture of different subjects and limitations related to different module detection algorithms. Thus, in order to completely and reliably characterize the brain’s functional modular structure, there is a need for an algorithm that considers not only the inter-subject or inter-scan variability in the functional architecture, but also handles dense and weighted functional connectome data. We are, therefore, first motivated to invent algorithms for detection of functional modules in the brain from neuroimaging data.

Detecting functional hubs of the brain

While modules support the functional specialization in the brain, neuroimaging studies have uncovered collections of regions in the brain, which are crucial for facilitating effective signaling and communication in neuronal populations, often called ‘brain hubs’ (see [45] for review). The brain hubs support information integration in the brain, whereby they enable multiple complex cognitive brain processes. Brain hubs have been delineated by their high functional connectivity to other brain ROI and also their central positioning in the network [46–48]. This also provides an intuition about the possible loss in information processing in the event of damage to these hubs [49]. Consequently, several studies have pointed out abnormal hub organization in several neurological and psychiatric brain disorders [50, 51]. Notably, multiple studies detecting brain hubs did not consider the effect of weak connections on hub detection and thresholding on the preservation of brain’s modular architecture [52]. This can lead to non-hub regions being erroneously classified as brain hubs. Since they play a central role in brain organization, it is valuable to derive methods to correctly identify brain hubs and use insights derived from them to study neurodegenerative and psychiatric disorders. These factors motivate us to find novel measures that find functional brain hubs from neuroimaging data.

Decoding task-general and task-specific brain architectures

While intrinsic activity in the brain supports its resting state architecture, the ability to perform complex tasks depends on brain's ability of processing specialized information among sets of brain regions in functional modules and information integration among distributed brain regions. One of the fundamental questions in cognitive neuroscience is to understand how the the resting brain reorganizes spatially and temporally to support the high dimensional and flexible behavior during task performance. Multiple studies have investigated the modulations in brain functional architecture during task performance. Initial studies suggested that although there are small-scale alterations in the brain functional architecture, there is an intrinsic stable functional backbone during both the rest and task state that remains unchanged [53–55]. While studying modulations related to task performance, previous studies only captured linear relationships between functional connectivity at rest-state and task-states, which can lead to incorrect conclusions in deriving task-related modulations [54]. These studies disregarded the effect of inter-subject variability in functional architecture by performing operations at an averaged group-level matrix. They also focused on detecting task-related modulations as a whole, whereas multiple studies have uncovered a set of task independent modulations known as ‘task-general’ modulations that are common across all the tasks. Therefore, we are motivated to devise an approach that not only considers subject-level functional MRI data and also allows for non-linear relationships between rest and task data, but also separates task modulations into task-general and task-specific modulations.

Optimizing brain encoders for detecting diseased brain states

Besides deciphering task related modulations in the brain's functional architecture, the comparison between functional connectivity of normal and diseased patients can be used to understand disease mechanisms [56, 57] but they can also be used to train machine learning (ML) models to diagnose psychiatric, developmental and neurodegenerative brain disorders [58, 59]. Besides diagnosis based on symptom scores from clinical interviews, currently there are no gold standards to clinically validate the diagnosis. This can result in conflicting diagnosis across clinicians, with poor patient outcomes. Data from multiple functional neuroimaging techniques can be used to characterize functional abnormalities in the brain and improve diagnosis and

treatment [60]. However, there are unique challenges associated with the use of functional connectivity as features for ML models to classify or predict brain disorders. Present neural network models using functional MRI data use functional connectivity as input features to classify diseased and control subjects [61, 62]. Such studies typically involve $\sim 10,000$ input features with $\sim 1,000$ subject scans, making the neural networks prone to overfitting. Coming up with techniques that can lead us to networks with fewer trainable parameters and better generalization capability, without any significant loss in classification performance thus becomes important for computer aided detection with functional MRI data. To this end, we develop algorithms that find leaner and compact deep neural network models for brain state classification.

1.3 Major contributions

The major contributions of the thesis reside in the study of functional connectome of the human brain during rest, task and diseased states. The specific contributions are as follows:

- We propose a novel algorithm, the *Iterative Consensus Spectral Clustering (ICSC)*, that detects the brain's modular structure at the subject and group level, while considering the inter-subject variability in modular structure of subjects and using dense, weighted functional connectivity matrices. The novelty of the ICSC algorithm comes from the consensus-cost function defining the similarities or agreements between group-level and individual-level modular structures. This allows both the common features and the variability in functional architecture across individuals to be preserved. We showed the application of the proposed algorithm on real data from the Human Connectome Project (HCP) and synthetic data with known ground-truth modular architecture. We used the subject and group level modularizations obtained from the data to study the inter-subject and inter-scan variability in brain functional architecture. We show an application of the proposed ICSC modularization method in detecting major depressive disorder, a disease that is characterized by disruption in modular structure.
- We use the subject level modularization derived from the ICSC algorithm and propose a novel measure, the *ambivert degree*, that considers the effect of weak edges in hub

detection. We argue that the continuous connection weights in brain functional networks results in misclassification of nodes as hubs. Previous studies have shown that the thresholding step used to remove weak connections does not take into account the preservation of brain functional modular structure. We show that the hubs detected by ambivert degree are significantly more crucial than those detected by previous measures. We found disruption in brain hubs for subjects suffering from Alzheimer’s disease (AD) and Autism Spectrum Disorder (ASD), and used ambivert hub measures as features for computer aided disease diagnosis. We found that the classifiers based on ambivert degree hubs gave a significantly higher classification accuracy than the traditionally used functional connectivity features.

- The functional connectome can capture the complex and dynamic interactions between brain regions during task performance. We propose a novel deep learning approach to decode brain activations and thereby study task related modulations in the brain’s functional architecture. We train a deep neural network (DNN) classifier on resting and task functional MRI data and decode the trained DNN to derive salience scores of the functional connections and modules distinguishing the rest and task brain states. Using the proposed approach, we not only uncover the task-general architecture that consists of a set of modulations common across all tasks, but also find the task-specific brain architecture that is specific to the task at hand. Furthermore, we study the task related modulations with respect to the inter-modular and intra-modular alterations in functional connectivity. We show that the proposed method detects task-general and task-specific architectures that are stable across subjects.
- Finally, we used neuroscientific insights regarding brain’s functional connectivity to propose algorithms to find leaner and compact neural architectures for encoding the brain states. We hypothesized that there exist correlations in brain functional connectivity features and use this to propose two techniques, *layerwise elimination of accessory nodes (LEAN)* and *correlation-based elimination of inputs (CLIP)*, for DNN pruning in neuroimaging applications. We remove redundant connectivity features using LEAN and CLIP and reduce the problem of overfitting when DNN models are used to classify neuroimaging datasets. The obtained DNNs maintained a similar performance as the full

network with only around 2% of the initial trainable parameters during classification of patients from healthy controls. We also decode the trained DNN to identify salient brain regions and connections from the functional connectome for multiple brain disorders. The proposed methods have cross-modal applications in obtaining leaner DNNs and neuroimaging biomarkers for diseases.

It is important to mention some assumptions that we made in this thesis. Although we considered inter-subject variability between subjects from multiple datasets, we assumed that the subjects in each dataset had a central tendency that is common across subjects in the dataset. This is a valid assumption for normal subjects in a narrow age range, but it may not cover the disease heterogeneity that is associated with disorders like AD [63], ASD [64], MDD [65], ADHD [66], et cetera. In order to obtain separate subject groups with common features, one can apply techniques like clustering [67], Latent Dirichlet Allocation [68], et cetera before applying the methods proposed in the thesis. In Chapters 2 and 3, where we studied the resting state brain functional architecture, we only considered the effect of positive correlations. The proposed methods can be extended to consider negative correlations for modularization [69] and hub detection [70].

1.4 Organisation of the thesis

In this thesis, we study the the brain’s functional connectome during resting, task, and diseased states. We discuss the existing methods in detail in the respective chapters. The remaining thesis consists of the following chapters:

In Chapter 2, we detail the proposed ICSC algorithm to detect subject and group level brain functional modules. It begins with an introduction about the relevance of modules in the brain functional connectome and details the previous studies in the area. The methods section provide the details about previous module detection algorithms, the proposed ICSC algorithm, how the ICSC algorithm can be used to detect group-level and subject-level modules, study inter-subject and inter-scan variability in the brain’s modular architecture, and its application for disease detection. This is followed by results of performing modularization with the dataset from the HCP, comparison with previous approaches, studying inter-subject and inter-scan variability of modularizations, experiments with synthetic networks with known ground-truth

and detecting major depressive disorder using modules derived from the ICSC algorithm. We conclude by discussing the results and their implications on the field.

In Chapter 3, we present the proposed novel measure, ambivert degree, to detect brain hubs and show that it can be used for diagnosing Alzheimer’s Disease (AD) and Autism Spectrum Disorder (ASD). In the methods section, we detail the different measures used to detect brain hubs (modular and connector hubs), the rationale behind the proposed ambivert degree measure, using path length to quantify the importance of different hubs, inter-subject variability of hubs, hub disruption in AD and ASD, and using hubs as features to detect AD and ASD. In the results section, we compare the hubs detected by ambivert degree with hubs detected by previous measures and show that ambivert degree hubs are the most crucial. We also compare ambivert hubs for the diseased and healthy controls for the AD and ASD datasets and find widespread disruption in brain hubs. We show that hubs can be used as features to diagnose these diseases. We discuss the results and scope of the proposed method on hub detection and disease classification.

In Chapter 4, we present a novel deep learning based approach to decode the functional connectivity and study task-general and task-specific functional brain architectures during task performance. In the methods section, we describe how we use the DNN to encode functional connectivity and decode the trained DNN by using model interpretability techniques. We describe the methods used to derive task-general and task-specific brain functional architectures by using the rest and task fMRI data from the HCP dataset. In the results section, we show the encoding accuracy obtained by the classifier for classifying rest and task samples, and describe the uncovered task-general and task-specific modulations in the motor and language task. We conclude by discussing the implications of the results, and the neuroscientific validity of the uncovered modulations.

In Chapter 5, we propose a novel approach to reduce the number of trainable parameters in DNN architectures for encoding and decoding the brain’s functional connectome. We start with discussing the gap in present techniques for diagnosing neurological diseases using functional connectivity data. In the methods section, we introduce the proposed two methods LEAN and CLIP that help in pruning the network while retaining a subset of correlated features. We show the application of the proposed methods on diagnosing AD, mild cognitive impairment (MCI), ASD, attention deficit hyperactivity disorder (ADHD), and major depressive disorder (MDD).

In the results section, we show that the proposed methods lead to a large drop in the number of trainable parameters with a minimal drop in classification performance. We also show that the proposed method gives us clinically relevant disease biomarkers. We conclude by discussing the applications of the proposed methods in disease diagnosis and biomarker detection.

In Chapter 6, we conclude by discussing about the proposed methods in the thesis. We discuss possible future work to enhance the scope of the proposed techniques.

Chapter 2

Detecting functional modules of the brain ¹

Brain’s drive for specialized processing is supported by subsets of functionally diverse brain regions organized as modules in the functional connectome. Despite prior evidence of pronounced inter-subject variability in the brain’s functional architecture, previous methods used to derive functional modules from the brain’s functional connectome ignore these differences and use incomplete functional connectivity information. We propose a novel algorithm, called Iterative Consensus Spectral Clustering (ICSC), that detects the most representative modules from individual dense weighted connectivity matrices derived from multiple scans. The ICSC algorithm derives group-level modules from individual-level modules by iteratively maximizing the similarity between the two. The performance of our algorithm was compared with existing functional brain modularization methods and we show that our method detects group-level modules that are more representative of the modules of multiple individuals. We demonstrate that the ICSC algorithm can be used to derive biologically plausible group-level (for multiple subjects) and subject-level (for multiple subject scans) brain modules, using resting-state fMRI scans of 589 subjects from the Human Connectome Project.

2.1 Brain functional modules and related work

Functional specialization is a fundamental organizational principle of the human brain. The functional connectome of the human brain can be decomposed into several subsets of brain regions (for review see [36]). These subsets of regions known as *functional modules* or clusters have been shown to be representatives of different systems performing specialized

¹The work in this chapter has been published in [71]

neurophysiological functions in the brain [14, 54]. It is important to reliably detect brain functional modules since they are used to detect cross-modal brain regions [45], known as brain hubs are implicated in various neurological diseases and brain injury [72]. Additionally, modular features can be used as biomarkers for detecting brain disorders since they have been found to be disrupted in neurological ailments like Alzheimer’s disease [73], major depressive disorder [74, 75], schizophrenia [76], post traumatic stress disorder [77] and chronic pain [78]. This has made it important to reliably detect brain modules to understand the dynamics between brain functional architecture and brain disease, prompting research to find methods that identify brain functional modules.

Numerous studies have shown that the brain functional connectome can be decomposed into modules containing functionally correlated regions. For example, Power et al [14] used the Infomap algorithm [79] to obtain modularizations for the averaged and thresholded connectivity matrix derived from resting-state fMRI data of multiple subjects. It was observed that regions from many of the detected communities were functionally linked but spatially distributed. This is a characteristic of connectivity based networks where connections signify statistical relationships between activations of brain regions instead of physical linkages as in the case of structural connectivity. Since, functional connectivity has no direct metabolic or material cost, forming strong long-distance functional connections is possible. Meunier et al [80] found a hierarchical modular structure from the thresholded and binarized brain functional connectome for each subject. From these subjects, they report the results for the modularizations for the subject that they found to be the most similar to the remaining subjects. Nicolini et al [43] proposed to use a new quality function, called Surprise, that tackles an inherent issue of resolution limit [81] with both Infomap and the commonly used modularity maximization. They show that the resolution limit affects the modularization capacity of both Infomap and modularity maximization for brain functional connectomes. They further proposed a weighted version of the quality function, Asymptotical Surprise, and showed its application on an averaged, thresholded adjacency matrix [42]. In Ji et al [41], the authors proposed to use the modularity maximization approach on an averaged weighted adjacency matrix along with an additional set of hard conditions that necessitated the detection of separate motor, visual and auditory networks.

In this chapter, we propose a novel technique to detect functional modules from the brain functional connectome obtained from fMRI BOLD time-series signal of functionally diverse

brain regions. Previous studies in the area [14, 41–43], have made an assumption that the modules obtained from applying module detection algorithms on an averaged connectivity matrix from multiple subjects reliably represent the modular structure of underlying subjects. This approach is flawed because averaging of subject connectivity matrices leads to the loss of important subject-level features [82–84] and ignores the heterogeneous inter-subject differences in brain functional architecture [85] that are consequential while detecting group-level brain functional modules. The variability in functional architecture across individuals may not be limited to variation in the function of brain regions but also grouping/sub-division of functional modules in individuals. Therefore, detecting group-level modules from multiple subjects in the ensemble should not only consider modular architecture for all the subjects but also allow for the variability in their modular structure.

We hypothesize that the most representative group-level modularization is maximally similar to the modularization of the individuals (note that these are different from the subject-level modularizations and these correspond to the modularizations for scans or subjects, as the case maybe. We refer to these as *individual-level modularizations*). Based on this hypothesis, we develop a new approach called *iterative consensus spectral clustering* (ICSC) based on consensus clustering [86] and multiclass spectral clustering [87]. We define the *consensus-cost* based on similarity between group-level and individual-level modularizations and iteratively refine both the individual-level and group-level modules making them more similar to each other. This procedure gives a reliable group-level representative modular structure by using the group-level modular information as a prior to refine individual-level modules. We allow variability in individual-level modularizations by not only allowing for variable modular structure but also varied number of modules for each individual. Recent studies have shown that weak edges predict differences in fluid intelligence [88, 89] and differentiate people suffering from schizophrenia from healthy controls [90], which necessitates that loss of information from steps like thresholding and binarizing should be avoided. We avoid loss of weak functional connections and connection weights by considering unthresholded (dense), weighted individual-level functional connectivity matrices. We show that the ICSC algorithm can compute biologically plausible modules without setting any hard conditions for specific modules to be detected (as in [41]). It also eliminates the need to vary parameters like network thresholds or resolution limit and specifying the number of modules. The ICSC algorithm found several smaller modules that were not detected by previous modularizations.

Using resting-state fMRI scans obtained from 589 subjects from the Human Connectome Project (HCP), we use ICSC algorithm to detect group-level modules from multiple subjects and subject-level modules from multiple scans of the subject. On computing variability using the group-level and subject-level modules, we found lower intra-subject (inter-scan) variability in comparison with inter-subject variability. The brain regions involved in elementary tasks like vision and motor were found to be stable both across subjects and across scans. We compared the modularization obtained from the ICSC algorithm with previous methods such as the Louvain [91], InfoMap [79], and Asymptotical Surprise [42] and found that these algorithms give a few large, functionally relevant set of modules, whereas the ICSC algorithm gives smaller biologically plausible modules, which are most similar to individual-level modularizations. Furthermore, we found that the iterative procedure in the ICSC algorithm, gives significantly more representative subject-level modules by using multiple subject scans from the HCP dataset. For a quantitative comparison, we generated synthetic networks having similar modular structure to resting-state functional connectivity networks obtained from human fMRI scans; and show that the ICSC algorithm was able to detect group-level modules stably across different noise levels and nodal variations. Finally, we show an application of the modularizations obtained using the ICSC algorithm for diagnosis of major depressive disorder using functional connectomes of healthy and diseased human subjects.

2.2 Previous methods to detect brain functional modules

Given a set of resting-state fMRI brain scans, we first derive functional connectivity matrices of the brain networks. For k th individual, the functional network $G^k = (\Omega, W^k)$ where Ω denotes the set of functionally relevant regions of interests (ROI) or nodes of the network and W^k denotes functional connectivity matrix derived from the fMRI scan.

2.2.1 Module detection at individual-level

A functional modularization (i.e., detecting the functional modules given the brain functional connectivity matrix) is a partitioning of Ω into disjoint sets or modules such that the connectivity among the nodes within a module is higher than the connectivity between the nodes of different modules. Given the functional brain network G^k for k th individual, a modularization finds a

vector $s^k = (s_i^k)_{i \in \Omega}$ where each element s_i^k corresponds to a module label assigned from a set $\{1, 2, \dots, L^k\}$ and L^k denotes the number of modules.

The challenge in finding individual-level functional modules is two folds. First, the number of modules L^k of an individual k is unknown. Second, the partitioning of the nodes at individual level, s^k , should be consistent with the consensus modules at the group-level. Given the number of modules L^k , we denote finding a modularization s^k as

$$s^k \leftarrow s^k(L^k) \quad (2.1)$$

where $s^k(L^k)$ denotes the node partition for k th individual containing L^k modules that can be derived from minimization of the cost function given by equation 2.4.

2.2.1.1 Modularity maximization

The most widely used method to perform modularization of the brain functional connectome is the modularity maximization method [92], that tries to maximize the modularity quality function Q given by:

$$Q = \sum_{i,j} \left(w_{ij}^k - r_{ij}^k \right) 1(s_i^k = s_j^k) \quad (2.2)$$

where w_{ij}^k represents the weights of edges between nodes i and j , r_{ij}^k represents the expected weights of edges (according to some user-defined null model), and $1(\cdot)$ denotes the identity function. Since optimizing Q is computationally intractable even for networks with few nodes (it is impossible to perform an exhaustive search for all possible node partitions), multiple heuristics have been proposed to approximate the partition that gives the global maxima for Q . Such heuristics include divisive algorithms [92], spectral decomposition [93], greedy methods [91, 94], et cetera.

A well-known issue with modularity maximization is the resolution limit of the algorithm, whereby it is impossible to detect modules below a certain scale (size), even if the all the nodes in the modules are connected to each other [81]. The resolution limit is a result of the trade-off between the number of modules and the modularity value Q , and the modularizations obtained for the maximum value of Q may not reflect the network's ground-truth modular structure. Therefore, the modules obtained from maximizing Q may contain several undetected

smaller modules. For circumventing this issue for modularity maximization, Reichardt et al [95] proposed to incorporate an additional resolution parameter in the formulation that scales the importance of the null model. Tuning the resolution parameter can help obtain modules with smaller (resolution parameter > 1) or larger (resolution parameter < 1) number of nodes. However, tuning the resolution parameter only gives modularizations containing different number of modules and still does not give information about the ground-truth modularization of the network [96].

2.2.1.2 Infomap

Infomap [79], another popular method used for modularization of the brain functional connectome is based on information theory. Infomap tries to find a compressed representation of a network such that the amount of information needed to describe information diffusion across the network is minimized. Information diffusion is modelled using a random walker that uses edges between nodes to traverse the network. The network is described using a two level procedure where each node is assigned a code based on its modular membership and a unique vertex name for the respective module, such that nodes across different modules can share the vertex name. The vertex name is derived using Huffman coding [97]. It is intuitive that a random walker would spend longer time within communities. Graph clustering is therefore formulated as finding the partition that leads to the minimum description length, known as the code length (C_L), of an infinite random walk. The code length can be expressed as the Shannon entropy of walks within and between the modules. Therefore, it becomes necessary to reduce the transitions of the random walker between the modules because every time there is a transition, the code for the respective module is added to the code length. The code length is minimized by combining greedy search with simulated annealing [79].

2.2.1.3 Normalized Cut Cost

A functional modularization s^k is also represented by a set of binary matrices as $X^k = \{x_l^k\}_{l=1}^{L_k}$ where l th module $x_l^k = (x_{li}^k)_{i \in \Omega}$ such that $x_{li}^k = 1(s_i^k = l)$. The degree matrix D^k of the connectivity matrix W^k defined as the diagonal matrix having diagonal elements as the degree of the corresponding nodes is given by

$$D^k = \text{diag}(W^k 1_{|\Omega|}) \quad (2.3)$$

where $1_{|\Omega|}$ is a vector of ones containing $|\Omega|$ elements.

The *normalized cut-cost* objective function is given by:

$$\text{cut-cost}(X^k, L^k) = \frac{1}{L^k} \sum_{l=1}^{L^k} \left(1 - \frac{(x_l^k)^T W^k x_l^k}{(x_l^k)^T D^k x_l^k} \right) \quad (2.4)$$

The minimization of the cut-cost is performed by multiclass spectral clustering (SC) [87]. This method gives an optimal solution and a modularization of G^k . Multiclass spectral clustering is fast and more robust than random initialization. Given the functional connectivity G^k and the number of modules L^k , we find the modularization denoted in (2.1) by using spectral clustering.

Modules from a network can be derived by optimization of any of the quality functions such as Newman's modularity [93], the Code length [79] or the normalized cut cost [98]. Besides Newman's modularity Q [81], previous studies have also shown an existence of a resolution bias for the modularization quality functions like Code length [99] and the normalized cut cost [100] function. This resolution bias has been shown to severely curtail the detection of small albeit functionally relevant brain modules [43]. However, it was pointed out that although these quality functions suffer from different resolution biases, they gave similar results when forced to converge to a specific number of modules [100]. Unlike the normalized cut-cost, the Code length and Newman's modularity objective functions cannot be forced to detect a fixed number of modules, they can be run multiple times to determine suitable parameters values that lead to convergence to the suitable number of modules. This is an important distinction between the objective functions for our case, because using prior neuroscientific knowledge, we can get a range of the number of functional modules that are present across subjects. Additionally, unthresholded resting-state functional connectivity networks worsen the module detection capability of Code length and Newman's modularity. On the other hand, multiclass spectral clustering [87] used to optimize the normalized cut-cost objective function solves the problem in eigenspace of the network Laplacian and is therefore not affected by the density of networks.

2.2.2 Group-level consensus modules

Earlier works have mostly focused on modular detection at the group-level [14, 42, 43] as a representation of the modules of the individuals. Such approaches perform module detection on

an averaged connectivity matrix ignoring the variability in functional architecture of individuals [83, 101, 102]. We propose to use a consensus clustering approach [86] since it considers the individual-level modularizations while deriving a group consensus matrix.

The individual-level modular matrix is a binary matrix where an element equals one if the corresponding nodes belong to the same modules and is zero, otherwise. For the k th individual, the individual-level modular matrix is given as $\Pi^k = \{\pi_{ij}^k \mid \pi_{ij}^k = 1(s_i^k = s_j^k)\}_{i,j \in \Omega}$. The *group consensus matrix* Π is computed by adding the individual-level modular matrices for all the individuals:

$$\Pi = \sum_k \Pi^k \quad (2.5)$$

The group consensus matrix is a multilevel affinity matrix and is used to find the group-level modularization. By using the spectral clustering on Π and assuming an L number of modules, we denote a group-level modularization as:

$$s \leftarrow s(L) \quad (2.6)$$

2.3 Iterative consensus spectral clustering (ICSC)

Iterative Consensus Spectral Clustering (ICSC) extends the multiclass spectral clustering to attain an individual-level modularizations close to the group-level consensus modularization. The ICSC iteratively maximizes the *consensus-cost* or minimizes the mean distance (similarity) between the group-level consensus modularization and the individual-level modularizations. We assume that in case of different subjects, the subject share common features in their functional architecture which can be inferred from the group level features, whereas in case of multiple scans from the same subject, the functional architecture across the scans is same but is distorted due to scanner artefacts or mind wandering during acquisition [85, 103–105].

The consensus cost of individual-level modularizations $\{s^k\}$ and group-level modularization s is given by

$$\text{consensus-cost}(s, \{s^k\}) = \sum_{k=1}^K \text{sim}(s, s^k) \quad (2.7)$$

where sim measures the similarity between two modularizations and K denotes the number of individuals.

We use the symmetric measure Adjusted Mutual Information (AMI) [106] to measure the similarity between the group-level and an individual-level modularization. The AMI is an extension to Mutual Information (MI) score and accounts for the chance. We avoid using MI and normalized MI (NMI) since they increase for two partitions with high number of modules, irrespective of the information shared [107]. Similarly, we also avoid using Adjusted Rand Index since it is preferred for modularizations with large similarly sized modules [108], whereas brain modules are non-uniformly sized [109]. The AMI between two modularizations s and s^k is given by:

$$\text{sim}(s, s^k) = \frac{MI(s, s^k) - \text{mean}(MI(s, s^k))}{\max(H(s), H(s^k)) - \text{mean}(MI(s, s^k))} \quad (2.8)$$

where H is the entropy of a partition and MI is the mutual information between the two partitions.

The ICSC optimizes the consensus-cost by alternatively (i) finding the number of modules L^k and L at the individual-level and group-level, respectively, and (ii) minimizing the normalized cut-cost(X^k) to attain individual-level modularizations. Finding the optimal number of clusters/modules for a given data is an open problem and there are various module quality indices available which aid in selecting a reasonable number of modules (see [110] for a review). For our data, we found that there was no agreement on the optimal number of modules suggested by different module quality indices. We, therefore, decided to approach the problem in terms of the information content of the group consensus matrix Π .

Given that there exists community structure in Π , then $L < |\Omega|$, which points to presence of redundant information in Π . For Π , the redundancy can be visualized as shared information between ROI which are part of the same modules across many individual-level modularizations which makes them linearly-dependent. Thus finding the number L of group-level modules can be visualized as the problem of finding the optimal number of dimensions that can balance the trade-off between independence and redundancy, minimizing the redundancy. This can be solved using the formulation given in [111]. We transform the ROI in the eigenspace, where the orthogonal eigenvectors are considered to represent modules. However, not all eigenvector directions contribute equally and the important eigenvectors can be computed from their eigenvalues. We estimate the number of dimensions by finding the ‘‘elbow’’ of the scree plot of the group consensus matrix, Π :

$$L = \text{elbow}(\Pi) \quad (2.9)$$

where the elbow function computes the point of maximum curvature of eigenvalues [112] in the range $L_{\min}$ and $L_{\max}$. The number of individual-level modules, L^k , are found by maximizing the similarity between the group-level s and individual-level modularizations, $\{s^k\}$, as

$$L^k = \operatorname{argmax}_l \operatorname{sim}(s, s^k(l)) \quad (2.10)$$

The minimum number of modules, $L_{\min}$, can be fixed on prior studies in the area. However, the choice for the maximum number of modules, $L_{\max}$, has a tradeoff in terms of the time taken for convergence and the choices available for l in equation 2.10 for maximization of consensus-cost. Lower than optimal values of $L_{\max}$ will have a sharp drop-off near $L_{\max}$ in the histogram of the distribution of number of individual-level modules, $\{L^k\}_{k=1}^K$. The sharp drop-off corresponds to individual-level modularizations that were assigned sub-optimal number of modules, L^k , which led to lower value for the consensus-cost.

The ICSC begins by performing individual-level modularization using SC and uses them to derive a group-level modularization. Thereafter, it iteratively finds the optimal numbers of modules L and L^k at the group-level and individual-level, respectively. First, it finds the individual-level modules s^k and then uses them to derive the group-level consensus matrix and finds the optimal number of group-level modules L , and so on until convergence. SC is used to obtain modularization at the individual-level and group-level, and the algorithm converges when no change in the *consensus-cost* is observed. In order to initialize $\{s^k\}$ and s , we compute the most suitable initial L^k and L using the elbow function with W^k and Π , respectively. The Algorithm 1 gives the complete ICSC procedure.

2.3.1 Inter-subject variability of functional modules

Several neuroimaging studies have shown pronounced inter-subject differences in the brain functional architecture distributed non-uniformly across the cortex [85, 113, 114]. Functional variability across subjects is a result of the combined effect of genetic and environmental factors that impact functional systems differently across subjects [115]. The variability of modular membership of brain regions across subject helps understand individual differences in cognition and behavior [84].

Using individual-level s^k and group-level modularizations s from the ICSC algorithm, the nodes with high variability in their modular assignment over individuals can be identified. For

Algorithm 1 Iterative consensus spectral clustering

Input: $\{G^k\}_{k=1}^K, L_{\min}, L_{\max}$

/ initial modularization */*

for every individual k **do**

$L^k \leftarrow \text{elbow}(W^k)$

$s^k \leftarrow s^k(L^k)$

$L \leftarrow \text{elbow}(\Pi)$

$s \leftarrow s(L)$

/ iterative refinement */*

repeat

/ individual-level modularization */*

for every individual k **do**

for every $l \in (L_{\min}, L_{\max})$ **do**

$\rho_l^k \leftarrow \text{sim}(s, s^k(l))$

$L^k \leftarrow \underset{l}{\text{argmax}} \rho_l^k$

$s^k \leftarrow s^k(L^k)$

/ group-level modularization */*

$L \leftarrow \text{elbow}(\Pi)$

$s \leftarrow s(L)$

until convergence

this, we first match the labels in the group-level and individual-level modularizations by using a Jaccard Index matching scheme described in [86] and then compute the metric *node purity* for node $i \in \Omega$ as

$$\psi_i = \frac{1}{K} \sum_k 1(s_i^k = s_i^*) \quad (2.11)$$

where s_i^* is the most frequently assigned module for node i across the individuals. We derive the purity of a module by averaging the purities of nodes belonging to the module. The module purity gives an idea about the variability of a module across the individuals.

2.3.2 Detecting subject-level functional modules using ICSC

Although group-level features render properties that are common to a group of individuals, finding neuro-functional features at the subject-level has significant applications in personalized neuroscience. Finding accurate modularization at subject-level is important to pinpoint susceptibility of the subject to diseases [76, 116, 117], and detect different brain states [118] and anatomical locations of the hubs in the subject's brain networks [42, 72]. However, it is also important to know if brain functional modules are reproducible across multiple scans of the same subject.

Previous studies show that there exist significant differences across scans of even the same subject [85, 103]. The inter-scan variability of the functional modules of the same individual arises due to noise in MRI acquisition and different functional activities during resting-state sessions [105]. The functions of brain regions (and therefore, functional modules) of a subject do not change much from one scan to the other and consequently the functional modules should have similar stability across scans.

We need a modularization method that considers each scan-level modular structure $s^{k\sigma}$ and uses it to derive the subject-level modularization s^k , where σ denotes a particular scan of the subject k . We propose the use of ICSC algorithm to derive a representative subject-level modularization from multiple scan-level modularizations. A good modularization algorithm should give less variability for intra-subject functional modules compared to the inter-subject variability of functional modules. We used the data from the HCP which involves multiple resting-states scans from the same subject. To calculate intra-subject variability, we obtain the purity ψ_i^k of node i across multiple scans σ of the subject k similar to (2.11). Then, for multiple subjects, we obtain the purity ψ_i of the node as $\psi_i = \frac{1}{K} \sum_k \psi_i^k$

2.3.3 Using ICSC modularizations to detect major depressive disorder

We simplify modeling of the brain functional architecture by combining the BOLD signal from multiple neurons in a brain ROI, by making an assumption that the neurons in the respective ROIs perform similar function. This can be taken one step further by combining regions in each functional module since the regions in each module represent a set of nodes that are involved in a similar specialized function. This becomes especially effective in case of neurological diseases where multiple studies investigating have implicated the change in inter-modular and intra-modular functional connectivity. Specific examples include the change in the connectivity of the salience module in Autism Spectrum Disorder [119], the default mode network in Alzheimer's Disease and Major Depressive Disorder (MDD) [74, 75, 120, 121], and the primarily areas in the medial prefrontal cortex of the salience module in Parkinson's disease [122].

Based on prior neuroscientific evidence, we propose the use of features based on brain modules for computer aided detection. We show that the functional modules derived from the ICSC algorithm can be used to successfully detect patients suffering from MDD from cognitively normal controls. We use the group-level modularizations s to derive a set of features for every subject sample based on its functional connectivity matrix W^k . We define the $L \times L$ edge counts matrix $E^k = \{e_{ij}\}_{i,j \in s}$, where e_{ij} represents the sum of functional connections between modules i and j for subject k . Since functional connectivity can be both positive and negative, we derive E^{k+} and E^{k-} by considering only the positive and negative edges separately. We obtain an edge counts matrix E^k for each subject, where lower triangular elements are taken from E^{k+} and upper triangular elements are taken from E^{k-} . For the intra-modular connections, we consider the positive and negative connections separately. We flatten E^k to give us a vector of features, e^k , for the classification model. We use different classification models including a feedforward neural network (FFN) and support vector machines. Considering a FFN with N layers, the output h_l of the hidden layer $l \neq N$ is given by:

$$h_l = f(V_l^T h_{l-1} + b_l) \quad (2.12)$$

where V_l and b_l denote the weights and biases of the layer l and f is the rectified linear unit (ReLU) activation function. For the input layer, $h_0 = e_k$.

The output layer $l = N$ is a softmax layer and the output probabilities y of the input e belonging to patient and healthy classes is given by

$$P(y = k|e) = \text{softmax}(V_N^T h_{N-1} + b_N) \quad (2.13)$$

The class label is assigned to the class with the highest probability. To learn the parameters of the network, the cross-entropy cost is defined as:

$$J(\theta) = -E_e[\log P(y = d|e, \theta)] \quad (2.14)$$

where E_e is the expectation taken over all scan samples e and $\theta = \{(V_l, b_l)\}_{l=1}^N$ denotes all the parameters in the network.

We used a two-class classifier for patients and healthy subjects. The weights and biases of the network are learnt using mini-batch stochastic gradient descent learning that minimises the cross-entropy cost. The networks were implemented in Python using the Keras [123] and Tensorflow library [124].

2.4 Results

2.4.1 Dataset

The primary data used in this study is from the Human Connectome Project (HCP) of the Washington University-Minnesota Consortium [125, 126] openly available at <https://www.humanconnectome.org>. The data was collected on a Siemens Skyra 3T scanner and included state-of-the-art whole-brain MRI acquisition with structural, functional, and diffusion-weighted imaging. Functional scans were acquired in two sessions with each session consisting of two 15 minutes resting-state run in different phase-encoding directions (left-right and right-left). We analyzed the scans of subjects who completed all four resting-state scan runs (two sessions, each with two encoding runs), which resulted in $n = 589$ subjects (males = 278, mean age = 28.33 years, range = 22 to 37 years). The HCP minimal processing pipeline [126, 127] was used for preprocessing, which included projection to the surface space, FWHM smoothing (2 mm), ICA+FIX denoising with minimal high-pass filtering and MSMall surface registration.

2.4.2 Brain Functional ROI

As functionally relevant regions of interest (ROI), we used the 264 brain regions identified by Power et al. [14]. The Power atlas contains both anatomically and functionally diverse brain regions and defines the functional role played by each ROI. We calculated the mean time series of all voxels within a sphere of radius 2.5 mm of each ROI. The Pearson correlation was used to compute the functional correlation between two ROI. We considered all the positive correlations between the ROIs, covering the entire cerebral cortex.

2.4.3 Group-level modules from multiple subject data

For each subject, the functional connectivity matrix was derived from the Pearson correlation of time-series between ROIs from each scan. The connectivity matrices from four scans were averaged to obtain the connectivity matrix for the subject. Power et al. found 14 modules and Yeo et al found 18 modules, and we assume that there are at least 5 functional systems discernible in the resting-state fMRI scans viz. sensory somatomotor, visual, default mode, cingulo-opercular task control, and fronto-parietal task control. We, therefore, fixed the minimum number of modules $L_{\min} = 5$ and started with $L_{\max} = 20$ for the ICSC algorithm. We increased $L_{\max}$ in steps of 5 and initiated 10 independent runs for each value. We observed a sharp drop in the histogram of $\{L^k\}$ distribution for $L_{\max} = \{20, 25\}$ and observed that for $L_{\max} = \{30, 35, 40\}$ the histogram for $\{L^k\}$ was smooth (Figure 2.4(a)). We also observed optimal consensus-cost at $L_{\max} = 30$ and therefore selected this value for the rest of the experiments. Refer to Appendix A (Figure A2) for more details.

We ran 100 independent runs of the ICSC algorithm until convergence on both the datasets. The number of iterations the ICSC procedure took to converge ranged from 4 to 19 (average 8.6 ± 3.5) for the HCP dataset. For the results of the algorithm to be reliable, the final group-level modules detected from the independent runs should be identical. We found that the AMI between the group-level modularizations of different runs was 0.84 ± 0.08 initially and 0.94 ± 0.02 at the convergence, which shows that the results are highly reproducible. The ICSC algorithm detected 20 group-level modules with the sizes of modules ranging from 6 to 18 (Figure 2.2). Out of the 100 independent runs, we selected the run with the maximum consensus-cost and study the individual-level and group-level modularizations for this run.

The group-level modularization found by the ICSC algorithm corresponds to previously defined functional networks of the human resting-state brain [14, 44, 128, 129]. We recovered several sub-modules from previously detected modules. We recovered the cingulo-opercular network containing regions from anterior cingulate areas and supplementary motor area (Figure 2.2(c)). The frontoparietal task control network was split into two subnetworks, left and right fronto-parietal task control networks (Figure 2.2(e) and Figure 2.2(n)). The ventral and dorsal attention networks were also found with the ventral attention network having regions from right superior temporal gyrus, right inferior and superior frontal gyrus (Figure 2.2(d)) and the dorsal attention network containing regions from the superior parietal lobe and superior precentral gyrus (Figure 2.2(q)). The default mode network (DMN) was split into three modules, a module with regions from the anterior and posterior cingulate (Figure 2.2(k)), a module with regions in the posterior cingulate cortex (PCC) and precuneus area (Figure 2.2(t)) and a module with regions in the dorso-lateral frontal, left middle temporal and lateral inferior parietal area 2.2(m)). The division of default mode network along medial and temporal regions is consistent with the prior literature [130]. The visual system was split into three modules, each representing a set of regions from medial (Figure 2.2(j)), lateral (Figure 2.2(h)) visual cortex and the occipital pole (Figure 2.2(r)), which is consistent with intensive anatomical differentiation of the visual cortex [131]. The medial and lateral visual modules are known to be involved in semantic language processing whereas occipital pole visual module is involved in space perception [132]. We also detected a language module containing regions from the left middle temporal and inferior frontal gyrus, dorsolateral superior frontal and cerebellum (Figure 2.2(f)). The sensory/somatomotor network was divided into different areas for feet (Figure 2.2(i)), hands (Figure 2.2(l)) and the mouth (Figure 2.2(s)) with regions lying in the precentral to postcentral gyrus. A module representing the memory retrieval system containing posterior medial regions from the cingulate gyrus and precuneus (Figure 2.2(v)) was also found. Finally, a module containing inferior regions from the temporal and frontal lobe (Figure 2.2(p)) was detected, whose function is unknown. We detected an independent module corresponding to the auditory system containing regions from superior temporal gyrus (Figure 2.2(j)). We also recovered a module containing subcortical regions from the cerebellum, thalamus, amygdala, and basal ganglia (Figure 2.2(g)). A module with regions from the middle frontal gyrus and anterior cingulate gyrus that form the salience network (Figure 2.2(o)) was also detected.

Fig. 2.1: The figure shows the similarity between the group-level modularizations computed for 3 folds of the HCP dataset and the whole HCP dataset. We use two different measures of similarity: the Adjusted Mutual Information (AMI) and the Normalized Mutual Information (NMI).

In order to understand the reproducibility of ICSC modularization on a subset of the dataset, we divided the HCP dataset into three folds, such that the average age of the subjects and the gender ratio in each fold is maintained. We ran 20 independent runs on the subject data in each fold and selected the run with the maximum consensus-cost. We found that the modularizations obtained from different folds were not only very similar to each other ($AMI = 0.87 \pm 0.0003$), but also to the one obtained with the whole dataset ($AMI = 0.92 \pm 0.01$). For details see figure 2.1.

We found considerable variation in the number of individual-level modules at convergence (mean = 20.91 ± 2.76), as shown in Figure 2.4(a), which needs to be taken into account while detecting group-level modular structure. Since many individual-level modularizations have greater or fewer number of modules than the group-level modules, we further studied which modules were different in individual-level modularizations. We found that the subcortical, language and ventral attention modules were subdivided into smaller modules for modularizations with high number of modules, whereas the PCC DMN, motor hand and feet, right fronto-parietal task control and salience modules were combined with other modules for modularizations with lesser number of modules.

Fig. 2.2: The twenty group-level modules detected by the ICSC algorithm on the resting-state fMRI scans from the HCP dataset. (a) The functional ROI colored according to the modules they belong to. (b) the group consensus matrix with reordered node indices to bring nodes in the same module together. (c) to (v) plotted using BrainNetViewer [133] show the ROIs belonging to each module ordered in the descending order of module size. The modules are given names based on functional networks identified by earlier studies or on the anatomical location of constituent regions.

2.4.4 Group-level modules from a validation dataset

We also used a validation dataset of 116 subjects provided by the Creativity and Affective Neuroscience Lab (led by Dr. Jiang Qiu), Brain Imaging Center of Southwest University. The dataset contained rs-fMRI scans collected from 3T MRI scanners for an 8-min period for 116 cognitively normal subjects with an average age of 23.54 ± 2.97 with 45 male participants at Xinnan (First Affiliated Hospital of Chongqing Medical School in Chongqing, China). We excluded the participants with age greater than 35 (to make it consistent with the HCP dataset). The preprocessing was performed using DPARSF [134] (<http://restfmri.net>), which is a toolbox based on the SPM8 software package. They discarded the first 10 EPI scans and the remaining scans underwent slice timing correction, motion correction, spatial normalization to standard Montreal Neurological Institute (MNI) space followed by spatial smoothing. To remove spurious correlations band-pass temporal filtering (0.01–0.1 Hz), nuisance signal removal from the ventricles and deep white matter and regression of 24 head motion parameters was applied. They also performed scrubbing movement correction [135] and removed subjects with mean framewise displacement greater than 10%. For more details on preprocessing, please refer to Cheng et al. [136].

We ran 100 independent runs of the ICSC algorithm on the validation dataset. The number of iterations the ICSC procedure took to converge ranged from 2 to 32 (average 9.3 ± 5.5). We select the run with the maximum consensus-cost and studied the group-level modularizations for this run. For the validation dataset also the ICSC algorithm detected 20 group-level modules with the sizes of modules ranging from 7 to 19. Although there are differences in the participant demography, scanning protocol and data preprocessing steps we obtained a similar modularization as with the HCP dataset (Figure A1 in Appendix A). We were able to detect sub-modules of the default mode network, the motor network, the visual network similar to the HCP dataset. We were also able to separate the auditory module from the motor module.

2.4.5 Inter-subject variability of functional modules

To study the inter-subject variability of brain functional modules, we computed nodal purities ψ_i using modularizations obtained from the ICSC algorithm. We found that nodal purity had a wide distribution (average 0.483 ± 0.184 , Figure 2.4(b)), showing that while some ROI are involved

Fig. 2.3: The inter-subject purity corresponding to different anatomical locations of the ROI (bigger size points to higher purity) along with their functional membership denoted by the ROI color.

in the same function across individuals, others are involved in varied individual-specific tasks. This is consistent with prior findings that inter-subject variability is non-uniformly distributed across the brain [101]. We found the regions in the occipital lobe, anterior/posterior central paracentral lobe have the highest purity while regions in the lateral frontal, subcortical, and anterior/post central cingulate gyrus have low purity (Figure 2.3). This is consistent from a brain evolutionary standpoint that the phylogenetically late developing frontal, temporal, and parietal association cortex areas involved in higher cognitive functions like reasoning lead to individual variability whereas the regions in the motor and visual cortex are consistent across individuals [137].

From a functional perspective, we observed that the ROI classified in the medial visual, occipital visual, and somatomotor hand modules had a consistently high purity whereas the ROIs in the fronto parietal, cingulo opercular and occipital default mode had high variation across individuals (Figure 2.4(c)). The low variability of visual and motor systems has been attributed to the high local information processing whereas modules like fronto-parietal, cingulo-opercular, and default mode are consistently involved in information processing between different functional systems [138]. The wide distribution in the purity of modules when studying inter-subject differences is consistent with previous findings on the variability in brain functional systems [85].

Fig. 2.4: Analysis of inter-subject and intra-subject variability of the ROI of the functions modules of the human brain. (a) shows the distribution of modules across subjects for group-level modularization. (b) shows the distribution of the purities of different ROI in the brain for inter-subject and intra-subject variability. (c) and (d) show the box plot for the purities of ROIs of different modules (green dotted line in the boxes shows the average purity for the module, blue dotted line shows the average for all the ROIs) for inter-subject and intra-subject variability, respectively.

2.4.6 Subject-level modules from multiple scans

To detect subject-level modules, we used multiple scans of the same subject from the HCP and computed the consensus modularization by using the ICSC algorithm. In this case, the individual modularizations, $s^{k\sigma}$, represent scan-level modules whereas the group-level modularization, s^k , represents subject-level modules. It took 1 to 4 (1.58 ± 0.7) iterations for the ICSC algorithm to converge for each subject in the HCP. We computed the similarity (using AMI) between scan-level modularizations of the same subject and between the scan-level and subject-level modularizations. The increase in similarity between scan-level modularizations at the start and convergence of ICSC algorithm was statistically significant (with a paired t-test p-value $< 10^{-51}$). Similarly, the increase in similarity between scan and subject-level modularizations at the start and convergence of ICSC algorithm was also statistically significant (with a paired t-test p-value $< 10^{-45}$). This shows that ICSC algorithm renders similar scan-level modularizations and a more representative subject-level modularization as a consensus among the scan-level modularizations.

For intra-subject variability of modules, we expected to see a high and similar purity values for the ROI since there is no variation of the regional functions across scans of the same subject. Correspondingly, we found a narrow spread of values for node purity (average 0.61 ± 0.09 as seen in Figure 2.4(b)). The spread of purity is higher for inter-subject variability than intra-subject variability. Intra-subject variability can be attributed to the noise artifacts in fMRI data and since ROI in the inferior temporal gyrus, inferior frontal gyrus and subcortical regions are known to have the highest noise confounds [85, 103], they should therefore have the lowest purity. Correspondingly, we found the subcortical and uncertain modules that contain nodes from these regions to have the lowest purity (Figure 2.4(d)).

2.4.7 Comparison with previous modularization methods

The performance of the ICSC algorithm was compared with previous approaches for detection of brain functional modules: InfoMap² that minimizes Code Length [79], Louvain³ that maximizes Newman's modularity [91, 93], and FAGSO⁴ that maximizes Asymptotical Surprise [42]. To

²www.mapequation.org

³<https://networkx.github.io/>

⁴<https://github.com/CarloNicolini/paco>

obtain group-level modules from different module detection methods, the consensus clustering for each method was used to obtain Π by obtaining s^k from 2000 iterations for each individual k . We reported the group-level modules s corresponding to the modularization with the optimal value of the fitness function: the modularity (Q) for Louvain algorithms, the Code length (C_L) for InfoMap, the Asymptotical Surprise (A_S) for FAGSO, and the consensus-cost for the ICSC algorithm.

For InfoMap, we thresholded the connectivity and retained a percentage of weighted edges in the connectivity matrices (from 5% to 20% at steps of 2.5%). On trying different thresholds to the data, we got the best value of InfoMap's objective function ($C_L = 7.185$ for retaining top 5 % edges). We obtained a modularization with 8 modules (containing more than one node) that represent motor merged with auditory system, visual, default mode, fronto-parietal task control and salience systems (Figure 2.5a). In case of the Louvain algorithm, we varied the resolution parameter from 0.3 to 1.5 (in steps of 0.1) and obtained $Q = 0.358$ for resolution parameter = 0.6, with 10 modules for the unthresholded weighted data. We observed that besides the motor, default mode, visual, fronto-parietal task control, and subcortical systems (Figure A5 in Appendix A), the Louvain algorithm was unable to detect other smaller brain functional modules. Even in the group consensus matrix detected using the Louvain algorithm (in Figure 2.5b), it can be seen that there are smaller clusters that are formed along the diagonal that the algorithm was unable to detect. The FAGSO algorithm found a group-level modularization (Figure 2.5c) consisting of four modules with more than one node and large number of isolated nodes, that did not provide much information about the brain's modular structure. We found that the ICSC algorithm was able to detect several functionally relevant, smaller modules that were found to be merged using these algorithms.

Previous studies in the area have used a thresholded and averaged group matrix to detect group-level modules. In order to replicate the previous methods reported in the literature, we performed modularizations on the group-level connectivity matrix obtained by averaging the thresholded individual connectivity matrices. Percolation analysis [139] was used for thresholding the averaged group matrices. We also used other different thresholds and performed modularization with Louvain, InfoMap and FAGSO. However, similar results were observed as with the matrix obtained from consensus clustering. The results are available in Appendix A (Figure A4).

Fig. 2.5: The group consensus matrices obtained for different modular detection approaches. Node indices were reordered to bring nodes in the same module together and the modules are arranged in descending order of their sizes. The red lines highlight the modules in the modularizations for the (a) InfoMap (5% top edges, removing isolated nodes), (b) Louvain (resolution parameter = 0.6), (c) Asymptotical Surprise, and (d) ICSC algorithms.

We calculated the average similarity between individual-level modules and group-level modules obtained from the group consensus matrix with the InfoMap, Louvain, FAGSO and ICSC algorithms (Figure 2.6(b)). The most representative s should have the highest similarity with $\{s^k\}$. Group-level modularization from the ICSC algorithm had the highest similarity to the individual-level modularizations. It is also to be noted that the variability in the similarity is the lowest for the ICSC algorithm, further indicating that the group-level modularization obtained from the ICSC algorithm is *consistently* similar to the individual-level modularizations.

We also compared the number of individual-level modules obtained with the best group partition for different modularization algorithms 2.6(c). The large number of individual-level modules found for FAGSO, InfoMap and Louvain are neuroscientifically implausible and contain a large number of isolated nodes. For ICSC, although we varied the individual-level modules s^k in a fixed range, we discovered a varied number of modules for individuals 2.4(a), which points to variability in individual modular structure.

2.4.8 Experiments with synthetic data

In order to get quantitative assessment of our methods, we generated synthetic data depicting human brain connectivity. We used the weighted Lancichinetti-Fortunato-Radicchi (LFR) benchmark networks [140, 141] that can mimic properties of resting-state brain networks such as scale-freeness, modules with heterogeneous sizes, and densely interconnected modules.

Fig. 2.6: Group-level modularizations for different module detection approaches. (a) the cost of ICSC algorithm after different iterations for 100 runs with random initialization, (b) the box plot of adjusted mutual information (AMI) between group-level modularizations and individual-level modularizations for the Asymptotical Surprise, Louvain, InfoMap and ICSC algorithms (green dotted line shows the average), and (c) the number of individual-level modules for different modularization algorithms.

To generate synthetic networks similar to resting-state brain networks, we studied the properties of the HCP data under consideration. The parameters such as modular sizes and distribution were derived from previously defined functional modules on resting-state fMRI data [14]. Accordingly, synthetic LFR benchmark networks were built with $|\Omega| = 264$, average degree = 25, maximum degree = 53, minimum degree = 1, degree distribution exponent = 1.13, module size distribution exponent = 1.36, weight distribution exponent = 1.5 and community sizes varying from 5 to 58, proportion of weights inside community $\mu_w = 0.5$, and proportion of inter modular edges $\mu_t = 0.5$. We get the Cholesky Decomposition, $C = LL^T$, of the nearest positive definite matrix, C , from LFR benchmark network. We generate $X = \{x_i^t\}_{i \in \Omega, 0 \leq t < T}$, a matrix with $T = 1200$ time points for each region using neuRosim R package and obtain $Y = LX$. We add Rician noise to Y and calculate correlation between each region. We varied the signal-to-noise ratio (SNR) by varying the amount of noise added to the network. Since the ground-truth data is available for synthetic data, we used Matthew Correlation Coefficient (MCC) to measure the level of agreement between the detected modules and the ground-truth modules. Given true positives (tp), true negatives (tn), false positives (fp), and false negatives (fn), the Matthew Correlation Coefficient (MCC) is given by

$$MCC = \frac{tp \times tn - fp \times fn}{\sqrt{(tp + fp)(tp + fn)(tn + fp)(tn + fn)}}$$

To study the accuracy of group-level modularizations, we generated synthetic brain network datasets representing multiple individuals. Inter-subject differences cause same regions in the brain to participate in different modules across different individuals. To simulate this, we varied the module memberships for nodes that are most likely to be in alternate modules. We identified such nodes by using their betweenness centrality score [142] that gives the fraction of shortest paths between all nodes passing through a node. We selected a fraction of nodes with high betweenness centrality and varied their modular membership across individuals. To choose an alternate modular assignment for a node, its clustering coefficient was used with a subgraph formed with each module. In addition, we varied the amount of Rician noise added to the network and evaluated the performance of the ICSC algorithm at different signal-to-noise ratios (SNR).

One hundred runs of the ICSC algorithm were launched on a set of 100 synthetic networks. Initially, a random number of modules were assigned to the individuals in each run and the MCC and the AMI values for the group and individual-level modularizations relative to the ground-truth were noted. We varied the amount of noise introduced in the network and measured the similarity between $\{s^k\}$ and s to the ground-truth. We observed that the individual-level modularizations $\{s^k\}$ improve with decreasing noise, while the group-level modules s were close to the ground-truth even for low SNR ($\text{SNR} \geq 30$, Figure 2.7(b) and 2.7(a)). We also observed that the number of group-level modules converged to exactly 14 (number of modules in ground-truth) at the final iteration for all SNRs, indicating the reproducibility of the ICSC algorithm.

To understand the effect of the variability in module assignment to nodes across individuals, we varied the average purity of the nodes. We observed that while the individual-level modularizations $\{s^k\}$ are close to ground-truth at only high purities, the group-level modularizations s have a high similarity to the ground-truth throughout (Figure 2.8(b) and 2.8(a)). This shows that the ICSC algorithm is able to detect group-level modularizations reliably irrespective of the purity of nodes in the network. As was the case with noise, it was seen that the ICSC algorithm was able to detect the true number of group-level modules in the data at convergence.

For quantitatively evaluating the reproducibility of the ICSC algorithm in deriving subject-level modules from multiple scans of the same subject, we generated synthetic data for subjects with varying number of scans. We assume that the ROI are involved in the same module in a

Fig. 2.7: Performances of the ICSC algorithm in deriving group-level and individual-level modules from synthetic data, where (a) shows the Adjusted Mutual Information and (b) shows the Matthew Correlation Coefficient between the ground truth and detected modularizations with varying purity for nodes at SNR = 100.

Fig. 2.8: Performances for the ICSC algorithm in deriving group-level and individual-level modules from synthetic data, where (a) shows the Adjusted Mutual Information and (b) shows the Matthew Correlation Coefficient between the ground truth and detected modularizations with varying amounts of noise added to the individual scans.

Fig. 2.9: Performances of the ICSC algorithm in deriving group-level and individual-level modules from synthetic data, where (a) shows the Adjusted Mutual Information and (b) shows the Matthew Correlation Coefficient for module detection with different number of scans.

Table 2.1: MDD subject details, Hamilton Depression Rating Scale (HAM-D)

	Cognitively Normal (CN)	Major Depressive Disorder (MDD)
No. of subjects	116	112
Male/Female	45/71	41/71
Age	23.54 ± 2.97	26.20 ± 5.28
HAM-D score	-	20.03 ± 5.38

subject but noise in the connectivity data was varied for different scans. To mimic functional connectivity data for a single subject having multiple scans, we generated 100 LFR benchmark networks (representing different subjects) and generated multiple networks (representing scan sessions) by adding noise to each benchmark network. We generated 2, 4, 6, 8, 12, 16 and 20 networks representing different number of scan sessions for each LFR benchmark network. We report the MCC of the detected modularizations with respect to the ground-truth. We observed that the detection of subject-level modularizations s^k improved as the number of scans per subject increased whereas the scan-level modularizations $s^{k\sigma}$ had similar performance irrespective of the number of scans (Figure 2.9(b) and 2.9(a)).

2.4.9 Using ICSC modularizations for disease detection

We hypothesize that there are widespread changes in between and within module functional connectivity in patients suffering from MDD and we can therefore use these for detecting MDD. We demonstrate this by using the MDD dataset provided by the Creativity and Affective Neuroscience Lab, led by Dr. Jiang Qiu, Brain Imaging Center of Southwest University. The normal subjects from this dataset were also used as validation dataset (described in section 2.4.4). We used the modules obtained from the CN subjects to obtain the edge counts matrices, E^k , for all the subjects. As a result, we obtained 20×20 matrices. Since E^k is symmetric, we formed a vector of its lower triangular elements (and the diagonal elements) used the flattened lower triangular matrix to derive the feature set for training the classifier.

We used the derived features to perform classification with FFN and Support Vector Machine (SVM) models. We divide the data into a train and test set in the ratio 4:1. We perform model selection on the train set by performing 5 fold cross validation on it. For the FFN, we vary the number of hidden layers, number of neurons in each layer, batch size, and learning rate in the experiments. We added dropouts to the hidden layers and imposed early stopping criteria to prevent overfitting. For SVMs, we experimented with a range of parameters ($C \in \{0.001, 0.01, 0.1, 1, 10\}$, $\gamma \in \{0.001, 0.01, 0.1, 1\}$) and different kernels. The best hyperparameters were selected by performing a grid search on the average accuracy of different folds obtained from nested 5-fold cross validation. The performance was also compared with a classifier trained on functional connectivity features from each subject's unthresholded connectivity matrix W . Since W is symmetric, we formed a vector of its lower triangular elements, which results in a vector of 34716 functional connectivity features for each subject.

For the FFN models based on modularization features, the highest accuracy was obtained with 3 hidden layers (100, 50, 10 neurons each), a batch size of 4, dropout of 0.1, and learning rate of 0.04. The performance of the FFN along with the best SVM model for different kernels are shown in Table 2.2. For the modularization features, we observed that the highest accuracy was obtained for the FFN using E^{k-} as features. For the FFN models based on functional connectivity features, the highest accuracy was obtained with 3 hidden layers (1000, 64, 32 neurons each), a batch size of 8, dropout of 0.1, and learning rate of 0.0001. We obtained a higher accuracy for SVM model (linear kernel, $C = 0.1$) for functional connectivity features. Although the functional connectivity based features gave a slightly higher accuracy than modularization

Table 2.2: Classification accuracies for MDD and CN subjects using deep neural networks and module features as features. * refers to a statistically significant result in comparison to all the others.

Model	Accuracy
Positive edge counts E^{k+} only	
FFN (3 hidden layer of 100, 50, 10)	61.7 %
SVM (Sigmoid, $C = 10$, $\gamma = 0.001$)	56.5%
Negative edge counts E^{k-} only	
FFN (3 hidden layers of 100, 50, 10)	64.5%
SVM (Sigmoid, $C = 10$, $\gamma = 0.001$)	63.1%
Both positive and negative edge counts E^k	
FFN (3 hidden layers of 100, 50, 10)	62.4%
SVM (Sigmoid, $C = 10$, $\gamma = 0.001$)	60.3%
Functional connectivity, W^k	
FFN (3 hidden layers of 1000, 64, 32)	64.8%
SVM (Linear, $C = 0.1$)	65.7%

features (65.7% vs 64.5%), the improvement in classification accuracy was not statistically significant (one tailed t-test, p-value = 0.67).

2.5 Discussion

2.5.1 Detecting brain functional modules by the ICSC

We demonstrated the use of ICSC algorithm for detecting brain functional modules on 589 healthy individuals from HCP at the group-level, subject-level, and scan-level. The functional modules were derived from dense and weighted functional connectivity matrices derived from resting-state fMRI scans. At the group-level, we obtained 20 functional modules of heterogeneous module sizes, that have close correspondence to the functional modules discussed in the literature. Not only did we detect previously reported major functional networks, we were able to detect the associated sub-modules in these networks. For example, we were able to differentiate sensory somatomotor network into well-defined modules for feet, hands and mouth; the visual network was also differentiated into three anatomically distinct modules;

we were also able to separate the auditory and motor modules that have been found to be clubbed together [44, 143] and we detected a separate module that corresponds to the language network of the brain. The derived group-level modularization by the ICSC algorithm is close to the modularization derived using meta-analyses of functional connectivity data [14]. For quantitative evaluation, we resorted to the use of synthetic networks with known underlying modular structure. To create synthetic data closest to human functional connectivity, we used the LFR networks and generated networks with heterogeneous nodal degree distributions [144] and modules with heterogeneous sizes [109]. The inter-subject variability was attributed to the modular membership of nodes across individuals and noise. The ICSC algorithm detected the ground-truth modularization with high accuracy and was robust to initialization. We found that the ICSC algorithm was able to detect group-level modularizations close to the ground truth when $\text{SNR} \geq 30$, for different amounts of inter-subject variability, whereas the HCP data has a mean $\text{SNR} = 728$ (minimum $\text{SNR} = 227.832$, $\text{SNR} = \text{ratio of mean and standard deviation of the time series signal strength}$), which gives further credibility to the group-level modularization.

We have shown that the derivation of group-level modules by ICSC is reproducible since multiple independent runs of the algorithm produced very similar group-level modules. It is to be noted that although the ICSC takes the range of modules (L_{min}, L_{max}) as user parameters, the upper bound L_{max} can be found systematically. We also observed that selecting L_{max} beyond a reasonable number (30, in this case) doesn't change the obtained group-level modularization, but only makes the modularization procedure slower. This makes the ICSC user-friendly since it doesn't require the user to try out different number of modules or thresholds and does not need to be stopped manually. The derivation of individual-level modularizations is the most expensive step in the algorithm, which can be run parallelly for different subjects, thus making the algorithm fast.

2.5.2 Inter-subject variability of functional modules

We used the individual-level and group-level modularizations to study the variability in functional architecture across subjects and comparing our results with prior work in the area. The variation in functional membership of brain regions in the frontal, temporal and parietal association cortices is believed to be due to a prolonged period of maturation of white and gray matter [145], which renders them susceptible to diverse extrinsic individual experiences during

a long period of high neuroplasticity [146]. On the other hand, early-developing brain regions that participate in elementary functions like somatomotor/sensory and vision have been found to be stable across individuals [44]. It is also to be noted inter-subject variability in functional connectivity due to aging are well-documented [147, 148], but our participants have a narrow age range (mean 28.33 ± 3.7) and therefore age-related differences can be disregarded. Upon investigation of the inter-subject variability of functional modules, we uncovered heterogeneous variability in functional membership of regions depending on their functions. We found the subcortical, auditory, dorsal and ventral attentional networks along with inferior brain regions to have high variability across subjects. The attentional networks are located in the association cortices and are therefore highly variable, whereas the variability of basal brain regions can be attributed to the noise in fMRI acquisition due to strong magnetic susceptibility artifacts. Although the high variability in auditory network has not been reported before, but previous studies have reported that auditory network is hard to detect and often fused with a ventral somatomotor network [44, 149]. In contrast, the three motor networks for hands, feet and mouth and the visual systems were found to be highly stable across individual-level modularizations. These findings are consistent with the earlier studies that derived inter-subject variability using the intraclass correlation measure [85, 103, 150].

2.5.3 Using ICSC algorithm for deriving subject-level modularization

Functional modularizations detected at subject-level have wide applications in clinical practice and personalized medicine. However, noise and different cognitive mind states during fMRI scan acquisition introduce variability across fMRI scans of a subject [85, 151]. We proposed to use the ICSC algorithm to detect a subject's modularization from multiple scans. We experiment with the HCP dataset comprising of four scans per subject and with synthetic datasets comprising of different numbers of scans per subject. We observed a statistically significant improvement in similarity between the subject-level and scan-level modularizations at the initial and final iterations of the ICSC algorithm, demonstrating how iterations in the ICSC algorithm improve the subject-level and scan-level modularizations. For the synthetic dataset, we found an increasing agreement between the ground-truth and the subject-level modularization as the number of scans per individual is increased. It is important to note that the initial iteration of the ICSC algorithm corresponds to simple consensus clustering [86] and

our experiments with HCP dataset depict how the multiple modular alignments in the ICSC give more representative subject-level modules from multiple fMRI scans of a subject.

2.5.4 Comparison of ICSC algorithm with previous studies and methods

We also performed a comparison of the performance of ICSC with modularization methods used previously. Previous studies have used thresholding [14, 42–44, 143], averaging [41], and/or binarizing of functional connectivity matrices for detection of functional modules, which leads to loss of information and introduces artifacts to their results. Although averaging of networks or connectivity matrices has traditionally been performed to compensate for the low SNR of fMRI acquisitions, it ignores the inter-subject and inter-scan variability and leads to spurious and missing functional connections. Present fMRI scans have high SNR due to long acquisition times and lower repetition time so much so that individual-level modularizations have been found to better agree with the individuals’ behavioral characteristics [152]. It is, therefore, necessary to use complete connectivity information of individuals to derive modularizations for functional brain connectivity data. Our results also show that although we do not impose explicit conditions on detection of specific brain functional modules as in [41], we were successfully able to detect separate modules for the auditory and somatomotor brain systems. Another approach to detect functional modules in the brain is by using time-series of pre-selected seed voxels against the time-series of all other voxels [153]. However, the output of the seed-based approaches are highly dependent on the number and location of initial selected seed regions. In contrast, our approach computes the number of functional modules in a data-driven manner at both individual and group-level and is highly reproducible over independent runs.

The difference in performance between ICSC and previously used algorithms can also be attributed to the quality function being optimized. While, the maximization of ICSC quality function focuses on finding the most representative group-level modularization, other optimization functions find optimized individual-level modularizations. We compared the performance of the ICSC algorithm with several previous used modular detection algorithm, where we derived group-level modularizations with both consensus clustering and averaged subject matrices. For dense unthresholded networks, InfoMap detected modules that correspond to known functional brain systems, but it was unable to identify smaller functional systems in the network. We observed that high density of connections lead to a very few known large

sized modules while low thresholds lead to many isolated nodes in the network. For Louvain, we varied the resolution parameter and optimized the modularity of group-level modules. We obtained group-level modules corresponding to known brain functional systems, but failed to resolve several smaller systems in the brain. As described before, this is a consequence of the resolution bias of the quality functions of these algorithms, which depends on the number of connections in the network [96, 99]. Although, the normalized cut objective function used in ICSC algorithm also suffers from this resolution bias, we were able to avoid it by specifying a uniform range (L_{min}, L_{max}) for expected number of modules for the modularization. Finding such a range is not trivial in case of Louvain and InfoMap because it would have to be computed for each individual functional connectivity matrix. Finally, we computed modularizations by using FAGSO algorithm that maximizes Asymptotical Surprise, which has shown to be free from resolution bias [42]. We obtained a group-level modularization containing a few large modules and several small modules composed of a few or isolated nodes.

It should be noted that although weak functional connections have been found to be predictive of intelligence and psychiatric illnesses [116–118], their effects on modular structure would not be considered due to inability of present module detection methods to work in their presence. The ICSC algorithm considers weak edges in detecting the modular structure and may therefore highlight anomalous functional systems in the brain. Besides showing the neuroscientifically plausibility of the group-level modules computed by ICSC algorithm, we show that the modularization was the closest to the individual-level modularizations, which shows that the ICSC algorithm gives the most representative modular structure. In this chapter, we showed that the modularizations derived from the ICSC algorithm can be used for detecting diseases. While most of the previous studies studied the disruption in modular structure during neurological disorders in qualitative terms [74, 75, 119–122], we use the modular structure to classify subjects as healthy and diseased using the change in their inter-modular and intra-modular connectivity. We found that although the coarse grained features obtained from brain modularization are much fewer in number (≈ 400), they give similar classification performance as whole brain functional connectivity features ($\approx 34,000$) for detecting major depressive disorder. In the light of the above, the ICSC algorithm becomes a powerful tool to study brain's functional architecture during different brain states.

2.6 Chapter Summary

In this chapter, we introduce the significance of module detection in the brain functional connectome. We proposed the ICSC algorithm that tackles the pitfalls of previously proposed algorithms in the domain. We employed a multipronged strategy to show the validity of the modularizations obtained from the ICSC algorithm. We show that the modules detected by the ICSC algorithm are biologically relevant on the HCP dataset and a validation dataset. The performance of our algorithm was compared with existing functional brain modularization methods and we show that our method detects group-level modules that are more representative of the modules of multiple individuals. We show a heterogeneous variability in the modular structure across subjects where modules involved in visual and motor processing were highly stable across subjects. Conversely, we found a lower variability across scans of the same subject. We also performed experiments on synthetic images to quantitatively demonstrate that the ICSC algorithm detects group-level and subject-level modules accurately under varied conditions.

Finally, we demonstrate an application of the detected brain modularizations for detecting neurological ailments. Therefore, besides identifying functional modules for a population of subjects, the proposed method can be used for applications in personalized neuroscience. While the present chapter focused on detecting brain functional modules that represent the functional specialization property of the brain, in the following chapter, we propose methods to detect brain hubs that represent functional integration in the brain.

Chapter 3

Ambivert degree for detecting functional hubs of the brain ¹

In the previous chapter we proposed an algorithm to detect the modules in the brain functional connectome. While the functional modules in the human brain support its drive for specialization, there are specific brain regions, known as brain hubs, that act as focal points for information integration by virtue of their large number of both within and between module connections. In this chapter, we propose a new measure called ambivert degree to detect brain hubs and use brain hubs as biomarkers to detect diseases that are characterized by hub disruption. We argue that weak connections in brain functional connectome lead to misclassification of brain regions as hubs. Ambivert degree considers the node's degree as well as connection weights in order to identify nodes with both high degree and high connection weights as hubs. Using resting-state fMRI scans from the Human Connectome Project, we show that ambivert degree identifies brain hubs that are not only crucial but also invariable across subjects. We hypothesize that nodal measures based on ambivert degree can be effectively used to classify patients from healthy controls for diseases that are known to have widespread hub disruption. Using patient data for Alzheimer's disease and Autism spectrum disorder, we show that the hubs in the patient and healthy groups are very different for both the diseases and deep feedforward neural networks trained on nodal hub features lead to a significantly higher classification accuracy with significantly fewer trainable weights compared to using functional connectivity features. Thus, the ambivert degree improves identification of crucial brain hubs in healthy subjects and can be used as a diagnostic feature to detect neurological diseases where hubs are disorganized.

¹The work in this chapter has been published in [109, 144, 154]

3.1 Existing work on brain's functional hubs

Multiple anatomical, physiological and neuroimaging studies have pointed to the coordination between functionally diverse specialized systems, resulting in cognition and behavior [155]. This allows the brain to coordinate and perform varied cognitive tasks such as language perception [156] and vision [157], that involve inputs from multiple brain regions. This coordination or integration of information between neurons in these regions is facilitated by a specific set of 'hub' regions that form the brain's information transmission backbone [158].

Studies indicate that functional connectome of the human brain incorporates properties that promote functional specialization with a modular structure (for review refer to [36]) and efficient communication by virtue of network hubs (for review refer to [45]). A recent study found that hubs are differentiated into three classes based on their connectivity to different functional modules [159] and were shown to modulate different tasks. The central role of hubs in information processing makes their identification an important research problem.

Brain hubs were initially characterized by the nodes having large numbers of connections to other networks, that is, nodes having high nodal degrees [46–48]. However, later studies argued that since the brain modules had heterogeneous sizes and nodes had high within-module connections, simply taking the nodal degree would lead to giving undue weightage to nodes from large modules (usually the default mode network) [72]. Therefore, the nodes with high degree in comparison to the nodes in the same module [72] were considered crucial for intra-modular communication and were termed as *modular hubs*. Conversely, the participation coefficient, a measure that quantifies the extent a node connects to other modules was proposed to detect heteromodal nodes facilitating inter-modular communication known as *connector hubs* [42, 72, 159]. The weighted average of scores derived from the two measures identified nodes with high intra-modular degree and heterogeneous connections often termed as *network hubs*.

Studies involving the functional connectome involve network sparsification steps that remove weak brain functional connections that are affected by experimental noise [160]. However, this affects not only the number of weak connections in the network but also the network's underlying modular structure [52]. This can in turn affect the detection of network hubs because (i) disruption of brain modular structure leads to spurious detection of hubs; and (ii)

a node may possess a very large number of weak connections and still be classified as a hub whereas a node with fewer but stronger connections may be ignored. In order to account for (i), we use the thresholding scheme proposed by Bordier et al. [52], which preserves the network's modular structure; and (ii), we propose a new measure called *ambivert degree* that considers both the nodal degree and the strength of connections. We also consider an extension of the participation coefficient, called *gateway coefficient*, that considers information about the importance of a node's neighbors in their module and the uniqueness of the node's inter-modular connections [161]. Considering both of these measures, we provide a new definition for brain functional hubs and study their importance in processing information in resting state brain functional connectomes of 589 subjects from the HCP. By considering the effect of inducing artificial lesions in the brain functional networks, we found that the ambivert degree gives the most crucial modular hubs and instead of the node's intra-modular degree (as previously thought), the ambivert degree combined with participation coefficient gives the most crucial network hubs.

Hubs are known to support brain information processing and alterations in the functional connectivity of hubs have global effect on brain network and function [162, 163]. Previous research has shown that various neurological ailments such as Alzheimer's disease (AD) [164], Schizophrenia [165, 166], and autism spectrum disorder (ASD) [167] affect brain functional hubs disproportionately and more than the other brain regions. These disruptions reflected in hub connectivity could be a result of the high information processing at the hub regions, causing high baseline activity and metabolic requirements of hub regions [168, 169] leaving the hub neurons vulnerable to metabolic stress or degeneration due to high activity [170]. In case of ASD, a developmental disorder characterized by a range of heterogeneous symptoms with varying degrees of severity, multiple resting state fMRI studies have found widespread impairment in intrinsic functional connectivity (for a review, see [171]), especially in the hubs [167]. Similarly, research suggests that highly active heteromodal regions are preferentially affected in Alzheimer's disease (AD). Buckner et al. [46] used a node's degree centrality to define brain hubs and found that degree hubs obtained from healthy human brains have a significant overlap with brain regions having higher Amyloid- β ($A\beta$) deposition AD patients. Willem et al. [172] found that excessive neuronal activity that led to degeneration provides a possible explanation for hub vulnerability in AD. The $A\beta$ cascade hypothesis suggests that $A\beta$

have a toxic effect on adjacent neurons and synapses, disrupting their normal functioning [173] in AD.

Based on prior studies, we hypothesized that nodal hub scores can be used to classify AD and ASD subjects from their respective healthy subject cohort. We apply our methods on resting state fMRI data for AD and ASD subjects to demonstrate how brain hubs can aid in diagnosis of brain diseases. We chose these AD and ASD for our analysis because of existing evidences of hub disruptions and availability of large sets of data for deep neural network modeling. We detected brain hubs from the functional scans and found that the brain hubs of AD and ASD patients were different from cognitively normal controls. Using both the nodal hub scores and functional connectivity as input features, we trained neural networks and SVM classifiers, and show that using hub scores gave a significantly higher classification accuracy with a much lower number of trainable weights.

3.2 Detecting brain's functional hubs

Let $G = (\Omega, W)$ denote the functional brain connectome where Ω denotes the set of brain regions of interest (ROI) or nodes and $W = \{w_{ij}\}$ is the matrix of edge weights of the network with $w_{i,j}$ denoting the weight between brain ROI i and j , and $i, j \in \Omega$.

3.2.1 Detecting brain functional modules

While some previous studies detected brain modules and hubs by using subject averaged functional connectivity matrix [42, 72, 174], there are multiple studies pointing out varied individual differences in the functional connectome [101]. Therefore, inter-subject differences in functional architecture need to be considered while detecting functional modules (and subsequently network hubs) and combined group results from all the subjects need to be reported [159]. We, therefore, use the Iterative Consensus Spectral Clustering (ICSC) algorithm proposed in Chapter 2 that maximizes the similarity between group-level and subject-level modules, iteratively making them more similar. The ICSC minimizes the normalized-cut objective function to detect subject-level modularizations S^k of subject k . Using the subject-level modularizations from ICSC, modular labels assigned to nodes for each subject were aligned with those at the group-level modularization. We align the subject-level modules, S^k , with the group-level modules S

by using the Jaccard Index (JI) matching scheme [86]. A module with index b in S^k is assigned to a group-level module label a if for both the group-level module a , $JI(a, b) > JI(a, c)_{c \in S^k \setminus b}$ and subject-level module b , $JI(b, a) > JI(b, c)_{c \in S \setminus b}$.

3.2.2 Thresholding functional connectivity

While our modularization method considers weighted unthresholded subject-level networks, weak connections that are most influenced by experimental noise [62, 160] need to be removed for further analysis. Researchers often select an *a priori* density for the network and compute a threshold that retains the target density of edges [72, 159, 174]. Analyses over a range of densities have been performed to show the stability of derived topological features over a range of thresholds. However, weak correlations could still contain significant functional information, and therefore it is important to use methods that identify the optimal trade-off between the information gain by the removal of noisy edges and the loss due to removal of potentially useful weak edges. Bordier et al. [52] used synthetic networks with known modular structure and noise similar to brain functional networks and showed that percolation analysis [139], retains the minimum number of edges to keep the network connected while maximizing information on brain functional networks' modular structure. Also, fixing the same edge density in brain functional networks of subject-groups may lead to inclusion of weak and spurious connections in a subject-group with disrupted functional connectivity and omission of strong connections in another subject-group. The latter becomes especially significant in our case since we perform analysis at the subject-level rather than at group-level. We perform percolation analysis on each subject's functional connectome to determine the optimal threshold t for weights and assign zero weights for connections with weights less than t .

3.2.3 Degree distribution of the modules

Nodal degree distribution of the brain functional connectome has been studied previously [144, 175] in order to understand the connectivity pattern and determine the frequency of high degree nodes in the whole network [176]. This helps in the identification of nodes with high connectivity at the network level and understanding the principles behind the evolution of the network structure. For hub detection, we do not consider highly connected nodes at the network level, but at the modular level. We, therefore, find the best fit nodal degree distribution for

the modules for each individual subject by fitting the power-law, exponential and stretched exponential probability distributions to get a sense of the nodal degree distribution at the modular level. The probability of finding a node with degree θ in a network with a power-law distribution is given by [177]:

$$p_{\text{power}}(\theta) = C_1 \theta^{-a}$$

where a is the scaling parameter, $C_1 = (a - 1) \theta_{\min}^{a-1}$, and $\theta_{\min}$ is the minimum degree that obeys the power-law.

The networks with exponential degree distribution are defined by

$$p_{\text{exponential}}(\theta) = C_2 \exp(-b\theta)$$

where b is the decay parameter and $C_2 = b \exp(b\theta_{\min})$.

The stretched exponential networks combine power-law and exponential degree distributions and are defined by:

$$p_{\text{stretched}}(\theta) = C_3 \theta^{a-1} \exp(-b\theta^a)$$

where $C_3 = ab \exp(b\theta_{\min}^a)$ and b is the decay parameter.

The log normal degree distributions are defined by:

$$p_{\text{log-normal}}(\theta) = C_4 \frac{1}{\theta} \exp\left(-\frac{\ln \theta - a}{2b^2}\right)$$

where $C_4 = \sqrt{\frac{2}{\pi b^2}} \left[\text{erfc}\left(\frac{\ln \theta_{\min} - a}{\sqrt{2}b}\right) \right]^{-1}$ where a and b^2 represent the mean and variance of the distribution, and erfc is the complement of the error function.

The parameters $\{\theta_{\min}, a, b\}$ were estimated using Kolmogorov Smirnov statistics and maximum likelihood estimates as described in [177]. We use bayesian information criterion (BIC) for model selection in order to account for model complexity and avoid overfitting:

$$BIC = N \ln(K) - \ln(p_{\text{model}}(\theta)) \quad (3.1)$$

where K is the number of nodes in the module, N is the number of parameters to be estimated in the model and p_{model} is the likelihood of the fitted model. We calculated the likelihood and BIC corresponding to each distribution for every module (for each subject) and used Tukey's pair t-test to find the significance of the results. Higher exponents (a , b and a in case of power law, exponential and stretched exponential, respectively) point to a more homogeneous degree distribution [178], and therefore fewer hubs. We take care of the fact that adequate number of samples (nodal degree values, in this case) are provided for the degree distribution fitting because too few samples can result in bias in the results.

3.2.4 Modular hubs

The brain regions most critical for information processing within a brain module are known as *modular hubs*. While previous studies have assumed that modular hubs are defined by their intra-modular degree, we wanted to empirically find modular hubs.

3.2.4.1 Intra-modular degree

The nodal degree of a node $i \in \Omega$ in the functional connectome is defined as

$$\theta_i = \sum_{j \in \Omega} w_{ij} \quad (3.2)$$

where functional connectivity matrix $W = \{w_{ij}\}_{i,j \in \Omega}$. The number of connections is given by $n_i = \sum_{j \in \Omega} 1(w_{ij} > 0)$. The nodal degree θ_i and the number of connections n_i represent a node's global network property.

With respect to a set of nodes in S , the degree $\theta_{i,S}$ for node i is given by:

$$\theta_{i,S} = \sum_{j \in S} w_{ij} \quad (3.3)$$

and $n_{i,S} = \sum_{j \in S} 1(w_{ij} > 0)$, where $1(\cdot)$ is the indicator function. The nodes in the set S can represent a module from the network or all the nodes in the network. When the node $i \in S$, we refer to $\theta_{i,S}$ as the *intra-modular degree*.

3.2.4.2 Ambivert degree

Nodes with high intra-modular degree are characterized by a large number of weak or strong intra-modular weights (determined by correlations of brain activations), or a combination of both. Although, nodes with a large number of weak correlations of brain activations do not convey meaningful information, under the existing techniques, these nodes are also classified as modular hubs (see Figure 3.1). These nodes can be thought to be *extroverts* in social networks, i.e., nodes with a large number of weak ties to other nodes. We, on the other hand, are interested in *ambiverts*, i.e., nodes that not only connect with a sufficiently large number of nodes but also share meaningful relationships with them. For this purpose, we propose a measure called *ambivert degree* that identifies high intra-modular degree nodes that are synchronized with

Fig. 3.1: Two cases where both the nodes have the same degree, but one node has a large number of connections with low weight while the other has fewer connections but with a higher weight.

multiple nodes in the same module. The ambivert degree of node $i \in S$ in module S is defined as

$$\alpha_{i,S} = \theta_{i,S} \left(\frac{\theta_{i,S}}{n_{i,S}} \right) \quad (3.4)$$

The ambivert degree is the intramodular degree weighted by average weight per connection.

3.2.4.3 Betweenness centrality

Besides the intra-modular degree, we also compute the intra-modular *betweenness centrality* $\beta_{i,S}$ of a node i in module S to find nodes that lie on the shortest paths for information processing [179]:

$$\beta_{i,S} = \sum_{j,k \in S} \frac{\sigma_{jk}(i)}{\sigma_{jk}} \quad (3.5)$$

where σ_{jk} measures the number of shortest paths between j and k and $\sigma_{jk}(i)$ measures the number of shortest paths between j and k , which pass through node i .

3.2.4.4 Detection of modular hubs

Modular hubs are characterised by nodes having significantly high intra-modular connections and detected by using intra-modular degree, ambivert degree, or betweenness centrality. We compute the z-score $\zeta(m_i)$ by considering a univariate normal distribution of different modular hub measures $m_i \in \{\theta_{i,S}, \alpha_{i,S}, \beta_{i,S}\}$ for a node $i \in S$ over the module S . The modular hubs are

characterized by the nodes that have significantly higher hub measures than the other nodes in the module.

3.2.5 Connector hubs

Modular hubs are central to the information processing within their respective modules. However, hubs not only play a role in intra-modular information processing but also act as bridges or *connectors* for inter-modular information transmission. The *connector hubs* refer to the nodes that represent heteromodal brain regions that participate in information processing between different functional modules. Traditionally, the participation coefficient has been used to measure the role of a node as a connector hub. Here we explore the participation coefficient as well as the *gateway coefficient*, a recent extension of the participation coefficient, to identify connector hubs.

3.2.5.1 Participation coefficient

Connector hubs are quantitatively identified using the *participation coefficient* p_i of the node i [180]:

$$p_i = 1 - \sum_S \left(\frac{\theta_{i,S}}{\theta_i} \right)^2 \quad (3.6)$$

3.2.5.2 Gateway coefficient

Although the participation coefficient has been widely used to characterize connector hubs, it discounts information about the importance of the neighbors of a node and the exclusivity of the node's connections to other modules. For example, two nodes can have the same number of intra-modular connections and participation coefficient, but one of them may be more important since it may have the only connection between two modules, or its neighbors in a module could be more important in terms of degree or betweenness. This was rectified by the *gateway coefficient* [161] that renders an importance score to connections of a node to each module. The gateway coefficient of a node is given by:

$$q_i = 1 - \sum_S \left(\frac{\theta_{i,S}}{\theta_i} \right)^2 \left(1 - \gamma_{i,S} \gamma_{i,S}^{intra} \right)^2 \quad (3.7)$$

where $\gamma_{i,S}$ measures the fraction of the edges of node i to the module S relative to other nodes in node i 's module; and $\gamma_{i,S}^{intra}$ measures the importance of nodes that are adjacent to node i . $\gamma_{i,S}$ and $\gamma_{i,S}^{intra}$ are given by:

$$\gamma_{i,S} = \frac{\theta_{i,S}}{\max_{S'} \sum_{i,j \in S'} \theta_{j,S}} \quad (3.8)$$

$$\gamma_{i,S}^{intra} = \frac{\sum_{j \in S} 1(w_{ij} > 0) \beta_{j,S}}{\max_{S'} \sum_{k \in S'} \beta_{k,S'}} \quad (3.9)$$

3.2.6 Network hubs

The modular hubs represent brain regions that participate in intra-modular information processing while the connector hubs represent heteromodal regions that facilitate inter-modular information processing. *Network hubs* represent key brain regions that are crucial for information processing both inside and outside functional modules. In order to identify network hubs, we propose to combine modular and connector hub measures, giving equal weights to the node's role in intra-modular and inter-modular information processing.

Given a node i , we define a network hub measure x_i to identify network hubs:

$$x_i = \zeta(m_i) + c_i \quad (3.10)$$

where modular hub measure $m_i \in \{\theta_{i,S}, \alpha_{i,S}, \beta_{i,S}\}$ and connector hub measure $c_i \in \{p_i, q_i\}$. We explore different combinations of modular and connector hub measures to detect network hubs.

In order to find the network hubs in a subject, we compute x_i for each node and classify it as network hubs if x_i is in the top 2.5% to 20% percentile of nodes [181, 182]. We investigated the efficacy of this threshold in steps of 2.5%.

3.2.6.1 Inter-subject variability of hubs

The heterogeneously distributed inter-subject variability of brain functional architecture [71, 85] introduces variability to the hubs detected across subjects. Since we detect hubs for individual subjects and report group-level hubs, it becomes imperative to study the consistency with which

a node is classified as a network hub over subjects. We define inter-subject variability λ_i of a hub node i as:

$$\lambda_i = 1 - \frac{k_i}{K} \quad (3.11)$$

where k_i is the number of subjects where node i is classified as a hub and K is the total number of subjects in the dataset.

3.2.6.2 Artificial lesioning: efficacy of hubs

The hubs detected using different hub measures represent key brain regions with specific roles in brain information processing, and in order to understand their effects in detecting hubs at the modular and network levels, we simulated brain lesions by removal of different categories of hubs and examining the effects on the network path length. The path length of a network represents the efficiency of information transmission of the network where networks with long path length are inefficient in terms of information transmission. Given the shortest path length d_{ij} between nodes i and j , the path length l_S of a module S is given by:

$$l_S = \frac{1}{n(n-1)} \sum_{i,j \in S', i \neq j} d_{ij} \quad (3.12)$$

If the module S is the brain network G , we refer to the path length as the global network path length and if the network is a brain module, we refer to the path length as modular path length.

We compute the change in path lengths after removal of hubs at both the network and modular level to understand the effect of removal of hubs on the network's path length. For comparison between different hubs, we used Tukey's paired t-test. Since there is no consensus on what percentage of nodes to denote as hubs, we repeat the hub removal for different percentages of nodes.

3.3 Analyzing disruption of brain's functional hubs in AD and ASD and using hubs as disease biomarkers

Previous studies have shown evidence of hub disruption in several brain diseases including AD [164] and ASD [167]. We hypothesized that the large-scale disruption of hubs and alterations in hub measures in these diseases can render hub measures as effective biomarkers for these diseases. In order to investigate this, we obtained data of patients suffering from

AD and ASD (along with their healthy age matched controls) and used deep neural networks (DNN) to classify patients and healthy subjects by using hub measures of functional nodes as input features. We compared classification performances achieved using nodal hub measures with those obtained using raw functional connectivity weights as features. The formulation for the deep neural network (DNN) used can be derived from section 2.3.3.

3.3.1 AD dataset

For AD, we used the ADNI dataset (URL: <http://adni.loni.usc.edu/>), which is the largest publicly available dataset for AD. Functional and structural MRI data were collected according to the ADNI acquisition protocol using a Philips 3 Tesla scanner. The resting state fMRI data for each subject consisted of 140 or 200 functional volumes, acquired with a repetition time (TR) = 3000 ms; echo time (TE) = 30 ms; flip angle = 80°; slice thickness = 3.313 mm; and 48 slices. Results included in this chapter come from data pre-processed using fMRIPrep [183] (details can be found in the appendix B). We obtained 88 subjects of which 49 were cognitively normal (CN, age: 76.6 ± 5.5) and 29 subjects suffering from AD (age: 75.4 ± 8.2). Subjects classified as AD have Mini-Mental State Examination (MMSE) scores between 15.7 and 26.9, a Clinical Dementia Rating (CDR) between 0.5 and 3.0 and fulfilled the criteria for AD laid down by National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (NINCDS/ADRDA). The CN subjects did not suffer from depression, cognitive impairment, or dementia, and had MMSE scores between 23.8 and 30.0 and a CDR of 0.0 and 1.0.

3.3.2 ASD dataset

For ASD, we used resting state fMRI data from the NYU Langone Medical Center available at Autism Brain Imaging Data Exchange (ABIDE URL: http://fcon_1000.projects.nitrc.org/indi/abide/ [184], a consortium of brain imaging data for sharing within the scientific community. Specifically, resting state fMRI data of 73 ASD (age 13.9 ± 6.3) and 88 CN (age 15.8 ± 6.3) subjects were used, removing subjects with mean framewise displacement greater than 0.2mm. The data downloaded was preprocessed using the C-PAC pipeline from the Preprocessed Connectomes Project. The preprocessing of the fMRI data includes correction of slice timing, realignment of motion, voxel intensity normalization, nuisance signal removal and band-pass filtering.

3.4 Results

3.4.1 HCP dataset

The dataset included 589 healthy adults (males = 278, mean age = 28.33 years, range = 22 to 37 years) from the S900 release of the HCP [185] available at <https://www.humanconnectome.org>. Details for the dataset can be found in section 2.4.1. The functional connectivity of the brain was calculated as Pearson correlations of fMRI time-series on the 264 functionally diverse ROI identified in the Power atlas [14]. The details are given in section 2.4.2.

3.4.2 Brain functional modules

We used the ICSC algorithm described in the previous chapter and ran 100 independent runs on unthresholded subject matrices until convergence. We selected the run with the highest value for ICSC quality function and used the subject-level and group-level modularizations for this run. The ICSC algorithm detected 19 group-level modules, with the number of nodes in the modules ranging from 6 to 18 (see Figure 2.2), which correspond to previously defined functional networks of the human resting-state brain [14, 44, 128, 129]. Using the Jaccard Index matching scheme, we assigned labels to subject-level modules and computed modular hub and network hub measures for network nodes.

3.4.3 Best fit for nodal degree distributions

Using python's powerlaw package [186], we found stretched exponential to be the best fit for the nodal degree distribution of the brain's functional connectome. Many detected functional modules had a small number of nodes in a subject and we therefore only considered modules having more than 10 nodes. We found that the stretched exponential is the best fit for degree distribution in all functional modules of the brain by comparing the likelihood values. We applied Tukey's paired t-test on BIC values for fits of different probability distributions detected modules. We found that the BIC values of fits for the stretched exponential distribution were statistically lower compared to power-law, exponential, and log-normal models (3.1). We summarize mean values of the parameter fits for the stretched exponential distribution in table 3.2.

Table 3.1: p – value for paired t-test on Bayesian Information Criterion for fits of power law, exponential, log-normal and stretched exponential degree distributions for different modules.

Functional module	$BIC_{power} > BIC_{stretched}$	$BIC_{expo} > BIC_{stretched}$	$BIC_{log-normal} > BIC_{stretched}$
Salience	$\approx 10^{-141}$	$\approx 10^{-60}$	$\approx 10^{-3}$
Subcortical	$\approx 10^{-38}$	$\approx 10^{-2}$	$\approx 10^{-9}$
Ventral Attention	$\approx 10^{-128}$	$\approx 10^{-57}$	$\approx 10^{-7}$
Motor Hand	$\approx 10^{-71}$	$\approx 10^{-43}$	$\approx 10^{-15}$
Medial Visual	$\approx 10^{-83}$	$\approx 10^{-48}$	$\approx 10^{-23}$
Motor Mouth	$\approx 10^{-88}$	$\approx 10^{-48}$	$\approx 10^{-4}$
Motor Feet	$\approx 10^{-72}$	$\approx 10^{-36}$	$\approx 10^{-14}$
Memory Retrieval	$\approx 10^{-84}$	$\approx 10^{-49}$	$\approx 10^{-2}$
Medial DMN	$\approx 10^{-152}$	$\approx 10^{-71}$	$\approx 10^{-19}$
Left FP	$\approx 10^{-122}$	$\approx 10^{-60}$	$\approx 10^{-16}$
Language	$\approx 10^{-131}$	$\approx 10^{-61}$	$\approx 10^{-15}$
PCC DMN	$\approx 10^{-123}$	$\approx 10^{-57}$	$\approx 10^{-3}$
DMN	$\approx 10^{-112}$	$\approx 10^{-53}$	$\approx 10^{-8}$
Right FP	$\approx 10^{-144}$	$\approx 10^{-69}$	$\approx 10^{-11}$
Lateral Visual	$\approx 10^{-74}$	$\approx 10^{-41}$	$\approx 10^{-19}$
CO Task Control	$\approx 10^{-150}$	$\approx 10^{-67}$	$\approx 10^{-31}$
Dorsal Attention	$\approx 10^{-63}$	$\approx 10^{-38}$	$\approx 10^{-11}$
Auditory	$\approx 10^{-42}$	$\approx 10^{-15}$	$\approx 10^{-6}$
Occipital Pole Visual	$\approx 10^{-45}$	$\approx 10^{-24}$	$\approx 10^{-3}$

Table 3.2: Fitting parameters of stretched exponential for brain functional network and modules.

Functional module	No. of nodes	b	a
Functional connectome	264	0.009	1.83
CO Task Control	20.6	0.001	7.46
Lateral Visual	18.6	0.001	8.64
Motor Feet	18.5	0.013	6.92
Medial Visual	18.3	0.0004	8.62
Motor Hand	17.9	0.002	7.93
Medial DMN	17.5	0.002	7.85
Subcortical	17.4	1.15	2.87
Language	17.2	0.263	7.60
Left FP	16.1	0.002	7.92
Ventral Attention	16.0	0.012	7.21
Right FP	15.9	0.001	7.83
DMN	15.3	0.618	7.60
Salience	14.8	0.005	6.99
Dorsal Attention	14.5	0.087	8.53
PCC DMN	14.0	0.563	7.26
Auditory	13.6	2.44	6.02
Motor Mouth	13.2	0.030	8.12
Occipital Pole Visual	10.8	0.007	8.01
Memory Retrieval	10.6	0.035	8.97

3.4.4 Artificial lesioning to identify the most crucial modular and network hubs

Previous studies assumed that the most crucial modular hubs are nodes with a high intra-modular degree [42, 72]. We tested this assumption since finding nodes that were more crucial to inter-modular communication could also lead to detection of crucial nodes responsible for overall network communication. We computed the intra-modular degree $\theta_{i,S}$, the ambivert degree $\alpha_{i,S}$, and the betweenness centrality $\beta_{i,S}$ for nodes in all subjects. Following the approach of [181, 182] we classified different percentile of nodes as modular hubs, ranging from top 2.5% to 20% (in steps of 2.5%) and inspected the change in the modular path length l on their removal for a thorough analysis. In order to determine which hubs were the most crucial, we performed Tukey's paired t-test on the path length of the network after removal of the hubs.

We performed the above set of experiments on all subjects after performing thresholding using percolation analysis. We found that for all the modules except the attention, motor, and salience modules, removal of nodes with a high ambivert degree led to a significantly higher increase in the modular path length (paired t-test $p < 0.05$) than removal of nodes with high intra-modular degree or betweenness centrality (refer Figure 3.2). After the ambivert degree, the intra-modular degree gave the most crucial modular hubs.

Since the ambivert degree gives more crucial modular hubs than those detected by intra-modular degree and betweenness centrality, we hypothesize that considering nodes with high ambivert degree, taking into weaker connections into considerations, as modular hub measure would yield more important overall brain network hubs. This hypothesis for each subject's network tested by computing network hub measures and removed the nodes with top 20% scores (in steps of 2.5%) and inspected the change in the network path length l upon their removal.

To understand which hub measure gives the most crucial network hubs, we first compared the path length of different types of connector hubs with the same modular hubs. For modular hubs detected by different measures, we found that participation coefficient had a similar or larger impact on the path length than different gateway coefficient measures (Appendix B, Figure B1). We, therefore, compared the effect of removal of modular hubs only along with their participation coefficient on the network path length since these represented the most crucial network hubs detected using different modular hub measures (Figure 3.3). We found that the

Fig. 3.2: The changes in path lengths of brain functional modules after artificial lesioning of modular hubs detected by different hub measures. Except the motor, attention and salience modules the removal of hubs detected by the ambivert degree led to significantly greater increase in path than the other modular hubs. The memory retrieval module was excluded from the analysis because it was found in fewer than 50 subjects.

Fig. 3.3: The increase in the path length after artificial lesioning of different percentiles of network hubs determined by different modular hub measures and their participation coefficient.

removal of network hubs detected by ambivert degree led to a statistically significant higher increase in network path length than the removal of hubs detected by betweenness centrality (paired t-test p -value $< 10^{-9}$), and intra-modular degree (paired t-test p -value $< 10^{-45}$).

We also computed the inter-subject variability of network hubs across subjects, detected by different modular hub measures. For a complete analysis, we determined the hub variability across different percentiles ranging from 80 to 97.5 (with steps of 2.5). On comparison of variability of hubs detected by ambivert degree with betweenness centrality and intra-modular degree, we found that ambivert degree found the hubs that are most invariable across subjects. The difference in inter-subject variability was statistically significant (p -value < 0.05) for hubs detected by betweenness centrality but insignificant for those detected by intra-modular degree (refer Figure 3.4).

We computed the top 10% hubs given by ambivert degree and participation coefficient for each subject and determined a consensus over all the subjects. We found that the regions in the middle and superior temporal gyrus; the cuneus and precuneus; inferior parietal lobe; superior frontal gyrus, precentral and postcentral gyrus, and posterior cingulate were most frequently identified by ambivert degree combined with participation coefficient (Figure 3.5).

3.4.5 Disruption of network hubs in AD and ASD

We ran 100 independent runs of the ICSC algorithm on unthresholded subject matrices for the CN and diseased subjects from the ADNI and ABIDE datasets. We selected the run with the

Fig. 3.4: The inter-subject variability of hubs detected by different hub measures and the participation coefficient at different percentiles.

Fig. 3.5: The axial and sagittal views of the top 10% of the regions identified as network hubs by ambivert degree and participation coefficient. The size and color of the node correspond to its inter-subject variability and its functional module, respectively.

highest value for the ICSC quality function for both and used the subject-level and group-level modularizations from this run. We used the Jaccard Index matching scheme described before and assigned labels to each subject's modules. We chose CN subjects taken from the respective datasets for comparison, because the subjects are age-matched (which removes the effect of age related changes), and the scans were collected on the same scanner. We observed that across thresholds hubs were disrupted in patients suffering from AD and ASD.

The top 10% ambivert participation coefficient hubs for the CN subjects were located in the cingulate gyrus, middle temporal pole (R: right), supplementary motor, precuneus, thalamus, calcarine (R), insula (R), superior medial frontal (L: left) and middle frontal (L) regions, whereas the hubs for the subjects suffering from AD were located in the middle cingulum, calcarine, anterior cingulate (L), inferior parietal, supplementary motor, precuneus, and fusiform gyrus. We observed that the CN hubs located in the middle temporal (R), insula (R), lingual (L), middle frontal (L), inferior temporal (R), thalamus (R), calcarine (R) and precuneus were disrupted in AD subjects (Figure 3.6(a)).

For the ASD subjects, the top 10% hubs detected by the ambivert degree were located in the supplementary motor area (L and R), anterior (L) and middle (R) cingulum, calcarine (L and R), angular gyrus (L), middle occipital (L and R) and insula (L and R); whereas the hubs for CN subjects were in supplementary motor area (L and R), frontal superior medial (L and R), anterior (R) and middle (L) cingulum, precuneus (L and R), middle occipital (R), and angular gyrus (L and R). We observed that the hubs located in the angular gyrus (R), anterior cingulum (L), precuneus (L), and medial superior frontal (L), and middle occipital (R) in normal subjects were disrupted in ASD subjects (Figure 3.6(b)). For other thresholds, please refer to Appendix B Figures B3 (for AD) and B4 (for ASD).

We studied the regions with the highest disruption for both AD and ASD patients in comparison to their normal counterparts. For this, we computed the difference between the nodal hub measures (average ambivert degree and participation coefficient) for patients and CN subjects. For AD, we found that the network hubs were disrupted for regions in the middle and medial temporal gyrus, frontal orbital lobe, and occipital lobe (table 3.3) whereas for ASD we found that regions in the thalamus, posterior cingulate cortex, and supplementary motor areas were most affected (table 3.4).

Table 3.3: Disrupted brain regions in AD. MNI coordinates of the most disrupted regions (R: Right, L: Left) detected by ambivert degree. Hub measures are given as the sum of ambivert degree and participation coefficient.

MNI			Hub measure		Anatomical Location
x	y	z	CN	AD	
50	3	-24	1.39	0.67	Temporal Pole (R)
-20	36	-15	0.55	-0.06	Frontal Mid Orbital (L)
53	-48	12	0.80	0.27	Temporal Mid (R)
44	-48	-15	0.89	0.42	Fusiform (R)
8	42	-9	0.86	0.40	Rectus (R)
35	-81	0	1.07	0.65	Occipital Mid (R)
-44	27	-9	0.42	0.01	Frontal Inferior Orb (L)
23	-87	21	0.59	0.19	Occipital Superior (R)
-31	-78	-15	0.99	0.59	Occipital Inferior (L)
-50	0	-24	0.93	0.55	Temporal Mid (L)

Table 3.4: Disrupted brain regions in ASD. MNI coordinates of the most disrupted regions (Right, L: Left) detected by ambivert degree. Hub measures are given as the sum of ambivert degree and participation coefficient.

MNI			Hub measure		Anatomical Location
x	y	z	CN	ASD	
8	-7	8	0.76	0.43	Thalamus (R)
-2	-16	13	0.63	0.31	Thalamus (L)
51	-45	22	0.37	0.68	Temporal Superior (R)
9	17	30	0.54	0.85	Cingulum Anterior (R)
-7	-72	38	1.18	0.88	Precuneus (L)
-53	-15	-9	0.50	0.80	Temporal Mid (L)
-2	10	45	1.44	1.14	Suppl Motor (L)
-1	25	30	1.09	1.38	Cingulum Anterior (L)
5	3	51	1.46	1.16	Suppl Motor (R)

Fig. 3.6: The differences between brain regions that are identified as top 10% network hubs, using ambivert degree and participation coefficient in cognitively normal (CN) and diseased subjects. For (a) the subjects suffer from Alzheimer's Disease (AD); and for (b) the subjects suffer from Autism Spectrum Disorder. The regions in red were identified as hubs in the diseased subjects but not in CN subjects whereas the regions in green were identified as hubs in CN subjects but not in the diseased subjects.

3.4.6 Classification of patients and healthy controls using nodal hub scores

We hypothesize that as there is widespread hub disruption in the patients suffering from AD and ASD, hub measures can be used as features to train deep neural networks to classify patients and healthy controls. We therefore used each node's ambivert degree (z-score) and participation coefficient as features to perform classification with deep neural networks (DNN). The performance was also compared with a classifier trained on functional connectivity features from each subject's unthresholded connectivity matrix W . Since W is symmetric, we formed a vector of its lower triangular elements, which results in a vector of 34716 functional connectivity features for each subject.

For DNN, we empirically set the number of hidden layers, number of neurons in each layer, batch size, and learning rate in the experiments. We added dropouts to the hidden layers and imposed early stopping criteria to prevent overfitting. For SVMs, we experimented with a range of parameters ($C \in \{0.001, 0.01, 0.1, 1, 10\}$, $\gamma \in \{0.001, 0.01, 0.1, 1\}$) and different kernels. The best hyperparameters were selected by performing a grid search on the average accuracy of different folds obtained from 5-fold cross validation.

The performance of the 5-layer DNN along with the best SVM model for different kernels are shown in Tables 3.5 and 3.6. All samples were included in the test set at least once and the experiment was repeated for 10 random seeds. We used four sets of features derived from hub scores: ambivert degree (264 features), participation coefficient (264 features), ambivert degree + participation coefficient (264 features) and ambivert degree scores concatenated with participation coefficient scores (528 features). We observed that the highest accuracy for both AD and ASD were obtained for a 3 hidden layer feedforward DNN using ambivert degree and participation coefficient as features. The SVM models gave a consistently poor classification accuracy. The improvement in classification accuracy by using hub score features in comparison to functional connectivity features was statistically significant for both AD (one tailed t-test, p-value = 0.0048) and ASD (one tailed t-test, p-value = 0.045). We also observed that usage of hub scores (that is, the sum of ambivert degree and participation coefficient) as features led to a significant increase in classification accuracy for ASD and AD in comparison to their individual usage, which means that both measures carry crucial information.

Table 3.5: Accuracies of classification of AD and NC using deep neural networks and nodal hub scores as features. * refers to a statistically significant result in comparison to all the others.

Model	Accuracy
Features: Ambivert degree + Part. coeff.	
DNN (3 hidden layer of 50,32,10)	79.4%
SVM (Linear, C = 0.1)	66.5%
Features: (Ambivert degree, Part. coeff.)	
DNN (3 hidden layers of 50,32,10)	81.2%*
SVM (Linear, C = 0.1)	64.5%
Features: Ambivert degree	
DNN (3 hidden layers of 50, 20, 10)	77.5%
SVM (Linear, C = 0.1)	67.1%
Features: Part. coeff.	
DNN (3 hidden layers of 40,20,10)	71.7%
SVM (Linear, C = 0.01)	62.8%
Features: Functional connectivity W	
DNN (3 hidden layer sizes: 500, 50, 10)	76.7%

Table 3.6: Classification of ASD and CN with deep neural networks with nodal hub measures and connectivity features. * refers to a statistically significant result in comparison to all the others.

Model	Accuracy
Features: Ambivert degree + Part. coeff.	
DNN (3 hidden layers of 40,20,10)	74.1% *
SVM (Poly, $C = 0.01$, $\gamma = 0.1$)	59.1%
Features: (Ambivert degree, Part. coeff.)	
DNN (3 hidden neurons layers of 50,32,10)	71.6%
SVM (Poly, $C = 0.001$, $\gamma = 0.1$)	57.9%
Features: Ambivert degree	
DNN (3 hidden neurons of 40,20,10)	71.3%
SVM (Poly, $C = 0.01$, $\gamma = 0.1$)	59.2%
Features: Part. coeff.	
DNN (3 hidden neurons layers of 40,20,10)	66.1%
SVM (Sigmoid, $C = 1.0$, $\gamma = 0.01$)	53.7%
Features: Functional connectivity W	
DNN (3 hidden layers of 1000,100,10)	72.1%

3.5 Discussion

3.5.1 Considering brain functional modules while detecting network hubs

Models of degree distribution suggest preferences in forming connections between nodes and number of high degree nodes in the network. Exponential degree distributions suggest random edge formation whereas power-law suggests preferential attachment model where edges are formed in proportion to the degree of incident node. Stretched exponential show intermediate properties between the two strategies. Consequently, stretched exponential distributions have “fatter” tails than exponential distributions but much less than naturally found power laws. We found that the nodes of the brain functional connectome and its modules follow a stretched exponential degree distribution. This suggests that while the modules have different functionalities but they follow the same model as the overall network. It also points to the fact that although the functional partitions have modular hubs but they’re expected to be few in number. Further, since the connections in the brain modules are not formed randomly or with the preferential attachment model, the degree of a node is not the only criteria while forming connections. This is intuitive since node function is an important criteria for forming relationships in the functional connectome.

Multiple studies have shown the modular nature of brain’s functional architecture with regions in the same modules having high synchronization with each other and thus high corresponding functional connectivity [36]. It has also been shown that modules in brain functional networks have heterogeneous sizes [43], leading to higher degree for nodes in a large module and vice versa [72]. Therefore, if only global degree of a node is considered to detect hubs, nodes in large modules would be invariably favored [72], making it imperative to detect hubs at a modular level, considering the inherent brain functional modular structure.

A weak connection signifies weak correlation between neuronal activity of the brain regions involved, which in turn points to low probability of signaling and communication between neurons in the regions. However, several studies have pointed out the importance of considering weak functional correlations while analyzing brain functional networks [88–90]. Therefore, weak connections should be given due importance while computing brain functional network topology. In order to account for this, previous studies detecting modular hubs have thresholded the networks such that only a fraction of initial connections in the network are retained. However,

thresholding strategies used in previous studies do not consider whether the underlying modular structure in functional networks is preserved. We took several measures to consider the modular architecture and considered the effect of weak connections on detection of hubs: (i) we used the ICSC algorithm which performs modularization on unthresholded networks; (ii) while computing network hub metrics, we used percolation analysis [139] to find the threshold which preserves modular structure in brain functional networks [52]; and (iii) we proposed a new measure called the ambivert degree that incorporates the average weight per connection information to the node's intra-modular degree.

Using the ICSC algorithm, we detected 19 group-level modules that corresponded to well-known functional systems. On performing percolation analysis, we found that even the functional networks derived from HCP subjects, which were in a narrow age range (mean 28.3 ± 3.7) and acquired on the same scanner with an average signal-to-noise ratio of 728, have a varied percentage threshold (mean 75.6 ± 10.3). Previous studies studying subject networks [159] did not consider this, since they performed their analysis using the same percentage threshold for all subjects.

3.5.2 Using ambivert degree to detect modular and network hubs

We compare the detection of modular hubs for within module processing by using intra-modular degree, ambivert degree and betweenness centrality. We systematically analyzed the importance of nodes for intra-modular processing and global network processing by performing artificial lesioning where we remove a range of hubs and observe changes in the modular or network path length. We found that ambivert degree gives significantly more crucial modular hubs across different thresholds. At the global network level, we found that hubs detected by combining ambivert degree and participation coefficients are significantly more crucial than network hubs detected by intra-modular degree and betweenness centrality. It is to be noted that the gateway coefficient, an extension of the participation coefficient introduced to detect better connector hubs [161] did not have a significant impact on detecting crucial network hubs. This means that as long as the degree and weight per connection are being considered to derive network hubs, the participation coefficient is sufficient to detect network hubs.

The variability in hub locations is a by-product of widespread inter-subject variation prevalent in brain functional architecture. However, since we considered fMRI scans of healthy

subjects in a narrow age range from the same scanner, the location of hubs should be relatively stable over subjects. On studying the variability of network hubs over multiple subjects, we found that while the hubs detected by ambivert degree hubs were more invariable than those detected by betweenness centrality significantly and intra-modular degree insignificantly.

Since many previous studies have reported nodes with high intra-modular degree and participation coefficient as functional brain hubs [42, 72], it is important to understand the differences in anatomical locations of these hubs and the network hubs characterized by ambivert degree and participation coefficients. We found that regions from the middle frontal gyrus (L) and superior temporal gyrus (R) which were classified as brain hubs using the old scheme are dropped, and instead regions from the cuneus (L), middle occipital gyrus (R) and inferior parietal lobe are added (Appendix B Figure B2). While newly detected regions are known to be primarily involved in the dorsal and the ventral stream visual processing [187], recent studies have also pointed to the role of the cuneus for internal directed attention activity in the brain during resting state [188, 189].

3.5.3 Disruption of network hubs in AD and ASD

Numerous studies have investigated the changes in the brain functional and structural networks caused by AD and Mild Cognitive Impairment [121, 190–192], specifically the disruption of brain hubs [46, 193, 194]. The changes in modular structure with cognitive decline and AD has been extensively studied with major alterations being reported in the default mode, the salience, and the fronto-parietal control networks [195–197]. Similarly, alterations in modular structure have been reported for ASD [198, 199]. Therefore, the modular structure of functional networks needs to be considered while detection of hubs, but previous studies investigating disruption of hubs in diseases either used traditional thresholding methods and did not consider preservation of network modular structure or used the global degree as a measure to find network hubs [46, 167, 193, 194], thus giving biased results. This warranted a fresh look at the disruption of hubs in the functional networks derived from AD and ASD subjects.

The hubs detected for the CN subjects from the disease datasets are different from the healthy subject group from the HCP dataset. Besides factors like different scanner acquisition protocol, preprocessing, and gender ratio of participants the difference in network hubs between HCP subject cohort and CN subject cohorts (from the ADNI and ABIDE) can be primarily attributed

to the difference in the age of the participants [200, 201]. The average ages for HCP, ADNI CN, and ABIDE CN participant cohorts are 28.3 years, 76.6 years, and 15.8 years, respectively. On computing hubs of subjects suffering from AD and ASD, we found that the network hubs for the patients were different from those for CN subject-group. Specifically, the hubs in the medial and inferior temporal lobe, precuneus, and posterior cingulate cortex in the CN subject group were found to be disrupted in the AD subject group. Atrophy in the medial temporal lobe has long been shown to be a biomarker for AD [202] whereas the functional connectivity of the precuneus with the default mode network and posterior cingulate cortex with the ventral attention network is known to be disrupted in AD patients [203, 204] which validates our findings. Similarly, our results for hub disruptions in ASD subject group were corroborated by multiple previous studies, where disruption in connectivity of the angular gyrus [198, 199, 205] and precuneus [167, 198, 199] regions participating in the DMN have been studied and alterations of both functional and structural fibers in the anterior cingulum [167, 206–209] have been reported.

3.5.4 Brain hubs as diagnostic features for AD and ASD

We hypothesized that the disruption of hubs in AD and ASD can be used as a potential diagnostic tool for detecting these diseases. Using multiple machine learning models, we found that a 3 hidden layer feedforward DNN using nodal hub measures as features were capable of detecting AD and ASD patients from CN subjects with an accuracy of 81% and 74%, respectively. We also found that combining both the inter-modular and intra-modular hub measures improves the classification accuracy.

Previous studies in the area have used functional connectivity measures [210–212] or additional graph measures [213, 214] to perform classification of AD and ASD. The large number of training features not only causes overfitting of the models because of few available training samples but also uses high resources due to the large number of trainable weights [190]. In this chapter, we employed novel hub features including ambivert degree features for AD and ASD and demonstrated their use as diagnostic features for diseases. We observed that the hub disruption was more widespread in case of AD than ASD (refer Figure 3.6), and correspondingly the difference in classification accuracy based on nodal hub measures and functional connectivity features was more for AD than ASD. While previous studies have used

functional connectivity features for disease state classification, but we show that using brain hubs as features is sufficient and more efficient than functional connectivity. For example: using just 0.09% (for AD) and 0.03% (for ASD) of the total number of trainable weights used in functional connectivity based models, we obtained a significantly higher classification accuracy. This work paves the way for using hub scores instead of functional connectivity as input features for diseases where hub disruption is a trademark.

3.6 Chapter Summary

In the present chapter we proposed a new measure, ambivert degree, to detect crucial brain regions known as brain hubs. We studied the degree distributions of the brain functional connectome and its modules and found a similar distribution across all modules. Using experiments with the ambivert degree hubs and hubs proposed by previously used measures, we show that removing ambivert hubs have a higher impact on the network efficiency than previously used measures. We also demonstrate that ambivert degree hubs are more stable across subjects. Therefore, ambivert degree detects more crucial and stable hubs.

Multiple studies have pointed the disruption of brain hubs during neurological diseases. Using ambivert degree hubs we first show that hub are indeed disrupted during AD and ASD, as shown by previous studies. We further show that nodal hub scores can be used as features for diagnosis of AD and ASD and give a significantly better diagnostic accuracy than the commonly used functional connectivity.

In the present and preceding chapters, we have developed methods to study key aspects of the brain's functional architecture during the resting state. While the work so far has been focused on studying the resting state brain, the brain states continuously transition between rest and task states. It is therefore important to study task related changes in the brain. Accordingly, in the next chapter, we propose methods to study the modulations in the brain functional architecture during task performance.

Chapter 4

Decoding task-general and task-specific brain architectures ¹

Studying task related alterations in the brain and understanding how different functional sub-systems interact with each other to support task performance is a fundamental problem in cognitive neuroscience. The complex and dynamic interactions between brain regions during task performance can be captured using fMRI scans. While the previous chapters in this thesis focused on studying the brain’s functional architecture during rest, this chapter aims to study the alterations related to task performance in the brain architecture. Task related alterations in the brain can be further sub-divided into task-specific and task-general alterations. Task general alterations in the brain are common across all tasks, while task-specific alterations are linked to the particular task under consideration. Using recent advances in *deciphering* deep learning models, we propose an approach that determines both task-specific and task-general architectures in the brain functional architecture. We use a reference-based decoder to uncover distinguishing task-related architectures from trained deep neural networks (DNNs). We demonstrate the performance of the proposed method on 12,500 rest and task functional connectivity samples taken from the Human Connectome Project (HCP). The discovered task-specific architectures for the motor and language tasks were validated with findings from previous studies and we pinpoint a set of just 34 salient functional connections that classified rest and task brain scans with almost perfect accuracy. The uncovered task-general brain architectures could serve as potential disease biomarkers as alterations in task-general brain architectures have been implicated in several neurological disorders.

¹The work in this chapter has been submitted for publication to the *Human Brain Mapping* journal.

4.1 Task related brain alterations and related work

The study of functional self-organization of the human brain during task performance is among widely researched areas in cognitive neuroscience. There is significant functional modulation in the brain during task performance [12]. Two types of task-evoked alterations are known to shape the brain task states: task-specific and task-general modulations [54, 215–217]. *Task-general* alterations are task-independent changes in the brain that occur across a set of tasks whereas *task-specific* alterations depend on the particular task being performed.

Aging [218] and multiple neurological disorders such as major depressive disorder (MDD) [219, 220], attention deficit hyperactivity disorder (ADHD) [221], mild cognitive impairment (MCI) [222] and Schizophrenia [223] have been characterized by anomalous task-general architecture. Task-general alterations are further classified as *task positive* and *task negative*, depending on whether the neurons and their connections in a region are activated or deactivated during the task performance. We model the alterations in the brain functional connectivity as *positive task-general* and *negative task-general* depending on the presence or absence of functional connectivity during task-states relative to the rest-state.

Advances in neuroimaging have made it easier to study dynamic *in vivo* changes in regional relationships during different tasks and the rest state. Multiple studies have investigated the modulations in brain functional architecture during task performance. Betti et al. [53] used group averaged connectivity from magnetoencephalography (MEG) and fMRI scans for twelve subjects to recover rest and task functional connectivity and study the changes during rest and task. They concluded that although the resting-state and task network interactions had similar network topography but clear differences were observed in the frequency domain. Krienen et al. [55] acquired task fMRI data for 48 subjects for 14 distinct sub-tasks (grouped into passive, sensory and motor tasks) and found that the task modulations varied around a common central tendency across all task states by performing correlation and seed-based analysis on the group averaged data. Cole et al. [54] used two task datasets, one containing 64 tasks and the other containing 7 tasks (from the Human Connectome Project), and found that the central tendency for each functional connection during varied tasks was the one represented during the rest state. They found that a stable intrinsic functional network architecture that is present across rest and tasks and a few but statistically significant alterations shape task states. In another study,

Cole et al. [224] selected the first principal component on task activation to detect task-general architecture. These studies suggest that although there are small-scale alterations in the brain functional architecture, there is an intrinsic stable functional backbone during both the rest and task state that remains unchanged [53–55].

Although multiple recent studies [82, 225, 226] have depicted the role of subject-level characteristics, most previous efforts studying modulations in functional connectivity involved studying group-level alterations. Gratton et al. [225] used data for 9 subjects with 10 hours of scan data per subject (10 scans of 1 hour each containing resting state and 4 task states) from the Midnight Scan Club [82] and studied the dependence of functional connectivity on task states, subjects, and scan sessions. They found that while the task states and the sessions shaped the functional connectivity patterns, common organizational principles and subject-specific features had a higher impact. Similarly, Xie et al [226] used task fMRI scans and used k-means clustering to separate the effects of subject-level functional connectivity and tasks to determine what shapes group functional connectivity. They found that subject-level functional connectivity was the most dominant factor in group-level functional connectivity variability. These results, combined with previous studies on resting state brain functional architecture pointing towards significant inter-subject variability [84, 85, 101, 153, 227] necessitate considering subject variability in the functional architecture while studying modulations during task performance. However, initial studies only considered group-level functional connectivity. In this study, instead of studying group level modulations, we account for inter-subject variability by explicitly modelling subject-level modulations.

Another concern with previous studies is that they only captured linear relationships between functional connectivity at rest or task. In a recent study, He et al. [228] used resting state functional connectivity to predict behavior using traditional machine learning algorithms and multiple deep neural networks (DNN) proposed for functional connectivity based analysis. They demonstrated that kernel regression applied on functional connectivity is sufficient and DNNs do not bring additional value as far as predicting individual behavior is concerned. We test this with classification of rest and task functional connectivity by using traditional supervised learning models along with the more complex DNNs. We train classifiers on both types of models on individual subject rest-state and task-state data and thereafter derive the distinguishing features from the trained model. We refer to identification of salient features of

a network as decoding (different from the decoding performed in autoencoders). We found that although all the classification models gave a high classification accuracy, DNNs were able to classify with a much better accuracy than the shallow models such as support vector machines (SVM). We uncover the task-general architecture by decoding a binary classifier trained for classifying rest and multiple task states. Using the obtained task-general architecture, we demonstrated decoding of task-specific brain architectures for the motor and language tasks.

Recent attempts have been made to evaluate the salience given to input features of a neural network model classifying neuroimaging data. Jang et al. [229] used the trained weights in different layers of a feedforward neural network and mapped them to four sensorimotor task-states. Several works using a gradient-based decoder with different neural network architectures have culminated. Using a gradient-based decoder on a trained long short-term memory recurrent neural network, Li et al. [230] identified functional signatures in rest and task-states. Floren et al. [231] identified brain regions involved in different visual processing tasks. However, these studies did not validate their derived feature set and issues like gradient saturation and discontinuities due to bias terms have not been addressed before using gradient-based decoders on neural networks [232].

Recent attempts such as Integrated Gradients [233], DeepLIFT [232] and SHAP [234] attempt to ameliorate the issues with gradient-based approaches. These approaches find contributions of the neurons at each layer from those at the output layer and then backpropagate the contributions of all the layers to the input layer. DeepLIFT is one of the approaches, which gives consideration to both negative and positive contributions, and computes the salience scores efficiently in a single pass, thus addressing the issues confounded in gradient based approaches. In this article, we use DeepLIFT to study the microscale and mesoscale brain changes shape task-related alterations in the brain [224]. We used the state-of-the-art resting state and seven-task fMRI data of 690 subjects from the Human Connectome Project (HCP) dataset.

Using functional connectivity between brain regions as inputs to train a deep neural network (DNN), we proposed a novel method that uncovers salient task-general and task-specific functional architectures of the brain. We derive sets of 34 connectivity features (instead of the initial 34,716 features) that represent the task-general and task-specific functional brain architectures. We validated the saliency of the detected features by using a small subset of

features identified by the method to perform classification. We show that although there is pronounced inter-subject variability in task-related brain architectures, the small subset of detected features identified by our method are able to provide near 100% classification accuracy while distinguishing tasks from the rest states. Using feature salience scores, we identified functional connections and functional modular interactions that characterize task-general, and motor and language task-specific architectures. We discovered widespread system reorganization during task performance and were able to classify the rest vs all tasks, rest vs motor task, and rest vs language task with high accuracies by using a 5-layer feedforward DNN. The high classification accuracy was maintained even while taking only a subset of features classified as highly salient by our decoder. We confirmed that the default mode network, commonly known for its deactivation during task performance, forms positive correlations with the salience network [235]. Similarly, we found an important role for the regions in the motor sub-systems for performing the motor task and left lateralized temporal regions in performing the language task [236].

4.2 Proposed methods to study task related modulations in the brain

Let brain functional network $G = (\Omega, W)$ where Ω denotes the set of brain regions-of-interests (ROI) or nodes and $W = \{w_{ij}\}_{i,j \in \Omega}$ denotes the adjacency matrix of functional connectivity between different ROI. We obtained a symmetric W for each fMRI scan and input features x for the feedforward DNN were obtained by taking only the lower diagonal elements of adjacency matrix. We used rest and task data gathered in HCP on healthy subjects.

4.2.1 Feedforward DNN

For each subject, we obtained (x, d) where $x = (x_i)$ denotes the input feature vector and $d \in D$ denotes the sample label, we used a feedforward neural network of L layers with the first $L - 1$ layers having rectified linear unit activation and a final softmax layer. The formulation of the feedforward DNN is the same as in section 2.3.3. The weights and biases of the layer l be given by V_l and b_l , respectively, such that $\theta = \{(V_l, b_l)\}_{l=1}^L$ denotes all the parameters in the network.

4.2.2 Saliency of input features

Let f be neural network function mapping input x to output y . Let g be a simpler *explanation model* that is interpretable and an approximation of the network mapping f . Let the number of neurons in layer l be n_l . Using an appropriate reference, let us assign to each neuron i at layer l its contribution $C_{\Delta h_{i,l} \Delta h_{k,l+1}}$ to the change in the output of neuron k at layer $l+1$. Then, the change $\Delta h_{i,l}$ in the activation of the i^{th} neuron of layer l due to the input relative to the reference is given by

$$\Delta h_{k,l+1} = \sum_{i \leq n_l} C_{\Delta h_{i,l}} \quad (4.1)$$

when $l \in \{0, \dots, L-2\}$ and

$$\Delta y = \sum_{i \leq n_{L-1}} C_{\Delta h_{i,L-1}} \quad (4.2)$$

when $l = L-1$.

In order to determine the feature saliency scores, we evaluate 3 types of references; viz. sample mean, sample mode and taking a sample which has the least L_2 norm from the rest of the samples. We empirically determined the mean reference $\bar{x}$ to be the most representative of the samples. Given the reference input $\bar{x}$ and the original input x , we can substitute $\Delta y = f(x) - f(\bar{x})$ and $g(x) = f(x)$, giving us an equation for the model g :

$$g(x) = f(\bar{x}) + \sum_{i,l} C_{\Delta h_{i,l} \Delta y} \quad (4.3)$$

where the contribution of neurons in each layer l to the output y is given by $C_{\Delta h_{i,l} \Delta y}$ [232]:

$$C_{\Delta h_{i,l} \Delta y} = \Delta h_{i,l} \sum_{k,l} \frac{C_{\Delta h_{i,l} \Delta h_{k,l+1}}}{\Delta h_{i,l}} \frac{C_{\Delta h_{k,l+1} \Delta y}}{\Delta h_{k,l+1}} \quad (4.4)$$

where $C_{\Delta h_{i,l} \Delta h_{k,l+1}}$ can be computed from the Linear, Rescale and RevealCancel rule [232]. We can get contribution of neurons in all layers $0 \leq l < L$ to the output y by backpropagating contributions of layers to the input, using the chain-rule given by equation (4.4).

The *saliency scores* c for the input layer are evaluated as $c = (C_{\Delta h_{i,0} \Delta y})$ such that each element gives the contribution of the corresponding input to the changes of the output. The DeepLIFT method [232] implements computation of feature saliency scores based on the changes of the output from a reference input, allowing information to propagate across the

network layers even when the gradient is zero. We compute feature importance score using equation (4.3):

$$c = \text{mean}_x \left(f(\bar{x}) - g(x) - \sum_{i,l \in \{1, \dots, L-1\}} C_{\Delta h_{i,l} \Delta y} \right) \quad (4.5)$$

We map the salience vector c to a salience matrix $T = \{t_{ij}\}_{i,j \in \Omega}$, where t_{ij} represents the salience score between ROI i and j .

4.2.3 Decoding functional brain connectivity

We define decoding functional brain connectivity as identification of salient functional connectivity features that are crucial for classifying different brain states. The classification models are built using functional connectivity features for rest and task data while the DeepLIFT algorithm is used to find the feature salience. Although there are a large number of functional connections and brain regions, only a few are consistently involved in task performance. Additionally, these distinguishing connections and regions vary according to the task and there is no definitive way to select an important subset from the given features. In our case, it is known that the functional connectivity features with a high positive salience score represent the features whose presence during the task state is important (and vice versa), but we need to identify a smallest subset of features that are able to correctly distinguish between rest and task states. For this purpose, we evolve a recursive feature elimination framework that takes into consideration the salience scores of the input features and choose only the salient features.

We perform classification between the task and rest samples by using a salient subset of input features identified by the decoder. This not only gives us the minimal subset of features needed for the classification but also tests the validity of the features identified by the decoder. We recursively removed input features that were insignificant such that salience $|t_{ij}| < \alpha$ of connection between brain regions i and j less than a threshold α , and retrained the neural network on the reduced feature set. The threshold was correspondingly set at the 90th percentile of significance. We repeated this experiment till only 0.1% initial features were left. We did not go below 0.1% as that corresponded to ≈ 34 features which is sufficiently small and gave an accuracy close to the base accuracy. It is important to note that obtaining a similar accuracy with a smaller feature set does means that a smaller set of features were sufficient for classification. This subset of features thus defines the decoded functional architecture of the brain.

4.2.4 Task-general brain architecture

The alterations between resting state and a task state are an amalgamation of both task-general and task-specific brain modulations. Additionally, task-general modulations are represented by a common set of features that distinguish the brain at rest and during varied task states. In our formulation, connectivity features are the inputs to the DNN classifier and the task general architecture is defined by salient input features classifying the rest task vs all the tasks. We hypothesize that if we classify the brain at rest and during multiple varied task-states, then the set of most salient features distinguishing rest vs all tasks represent the task-general architecture of functional connectivity. To test this hypothesis, we identified the features with high salience scores distinguishing between rest and multiple dissimilar task-states.

The sign of salience t_{ij} of a connection determines whether the *presence* or *absence* of functional connectivity between brain regions i and j leads to the differentiation of the rest-state from the task-states. This information is in turn used to derive positive and negative task-general architectures such that the presence of a positive task-general connection during tasks is crucial to differentiate the task states from the rest state and vice-versa. We derived the task general architecture defined by the seven task in the HCP data.

4.2.5 Task-specific brain architecture

Once the task-general architecture is derived, we obtain the task-specific architecture with salient features obtained from a DNN trained to classify rest and the task under investigation. The salient features for rest vs single task classification represent both task-general and task-specific brain modulations. We hypothesize that the salient connectivity features distinguishing a particular task from the rest represents both the task-general and task-specific architectures. Therefore, salient features from the DNN classifying a task and the rest state and the task-general architecture, we can derive the task-specific architecture.

In order to demonstrate determination of task-specific architecture, we used the subject data from the motor and language tasks and performed classification for rest vs motor and rest vs language tasks. Let $T_{rest/tasks}$ denote the salience of connectivity features representing the task-general architecture. If the salience of features in rest vs motor and rest vs language (containing both task-general and task-specific modulations) are denoted by $T_{rest/motor}$ and $T_{rest/language}$, respectively, then the motor and language task-specific architectures are given by the salient features highlighted by $T_{rest/motor} - T_{rest/tasks}$ and $T_{rest/language} - T_{rest/tasks}$, respectively.

4.2.6 Salient interactions between functional modules

Changes in brain functional architecture are often discussed in terms of changes between different brain functional modules. The brain is organized into functional modules consisting brain regions involved in specialized functional tasks. We quantify the importance of interactions among different brain functional modules by *modular salience maps*. Given functional modules a and b , the salience of interactions $m_{a,b}$ between them is obtained by

$$m_{a,b} = \text{mean}_{i \in a, j \in b} \{t_{i,j}\} \quad (4.6)$$

The modular salience map $\{m_{a,b}\}$ represents the average salience of the nodal connections within (when $a = b$) and between (when $a \neq b$) brain functional modules. The modular salience map provides a powerful framework to interpret and understand brain functional changes in terms of how the interactions between different functional modules vary during different brain states (in this case, tasks).

4.2.7 Inter-subject variability in the task-related functional architectures

Multiple studies have pointed out the pronounced inter-subject variability in the brain functional architecture both during rest [84, 85, 101, 153, 227] and task performance [82, 225, 226]. It is, therefore, important that the detected task-general and task-specific architectures consist of features that are stable across subjects. We measure the inter-subject variability in the feature salience scores for different subjects as:

$$c_{std} = \text{std}_x \left(f(\bar{x}) - g(x) - \sum_{i,l \in \{1, \dots, L-1\}} C_{\Delta h_{i,l} \Delta y} \right) \quad (4.7)$$

where std measures the standard deviation of the salience scores over multiple subjects. The mean of the salience scores is given by salience vector c (given by equation 4.5). We perform regression on the mean, c , and the standard deviation, c_{std} , of the salience scores to understand the relationship between the two sets of values.

4.3 Experiments and results

4.3.1 Dataset

We used data from the Human Connectome Project (HCP) [125], comprising fMRI data from 897 healthy adults (mean age = 28.1 years, 390 females). The Human Connectome Project (HCP) currently hosts one of the largest open databases of fMRI data. We used both the task-evoked fMRI (tfMRI) for 7 tasks and the rest-state fMRI (rs-fMRI) data of subjects, collected from a Siemens 3T Skyra scanner with TR = 720ms, TE = 33.1ms, flip angle = 52°, FOV = 208 x 180 mm, 2mm x 2mm isotropic voxel and 72 slices. We used the preprocessed data release that includes removal of spatial artifacts, motion correction, intensity normalization and denoising [127]. We further processed the task data to fit a finite impulse response (FIR) model to each of the 24 subtasks for the 7 tasks (emotion processing, gambling, language, motor, relational processing, social cognition, and working memory) and removed the mean task evoked activations from time-series data. This step is necessary as these task related activations have been found to inflate task-state functional connectivity inappropriately [237], especially with fMRI data. We also removed the rest periods from the task related time series for each region before computing functional correlations.

Table 4.1 summarizes the available data for each task and the rest-state. The average number of time points for the task fMRI scans was 277. We divided each fMRI resting state BOLD scan of 1200 time points into 4 scans of 300 time points each. From each of these truncated BOLD time series, we obtained 4 functional connectivity matrices. We obtained 690 subjects having scans for all 7 tasks and the resting state since not all subjects had been able to complete every task and resting state scans. For each subject, by considering one for each encoding direction, we obtained 8 resting state functional connectivity matrices and 14 task state matrices.

A functional network was derived for each fMRI scan. We selected anatomically and functionally diverse 264 ROIs identified by Power et al. [14]. The details are given in section 2.4.2.

4.3.2 Building the classifiers

Using the functional connectivity between 264 ROIs identified by Power et al. [14] as inputs, we trained classifiers on both task-fMRI (tfMRI) and rest-state (rfMRI) data. We ensured that

Table 4.1: Summary of HCP task and resting state data

Condition	Number of Subjects	Number of Time Points	Description
Resting	756	1200	Participants were asked to lie flat, with eyes open and maintain a relaxed fixation.
Emotion processing	806	176	Participants were tasked to match two simultaneously presented faces or shapes on two screens.
Gambling	829	253	Participants were given incentive to guess the number of a mystery card.
Language	805	316	Participants are presented with a question on short narrated stories.
Motor	826	284	Participants were asked to either move their fingers, toes or their tongue on cue.
Relational processing	802	232	Participants are presented with relational and control matching task.
Social cognition	807	274	Participants were presented with short videos and decided if the object movement were inter-related.
Working Memory	828	405	Participants were asked to complete a two-back and a zero-back working memory tasks.

all the resting state and task scans of a subject are either in the test or the train set. We used Tensorflow [124] and Keras [123] python libraries to implement the DNN. Since, there is a class imbalance with ≈ 9650 task samples and ≈ 8300 resting state samples, we also computed the mean absolute error (MAE) that is the mean of the absolute errors between the predicted outputs and the ground-truth labels. We first split the dataset into a train and test set in the ratio 4 : 1 and then perform 5-fold cross validation to select the best model parameters on the train set. We ensured that all the samples for each subject were either in the train or test set (and note that the samples from the same subject were in the same fold during 5-fold cross validation).

We implemented different classifiers using the feedforward neural networks (FFN), convolution neural network (CNN) [192, 238] and support vector machines (SVM). We empirically tune the classifier parameters to determine the best model. For the FFN, we varied the number of hidden layers $L \in \{1, 2, 3\}$, the number of hidden neurons in the first hidden layer in $\{5, 10, 20, 50, 100\}$. For the CNNs, the input was the connectivity matrix and we varied the number of filters in each layer based on the number of weights in the corresponding FFN such that the number of trainable parameters (from 1.7×10^4 to 3.5×10^6) were same in both the architectures. The layers were similar to the ones implemented in [192, 238] where we had architectures with and without the Hadamard layer and with the edge-to-node and node-to-graph filters. For SVMs, we varied the parameters ($C \in \{0.001, 0.01, 0.1, 1\}$, $\gamma \in \{0.001, 0.01, 0.1, 1\}$) and tried different kernels such as linear, polynomial, radial basis function and sigmoid. For the FFN, the number of hidden layers and the number of neurons in each hidden layer were varied. We obtained the final architecture by using grid search from the accuracies of 5-fold cross validation. The details of the different configurations for the FFN and CNN models can be found in the appendix C.

For all the experiments (rest vs all tasks and rest vs specific tasks), we found that all the classifiers achieved classification accuracy above 90%. However, SVM based classifiers gave lower accuracies while the FFN and CNN based classifiers gave similar accuracies for the same number of parameters. We, therefore, used FFN based classifiers with optimal hyper parameters for feature selection and decoding. We used the best FFN model after performing experiments on the train and test set for five different FFN parameter initializations. Since the accuracies for all the classifiers were high, we performed additional experiments to determine the time taken for error convergence for the different classifiers and their performance for different subset of

Table 4.2: Performance of different classification models on the HCP data with 5-fold cross-validation.

Classes	FFN	CNN	SVM
rest vs all tasks	99.75	99.98	98.26
emotion	99.70	99.93	97.06
gambling	99.78	99.91	97.31
language	99.88	99.96	98.78
motor	99.63	99.94	97.25
relational	99.57	99.97	97.08
social	99.72	99.92	97.40
working memory	99.53	99.95	97.18

training samples (refer figures 4.1, C2 and C3). We found that the test performance for the FFN converged faster than the corresponding SVM classifier for all the cases.

4.3.3 Decoding the task-general brain architecture

We obtained the best set of hyperparameters for the rest vs all tasks classification as described in the previous section. For the best FFN model, the configuration is given in table 4.3. Using the trained classifier, we computed the salience for the input connectivity features by using DeepLIFT. The salience scores were computed using samples from the train set. It was observed that the majority of features had positive salience (number: 31381, range: 0 to 1891.6, average: 81.96 ± 124.2) in comparison with features with negative salience (number: 3335, range: 0 to -259.60, average: -29.62 ± 31.42). The histogram for distribution of salience scores and the feature salience score matrix $T_{rest/tasks}$ are available in the appendix C (figure C1a and figure C1b).

In order to decode the task-general architecture, we recursively eliminated the input features, using the salience scores obtained from the FFN on the train set. We observed that although the classifier performance slightly deteriorated, it was able to achieve more than 99% accuracy even with just 0.1% of initial number of features. For SVM, we also performed recursive feature elimination based on the corresponding weights obtained for the best SVM model (linear kernel, $C = 0.001$). The SVM classification accuracy suddenly drops when there are just 0.1% features left. The respective MAE and accuracies for both FFN and the SVM classifiers can be discerned from table 4.4. The recursive feature elimination experiments validate the correctness of the

Fig. 4.1: The accuracy and time taken for convergence of different subsets of samples for rest vs all tasks ((a) and (b)); rest vs motor task ((c) and (d)); and rest vs language task ((e) and (f)).

Table 4.3: Parameters of the 4 layer FFN used for classification of rest vs all tasks and rest vs language task.

Parameter	Value
Input Nodes	34716
Hidden Nodes (each layer)	34716-100-32-2
Learning Rate	0.0001
Batch size	8
Optimizer	Adam
Dropout	0.1

Table 4.4: Performance of the FFN and the SVM (on the test set) for rest vs all task classification using different fractions of features. MAE denotes mean absolute error.

% features	FFN		SVM	
	Accuracy	MAE	Accuracy	MAE
100%	99.81	0.0048	98.35	0.0164
10%	99.88	0.0022	98.8	0.0118
1%	99.98	0.0004	98.38	0.0161
0.1%	99.98	0.0011	93.43	0.0657

salience scores derived from our decoder. More importantly, they indicate that a small subset of features that is sufficient to distinguish between the rest and task states across subjects, representing the task-general brain architecture. These subset of connections form the brain’s task-general architecture. For the FFN, we obtained the consensus of the 0.1% (34 connections) remaining connections from each run and show the decoded features in figure 4.2. From figure 4.2, it can be seen that the top 0.1% task-general salient connections were among the default mode network, fronto-parietal (FP) task-control, salience network, cingulo-opercular (CO) task-control, and somatomotor hand modules.

The modular salience maps highlighting the task-general architecture of the brain were derived using DeepLIFT (refer figure 4.3b and 4.3a), where the functional modules were used from Power et al. [14]. Connections with positive and negative salience scores in $T_{rest/tasks}$, correspond to positive task-general and negative task-general connections, respectively. We compute the modular salience maps for task-positive and task-negative brain architectures by using equation 4.6 (refer figure 4.3b and 4.3a). Using the salient modular interaction scores, we computed statistically significant interactions having p -value < 0.05 (assuming a

Fig. 4.2: The decoded task-general brain architecture obtained from consensus over multiple runs. Figure corresponds to the decoded top 0.1% task general brain architecture. Edge thickness corresponds to the number of runs these connections were classified as relevant for distinguishing rest and task states, and the node color corresponds to the node's modular membership.

Fig. 4.3: The task general architectures in the brain at the modular level. (a) and (b) depict the modular salience matrices for negative and positive task-general connections, respectively.

Table 4.5: Significant task-general modular interactions.

Module	Module	p -value
Positive task-general		
Motor Mouth	Motor Mouth	$< 10^{-3}$
Visual	Visual	$< 10^{-3}$
Memory retrieval	Memory retrieval	$< 10^{-3}$
Negative task-general		
Cerebellar	Salience	$< 10^{-3}$
Salience	Salience	$< 10^{-2}$
Salience	FP Task Control	< 0.02
Motor Hand	Motor Hand	< 0.02
Default mode	FP Task Control	< 0.02
Cerebellar	FP Task Control	< 0.03
FP Task Control	Motor Hand	< 0.03
Default mode	Motor Hand	< 0.04
Default mode	Ventral attention	< 0.05
CO Task Control	FP Task Control	< 0.05
FP Task Control	FP Task Control	< 0.05

normal distribution of scores). We report significant task-positive and task-negative modular interactions in table 4.5. It can be seen that the default mode network, FP task-control, and dorsal attention networks are involved in both task-positive and task-negative interactions.

4.3.4 Decoding motor task-specific brain architecture

In the motor task (adapted from Buckner et al. [239]), participants were asked to either tap their left or right fingers, or squeeze their left or right toes, or move their tongue on cue. Each block of movement style lasted 12s and is preceded by a 3s cue. In each run, there were thirteen blocks of which there were two tongue movements, four hand movements, four foot movements and three 15s fixation blocks.

The configuration of the best FFN model derived for classifying rest vs motor task is given in table 4.6. For the SVM, the best accuracy was obtained with a linear kernel and $C = 0.001$. We achieved an average accuracy of 99.65% and 97.82% using all the features for the rest vs motor task with the FFN and SVM models, respectively. We show the results for the recursive feature elimination experiments with both FFN and SVM in table 4.7. For the FFN, the MAE

Table 4.6: Parameters of the 5 layer FFN used for classification of rest vs motor task

Parameter	Value
Input Nodes	34716
Hidden Nodes (each layer)	34716-20-10-8-2
Learning Rate	0.0001
Batch size	8
Optimizer	Adam
Dropout	0.1

Table 4.7: Performance of the FFN and the SVM for rest vs motor task classification using different fraction of features. MAE is mean absolute error.

% features	FFN		SVM	
	Accuracy	MAE	Accuracy	MAE
100%	99.65	0.0051	97.82	0.0217
10%	99.63	0.0042	98.14	0.0185
1%	99.81	0.0025	98.51	0.0149
0.1%	99.94	0.0010	97.58	0.0242

improved with fewer features and the accuracy for classification with just 0.1% features was the highest (and close to 100%). For the SVM, the accuracy and MAE improved initially, but there was a deterioration when only 0.1% features were left.

In order to obtain the task specific architecture, we used the salience scores of the connections for the rest vs motor task and subtracted the task-general scores. We averaged the feature salience scores $T_{rest/motor}$ for all the seeds and subtracted $T_{rest/tasks}$ from $T_{rest/motor}$. We computed the top 0.1% relevant task specific features corresponding to the motor task. A consensus of these connections over multiple runs is given in figure 4.4. As a result, we were able to decode a set of 34 task-specific connections from the FFN that were important for the classification task. The motor-task specific architecture was represented by correlations between the somatomotor hand, salience network, dorsal attention, FP Task control and default mode network (pre-frontal cortex) subsystems.

Using the task specific salience scores, we obtained the module salience maps for both the positive and negative task-specific features for the motor task (refer appendix figure C4(b) and C4(a)) and obtained the significant modular interactions (with p -value < 0.05). The significant

Fig. 4.4: The motor task-specific architecture in the brain. Figure shows the top 0.1% task-specific connections for the motor task. The edge thickness corresponds to the saliency score and the node color corresponds to the node's modular membership.

task-specific motor modular interactions involved the somatomotor mouth, somatomotor hand, CO Task Control, Dorsal attention and the FP Task Control modules (table 4.8).

4.3.5 Decoding language task-specific brain architecture

The language task was proposed by Binder et al. [236], where participants were presented with short narrated stories and given a two-alternative forced-choice question about the topic of the story. The 2 runs that were interleaved with 4 blocks of a story task and 4 blocks of a math task. Story blocks presented brief sentences from Aesop's fables, followed by 2-choice questions querying the topic of the story, whereas math blocks were presented aurally and required the subject to compute addition and subtraction problems and chose the result among the 2 choices by a finger tap feedback.

For the rest vs language task the best FFN configuration was the same as rest vs all tasks (given in table 4.3). For 5 seeds, we achieved an average accuracy of 99.86% using all the features for the language task classification. The accuracy for classification with just 0.1% features was higher than the accuracy for 100% features (refer table 4.9). For the selected SVM model (linear kernel, $C = 1.0$), we were able to achieve a high classification accuracy with the

Table 4.8: Significant modular interactions specific to the motor task.

Module	Module	p -value
Positive motor task-specific		
Motor Hand	FP Task Control	$< 10^{-3}$
Motor Mouth	Motor Mouth	$< 10^{-3}$
Dorsal attention	Default mode	$< 10^{-3}$
Motor Hand	Default mode	$< 10^{-2}$
Memory retrieval	Memory retrieval	$< 10^{-2}$
CO Task Control	Default mode	$< 10^{-2}$
Visual	Visual	$< 10^{-2}$
Motor Hand	Salience	$< 10^{-2}$
CO Task Control	FP Task Control	< 0.05
Negative motor task-specific		
Visual	Visual	$< 10^{-3}$
FP Task Control	FP Task Control	$< 10^{-2}$
Dorsal attention	Dorsal attention	$< 10^{-2}$
Motor Mouth	Auditory	$< 10^{-2}$
Memory retrieval	Memory retrieval	$< 10^{-2}$
Motor Mouth	Motor Mouth	< 0.02
Motor Mouth	Motor Hand	< 0.03
Motor Mouth	FP Task Control	< 0.05

Fig. 4.5: The language task-specific architecture in the brain. Figure shows the top 0.1% task-specific connections for the language task. The edge thickness corresponds to the salience score and the node color corresponds to the node’s modular membership.

whole feature set. However, the performance with SVM deteriorated with fewer features both in terms of accuracy and MAE. We averaged the feature salience scores $T_{rest/language}$ for all the seeds and obtained the language task-specific feature salience scores by subtracting $T_{rest/tasks}$ from $T_{rest/language}$. We computed the top 0.1% relevant task specific features (shown in figure 4.5) corresponding to the language task and observed that the connections from the auditory (mostly left: L temporal), FP task-control (angular L and right: R), default mode network (inferior parietal L and R, and frontal), salience network (anterior cingulum), somatomotor hand (post central gyrus) and visual (occipital lobe) were language task specific. Since the language network is not delineated clearly in the Power Atlas networks, we also computed the salience of connections between different anatomical regions by mapping regions from the Power atlas to regions in the Crossley atlas [15], with known anatomical labels, using their Euclidean distance. We computed the salience between the functional connectivity of different regions by averaging the salience for connections between multiple Power atlas regions mapped to the same region from the Crossley atlas (shown in figure 4.6).

Using the task specific salience scores, we obtained the module salience maps for both the positive and negative task-specific features for the language task (refer appendix figure C4(d) and C4(c)) and obtained the significant modular interactions (with p -value < 0.05). The

Table 4.9: Performance of the FFN and the SVM for rest vs language task classification using different fraction of features. MAE is mean absolute error.

% features	FFN		SVM	
	Accuracy	MAE	Accuracy	MAE
100%	99.86	0.0067	99.47	0.0082
10%	99.79	0.0098	99.51	0.0098
1%	99.82	0.0055	99.34	0.0064
0.1%	99.92	0.0030	98.10	0.0189

Fig. 4.6: The salient anatomical connections for the task-specific language architecture in the brain. (a) and (b) depict the salient connections between regions for the positive and negative language task-specific connections, respectively.

Table 4.10: Significant modular interactions specific to the language task.

Module	Module	p -value
Positive language task-specific		
Motor Hand	Motor Hand	$< 10^{-3}$
FP Task Control	FP Task Control	$< 10^{-3}$
Dorsal attention	Dorsal attention	$< 10^{-3}$
Motor Hand	Motor Mouth	$< 10^{-3}$
Salience	Salience	$< 10^{-2}$
Ventral attention	Ventral attention	< 0.03
Motor Hand	CO Task Control	< 0.03
Motor Hand	Dorsal attention	< 0.04
Negative language task-specific		
Motor Mouth	Motor Mouth	$< 10^{-3}$
Visual	Visual	$< 10^{-3}$
Dorsal attention	Default mode	$< 10^{-3}$
Memory retrieval	Memory retrieval	$< 10^{-2}$
CO Task Control	Default mode	$< 10^{-2}$
Motor Hand	FP Task Control	< 0.02
Motor Hand	Default mode	< 0.03
Motor Hand	Salience	< 0.04

significant task-specific language modular interactions involved the visual module (regions from the occipital lobe), the salience network (anterior insula and dorsal anterior cingulate cortex), somatomotor hand, somatomotor mouth, dorsal and ventral attention modules, memory retrieval (precuneus and posterior cingulate cortex) modules (table 4.10).

4.3.6 Inter-subject variability in task-related architectures

We measured the inter-subject variability in the feature salience scores by considering the mean of the salience scores, c , and the inter-subject variability in the salience scores c_{std} . The scatter plots for the normalized mean and standard deviation scores are given in figure 4.7. While the green markers represent all the features, the decoded task-general (figure 4.7(a)) and task-specific architectures (figure 4.7(b) for motor and (figure 4.7(c) for language task) are represented by features in blue color. We performed polynomial regression for the scatter plots and the best fits for a quadratic polynomial are shown by a red curve in the figures. On computing the

correlation between the mean c and the inter-subject variability c_{std} in the salience scores, we found that these two are weakly but positively correlated (Pearson’s correlation for task-general = 0.43, motor task-specific = 0.56, and language task-specific = 0.21). It can be observed that the decoded task-related architectures consist of features that have a high salience score and low variability across subjects. However, some features that were on the lower end of the salience scores and had a low variability across subjects were also included. This can be discerned from the fact that most of the decoded features are below the quadratic polynomial curve fit for the respective task.

4.4 Discussion

4.4.1 Classification of rest and task brain functional scans

We performed classification for rest vs all tasks and rest vs individual tasks (for tasks besides motor and language, refer to appendix C). We found that although shallow models such as SVM give a high classification performance but deeper models are able to achieve a higher accuracy even with a subset of input features. This points to the presence of non-linear and hierarchical relationships between functional connectivity and the task labels. The decoded task-related functional brain architecture is composed of a small set of functional connectivity features. This was validated by our feature elimination experiments, where a small fraction of salient connections between brain regions were found to accurately distinguish brain resting and task states.

Recursive feature elimination experiments with SVM models show a deterioration in classification performance with a small set of connectivity features. This could be due to the fact that SVM models are not able to distinguish between the rest and task architectures as well as the DNN even with a full feature set. Our results differ from the results from He et al. [228] which showed that simple models trained on functional connectivity can achieve comparable accuracy as complex deep models for a behavior prediction task. In the succeeding chapter, we will use functional connectivity features for disease diagnosis and see whether simple models are able to perform as well as DNNs.

Although we removed the resting state time segments from the time series of the task data, a limitation of this study arises from the fact that we did not consider the different sub-tasks

(a) Task-general features

(b) Motor task-specific features

(c) Language task-specific features

Fig. 4.7: Inter-subject variability in task-related architectures. Figures (a), and (b), (c) shows the scatter plot for the mean and variability in the task general and task specific salience scores across subjects. The blue markers represent the decoded task-related features selected and the red curve represents the quadratic polynomial fit by regression.

in each task state separately. The sub-tasks in each state were assumed to be similar to each other and involved similar cognitive processing. But these sub-tasks may invoke different brain regions. If we consider each sub-task separately, we are left with fewer time points that leads to poor signal-to-noise ratio and ultimately degraded detection of functional activity [240]. We, therefore, used the whole task scan as one specific task for our modelling.

From our results, we allude that although healthy subjects have pronounced inter-subject variability in functional architecture [85], a small set of common distinguishing features characterize task related alterations across subjects. From our experiments, we inferred that there were two roughly two sets of features in the selected task-related architectures, one that had a high salience and high variability and another that had a moderate salience and low variability. We inferred that these two together help classify the rest and task states with an almost perfect accuracy, with the former handling corner cases, while the latter form a stable core across all tasks. This has an important implication in aging [218] and detection of various neurological disorders that are characterized by impairments/alterations in the task-general architecture.

4.4.2 Task-general architecture

The positive and negative task-general alterations found by our approach do not refer to activations/deactivations of regions, but to the presence or absence of the connectivity during task performance with respect to the rest state. The regions can have positive/negative/weak correlation during the rest state. We found that the positive task-general modulations are more wide-spread and apparent than negative task-general modulations. Similarly, different interactions from the same module can be part of the negative and positive task-general alterations. For negative task-general alterations, we found that it was primarily the connections of the regions in the default mode network (posterior cingulate cortex, precuneus, medial prefrontal cortex, middle temporal gyrus), the fronto parietal task control (prefrontal and intra-parietal sulcus) and the salience (anterior insula and dorsal anterior cingulate cortex) modules with other modules, which were altered. However, for positive task-general, it is primarily the connections with the somatomotor mouth, the visual and the memory retrieval modules that were altered. It is well-known that the cerebellum contains multiple somatomotor representations and relays information between the sensory neurons to the corresponding motor neurons in the cerebrum [239]. Since all tasks involved visually presented cues and the participant responses

by motor movement (finger tap in all except social cognition task which involved a verbal response), our results on the task-general architecture have invariably included the modulations between the cerebellar, motor and visual systems. The default mode network has been characterized as a “task-negative” network and it plays a role in extended support of internal mentation processes [12, 241] and observing internal and/or external environment [242]. The CO task control, FP task control and the salience networks help in configuring the brain to different brain states, and configuring information processing in response to different task demands. These networks thus support varied cognitive functions including sensory perception and motor control. The FP task control network acts as a flexible hub for cognitive control facilitating special attention to trial specific information [243] whereas the CO task control network maintains alertness during task performance and correction based on feedback from each trial [244]. The salience network determines the importance given to different stimuli for the brain to act and is known to be widely activated across different tasks [245]. Our results are thus in line with the widely-known task-related deactivations that are observed in the default mode network [246] and the memory retrieval systems [12, 14] while the the FP task-control, CO task-control and ventral attention systems are known for their task-related activations [14, 247, 248].

4.4.3 Task-specific architecture

We investigated task-specific modulations for tasks that involved more perceptual and fewer cognitive processing systems, so that we cannot only identify task-specific modulations but also verify them. The task-specific alterations for both the motor and the language tasks were found to be limited to primarily the somatomotor, attentional, the task control, and the salience systems.

It can be seen that the motor task-specific modulations within visual and motor mouth modules are reported as both positive and negative, which points to the differing role of connections within the same module. Although some of these alterations were also reported to be task-general alterations, but the direction (positive/negative) is different in case of task-general and task-specific alterations. Besides moving the hand and feet, the motor task also involved tongue movements. These are reflected in the connectivity of multiple systems with somatomotor hand and mouth systems being classified as crucial. The dorsal attention network is involved in voluntary attention and is altered when directed attention is required [248]. For

the language task, we found that the most distinguishing connections were from regions in the supramarginal gyrus, occipital lobe, angular gyrus, lingual gyrus, cuneus, temporal gyrus, and the frontal orbital lobe. While the supramarginal gyrus, angular gyrus, lingual gyrus and cuneus are well known for their role in language and visual processing tasks, multiple studies [236, 249–251], have reported extensive activations in different areas of the temporal lobe along with areas in the lateral occipital lobe and the frontal lobe during language processing. Thus, we have not only shown that a small set of features is able to classify the rest and task states, but also verified that these feature sets have been deemed important by previous studies.

4.5 Chapter Summary

In this chapter, we detected task-general and task-specific brain architectures by decoding the DNNs trained to classify rest-state and cognitive task-states fMRI data. Using functional connectivity between a set of functionally diverse regions as features for a 5-layer feedforward DNN, we were able to reliably classify rest-state and 7 tasks and rest-state from motor task scans of 690 subjects from the HCP with nearly 100% accuracy. Using salience backpropagation based decoder, we identified connections that form the task-general and task-specific brain architecture in the brain. We also studied what modular interactions are changed during task performance. We validated the results of our decoder by showing that a small subset of features deemed important by the decoder gives us a high accuracy. We studied the inter-subject variability in the task-related functional architectures.

The proposed methods have strong potential of replication in brain state classification and identification of distinctive features between the healthy and diseased brain states. In the next chapter, we extend the proposed techniques to derive more generalizable classification models for detecting brain ailments and also detect crucial disease biomarkers using the reference based decoder.

Chapter 5

Optimizing brain encoders for detecting diseased brain states ¹

In the previous chapters, we studied the brain functional connectome during rest and the changes in the functional connectome caused by varied tasks. We suggested using an encoder and a decoder in order to study the alterations in functional connectivity during task performance. However, for computer aided detection of diseases, the focus is not just on the decoding but also on efficient encoding. In this chapter, we remove the redundancies in both the functional connectivity and the encoder parameters by using outputs of the decoder. These redundancies are removed to mitigate the problem of overfitting that occur when deep neural network (DNN) models are used to classify neuroimaging datasets. We propose the LEAN algorithm that removes insignificant nodes of DNNs in a layerwise manner and the CLIP algorithm that adds a subset of correlated features in a single shot. When performing experiments with functional MRI datasets for classifying patients from healthy controls, we were able to obtain simpler and more generalizable DNNs. The obtained DNNs maintained a similar performance as the full network with only around 2% of the initial trainable parameters. Further, we used the trained network to identify salient brain regions and connections from functional connectome for multiple brain disorders. The identified biomarkers were found to closely correspond to previously known disease biomarkers. The proposed methods have cross-modal applications in obtaining leaner DNNs that fit the data better.

¹The work in this chapter has been published in [190,252]

5.1 Disease detection using functional connectivity

Deep neural networks (DNN) have been successfully applied to a wide variety of classification problems, ranging from image recognition to language processing [253, 254]. In these applications, the choice of network architecture and number of trainable parameters affects network performance significantly. Larger networks learn more complex and accurate mappings, which comes with increased computational complexity and possibilities of overfitting [255]. On the other hand, having insufficient number of parameters limits the network's ability to learn the correct mapping [256]. Therefore, finding the optimal neural network architecture is an important and challenging problem. One potential solution involves removing redundant nodes and connections from a large neural network. This has been shown to produce comparable performance with a smaller network but with better generalization capability [257, 258].

Neural networks and machine learning techniques have been successfully applied to analysis of neuroimaging data. Non-invasive neuroimaging techniques are used to characterize functional and structural anomalies in the brain and aid in better diagnosis and treatment [60]. Functional and/or structural connectivity derived from these techniques are used as features for the neural network models to classify diseased and normal subjects [61, 62]. However, such studies typically involve many more input features ($\sim 10,000$) than subject scans ($\sim 1,000$), making the neural networks prone to overfitting. Also, experimental noise introduces systematic connectivity [259], which should be ignored during the classification task. Most crucially, overfitting is exacerbated by the presence of redundancies: not all functional connectivity features are important for differentiating between scan samples from two subject groups.

In the previous chapter, we elucidated how model interpretability techniques can be used to find salience of input features. In order to find salient features, DeepLIFT [232] was used to compute salience scores of the input features. DeepLIFT is one of the recent attempts (besides Integrated Gradients [233], Layerwise Relevance Propagation [260] and SHAP [234]) that circumvent the issues with gradient-based approaches (such as zero gradients or discontinuities). These approaches find the contribution of nodes at each layer by propagating the contributions from the output layer. A major differentiating factor of DeepLIFT is that it gives consideration to both negative and positive contributions and computes the relevance scores efficiently in a single pass.

A preliminary work to arrive at much smaller models would be to recursively remove a fraction of the least salient input features, trying to reduce the problem of overfitting without compromising on classification accuracy. However, one limitation of the recursive feature elimination approaches is their efficiency: they involve multiple elimination and fine-tuning steps to arrive at a smaller neural network. In this chapter, we propose an algorithm called Layerwise Elimination of Accessory Nodes (LEAN) that performs the elimination of accessory nodes in a single shot. Accessory nodes are defined as nodes that do not influence the classification task significantly - such nodes will be assigned a low importance score by a valid decoder. LEAN uses node salience scores derived from a valid DNN decoder to eliminate accessory nodes in the network, producing a leaner, more computationally efficient, and more generalizable deep neural network that is derived in a single shot.

LEAN leads to a drastic reduction in parameters, including the removal of a large number of input features. While powerful and efficient, this also leads to the loss of whole groups of correlated features. In neuroimaging datasets, such correlations exist due to the *specialization* property of the brain, whereby distinct sub-systems in the brain perform specialized tasks [71]. To tackle this, we introduce another algorithm called Correlation-based eLImination of InPUts (CLIP) that identifies and retains a subset of correlated input features. When these features are combined with the remaining input features from LEAN, it leads to an improvement in the DNN performance as compared to LEAN. Combining LEAN and CLIP leads to a model that provides the best of both worlds: low number of parameters with minimal accuracy drop.

While encoding the functional connectivity serves the goal of getting an efficient classifier with a higher accuracy and generalization ability, a relevant and important research direction in neuroscience is to find the biological markers associated with neurological diseases. We denote identification of brain regions and connections that are associated with a particular brain state as brain decoding. Traditional methods for brain decoding include multivariate pattern analysis [261], sparse networks based feature selection [262], latent Dirichlet allocation [263], and the use of nodal features [264]. Such methods are however based on simple or linear models. On the other hand, DNN-based approaches build deep and hierarchical models and represent key patterns underlying brain activation in its nodes and connection weights [262]. Deep learning techniques make no assumptions about application specific a priori knowledge and therefore give consistent and unbiased decoding based on neuroimaging data. Recent

applications of neural networks on neuroimaging data have derived the importance of input features in the classification task to uncover biomarkers for neurological diseases [265] or reveal task-related brain functional alterations [266]. Identifying disease-specific biomarkers aids in building models that aid in classifying and predicting the progression or occurrence of unknown diseases in individuals. To do so, we use the saliency scores given by DeepLIFT to determine salient brain connections and regions for Alzheimer’s disease (AD), mild cognitive impairment (MCI), attention deficit hyperactivity disorder (ADHD), major depressive disorder (MDD), and Autism Spectrum Disorder (ASD) patients using their functional MRI scans.

In this chapter, we proposed an efficient layerwise node elimination approach, LEAN, that removes accessory input features and hidden layer nodes from a DNN in a single iteration. We further adapted it for brain decoding by proposing CLIP to add in a subset of correlated features, leading to a model with minimal number of connection while maintaining the accuracy. The proposed approach finds salient functional connections and identifies brain regions that are responsible for classification of brain disease. The proposed methods were applied on multiple brain disorders for classification of patients from healthy controls and identification of disease biomarkers.

5.2 Methods to optimize brain encoding and decoding

5.2.1 Feedforward DNN encoder

The formulation of the feedforward DNN is the same as in section 2.3.3. Given a pair (x, d) where $x = (x_i)$ denotes the input feature vector and $d \in D$ denotes the sample label, we used a feedforward neural network of L layers with the first $L - 1$ layers having rectified linear unit activation and a softmax layer as the last layer of the network.

5.2.2 Computing the saliency of nodes in network layers

We use the reference based decoder, DeepLIFT, that has been described in section 4.2.2 to derive the contribution of each neuron i in layer l a contribution $C_{\Delta h_{i,l} \Delta h_{k,l+1}}$ to the change in the output of neuron k in layer $l + 1$:

$$\sum_{i \leq n_l} C_{\Delta h_{i,l} \Delta h_{k,l+1}} = \Delta h_{k,l+1} \quad (5.1)$$

when $l \in \{0, \dots, L - 2\}$ and

$$\sum_{i \leq n_{L-1}} C_{\Delta h_{i,L-1} \Delta y} = \Delta y \quad (5.2)$$

when $l = L - 1$, where $\Delta h_{i,l}$ is the change in the activation of the i^{th} neuron of layer l due to the input relative to the reference. $C_{\Delta h_{i,l} \Delta h_{k,l+1}}$ is computed from the Linear, Rescale and RevealCancel rule from [232]. We obtain the difference in the output with respect to the reference $\bar{x}$ as $\Delta y = f(x) - f(\bar{x})$.

Contribution of nodes in all layers $0 \leq l < L$ to the output y is obtained by backpropagating contributions of layers to the output. For nodes in each layer $l \in \{0, \dots, L - 1\}$, we compute the *saliency score vector* c_l given by:

$$c_l = (C_{\Delta h_{i,l} \Delta y})_{i < n_l} \quad (5.3)$$

where $C_{\Delta h_{i,l} \Delta y}$ is derived from Equation 4.4. The DeepLIFT method [232] computes the layer's saliency scores based on change in the output from a reference, allowing information to propagate across the network layers even when the gradient is zero. We compute the average contributions for all layers over multiple test samples to get the final saliency score for each feature.

5.2.3 Layerwise elimination of accessory nodes (LEAN)

A decoder can identify a subset of distinguishing input features and weights that are important to classify samples. We propose a brain decoding strategy where we obtain the saliency scores for input features and hidden layer nodes. A popular approach for feature reduction in machine learning is recursive feature elimination, whereby a fraction of least important input features are recursively removed from the feature set. This can be translated to DNNs whereby, a fraction μ of nodes with the lowest saliency scores c_k from layers $k \in \{0, \dots, L - 1\}$ are removed iteratively and the pruned model is fine-tuned at each step. Although the proposed strategy can remove redundant features and trainable weights, the process involved multiple iterations of elimination and fine-tuning that is time-consuming and cumbersome. Thus, we propose an efficient strategy to find a layerwise saliency score threshold, below which nodes from the layers (input and hidden) of the DNN are removed.

For the salience score vector c_l in each layer, we determined the best fit distribution by computing the log likelihood of the model fit. We did not make any assumption about the best fit distribution and used multiple distributions (viz; power-law, log-normal, exponential and stretched exponential) to find the best fit for the salience score distribution at each layer. The power-law distribution is defined by the distribution of the salience scores c_{il} for nodes $i < n_l$ in layer l [177, 267]:

$$p_{\text{power}}(c_{il}) = k c_{il}^{-a}$$

where a is the scaling parameter, $k = (a - 1) c_{\min}^{a-1}$, and $c_{\min}$ is the minimum degree that obeys the power-law. The distribution parameters were computed using maximum likelihood estimation [177]. We compute statistical significance for the salience scores of nodes in their respective layers and remove the nodes with $p - \text{value} \geq 0.95$. The complete overview of the layerwise elimination of nodes is given in Algorithm 2.

Algorithm 2 LEAN: Layerwise elimination of accessory nodes

Input: DNN with layers $\{l_i\}_{i=0}^L$
Output: Reduced DNN with layers $\{l'_i\}_{i=0}^L$

Train DNN with layers $\{l_i\}_{i=0}^L$
for each layer l_i in $\{l_i\}_{i=0}^{L-1}$ **do**
 $c_i \leftarrow C_{\Delta h_{*, l_i} \Delta y}$
 $l'_i \leftarrow \{c_{ik} \mid p - \text{value}(c_{ik}) \geq 0.95\}$

5.2.4 Correlation-based elimination of inputs (CLIP)

In Chapter 2, we showed that the brain functional connectomes are modular in nature. Different modules correspond to different sub-systems in the brain, performing a specific function. Examples of these modules are shown in Figure 5.1(a). Modules or clusters have stronger connectivity between nodes within the module and weaker connectivity with nodes outside the module. This gives rise to the intuition that the functional connectivity for some regions of interest (ROI) - especially ROIs within the same module - are correlated with each other.

With our experiments on inter-subject variability of functional modules (section 2.4.5), we showed that while the modules involved in cognition and decision-making vary across subjects,

Fig. 5.1: Visualization of the functional connectivity matrices. (a) The matrix has its rows rearranged such that nodes in the same functional clusters are placed consecutively. The labels of different functional modules are given by the left color bar. (b) A magnified view of a subset of the correlation matrix is provided. Cells colored white have a higher weight than the cells colored blue. The white patches along the left diagonal represent intra-cluster connections, while the blue patches along the right diagonal represents inter-cluster connections. Intra-cluster connections have high and inter-cluster connections have low values across subjects and are therefore positively correlated. (c) A schematic of how some of the clusters are similar across subjects.

others involved in functions related to perception and motor control are stable across subjects. This gives rise to inter-subject relationships among functional connectivity features. While the extent of the correlations is difficult to quantify, the correlation between pairs of inter-cluster connections would likely be lower than the correlation between pairs of intra-cluster connections. This is because of the sparser and variable inter-cluster connections, and the relatively denser and more stable intra-cluster connections due to similar modularizations as shown in Figure 5.1(b) and 5.1(c). For example, the dense connections within the visual module, which has a low inter-subject variability, should be highly correlated across subjects.

5.2.4.1 Finding clusters of correlated features

While there are other dimensionality reduction approaches such as Principal Component Analysis or Independent Component Analysis that reduce the problem of multicollinearity [268], such approaches require either prior knowledge of the number of components to be used, or extensive experimentation to arrive at the optimal number. Additionally, techniques like PCA project the input feature space into a latent space that is hard to interpret in the neuroanatomical space. This makes it difficult to use for decoding or deriving biologically interpretable disease

Fig. 5.2: Schematic of how the features are distributed in different clusters throughout the distribution of salience scores and how salience scores consider only features from a few clusters.

biomarkers. Alternatively, a clustering-based approach presented below provides a more principled solution. Figure 5.2 provides a schematic of the intuition behind finding such clusters that are spread across the distribution of salience scores.

We formulate this problem in terms of finding clusters/communities of correlated features, where we compute the similarity matrix, $M = \{m_{ij}\}$, where m_{ij} denotes the correlation between features i and j from the training data:

$$m_{ij} = \frac{\text{Cov}(x_i, x_j)}{\sigma(x_i)\sigma(x_j)} \quad (5.4)$$

where Cov finds the covariance between two variables and σ finds the standard deviation, and x_k denotes the vector containing values of feature k across the samples.

We perform clustering of the similarity matrix M to obtain a cluster vector $s = (s_k)$ where each element s_k corresponds to the cluster label assigned to each voxel k from a set S where $|S|$ denotes the number of clusters. This is done by the minimization of the normalized cut cost, $\text{cut-cost}(M, |S|)$, given by equation 2.4, where the the diagonal degree matrix is derived from M .

Fig. 5.3: Overview of the feature selection process. To generate M , the vector of features representing each subject is concatenated and the similarity matrix is computed. Thereafter, clustering is performed on M and a subset of nodes with the highest intra-cluster degree each cluster are retained, while the rest are discarded. The matrix on the right has been rearranged such that nodes in the same cluster are placed consecutively.

The minimization of the cut-cost is performed by multi-class spectral clustering [87]. The number of clusters is found by computing the elbow point [112] (i.e. the point of maximum curvature) of the scree plot of the eigenvalues of similarity matrix M . However, since the number of features $|x|$ are often large, the similarity matrix has a large size and computing its eigendecomposition becomes expensive. Therefore, we sparsified the matrix M by thresholding correlations with values less than 0.3 (which is a recognised threshold for low correlation values [269]). Thereafter, we used the Implicitly Restarted Arnoldi Method [270] to perform eigendecomposition on the sparse matrix to obtain the eigenvalues.

5.2.4.2 Selecting subsets of correlated features

For each of the computed clusters, a subset of features that have the highest correlation with features from the same cluster was selected. We do this by using the *intra-cluster degree*, θ_{i,s_i} , of feature i in cluster s_i given by

$$\theta_{i,s_i} = \sum_{s_j=s_i} s_{ij} \quad (5.5)$$

We select a fraction of nodes with the highest θ_{i,s_i} values from each cluster and include them as features in addition to the ones selected from one-shot elimination pruning. The whole process for removal of subsets of correlated features is summarized in Figure 5.3 and Algorithm 3.

Algorithm 3 CLIP: Correlation-based elimination of inputs

Input: Concatenated features from training data: $\{x\}$, k Output: Reduced feature set l'_0

$$M \leftarrow \frac{\text{Cov}(x_i, x_j)}{\sigma(x_i)\sigma(x_j)}$$

$$S \leftarrow \text{elbow}(M)$$

$$m \leftarrow \text{minimize cut-cost}(M, S)$$

$$l'_0 \leftarrow l'_0 \cup \text{top } k\% \text{ features from each module in } s$$

5.2.5 Combining LEAN and CLIP

We argue that the presence of correlated features leads to a sudden drop in accuracy with LEAN because while performing recursive elimination, we retain a fraction of features from clusters of correlated features at different thresholds. However, with LEAN, entire clusters of correlated features are lost at once. This problem has been studied previously for various linear and non-linear classifiers [271] by using different feature selection methods [272]. For instance, in ordinary least square regression, the existence of multicollinearity leads to a larger standard error [273] and affects the interpretation of feature salience - a feature that would have been deemed as important is no longer significant when another correlated feature is present. We hypothesize that LEAN discards too many input features at once and using the additional subsets of correlated features (identified by Algorithm 3) will improve the model accuracy. Thus, while hidden layers nodes are eliminated only using LEAN (Algorithm 2), the input features are first eliminated using LEAN and then some are added back from the subsets of correlated features via CLIP (Algorithm 3).

5.2.6 Decoding the brain functional connectome associated with brain disease

Brain decoding is defined as the identification of brain activity patterns that emerge from activations as well as interactions among specific brain regions and connections, which distinguish one brain state from another. Decoding the DNN trained on connectome features translates to identifying salient features of the DNN, which correspond to biomarkers (i.e. key brain connections and ROI) associated with the brain state. We demonstrated this in the previous chapter where we successfully adopted DeepLIFT to decode brain functional connectivity for

task states. In this chapter, we propose a combination of LEAN and CLIP to not only decode the input features but also finds leaner DNN model for efficient classification without the loss of accuracy. By using LEAN and CLIP on resting-state fMRI brain scans gathered in brain disease, we achieve state-of-art accuracies for disease classification with leaner DNN models and salient input features. The decoded input features correspond to brain connections that are associated with brain disease.

5.3 Results

5.3.1 Datasets

We downloaded resting state functional MRI scans for AD and MCI from the Alzheimer’s Disease Neuroimaging Initiative (ADNI); for ADHD from the International Neuroimaging Datasharing Initiative (INDI) [274]; for MDD from the data provided by the Creativity and Affective Neuroscience Lab in the Brain Imaging Center of Southwest University and for ASD from the Autism Brain Imaging Data Exchange (ABIDE). The details of the preprocessing pipelines are attached in Appendix D.

5.3.1.1 ABIDE

We used neuroimaging datasets from ABIDE [184]. ABIDE involves 19 sites with over 1000 fMRI scans datasets from ASD and 593 normal controls (age range: 5-64 years) containing phenotypic characterization, particularly in regard to measures of core ASD and associated symptoms. In this study, we used rs-fMRI data of 387 ASD and 436 control subjects. We removed subjects with mean framewise displacement greater than 0.2mm and used the data preprocessed using the C-PAC pipeline from the Preprocessed Connectomes Project.

5.3.1.2 ADNI

Functional and structural MRI data used from ADNI were obtained from the ADNI database (adni.loni.usc.edu). The acquisition details of the ADNI dataset can be found in section 3.3.1. The preprocessing details for the dataset can be found in the appendix B. As there are too few scans available for MCI, we combined early MCI, MCI and late MCI into one category (refer table 5.2). Subjects classified as AD fulfilled the criteria for AD laid down by National

Table 5.1: ASD subject details, FIQ Score

	Cognitively Normal (CN)	Austism Spectrum Disorder (ASD)
No. of subjects	436	387
Male/Female	354/82	341/46
Age	16.40 ± 6.97	17.86 ± 9.11
FIQ Score	111.25 ± 12.51	105.70 ± 16.76

Table 5.2: ADNI subject details, including Mini-Mental State Examination (MMSE) and Clinical Dementia Rating (CDR) scores

	Cognitively Normal (CN)	Mild Cognitive Impairment (MCI)	Alzheimer’s Disease (AD)
No. of subjects/scans	49/196	90/356	29/103
Male/Female	22/27	46/44	15/14
Age	76.6 ± 5.5	74.4 ± 6.7	75.4 ± 8.2
MMSE score	28.9 ± 1.1	27.3 ± 2.4	21.2 ± 2.6
CDR score	0.13 ± 0.3	0.68 ± 0.4	1.13 ± 0.6

Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer’s Disease and Related Disorders Association. Subjects classified as MCI had a memory complaint, objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical Memory II, absence of dementia and significant levels of impairment in other cognitive domains. The CN subjects did not suffer from depression, cognitive impairment, or dementia. Table 5.2 gives a summary of the gender and age distribution of the dataset, and also includes metrics to measure cognitive function.

5.3.1.3 ADHD

We obtained the preprocessed ADHD data from the consortium of INDI, which contains aggregated data from collaboration of 8 international imaging sites [274]. We used the data of subjects preprocessed using the Athena pipeline from the New York University Child Study Center and retained the subjects that were classified as normal and ADHD Combined and had no secondary diagnosis.

Table 5.3: ADHD subject details, Conners’ Parent Rating Scale-Revised, Long version (CPRS-LV)

	Cognitively Normal (CN)	Attention Deficit Hyper. Disorder (ADHD)
No. of subjects/scans	99/184	123/212
Male/Female	48/51	97/26
Age	12.15 ± 3.13	11.13 ± 2.69
CPRS-LV	45.2 ± 6.01	71.1 ± 8.71

Table 5.4: MDD subject details, Hamilton Depression Rating Scale (HAM-D)

	Cognitively Normal (CN)	Major Depressive Disorder (MDD)
No. of subjects	168	289
Male/Female	66/102	99/190
Age	31.01 ± 11.79	38.74 ± 13.65
HAM-D score	-	20.78 ± 5.87

5.3.1.4 MDD

We used the preprocessed data provided by the Creativity and Affective Neuroscience Lab (led by Dr. Jiang Qiu), Brain Imaging Center of Southwest University. The dataset contained rs-fMRI scans collected from 3T MRI scanners for an 8-min period for 289 diseased 168 CN subjects at Xinan (First Affiliated Hospital of Chongqing Medical School in Chongqing, China). The participants were diagnosed according to the Diagnostic and Statistical Manual of Mental Disorder-IV criteria for MDD and depression severity and symptomatology were evaluated using the Hamilton Depression Rating Scale (HAMD, 17 items) [275] and the Beck Depression Inventory (BDI) [276].

5.3.2 Features for classification

We used the Power atlas [14] to obtain functionally diverse 264 ROIs spanning the entire cerebral cortex. The details are given in section 2.4.2. The functional connectivity matrix for each scan was generated by computing the Pearson correlation coefficient (without thresholding) for each pair of ROI. Since the matrices are symmetrical, we consider only the upper triangular

connectivity matrix and flatten it into an input vector for the network. This resulted in an input vector with 34716 elements.

5.3.3 Classification for full feature set

Both the encoder and decoder were implemented in Python using the Keras, Tensorflow and DeepLIFT libraries. For datasets with more than one scan per subject, we ensured that all the subject scans were either in the training or test set. For all datasets, a batch size of 8 was used along with a learning rate of 10^{-4} .

Besides the feedforward neural networks (FFN), different convolution neural network (CNN) architectures [192, 238] and support vector machines (SVM) were implemented with different parameters. For CNNs, we gave the connectivity matrix as an input, and varied the number of filters and the number of layers. The number of filters in each layer was varied based on the number of weights in the corresponding FFN such that the number of trainable parameters were same in both the architectures. The number of trainable weights were changed from 1.7×10^4 to 3.5×10^7 . For SVMs, we varied the parameters ($C \in \{0.001, 0.01, 0.1, 1, 10\}$, $\gamma \in \{0.001, 0.01, 0.1, 1\}$) and tried different kernels such as linear, polynomial, RBF and sigmoid. For the FFN, the number of hidden layers and the number of neurons in each hidden layer were varied. The final architecture for the full feature set was obtained using grid search from accuracies obtained from stratified 5-fold cross validation. We added dropout of 0.1 to the hidden layers and imposed early stopping to prevent overfitting. The details of the different configurations for the DNNs can be found in appendix D. We observed that the highest accuracies were still given by FFN (except in case of CN vs AD, where the difference was insignificant). The average accuracies along with the standard deviation for the SVM, CNN and FFN models are reported in table 5.5 and the corresponding architectures are reported in appendix D table D1. These accuracies are obtained from repeating the experiments using 10 different seeds. For each seed, 5-fold cross validation was performed.

5.3.4 Classification with feature subsets from CLIP and LEAN

We tried 3 different elimination approaches: ‘LEAN’, ‘CLIP + LEAN (Inputs Only)’ and ‘LEAN + CLIP’. ‘LEAN’ (described in Algorithm 2) involves keeping only statistically significant features at each layer; ‘CLIP + LEAN (Inputs Only)’ includes subsets of correlated features

Table 5.5: Performances of the best models of different models on the neuroimaging datasets.

Disorder	SVM	CNN	FFN
ASD	68.0% $\pm$ 3.3%	70.6% $\pm$ 2.9%	70.9% $\pm$ 4.2%
AD	71.3% $\pm$ 6.5%	78.1% $\pm$ 5.3%	77.4% $\pm$ 4.6%
ADHD	54.8% $\pm$ 5.5%	60.5% $\pm$ 4.5%	61.1% $\pm$ 4.2%
MCI	64.6% $\pm$ 4.3%	66.5% $\pm$ 3.8%	67.5% $\pm$ 4.1%
MDD	73.3% $\pm$ 3.9%	77.1% $\pm$ 3.6%	78.3% $\pm$ 3.2%

(Algorithm 3) in addition to DeepLIFT features, but no elimination is performed for the hidden layers; and ‘CLIP + LEAN’ goes even further to prune the hidden layers, as described in Section 5.2.5. For CLIP, we generated the similarity matrix (described in Section 5.2.4.1) to derive a subset of correlated features that are combined with features obtained from LEAN. We did this for each of the seeds and folds and retained 5% of the top features for each cluster. Table 5.6 summarises the changes in classification accuracy and remaining number of trainable parameters (compared to the the original model) for different approaches on the 3 datasets.

As seen in table 5.6, we observe that relative to the full feature set, LEAN is able to reduce the model parameters drastically to around 0.5% to 2.1% of the original number of parameters. Intuitively, keeping the most important features should lead to a model with minimal drop in accuracy. However, we found that the drop in accuracy is large (1.7% in ADHD to 3.1% in MDD) and thus LEAN cannot be directly used. One can argue that a different p-value threshold can be used to allow more features to be included, but it will take several iterations to arrive at the threshold which gives the optimal model. Furthermore, such a tuned threshold is unlikely to generalize to other situations.

We reduced the drop in accuracy by identifying subsets of correlated features in ‘CLIP + LEAN (Inputs Only)’. Evidently, this approach leads to a smaller drop in accuracy as compared to the LEAN approach (and even lead to an increase in accuracy for CN/AD). Although we obtain similar classification performance when we only prune the inputs (and not the hidden layers), the number of remaining parameters (7% to 11% of the original) are much larger than LEAN. By extending the node elimination process to the hidden layers in ‘CLIP + LEAN’, we further reduced the number of parameters involved to levels quite close to the LEAN approach, with similar accuracies (and with consistently higher accuracies than LEAN).

Table 5.6: Comparison of functional connectivity based classifiers. Params represent the number of parameters left in the reduced model, relative to the model that uses the full feature set.

	All Features	LEAN		CLIP + LEAN (Inputs Only)		CLIP + LEAN	
Disorder	Accuracy	Acc	Params	Acc	Params	Acc	Params
ASD	70.9%	68.1%	0.5%	69.3%	9.5%	68.9%	1.1%
AD	77.4%	75.8%	0.7%	78.1%	7.7%	78.1%	2.0%
ADHD	61.1%	59.0%	2.1%	59.3%	9.9%	59.4%	5.7%
MCI	67.5%	64.5%	0.6%	65.3%	6.9%	66.2%	2.0%
MDD	78.3%	75.2%	1.2%	76.4%	10.9%	76.0%	2.6%

Table 5.7: The differences of test and train losses computed using cross-entropy.

Disorder	All Features	LEAN	CLIP + LEAN
ASD	0.50	0.28	0.20
AD	0.34	0.14	0.13
ADHD	0.29	0.06	0.01
MCI	0.13	0.002	0.02
MDD	0.61	0.31	0.36

Comparing the results of ‘CLIP + LEAN (Inputs Only)’ and ‘CLIP + LEAN’, it is seen that there is no significant trend in terms of change in accuracy. This shows that we can safely remove hidden nodes to obtain a leaner model. Thus, our subsequent analysis will be focused on LEAN and CLIP + LEAN.

5.3.4.1 Reduction in overfitting

Overfitting happens during model training when the test loss increases after reaching a minimum [256]. If a model overfits, it fails to generalise and remembers only the training data. To evaluate overfitting, we use the difference between the test set loss and train set loss at the epoch as evaluated by cross-entropy when the best model was found. Because of the small size of the datasets, the difference of train and test losses renders a stable measure of overfitting [277].

The model was first trained with early stopping, such that the best model was chosen by looking at the test accuracies of each of the epochs. The best model is chosen at the epoch where test accuracy is highest. Table 5.7 shows the extent of overfitting for each of the datasets with

Table 5.8: Accuracies of models produced by keeping X% most important features from SVM and logistic regression models. C + L = CLIP + LEAN.

Disorder	FFN	C + L		1%	5%	10%	100%
ASD	70.9%	68.9%	SVM	53.0%	59.0%	67.7%	68.0%
			Logreg	63.1%	67.0%	66.6%	67.9%
AD	77.4%	78.1%	SVM	66.8%	70.2%	71.6%	71.3%
			Logreg	70.9%	72.7%	72.2%	71.3%
ADHD	61.1%	59.4%	SVM	53.2%	51.6%	52.0%	53.1%
			Logreg	52.5%	51.9%	52.1%	53.2%
MCI	67.5%	66.2%	SVM	64.6%	62.9%	60.1%	58.7%
			Logreg	56.1%	57.9%	58.4%	58.7%
MDD	78.3%	76.0%	SVM	69.4%	72.5%	72.9%	73.3%
			Logreg	70.5%	72.6%	72.9%	73.3%

cross-entropy being the metric used to compute the loss. The reported scores were calculated by averaging over 5 folds. As shown in the table, there is a clear reduction of overfitting across all datasets for both LEAN and CLIP + LEAN, after these algorithms were used to generate the leaner model.

5.3.4.2 Comparison with other redundancy removal methods

We compared performances of CLIP + LEAN with those of SVM and logistic regression, including their own feature importance methods. Experiments were conducted, keeping 1%, 5% and 10% of features as this is the same range of remaining features obtained from our proposed algorithms. For SVM, only the linear kernel is able to provide importance scores and for this comparison, the best model with a linear kernel was chosen for each dataset. As seen in table 5.8, as the number of features remaining increases, model accuracy for logistic regression generally increases but such a trend is less pronounced for SVM. Ultimately, CLIP + LEAN still generally outperforms both the SVM and logistic regression models.

Additionally, since traditional models like SVM and logistic regression assume that the used features are uncorrelated, we used CLIP to retain a subset of uncorrelated features and compared the performance of the resulting models. As seen in table 5.9, the classification accuracy usually suffers when a feature subset is used. However, the accuracy is significantly lower ($p - value < 10^{-3}$) than the corresponding FFN in all the cases (except for MCI).

Table 5.9: Accuracies of baseline models before and after CLIP was applied.

Disorder	SVM		Logistic Regression	
	Before	After	Before	After
ASD	68.0% $\pm$ 3.3%	62.0% $\pm$ 2.9%	67.9% $\pm$ 3.1%	61.9% $\pm$ 3.1%
AD	71.3% $\pm$ 6.5%	67.6% $\pm$ 6.4%	71.3% $\pm$ 6.6%	70.7% $\pm$ 7.0%
ADHD	54.8% $\pm$ 5.5%	55.9% $\pm$ 4.8%	53.2% $\pm$ 5.7%	51.6% $\pm$ 5.4%
MCI	64.6% $\pm$ 4.3%	64.6% $\pm$ 4.3%	58.7% $\pm$ 5.4%	56.3% $\pm$ 5.5%
MDD	73.3% $\pm$ 3.9%	71.9% $\pm$ 5.0%	73.3% $\pm$ 4.3%	68.9% $\pm$ 3.7%

Table 5.10: Average time taken per epoch (in seconds) for each variant of the proposed algorithms. mil = million.

Disorder	Dataset Size	Initial Params	All Features	LEAN	CLIP + LEAN
ASD	823	1.74 mil	0.71	0.53	0.52
AD	299	1.74 mil	0.36	0.24	0.24
ADHD	396	1.74 mil	0.47	0.34	0.37
MCI	554	0.69 mil	0.53	0.42	0.41
MDD	457	34.78 mil	2.09	0.42	0.54

5.3.4.3 Comparison of algorithm runtime

In order to evaluate time efficiency of different algorithms, average time taken per epoch to train the neural network were computed. As shown in table 5.10, all variants of our proposed redundancy reduction algorithms lead to a reduction of time taken - the increase being greater especially if the original ‘All Features’ model used was large like in the case of MDD. Here, we also specified the time taken for the one time operation to compute the feature clusters in CLIP. On average (over datasets and seeds), pre-processing took 135 minutes where the eigenvalue computation of the similarity matrix S is the source of the bottleneck. This can be significantly reduced by lowering the number of eigenvalues to be computed - in our experiments, we defined it to be 1000. Also, importance score computation took, on average, a range of approximately 20s for AD to approximately 2 minutes for MDD (which starts off from a large model).

Comparing the proposed algorithms with each other, there are small variations in the timings but the differences are largely insignificant (except for the case of MDD) and both variants take approximately the same amount of time to run. We also concluded that the pruned models,

LEAN and CLIP + LEAN, take a significantly shorter amount of time to run as compared to the full model ($p - value < 10^{-4}$). When a relatively small model is used, both variants take approximately the same amount of time to run.

5.3.5 Salient brain features

Thus far, we have seen how salience scores help to reduce the number of parameters involved, leading to a leaner model with minimal reduction in accuracy (or at times, even leading to an increase in model accuracy). We also use the salience scores to identify important features and ROI for the different diseases. The input fed into our neural network model is a 34716-dimensional vector and each element in the vector represents a correlation score between a pair of ROI (i.e. strength of a functional connection). We analyse the importance of the functional connections by using the salience scores and the importance of ROI by computing the sum of salience scores of the functional connections incident on the ROI. Since the Power atlas [14] does not provide anatomical labels, the Crossley atlas [15] (which has anatomical labels) was used to map ROI from the Power atlas on the basis of Euclidean distance. Then, ROI with the top 10% highest scores are visualised using Nilearn.

From Figure 5.5(a), we found that the salient connections for MDD lie between the transverse temporal gyrus and lingual gyrus; the transverse temporal gyrus and thalamus; superior occipital gyrus and precuneus, and between extra-nuclear and inferior temporal gyrus. For AD, we found that the connections between the superior temporal and subcallosal; superior temporal and extra nuclear, and between the uncus and inferior temporal regions were salient (Figure 5.4(a)). For MCI, however, we found that only the connections between the uncus, subcallosal and uncus inferior temporal were salient (Figure 5.6(a)). For ADHD, the salient connections were found between the paracentral and superior temporal; paracentral and inferior frontal; inferior temporal and superior occipital; supramarginal and sub-gyral; and supramarginal and extra nuclear regions (Figure 5.8(a)). For ASD, salient connections were found between regions in the frontal inferior orbital to the putamen, frontal superior orbital, and frontal inferior operculum; the putamen and temporal pole superior; and cerebellum and the vermis (Figure 5.7(a)).

For all the diseases, ROI in the inferior temporal gyrus and paracentral lobule were common among the salient regions. The other most distinctive ROI for MDD were located in the

subcallosal gyrus, supplementary motor area, the medial frontal gyrus, the parahippocampal gyrus and the posterior cingulate (Figure 5.5(b)); for ADHD in the postcentral gyrus, inferior frontal gyrus, transverse temporal gyrus, superior occipital gyrus, the insula and the medial frontal gyrus (Figure 5.8(b)). For both AD and MCI, the salient regions were found to be in the cerebellum, temporal pole middle, frontal superior medial, parahippocampal and fusiform gyrus (Figure 5.4(b) and figure 5.6(b), respectively). For ASD, salient regions were found in the inferior temporal lobe, the rectus, Heschl's gyrus, cerebellum, and the paracentral lobule (figure 5.7(b)).

5.4 Discussion

5.4.1 DNN models for full feature set

As seen in table 5.5, the feedforward DNN using the entire feature set is generally able to achieve better classification accuracies than CNNs and SVMs. The rather high standard deviation is attributable to the size of the datasets [278], which typically do not contain more than a few hundred scans. The low accuracy for ADHD is attributable to the mismatch between the age of the subjects and the age of the subjects used to derive the Power atlas [14]. Also, more subtle connectivity differences with respect to CN subjects makes it harder to classify MCI as compared to AD.

We found that CNNs and DNNs perform consistently better than SVMs. However, the difference in performance between CNNs and feedforward DNNs are subtle: except for CN-AD, feedforward DNNs did better than CNNs but the difference is rarely more than 1%. From these results, CNNs - even the customised ones - do not seem to be getting any additional significant information for functional connectivity. Although CNNs have desirable features such as parameter sharing, they only capture information within their local receptive field (usually a small square-shaped subset, or a row or column along a matrix). Such information is not sufficient to capture the functional connectivity present in brains as these receptive fields do not consider global connectivity patterns. On the other hand, the fact that CNNs have fewer parameters than feedforward DNNs shows the possibility of redundancies in feedforward DNNs which makes it important to eliminate accessory nodes from feedforward DNNs. With LEAN, we are able to get a leaner model while methodically removing less salient features and nodes (instead of being constricted by the spatial limitations of local receptive fields).

Fig. 5.4: The figure shows the salient brain functional regions and connections while classifying normal and Alzheimer's Disease patients. For the regions we show the axial, sagittal and coronal views of top 10% salient ROI differentiating normal and Alzheimer's Disease patients. Mid: Middle, Inf: Inferior, Sup: Superior, G: Gyrus

(a) Salient functional connectivity

(b) Salient brain regions

Fig. 5.5: The figure shows the salient brain functional regions and connections while classifying normal and major depressive disorder patients. For the regions we show the axial, sagittal and coronal views of top 10% salient ROI differentiating normal and major depressive disorder patients. Mid: Middle, Inf: Inferior, Sup: Superior, G: Gyrus

(a) Salient functional connectivity

(b) Salient brain regions

Fig. 5.6: The figure shows the salient brain functional regions and connections while classifying normal and mild cognitive impairment patients. For the regions we show the axial, sagittal and coronal views of top 10% salient ROI differentiating normal and mild cognitive impairment patients. Mid: Middle, Inf: Inferior, Sup: Superior, G: Gyrus

(a) Salient functional connectivity

(b) Salient brain regions

Fig. 5.7: The figure shows the salient brain functional regions and connections while classifying normal and autism spectrum disorder patients. For the regions we show the axial, sagittal and coronal views of top 10% salient ROI differentiating normal and autism spectrum disorder patients. Mid: Middle, Inf: Inferior, Sup: Superior, G: Gyrus

Fig. 5.8: The figure shows the salient brain functional regions and connections while classifying normal and attention deficit hyperactivity disorder patients. For the regions we show the axial, sagittal and coronal views of top 10% salient ROI differentiating normal and attention deficit hyperactivity disorder patients. Mid: Middle, Inf: Inferior, Sup: Superior, G: Gyrus

5.4.2 LEAN + CLIP: high accuracy with few parameters

From the results in table 5.6, we observe that LEAN has the largest accuracy drop but the smallest number of remaining parameters; but, ‘CLIP + LEAN (Inputs Only)’ results in the small drop in accuracy but retains the largest number of parameters. However, ‘CLIP + LEAN’ gives the best of both worlds - the number of weights retained is close to levels from LEAN, but the accuracy drop is similar to that of ‘CLIP + LEAN (Inputs Only)’. An implication of our results is the presence of correlated features in functional connectomes, which can be exploited to reduce input feature set in functional/structural connectomes. In our case, we observed that LEAN led to drastic reduction in the input feature set, thereby removing sets of correlated features entirely. Our results show that adding a subset of the correlated features in ‘CLIP + LEAN’ (and also in ‘CLIP + LEAN (Inputs Only)’) led to an improvement in the performance of the classifier.

Crucially, our results show that there is a large amount of redundancy in neural network models. Our experiments with quantification of overfitting and time taken per epoch for the proposed model show that the proposed model not only leads to a reduction in overfitting but is also faster. The latter was expected since there is a significant reduction in the number of trainable network parameters. Feedforward networks have been used by researchers in the past [265, 266, 279] to perform classification on functional connectivity data. However, these works often use the full set of features and overfitting is a key limitation in such approaches. In our work, we have proposed a way to reduce the effects of overfitting and showed how even as it leads to a drastic reduction of parameters, the drop in accuracy is minimal. Thus, just a simple feedforward DNN is sufficient to capture crucial patterns in the data to classify healthy and diseased brains. Adding more nodes in the hidden layers does not improve model accuracy much (and in some cases such as CN-AD classification, doing so even leads to poorer generalization).

Although we have only tested our approaches on neuroimaging datasets, they are also widely applicable to other datasets. LEAN (Algorithm 2) is applicable in settings where the number of features greatly outnumber the size of the dataset, or in cases where the dataset is small and will benefit from using fewer features so as to increase the generalizability of the model. CLIP (Algorithm 3) is applicable to datasets where there are features that have strong correlations with each other.

5.4.3 Identified salient features are clinically relevant

For our algorithm to perform well, it is important that the features classified as salient by our decoder have clinical relevance to the respective condition. To verify this, we compare the results of our decoder with results from previous studies finding disease biomarkers.

For AD/MCI, previous studies have consistently pointed out to changes in the hippocampal and medial temporal lobe [280–282]. Accordingly, we found alterations in the inferior temporal, parahippocampal [283], fusiform [280], and paracentral lobule [282]. We observed that the most salient connections were from the uncus for both AD and MCI, an anterior extremity of the parahippocampal gyrus known for disruption during both AD [284] and MCI [285].

Multiple functional networks such as the fronto-parietal task-control network, involved in attention and emotion regulation; the default mode network, for internal medication; the dorsal attention network, for directing external attention; and the salience network, which helps in emotion processing or monitoring salient events have been implicated in MDD [286, 287]. Deficits in cognitive control can be traced to the anomalies in the fronto-parietal task-control network, whereas too much internal rumination (and lesser engagement with the external world) may be reflected in the aberrant connectivity of the default mode network with other goal-directed networks. The amygdala [288, 289] (part of the subcallosal region), regions in the inferior temporal [290], the medial frontal lobe [290], posterior cingulate cortex [291] and supplementary motor area [292] have been implicated in MDD. In [293], disruption between the connections between regions in the thalamus and the transverse temporal gyrus were reported, which are consistent with our results.

Multiple studies have reported anomalies in the default-mode and dorsal attention networks in children and adults suffering with ADHD suggesting altered functional connectivity with attention and cognitive processing [294–297]. The postcentral gyrus and paracentral lobule areas involved in motor functioning [294, 295], medial and inferior frontal gyrus [294, 298], superior occipital gyrus [297], and the insula [296] have been found to be different for children/young adolescents having ADHD, which are consistent with our results.

Likewise, the frontal orbital regions and Heschl’s gyrus are involved in sensory integration, speech processing and decision-making [299, 300], the putamen responsible for focusing attention [301], the cerebellar (and vermis) involved in motor regulation [302] have been previously identified as biomarkers for ASD.

5.5 Chapter Summary

In this chapter, we have proposed two algorithms - LEAN and CLIP - to reduce overfitting in DNNs making them more generalizable than models that use the entire feature set. Our single shot approach leads us to a neural network architecture which is more efficient than previous methods that relied on recursive removal of features and nodes. Our experiments show that using both 'LEAN + CLIP' took into account both redundancy and correlation in input features and gave the best balance between drop in accuracy and reduction in trainable weights. We successfully applied the proposed algorithms on 4 datasets (and 5 different neurological disorders), showing its application in brain decoding and biomarker identification. The proposed approach has applications to both structural and functional neuroimaging datasets and to investigate brain disease.

Chapter 6

Conclusion and Future Work

Computer science models are used to understand functional Magnetic Resonance Imaging (fMRI) data because the large number and complex features cannot be deciphered manually. The brain's functional data can be modeled as a network of nodes and connections formed from the brain regions and correlation between temporal activations of the regions. This network, known as the functional connectome, has been studied during different brain states. In this thesis, we used the brain functional connectome to extract meaningful knowledge from fMRI scans to understand the complex neural underpinnings of brain states during rest, task and disease.

Since modules and hubs in the brain functional connectome shape brain function during different brain states, it is important to formulate methods to detect them non-invasively. In this thesis, we highlighted several issues with the previous approaches used to detect modules and hubs from the functional connectome.

We formulated the problem of modularization as deriving a representative modularization for brain functional networks using functional connectivity between different brain regions. We proposed the ICSC algorithm that takes into consideration multiple dense and weighted functional connectivity matrices and derives a group modularization that best represents the individual modularizations. We derived a group-level modularization for multiple subjects and subject-level modularizations from multiple scans of the same subject. Using the proposed algorithm we studied the inter-subject and inter-scan variability in brain's functional architecture and found marked variability in brain modular distribution across subjects and a lower inter-scan variability. Besides the data from the Human Connectome Project (HCP), we used a real validation dataset and generated synthetic networks with varying amounts of inter-subject

variability and noise to mimic the properties of the brain functional connectome. We also show an application of the derived modularizations for detecting neurological diseases that are known for disruption in modular architecture.

We used the ICSC modularization algorithm, proposed in this thesis, to detect crucial brain hubs in the functional connectome. Previous studies in the area did not consider the effect of weak connections and modular structure for detection of functional hubs. We proposed ambivert degree to detect brain hubs and show that the proposed measure detects more crucial hubs than previously used schemes. We use the ambivert degree to find widespread perturbation of hubs in AD and ASD patients. We used this property to develop effective classifiers for these diseases and showed that these classifiers perform significantly better than functional connectivity based classifiers, with negligible number of trainable weights in comparison. Thus, we show that the affected hubs detected by our methods could be used as biomarkers for neurological diseases where hub disruption is a hallmark.

An overarching goal of cognitive neuroscience is to study how task performance shapes the brain functional architecture. While the first two chapters of this thesis focused on studying the functional connectome at rest, we also studied modulations related to task performance in the functional connectome. Most of the previous studies group task related changes and ignored the inter-subject differences in task related changes. In this thesis, we segregated task related changes into task-general and task-specific modulations while taking each individual into consideration. We proposed a novel approach where we encode the functional connectome by training neural network classifier to classify rest and task functional connectome and decode the trained classifier to uncover distinguishing features between the two states. Using rest and task data for 7 tasks from the HCP, we uncovered both the task general and task specific architectures. We validated the results of our decoder by showing that a small subset of features deemed important by the decoder gives us a high accuracy. Our method has a strong potential of replication in brain state classification and identification of distinctive features between healthy and diseased brain states.

Uncovering the anomalies in the brain's functional connectome caused by neurological ailments can not only help in understanding the evolution of the disease but also in improving the computer aided diagnosis. In this thesis, we propose a solution that uncovers disease related biomarkers and uses them to obtain more generalizable disease classifiers. We proposed two

algorithms - LEAN and CLIP - to reduce overfitting in deep neural networks making them more generalizable than models that use the entire feature set. Our approach leads us to an optimal neural network architecture in a single shot that is more efficient than recursive feature removal. Furthermore, via CLIP, the approach is customised to reduce the effects of correlated features that are present in neuroimaging datasets. Our experiments show that using both ‘LEAN + CLIP’ took into account both redundancy and correlation in input features and gave the best balance between drop in accuracy and reduction in trainable weights. We successfully applied the proposed algorithms on five different neurological disorders, showing its application in brain decoding and biomarker identification. The proposed approach has applications on both structural and functional neuroimaging datasets.

In summary, we have developed novel methods to analyze the brain’s functional connectome during rest, task and disease in this thesis. The proposed methods have wide applications in not just deriving the features of the functional connectome during rest, but also studying the modulations in the features during varied brain states. The methods that have been proposed for computer aided detection of diseases can also be extended to predict treatment outcomes for patients, or the cognitive performance scores for different patients.

6.1 Future Work

6.1.1 Hierarchical modularization

Multiple previous studies [80, 303, 304] have elucidated the hierarchy in brain’s functional modules i.e. these modules have submodules within the modules and sub-submodules within submodules and so on. Although the recent work by Akiki et al. [303] considers hierarchical individual-level information to derive consensus group-level modularizations [305], it does not improve the similarity between the two. An extension to our work on module detection, would be to derive hierarchical modular structure using scans from multiple subject (or multiple scans from a single subject) and improve the similarity between the two by the ICSC framework derived in this work. Besides using hierarchical clustering methods for individual matrices, it would be important to derive consensus matrices which represent hierarchical information. This can be achieved with replacing the binary distances in the consensus matrix with dendrogram distances [306]. The similarity measure used with ICSC to derive consensus-cost will also

have to be replaced by a measure that considers hierarchical clusters (possible choice is the Fowlkes–Mallows Index [307]). This will help characterize the brain’s functional architecture in more detail because we will also get information on the relationship between different modules.

6.1.2 Studying modular disruption during diseases

We have shown an application of group-level brain modularization for diagnosing diseases that are characterized by modular disruption. This work can be extended to a meta-study containing (1) an analysis of modular disruption for different diseases; and (2) developing methods to use features derived from brain modules for disease diagnosis. In the present work, we consider only the group-level modular structure for disease diagnosis, whereas there exists heterogeneity in functional architecture [84, 85, 101, 153, 227] and diseases [63, 308, 309] across subjects. Therefore, it is imperative to consider subject-level modularizations while diagnosing diseases. This can be achieved by the ICSC algorithm, but there are several challenges on how to obtain subject-level features that are consistent across subjects. The number or the size of modules across subjects vary, whereby there is inconsistency in the derived features across subjects. Additionally, assigning labels to modules across subjects is not a trivial task, because some modules are divided into sub-modules in some subjects while others are combined to form bigger modules in some subjects. We plan to tackle it by using a customized subject-level module matching scheme that considers the variations in modularizations across subjects and derives features accordingly.

6.1.3 Mixing topological features for disease classification

In this thesis, we have shown that in addition to functional connectivity, modules and hubs can be used for disease diagnosis since they are known to be disrupted in multiple neurological disorders. These features have been shown to give a similar or better performance for disease diagnosis than the traditionally used functional connectivity features. While multiple previous studies have either used network topological [57, 310–312] or functional connectivity (for review refer to [61]) for disease classification, they have not combined the two. The features derived from functional connectome such as its hubs, modules and other topology measures like network motif counts, efficiency, clustering coefficient, et cetera may contain additional information that is not captured by functional connectivity. Therefore, a direct extension of

the present work would involve including network topology features along with functional connectivity features to improve disease diagnosis. In order to make the number of features contributed by the network topology comparable to the functional connectivity features, we can reduce the number of functional connectivity features by the LEAN and CLIP approaches. So far, we have only studied the classification between healthy controls and diseased patients. This can be extended by classifying patients sufferings from different disorders from each other, for example, patients suffering from AD, MCI and PD (neurodegenerative disorders) can be distinguished from each other.

6.1.4 Sub-task detection from task fMRI data

In the present thesis, the focus has been on studying the brain's functional architecture using the features derived from the brain functional connectome. From our chapter on task related brain modulations, we were able to infer that a small set of functional connectivity features were able to classify rest and task states. This can be extended to mean that the BOLD time series signal from only a subset of brain regions are sufficient for classification of different rest and task states. This insight can be used to design better recursive neural network encoders for real-time brain state classification [230, 313], where instead of considering the BOLD signal from all brain regions, only a subset of brain regions can be used.

Appendices

Appendix A

Appendix for Chapter 2

Validation on an independent dataset

We used the preprocessed data provided by the Creativity and Affective Neuroscience Lab (led by Dr. Jiang Qiu), Brain Imaging Center of Southwest University. The dataset contained rs-fMRI scans collected from 3T MRI scanners for an 8-min period for 116 cognitively normal subjects from Xinan (First Affiliated Hospital of Chongqing Medical School in Chongqing, China). We ran 100 independent runs of the ICSC algorithm on the validation dataset until convergence. The number of iterations the ICSC procedure took to converge ranged from 2 to 32 (average 9.3 ± 5.5). Out of the 100 independent runs, we selected the run with the maximum consensus-cost and study the individual-level and group-level modularizations for this run. For the validation dataset also the ICSC algorithm detected 20 group-level modules with the sizes of modules ranging from 7 to 19. Although there are differences in the participant demography, scanning protocol and data preprocessing steps we obtained a similar modularization as with the HCP dataset (Fig. A1). We were able to detect sub-modules of the default mode network (Figures A1(b), A1(n) and A1(u)), the motor network (Figures A1(c), A1(i) and A1(j)), the visual network (Figures A1(d), A1(g) and A1(t)), similar to the HCP dataset. We were also able to separate the auditory module from the motor module (Figure A1(l)).

Choosing L_{max} for ICSC

The ICSC algorithm expects the user to know the range of expected number of modules given by (L_{min}, L_{max}) . While the lower limit L_{min} can be fixed based on prior scientific knowledge, the choice of L_{max} has a trade-off in terms of time and optimization of the quality function

Fig. A1: The twenty group-level modules detected by the ICSC algorithm on the resting-state fMRI scans from the validation dataset. (a) the group consensus matrix with reordered node indices to bring nodes in the same module together. (b) to (u) plotted using BrainNetViewer [133] show the ROIs belonging to each module ordered in the descending order of module size. The modules are given names based on functional networks identified by earlier studies or on the anatomical location of constituent regions.

Fig. A2: Figures (a) to (e) show the distribution of number of individual-level modules $\{L^k\}_{k=1}^K$ for different values of $L_{\max}$. Figure (f) shows the value of the consensus-cost objective function for different values of $L_{\max}$.

associated with it. While smaller than optimal values of $L_{\max}$ will limit the options available for individual-level modularizations generating sub-optimal group-level partitions, larger than optimal values of $L_{\max}$ will lead to more $s^k(l)$ being evaluate and take a longer time to converge.

We demonstrate this by running 10 independent runs with varying $L_{\max} = \{20, 25, 30, 35, 40\}$, and observing the consensus-cost and distribution of number of individual-level modules $\{L^k\}_{k=1}^K$. For $L_{\max} = \{20, 25\}$, we found that there is a sharp drop in the $\{L^k\}$ histogram pointing to several individual-level modularizations taking the sub-optimal values for L^k . For $L_{\max} = 30, 35, 40$, we observed a smooth distribution of $\{L^k\}$. The consensus-cost was found to be optimal for $L_{\max} = 30$, with the average number of iterations ranging from 7.0 ± 2.9 for $L_{\max} = 20$ to 32.4 ± 20 for $L_{\max} = 40$.

Fig. A3: The intra-subject purity corresponding to different anatomical locations of brain ROIs. The sizes of the nodes correspond to the values of purity and the colors denote the functional modules.

Intra-subject variability of functional modules

Studies have shown variability in regional functions across multiple scans of the same individual, which is attributable to noise and different states of mind during fMRI acquisition. We measured the variability of the modular memberships of brain regions across multiple scans of the same subject by using the nodal purity. We found a narrow spread of values for nodal purity (average 0.501 ± 0.043) across the subjects. The intra-subject nodal purities corresponding to different brain anatomical locations are shown in figure A3, where the sizes of the nodes correspond to the values of purity. We found that all the nodes had similar purity in contrast to the inter-subject nodal purity where nodes had pronounced differences in purity.

Group-level modules from averaged and thresholded connectivity matrices

Previous studies in the area have used a thresholded and averaged group matrix to detect group-level modules. Thresholding removes weak edges and makes the graph sparser, which reduces the computational complexity but loses vital information related to the functional modules. We studied the group-level modularization on the HCP resting state data by using module detection approaches on an averaged and thresholded group-level matrix. For thresholding, we used percolation analysis [314], which iteratively removes weak edges till the connectedness in the network is preserved. Percolation analysis has been applied for thresholding networks for modularization in studies involving humans before [42]. The percolation threshold is where

nodes start getting disconnected from the largest component. For our data, this was around 90%. For the sake of completeness, we computed the group-level modules at two other thresholds (10% and 40%) besides the percolation threshold (Fig. A4a).

For the Louvain algorithm, we also varied the resolution parameter, γ between 0.3 and 1.3 and chose the modularization with the highest modularity. For low thresholds, we observed isolated nodes in the network. Across different thresholds, we observe that Infomap and Louvain algorithms give a few number of modules, while Asymptotical Surprise gives two large modules (module 1 containing >100 nodes and module 2 containing > 50 nodes) and multiple smaller modules composed of 7 nodes or less.

Fig. A4: The group-level modular structure obtained with different module detection algorithms at different thresholds of connectivity. The group matrix was obtained from an average of thresholded subject connectivity matrices, where thresholding was performed by percolation analysis. Isolated nodes that result from thresholding are not shown.

Fig. A5: The 7 group-level modules detected by the Louvain algorithm on the resting-state fMRI scans from the HCP dataset.

Appendix B

Appendix for Chapter 3

Functional MRI Preprocessing Pipeline for ADNI dataset

For each of the fMRI scans for a subject, the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated and the BOLD reference was co-registered to the T1w reference using `bbregister` (FreeSurfer). Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using `mcflirt` [315]. BOLD runs were slice-time corrected using `3dTshift` from AFNI [316]. The BOLD time-series were resampled onto their original, native space by applying a single, composite transform to correct for head-motion and susceptibility distortions. The BOLD time-series were resampled to MNI152NLin2009cAsym standard space, generating a preprocessed BOLD run in MNI152NLin2009cAsym space. Principal components are estimated after high-pass filtering the preprocessed BOLD time-series. The BOLD time-series, were resampled to the `fsaverage5` surface. Gridded (volumetric) resamplings were performed using `antsApplyTransforms` (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels [317]. Non-gridded (surface) resamplings were performed using `mri_vol2surf` (FreeSurfer).

Change in path length for different network hubs

For the nodes in each subject's network, we computed the ambivert degree, betweenness centrality, intra-module degree z-score, participation coefficient, gateway betweenness coefficient,

gateway ambivert coefficient and gateway degree coefficient to detect different types of hubs. To compute network hubs, we give equal weights to the node's modular hub scores and the respective connector hub score. Accordingly, we removed the nodes with top 20% scores (in steps of 2.5%) and inspect the change in the network path length l on their removal. We performed comparison for change in path length for removal of different types of connector hubs for a given modular hub for varied percentiles (B1). For all the modular hubs, we found that the different gateway coefficients (betweenness, ambivert, degree) have a very similar increase in path length, which was in all the cases either similar to or lower than increase in path length with participation coefficient.

We plot the difference in the anatomical locations of ambivert degree and participation coefficient hubs, and intra-modular degree and participation coefficient hubs in Figure B2.

Disruption of brain hubs in Alzheimer's Disease and Autism Spectrum Disorder

We computed the network hubs given by ambivert degree and participation coefficient hubs for the diseased (ASD and AD) subjects and their healthy counterparts. We noticed widespread disruption in the hub locations in case of both AD and ASD as compared to the normal controls at different thresholds (refer figures B3 and B4).

Acknowledgment

The primary aim for launching the Alzheimer's Disease Neuroimaging Initiative (in 2003) was to test whether medical imaging technologies like MRI and PET, biological markers, and clinical and neuropsychological assessments can be together used to quantify the progression of MCI and early AD. It is run as a public-private partnership, with Michael W. Weiner, MD as the Principal Investigator .

The National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company

(a) Intra-modular degree

(b) Ambivert degree

(c) Betweenness

Fig. B1: The increase in the path length after artificial lesioning of different percentiles of network hubs determined by (a) intra-modular degree modular hubs, (b) ambivert modular hubs and (c) betweenness centrality modular hubs along with different connector hub measures.

Fig. B2: The figure depicts the different nodes that are identified as hubs using ambivert degree and participation coefficient, and using intra-modular degree and participation coefficient. The regions in red were identified using intra-modular degree and participation coefficient hubs but not with ambivert degree and participation coefficient hubs, and vice versa for the regions in green.

Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics fund ADNI. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Fig. B3: The figure depicts the difference between nodes that are identified as hubs using ambivert degree and participation coefficient in cognitively normal (CN) and AD subjects for different percentiles. The regions in red were identified as hubs in the AD subjects but not in CN subjects, whereas the regions in green were identified as hubs in CN subjects but not in the AD subjects.

Fig. B4: The figure depicts the difference between nodes that are identified as hubs using ambivert degree and participation coefficient in cognitively normal (CN) and ASD subjects for different percentiles. The regions in red were identified as hubs in the ASD subjects but not in CN subjects, whereas the regions in green were identified as hubs in CN subjects but not in the ASD subjects.

Appendix C

Appendix for Chapter 4

Distribution of task-general salience scores

We trained a model to distinguish task and rest scans, and decoded the trained model to compute the salience for the input connectivity features. Majority of features had positive salience (number: 31381, range: 0 to 1891.6, average: 81.96 ± 124.2) in comparison with features with negative salience (number: 3335, range: 0 to -259.60, average: -29.62 ± 31.42) as can be seen from Fig. C1a. The feature salience score matrix $S_{rest/tasks}$ along with the corresponding functional modular labels is shown in Fig. C1b.

Finding optimal hyperparameters for the models

For the FFN, we varied the number of hidden layers and the number of neurons in each hidden layer. For the CNN we tried different architectures (with different filters), and varied the number of filters in each layer. The number of filters (in case of CNN) and the number of neurons (in case of FFN) were fixed such that the number of trainable parameters were approximately the same. For FFN we varied the number of hidden layers from 1 to 3, and the number of hidden layer neurons from 5 to 100. For SVM models, we varied the kernel type, the C and the parameter γ .

For CNN, we defined 3 types of convolution layers:

1. A ‘closeness layer’ containing convolution filters of size 2×2 .
2. A layer containing filters of shape $1 \times |ROI|$.

Fig. C1: The feature salience scores for rest vs all tasks. (a) shows the distribution of salience scores for functional connections; and (b) shows the salience matrix $S_{rest/tasks}$ highlighting the task-general architecture of the brain. The functional ROIs are ordered according to their functional modules, with the colors on the y-axis delineating different modules.

3. A layer containing filters of shape $|ROI| \times 1$.

We tried 2 different architectures, where we tried all the layers listed above (which we termed as funcNetCNN), and removed the closeness layer (which we termed as funcNetCNN_C). This was based on the previously proposed architectures by [192, 238]. The number of filters in the layers were varied such that they had approximately the same number of parameters as the other CNN models and the corresponding FFN models. The convolution layers were followed by a fully connected layer of 32 neurons and a softmax layer for classification.

Performance of classifiers with data subsets

Since we found that all the classifiers achieved classification accuracy above 90% for the classification tasks, it became important to understand the performance of classifiers for subsets of training data. We also performed additional experiments to determine the time taken for error convergence for the different classifiers (refer to figures C2 and C3).

Fig. C2: The accuracy and time taken for convergence of different subsets of samples for rest vs emotion task ((a) and (b)); rest vs gambling task ((c) and (d)); and rest vs relational task ((e) and (f)).

Table C1: Best architectures of SVM Models on the neuroimaging datasets

Dataset	SVM Model (Kernel, C)	CNN Model	FFN Model (hidden layer config)
rest vs all tasks	Linear , 0.001	funcNetCNN_C100	100-32
rest vs language	Linear , 1.0	funcNetCNN100	100-32
rest vs social	Linear , 0.001	funcNetCNN100	100-32
rest vs gambling	Linear , 0.1	funcNetCNN_C20	10-8-4
rest vs relational	Linear , 1.0	funcNetCNN_C50	10-8-4
rest vs working memory	Linear , 0.01	funcNetCNN_C20	10-8-4
rest vs motor	Linear , 0.001	funcNetCNN20	20-10-8
rest vs emotion	Linear , 0.1	funcNetCNN10	5-4-4

(a) Rest vs Social (accuracy)

(b) Rest vs Social (time)

(c) Rest vs Working Memory (accuracy)

(d) Rest vs Working Memory (time)

Fig. C3: The accuracy and time taken for convergence of different subsets of samples for rest vs social task ((a) and (b)); and rest vs working memory task ((c) and (d)).

Decoding task-specific features of motor and language tasks

From the multiple models trained to classify the motor and language task scans from the resting state scans, the feedforward and the convolutional neural networks had a similar performance. The SVM had a lower classification performance. We used the trained feedforward network classifier to compute the feature salience scores. From the task general architecture and the computed salience scores for the respective rest vs task classification, we obtained the modular salience maps for the motor and language tasks. The modular salience maps for the motor and language tasks are shown in Fig. C4.

Decoding task-specific features of the remaining tasks

Table C2 the accuracies of models classifying the resting state scans from the individual task with the complete set of features and different subsets of input features selected by the decoder. Further we computed the statistically significant salient modular interactions between different brain functional modules for the tasks under consideration. These can be discerned from the tables C3 and C4.

Emotion

This is an adaptation of the task proposed by Hariri et al. [318] where participants were tasked to match two simultaneously presented faces or shapes on two screens: top and bottom. The faces show either an angry or fearful expression. A single run presented trials in blocks of 6 where each trial of a block was the same face or shape. Trial blocks were preceded with a 3000ms cue of either ‘task’ or ‘shape’. Then stimulus was presented for 2000ms followed by an interaction time of 1000ms (to match faces). 8 seconds of fixation were used to conclude the run as a control measure. Each of the 2 runs consisted of 3 face and 3 shape block trials.

Gambling

This task is an adaptation of one proposed by Delgado et al [319]. Participants were tasked to guess the number of a mystery card with the incentive of gaining money on a correct guess or the loss of money for the opposite. Participants were informed of potential card numbers

(a) Negative motor task-specific

(b) Positive motor task-specific

(c) Negative language task-specific

(d) Positive language task-specific

Fig. C4: (a), (c) and (b), (d) depict the modular salience matrices for negative and positive task-specific connections related to the motor and language tasks, respectively.

Table C2: Results of recursive feature elimination for the Emotion, Gambling, Relational, Social and Working Memory tasks

Task	% features	FFN		SVM	
		Accuracy	MAE	Accuracy	MAE
Emotion	100%	99.52	0.0051	97.26	0.0273
	10%	99.40	0.0065	97.99	0.0201
	1%	99.32	0.0074	98.39	0.0161
	0.1%	99.54	0.0061	97.74	0.0225
Gambling	100%	99.74	0.0035	97.70	0.0229
	10%	99.62	0.0041	98.11	0.0189
	1%	99.60	0.0045	98.55	0.0144
	0.1%	99.93	0.0016	97.66	0.0233
Relational	100%	99.55	0.0049	97.50	0.0249
	10%	99.33	0.0071	97.91	0.0209
	1%	99.33	0.0074	98.55	0.01449
	0.1%	99.85	0.0023	98.11	0.0189
Social	100%	99.72	0.0033	97.74	0.0225
	10%	99.79	0.0051	98.23	0.0177
	1%	99.88	0.0018	98.15	0.0185
	0.1%	99.98	0.0005	96.94	0.0306
Working Memory	100%	99.54	0.0049	97.62	0.0237
	10%	99.31	0.0073	98.19	0.0181
	1%	99.53	0.0052	98.47	0.0153
	0.1%	99.87	0.0022	97.42	0.0258

ranging from 1 to 9 and asked to indicate if the number was more or less than 5 through a response box. A program would feedback the true number on the mystery card along with 3 possible responses: 1) \$1 for reward trial; 2) -\$0.50 for loss trials; or 3) double gray arrow for neutral trial. The trial was structured as follows. The ‘?’ mystery cue is presented for 1500ms where a quick response would render a fixation cross for the remaining time until 1500ms was fulfilled. Feedback showing the result of the gamble is presented for 1000ms followed by 1000ms interval between tasks. Trials are presented in blocks of 8 that are either mostly reward trials or mostly loss trials. A single run has 2 blocks of mostly reward and 2 mostly loss while 4 fixation blocks were interleaved.

Relational

This task was an adaptation of one proposed by Smith et al. [320]. Participants were presented with 2 conditional tasks: relational condition and control matching condition. In the control matching condition, participants were presented with 2 pairs of objects, 2 objects on a top of a screen, 1 object on the bottom; with the dimension (shape or texture) presented in the middle of the screen. Participants were asked to decide on the dimension (shape or texture) that separated the 1st pair of objects then decide if the 2nd pair of objects differ on that same dimension. They will respond 'yes' or 'no' using a button for each response. In the relational condition, one pair of objects were presented at the top of the screen, another pair at the bottom. Participants were asked to decide if the top pair differed in either shape or texture, then decided if the bottom pair also differed in the same dimension. In each run of this task, 3 relational blocks, 3 matching blocks, and 3 16-second fixation blocks were presented.

Social

This task uses videos developed by Castelli et al. [321]. Participants were presented with short video clips of shaped objects that interacted in some way or randomly. They were then asked to decide if the movement of the objects were related with another object in any way. They were asked to give their response in the form of three choices, 'Yes', 'Not Sure', or 'No'. Each run of this task spans five video blocks, (2 mental and 3 random or 3 mental and 2 random) and five 15-second fixation blocks.

Working Memory

The method for this task was adapted from Drobyshevsky et al. [322] Participants were asked to complete a two-back working memory tasks and a zero-back working memory task which comprised of pictures related to places, tools, faces or body parts. Each run presents 8 task blocks and 4 15-second fixation blocks. 4 blocks use a 2-back working memory task and 4 blocks use a 0-back working memory task. At each block, 10 trials are presented where stimulus is presented for 2 seconds, followed by 500ms interval between tasks.

Table C3: Significant task specific modular interactions corresponding to the Emotion, Gambling, Relational, and Social tasks.

Task	Module	Module	p -value
Emotion	Motor Mouth	Motor Mouth	$< 10^{-3}$
	Memory retrieval	Motor Mouth	< 0.05
	Salience	FP Task Control	$< -10^{-3}$
	CO Task Control	Default mode	$< -10^{-3}$
	Cerebellar	FP Task Control	$< -10^{-3}$
	Cerebellar	Dorsal attention	$< -10^{-3}$
	Cerebellar	Cerebellar	$< -10^{-3}$
	CO Task Control	FP Task Control	$< -10^{-2}$
	Salience	Default mode	$< -10^{-2}$
Gambling	Motor Hand	Visual	$< 10^{-3}$
	Motor Mouth	Motor Mouth	$< 10^{-3}$
	Memory retrieval	Dorsal attention	$< 10^{-3}$
	Memory retrieval	Motor Mouth	$< 10^{-3}$
	FP Task Control	Visual	$< 10^{-2}$
	Salience	Visual	$< 10^{-2}$
	Dorsal attention	Default mode	< 0.02
	CO Task Control	FP Task Control	$< -10^{-3}$
	Cerebellar	Cerebellar	$< -10^{-3}$
	Motor Mouth	Motor Hand	$< -10^{-3}$
	Salience	FP Task Control	$< -10^{-2}$
	Salience	Default mode	< -0.02
	Salience	Motor Hand	< -0.02
	Cerebellar	CO Task Control	< -0.04
Relational	Cerebellar	Visual	$< 10^{-3}$
	FP Task Control	Visual	$< 10^{-3}$
	Dorsal attention	Visual	$< 10^{-3}$
	Dorsal attention	FP Task Control	$< 10^{-2}$
	Motor Mouth	Motor Mouth	< 0.03
	Motor Hand	Visual	< 0.04
	Salience	FP Task Control	$< -10^{-3}$
	Motor Mouth	Motor Hand	$< -10^{-3}$
	Visual	Visual	$< -10^{-2}$
	Salience	Default mode	$< -10^{-2}$
	Salience	Salience	$< -10^{-2}$
	CO Task Control	FP Task Control	$< -10^{-2}$
	Motor Mouth	Auditory	$< -10^{-2}$
	Cerebellar	Cerebellar	$< -10^{-2}$
	FP Task Control	Default mode	< -0.04

Table C4: Significant task specific modular interactions corresponding to the social and working memory tasks.

Task	Module	Module	<i>p</i> -value
Social	Dorsal attention	Visual	$< 10^{-3}$
	Dorsal attention	Dorsal attention	$< 10^{-3}$
	Ventral attention	Dorsal attention	$< 10^{-3}$
	Ventral attention	Visual	< 0.04
	Visual	Visual	< 0.04
	Salience	FP Task Control	$< -10^{-3}$
	Cerebellar	Dorsal attention	$< -10^{-3}$
	Memory retrieval	Motor Hand	< -0.02
	Memory retrieval	Dorsal attention	< -0.02
	Cerebellar	Visual	< -0.03
	Cerebellar	Cerebellar	< -0.03
	Dorsal attention	FP Task Control	< -0.04
	Dorsal attention	Salience	< -0.05
	CO Task Control	FP Task Control	< -0.05
Working Memory	Dorsal attention	Default mode	$< 10^{-3}$
	Motor Mouth	Motor Mouth	$< 10^{-3}$
	Memory retrieval	Dorsal attention	$< 10^{-3}$
	FP Task Control	Visual	$< 10^{-2}$
	Auditory	Visual	$< 10^{-2}$
	Dorsal attention	FP Task Control	$< 10^{-2}$
	Memory retrieval	Ventral attention	$< 10^{-2}$
	Motor Hand	Visual	< 0.03
	FP Task Control	FP Task Control	< 0.03
	CO Task Control	Default mode	< 0.04
	Motor Mouth	Visual	< 0.05
	Motor Mouth	Motor Hand	$< -10^{-3}$
	Cerebellar	Dorsal attention	$< -10^{-3}$
	Cerebellar	Cerebellar	$< -10^{-3}$
	CO Task Control	FP Task Control	$< -10^{-2}$
	Cerebellar	CO Task Control	$< -10^{-2}$
	Cerebellar	Visual	< -0.02
	Salience	Motor Hand	< -0.05

Appendix D

Appendix for Chapter 5

Functional MRI Preprocessing

ABIDE dataset

In this study, we used rs-fMRI data of 387 ASD and 436 control subjects from ABIDE [\[184\]](#), a consortium of brain imaging data for sharing within the scientific community. We removed subjects with mean framewise displacement greater than 0.2mm and used the data preprocessed using the C-PAC pipeline from the Preprocessed Connectomes Project. The preprocessing of the fMRI data includes correction of slice timing, realignment of motion, voxel intensity normalization, nuisance signal removal and band-pass filtering.

ADNI dataset

The preprocessing details for the ADNI dataset are the same as in appendix [B](#), section [B](#).

ADHD dataset

Athena's rs-fMRI pipeline involved removing the first four volumes, slice timing correction, realignment for motion correction and linear transformation between the subject's functional and structural MRI image. This was followed by computation of WM and CSF signals and regressing 6 head motion parameters and variation due to physiological noise, head motion, and scanner drifts from the time series. Band-pass filtering and spatial smoothening with a 6 mm FWHM Gaussian filter was applied further. Further details can be seen in [\[274\]](#).

Table D1: Best architectures of SVM Models on the neuroimaging datasets

Dataset	SVM Model (Kernel, C, γ)	CNN Model	FFN Model (hidden layer config)
ABIDE	Linear, 0.001, -	funcNetCNN_C1000	50-10
AD	Linear, 0.001, -	funcNetCNN_C50	50-32-32
ADHD	Polynomial, 1.0, 0.01	funcNetCNN_500	50-32-32
MCI	Sigmoid, 0.1, 0.01	funcNetCNN_C100	20-32-32
MDD	Linear, 0.1, -	funcNetCNN_50	1000-64-32

MDD

We used the preprocessed data provided by the Creativity and Affective Neuroscience Lab (led by Dr. Jiang Qiu), Brain Imaging Center of Southwest University. The dataset contained rs-fMRI scans collected from 3T MRI scanners for an 8-min period for 289 diseased 168 CN subjects at Xinan (First Affiliated Hospital of Chongqing Medical School in Chongqing, China). The preprocessing was performed using DPARSF [134] (<http://restfmri.net>), which is a toolbox based on the SPM8 software package. They discarded the first 10 EPI scans and the remaining scans underwent slice timing correction, motion correction, spatial normalization to standard Montreal Neurological Institute (MNI) space followed by spatial smoothing. To remove spurious correlations band-pass temporal filtering (0.01–0.1 Hz), nuisance signal removal from the ventricles and deep white matter and regression of 24 head motion parameters was applied. They also performed scrubbing movement correction Power et al. [135] and removed subjects with mean framewise displacement greater than 10%. For more details on preprocessing please refer to [136].

Finding optimal hyperparameters for the models

We used the hyperparameter optimization scheme described in Appendix C, section C. The FFN, whereby we and CNN architectures were also similar, except that we tried a wider range of parameters for the FFN and CNN by varying the number of hidden layer neurons from 5 to 1000. The accuracies for different models (FFN, SVM and the CNN) are given in Figure D1.

Fig. D1: Accuracy plots of SVM, FFN and CNN models applied on the various neuropsychiatric diseases. The black dotted lines show the results for SVM (a straight line since there are no neurons in SVMs) and the legend shows the exact configuration used).

References

- [1] 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*, vol. 34, no. 4, pp. 537–541, 1995.
- [2] B. B. Biswal, J. V. Kylen, and J. S. Hyde, "Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps," *NMR in Biomedicine*, vol. 10, no. 4-5, pp. 165–170, 1997.
- [3] D. Cordes, V. M. Haughton, K. Arfanakis, G. J. Wendt, P. A. Turski, C. H. Moritz, M. A. Quigley, and M. E. Meyerand, "Mapping functionally related regions of brain with functional connectivity MR imaging," *American journal of neuroradiology*, vol. 21, no. 9, pp. 1636–1644, 2000.
- [4] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, "Functional connectivity in the resting brain: a network analysis of the default mode hypothesis," *Proceedings of the National Academy of Sciences*, vol. 100, no. 1, pp. 253–258, 2003.
- [5] P. J. Koopmans, M. Barth, S. Orzada, and D. G. Norris, "Multi-echo fMRI of the cortical laminae in humans at 7 t," *NeuroImage*, vol. 56, no. 3, pp. 1276–1285, 2011.
- [6] W. D. Penny, K. J. Friston, J. T. Ashburner, S. J. Kiebel, and T. E. Nichols, *Statistical parametric mapping: the analysis of functional brain images*. Elsevier, 2011.
- [7] R. W. Cox, "Afni: software for analysis and visualization of functional magnetic resonance neuroimages," *Computers and Biomedical Research*, vol. 29, no. 3, pp. 162–173, 1996.
- [8] S. Whitfield-Gabrieli and A. Nieto-Castanon, "Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks," *Brain connectivity*, vol. 2, no. 3, pp. 125–141, 2012.
- [9] C. F. Beckmann and S. M. Smith, "Probabilistic independent component analysis for functional magnetic resonance imaging," *IEEE Transactions on medical imaging*, vol. 23, no. 2, pp. 137–152, 2004.
- [10] O. Esteban, R. Blair, and C. J. e. a. Markiewicz, "Fmriprep 1.1.2," *Software*, 2018.
- [11] R. A. Poldrack, "Region of interest analysis for fMRI," *Social cognitive and affective neuroscience*, vol. 2, no. 1, pp. 67–70, 2007.
- [12] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, "The human brain is intrinsically organized into dynamic, anticorrelated functional networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 27, pp. 9673–9678, 2005.
- [13] V. D. Calhoun, T. Adali, G. D. Pearlson, and J. J. Pekar, "A method for making group inferences from functional MRI data using independent component analysis," *Human Brain Mapping*, vol. 14, no. 3, pp. 140–151, 2001.
- [14] J. D. Power, A. L. Cohen, S. M. Nelson, G. S. Wig *et al.*, "Functional network organization of the human brain," *Neuron*, vol. 72, no. 4, pp. 665–678, 2011.
- [15] N. A. Crossley, A. Mechelli, P. E. Vértes, T. T. Winton-Brown, A. X. Patel, C. E. Ginestet, P. McGuire, and E. T. Bullmore, "Cognitive relevance of the community structure of the human brain functional coactivation network," *Proceedings of the National Academy of Sciences*, vol. 110, no. 28, pp. 11 583–11 588, 2013.
- [16] M. F. Glasser, T. S. Coalson, E. C. Robinson, C. D. Hacker, J. Harwell, E. Yacoub, K. Ugurbil, J. Andersson, C. F. Beckmann, M. Jenkinson *et al.*, "A multi-modal parcellation of human cerebral cortex," *Nature*, vol. 536, no. 7615, pp. 171–178, 2016.

REFERENCES

- [17] H. E. Wang, C. G. Bénar, P. P. Quilichini, K. J. Friston, V. K. Jirsa, and C. Bernard, "A systematic framework for functional connectivity measures," *Frontiers in Neuroscience*, vol. 8, p. 405, 2014.
- [18] Y. Zang, T. Jiang, Y. Lu, Y. He, and L. Tian, "Regional homogeneity approach to fMRI data analysis," *NeuroImage*, vol. 22, no. 1, pp. 394–400, 2004.
- [19] F. T. Sun, L. M. Miller, and M. D'Esposito, "Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data," *NeuroImage*, vol. 21, no. 2, pp. 647–658, 2004.
- [20] A. Roebroeck, E. Formisano, and R. Goebel, "Mapping directed influence over the brain using granger causality and fmri," *Neuroimage*, vol. 25, no. 1, pp. 230–242, 2005.
- [21] A. K. Seth, "A matlab toolbox for granger causal connectivity analysis," *Journal of neuroscience methods*, vol. 186, no. 2, pp. 262–273, 2010.
- [22] G. Nolte, A. Ziehe, V. V. Nikulin, A. Schlögl, N. Krämer, T. Brismar, and K.-R. Müller, "Robustly estimating the flow direction of information in complex physical systems," *Physical review letters*, vol. 100, no. 23, p. 234101, 2008.
- [23] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *Neuroimage*, vol. 19, no. 4, pp. 1273–1302, 2003.
- [24] A. Hyvärinen and S. M. Smith, "Pairwise likelihood ratios for estimation of non-gaussian structural equation models," *Journal of Machine Learning Research*, vol. 14, no. Jan, pp. 111–152, 2013.
- [25] J. A. Mumford and J. D. Ramsey, "Bayesian networks for fmri: a primer," *Neuroimage*, vol. 86, pp. 573–582, 2014.
- [26] S. Achard and E. Bullmore, "Efficiency and cost of economical brain functional networks," *PLoS Comput Biol*, vol. 3, no. 2, p. e17, 2007.
- [27] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *Journal of Neuroscience*, vol. 26, no. 1, pp. 63–72, 2006.
- [28] X. Liu and H. Huang, "Alterations of functional connectivities associated with autism spectrum disorder symptom severity: a multi-site study using multivariate pattern analysis," *Scientific Reports*, vol. 10, no. 1, pp. 1–13, 2020.
- [29] T. Xie and Y. He, "Mapping the Alzheimer's brain with connectomics," *Frontiers in psychiatry*, vol. 2, p. 77, 2012.
- [30] J. Zhang, J. Wang, Q. Wu, W. Kuang, X. Huang, Y. He, and Q. Gong, "Disrupted brain connectivity networks in drug-naive, first-episode major depressive disorder," *Biological psychiatry*, vol. 70, no. 4, pp. 334–342, 2011.
- [31] A. Di Martino, X.-N. Zuo, C. Kelly, R. Grzadzinski, M. Mennes, A. Schvarcz, J. Rodman, C. Lord, F. X. Castellanos, and M. P. Milham, "Shared and distinct intrinsic functional network centrality in autism and attention-deficit/hyperactivity disorder," *Biological psychiatry*, vol. 74, no. 8, pp. 623–632, 2013.
- [32] M. Cao, J.-H. Wang, Z.-J. Dai, X.-Y. Cao, L.-L. Jiang, F.-M. Fan, X.-W. Song, M.-R. Xia, N. Shu, Q. Dong *et al.*, "Topological organization of the human brain functional connectome across the lifespan," *Developmental cognitive neuroscience*, vol. 7, pp. 76–93, 2014.
- [33] W. M. Otte, E. van Diessen, S. Paul, R. Ramaswamy, V. S. Rallabandi, C. J. Stam, and P. K. Roy, "Aging alterations in whole-brain networks during adulthood mapped with the minimum spanning tree indices: the interplay of density, connectivity cost and life-time trajectory," *NeuroImage*, vol. 109, pp. 171–189, 2015.
- [34] S. A. Huettel, A. W. Song, and G. McCarthy, *Functional magnetic resonance imaging*. Sinauer Associates Sunderland, MA, 2004, vol. 1.
- [35] G. L. Baum, R. Ciric, D. R. Roalf, R. F. Betzel, T. M. Moore, R. T. Shinohara, A. E. Kahn, S. N. Vandekar, P. E. Rupert, M. Quarmley *et al.*, "Modular segregation of structural brain networks supports the development of executive function in youth," *Current Biology*, vol. 27, no. 11, pp. 1561–1572, 2017.
- [36] O. Sporns and R. F. Betzel, "Modular brain networks," *Annual review of psychology*, vol. 67, pp. 613–640, 2016.
- [37] T. O. Laumann, E. M. Gordon, B. Adeyemo, A. Z. Snyder, S. J. Joo, M.-Y. Chen, A. W. Gilmore, K. B. McDermott, S. M. Nelson, N. U. Dosenbach *et al.*, "Functional system and areal organization of a highly sampled individual human brain," *Neuron*, vol. 87, no. 3, pp. 657–670, 2015.

REFERENCES

- [38] C. Espinosa-Soto and A. Wagner, "Specialization can drive the evolution of modularity," *PLoS computational biology*, vol. 6, no. 3, p. e1000719, 2010.
- [39] A. Nematzadeh, E. Ferrara, A. Flammini, and Y.-Y. Ahn, "Optimal network modularity for information diffusion," *Physical Review Letters*, vol. 113, no. 8, p. 088701, 2014.
- [40] S. Fortunato and D. Hric, "Community detection in networks: A user guide," *Physics Reports*, vol. 659, pp. 1–44, 2016.
- [41] J. L. Ji, M. Spronk, K. Kulkarni, G. Repovš, A. Anticevic, and M. W. Cole, "Mapping the human brain's cortical-subcortical functional network organization," *NeuroImage*, vol. 185, pp. 35–57, 2019.
- [42] C. Nicolini, C. Bordier, and A. Bifone, "Community detection in weighted brain connectivity networks beyond the resolution limit," *NeuroImage*, vol. 146, pp. 28–39, 2017.
- [43] C. Nicolini and A. Bifone, "Modular structure of brain functional networks: breaking the resolution limit by surprise," *Scientific reports*, vol. 6, no. 1, pp. 1–13, 2016.
- [44] B. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni *et al.*, "The organization of the human cerebral cortex estimated by intrinsic functional connectivity," *Journal of neurophysiology*, vol. 106, no. 3, pp. 1125–1165, 2011.
- [45] M. P. van den Heuvel and O. Sporns, "Network hubs in the human brain," *Trends in Cognitive Sciences*, vol. 17, no. 12, pp. 683–696, 2013.
- [46] R. L. Buckner, J. Sepulcre, T. Talukdar, F. M. Krienen, H. Liu, T. Hedden, J. R. Andrews-Hanna, R. A. Sperling, and K. A. Johnson, "Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease," *Journal of Neuroscience*, vol. 29, no. 6, pp. 1860–1873, 2009.
- [47] M. W. Cole, S. Pathak, and W. Schneider, "Identifying the brain's most globally connected regions," *NeuroImage*, vol. 49, no. 4, pp. 3132–3148, 2010.
- [48] D. Tomasi and N. D. Volkow, "Functional connectivity hubs in the human brain," *NeuroImage*, vol. 57, no. 3, pp. 908–917, 2011.
- [49] D. E. Warren, J. D. Power, J. Bruss, N. L. Denburg, E. J. Waldron, H. Sun, S. E. Petersen, and D. Tranel, "Network measures predict neuropsychological outcome after brain injury," *Proceedings of the National Academy of Sciences*, vol. 111, no. 39, pp. 14 247–14 252, 2014.
- [50] E. Bullmore and O. Sporns, "The economy of brain network organization," *Nature Reviews Neuroscience*, vol. 13, no. 5, pp. 336–349, 2012.
- [51] D. S. Bassett and E. T. Bullmore, "Human brain networks in health and disease," *Current opinion in neurology*, vol. 22, no. 4, p. 340, 2009.
- [52] C. Bordier, C. Nicolini, and A. Bifone, "Graph analysis and modularity of brain functional connectivity networks: searching for the optimal threshold," *Frontiers in Neuroscience*, vol. 11, p. 441, 2017.
- [53] V. Betti, S. Della Penna, F. de Pasquale, D. Mantini, L. Marzetti, G. L. Romani, and M. Corbetta, "Natural scenes viewing alters the dynamics of functional connectivity in the human brain," *Neuron*, vol. 79, no. 4, pp. 782–797, 2013.
- [54] M. W. Cole, D. S. Bassett, J. D. Power, T. S. Braver, and S. E. Petersen, "Intrinsic and task-evoked network architectures of the human brain," *Neuron*, vol. 83, no. 1, pp. 238–251, 2014.
- [55] F. M. Krienen, B. T. Yeo, and R. L. Buckner, "Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 369, no. 1653, p. 20130526, 2014.
- [56] B. Birur, N. V. Kraguljac, R. C. Shelton, and A. C. Lahti, "Brain structure, function, and neurochemistry in Schizophrenia and bipolar disorder—a systematic review of the magnetic resonance neuroimaging literature," *npj Schizophrenia*, vol. 3, no. 1, pp. 1–15, 2017.
- [57] A. Khazaee, A. Ebrahimzadeh, and A. Babajani-Feremi, "Identifying patients with Alzheimer's disease using resting-state fMRI and graph theory," *Clinical Neurophysiology*, vol. 126, no. 11, pp. 2132–2141, 2015.
- [58] M. R. Arbabshirani, K. Kiehl, G. Pearlson, and V. D. Calhoun, "Classification of Schizophrenia patients based on resting-state functional network connectivity," *Frontiers in neuroscience*, vol. 7, p. 133, 2013.

REFERENCES

- [59] K. E. Stephan, F. Schlagenhauf, Q. J. Huys, S. Raman, E. A. Aponte, K. H. Brodersen, L. Rigoux, R. J. Moran, J. Daunizeau, R. J. Dolan *et al.*, “Computational neuroimaging strategies for single patient predictions,” *NeuroImage*, vol. 145, pp. 180–199, 2017.
- [60] G. Deco and M. L. Kringelbach, “Great expectations: using whole-brain computational connectomics for understanding neuropsychiatric disorders,” *Neuron*, vol. 84, no. 5, pp. 892–905, 2014.
- [61] Y. Du, Z. Fu, and V. D. Calhoun, “Classification and prediction of brain disorders using functional connectivity: promising but challenging,” *Frontiers in neuroscience*, vol. 12, 2018.
- [62] R. M. Birn, “The role of physiological noise in resting-state functional connectivity,” *NeuroImage*, vol. 62, no. 2, pp. 864–870, 2012.
- [63] J. L. Cummings, “Cognitive and behavioral heterogeneity in alzheimer’s disease: seeking the neurobiological basis,” *Neurobiology of aging*, vol. 21, no. 6, pp. 845–861, 2000.
- [64] H. Bruining, L. De Sonneville, H. Swaab, M. De Jonge, M. Kas, H. van Engeland, and J. Vorstman, “Dissecting the clinical heterogeneity of autism spectrum disorders through defined genotypes,” *PLoS One*, vol. 5, no. 5, p. e10887, 2010.
- [65] R. E. Peterson, N. Cai, A. W. Dahl, T. B. Bigdeli, A. C. Edwards, B. T. Webb, S.-A. Bacanu, N. Zaitlen, J. Flint, and K. S. Kendler, “Molecular genetic analysis subdivided by adversity exposure suggests etiologic heterogeneity in major depression,” *American Journal of Psychiatry*, vol. 175, no. 6, pp. 545–554, 2018.
- [66] T. Wolfers, C. F. Beckmann, M. Hoogman, J. K. Buitelaar, B. Franke, and A. F. Marquand, “Individual differences v. the average patient: mapping the heterogeneity in adhd using normative models,” *Psychological Medicine*, vol. 50, no. 2, pp. 314–323, 2020.
- [67] K. M. Gates, P. C. Molenaar, S. P. Iyer, J. T. Nigg, and D. A. Fair, “Organizing heterogeneous samples using community detection of gimme-derived resting state functional networks,” *PloS one*, vol. 9, no. 3, p. e91322, 2014.
- [68] X. Sui, J. C. Rajapakse, A. D. N. Initiative *et al.*, “Profiling heterogeneity of alzheimer’s disease using white-matter impairment factors,” *NeuroImage: Clinical*, vol. 20, pp. 1222–1232, 2018.
- [69] J. Kunegis, S. Schmidt, A. Lommatzsch, J. Lerner, E. W. De Luca, and S. Albayrak, “Spectral analysis of signed graphs for clustering, prediction and visualization,” in *Proceedings of the 2010 SIAM International Conference on Data Mining*. SIAM, 2010, pp. 559–570.
- [70] W.-C. Liu, L.-C. Huang, C. W.-J. Liu, and F. Jordán, “A simple approach for quantifying node centrality in signed and directed social networks,” *Applied Network Science*, vol. 5, no. 1, pp. 1–26, 2020.
- [71] S. Gupta and J. C. Rajapakse, “Iterative consensus spectral clustering improves detection of subject and group level brain functional modules,” *Scientific Reports*, vol. 10, no. 1, pp. 1–15, 2020.
- [72] J. D. Power, B. L. Schlaggar, C. N. Lessov-Schlaggar, and S. E. Petersen, “Evidence for hubs in human functional brain networks,” *Neuron*, vol. 79, no. 4, pp. 798–813, 2013.
- [73] B. M. Tijms, A. M. Wink, W. de Haan, W. M. van der Flier, C. J. Stam, P. Scheltens, and F. Barkhof, “Alzheimer’s disease: connecting findings from graph theoretical studies of brain networks,” *Neurobiology of Aging*, vol. 34, no. 8, pp. 2023–2036, 2013.
- [74] Y. I. Sheline, J. L. Price, Z. Yan, and M. A. Mintun, “Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus,” *Proceedings of the National Academy of Sciences*, vol. 107, no. 24, pp. 11 020–11 025, 2010.
- [75] Q. Ma, Y. Tang, F. Wang, X. Liao, X. Jiang, S. Wei, A. Mechelli, Y. He, and M. Xia, “Transdiagnostic dysfunctions in brain modules across patients with Schizophrenia, bipolar disorder, and major depressive disorder: a connectome-based study,” *Schizophrenia Bulletin*, vol. 46, no. 3, pp. 699–712, 2020.
- [76] A. Alexander-Bloch, R. Lambiotte, B. Roberts, J. Giedd, N. Gogtay, and E. Bullmore, “The discovery of population differences in network community structure: new methods and applications to brain functional networks in Schizophrenia,” *NeuroImage*, vol. 59, no. 4, pp. 3889–3900, 2012.
- [77] T. J. Akiki, C. L. Averill, K. M. Wrocklage, J. C. Scott, L. A. Averill, B. Schweinsburg, A. Alexander-Bloch, B. Martini, S. M. Southwick, J. H. Krystal *et al.*, “Default mode network abnormalities in posttraumatic stress disorder: A novel network-restricted topology approach,” *NeuroImage*, vol. 176, pp. 489–498, 2018.
- [78] P. Balenzuela, A. Chernomoretz, D. Fraiman, I. Cifre, C. Sitges, P. Montoya, and D. R. Chialvo, “Modular organization of brain resting state networks in chronic back pain patients,” *Frontiers in Neuroinformatics*, vol. 4, p. 116, 2010.

REFERENCES

- [79] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proceedings of the National Academy of Sciences*, vol. 105, no. 4, pp. 1118–1123, 2008.
- [80] D. Meunier, R. Lambiotte, A. Fornito, K. Ersche, and E. T. Bullmore, "Hierarchical modularity in human brain functional networks," *Frontiers in Neuroinformatics*, vol. 3, p. 37, 2009.
- [81] S. Fortunato and M. Barthélemy, "Resolution limit in community detection," *Proceedings of the National Academy of Sciences*, vol. 104, no. 1, pp. 36–41, 2007.
- [82] E. M. Gordon, T. O. Laumann, A. W. Gilmore, D. J. Newbold, D. J. Greene, J. J. Berg, M. Ortega, C. Hoyt-Drazen, C. Gratton, H. Sun *et al.*, "Precision functional mapping of individual human brains," *Neuron*, vol. 95, no. 4, pp. 791–807, 2017.
- [83] J. Dubois and R. Adolphs, "Building a science of individual differences from fMRI," *Trends in Cognitive sciences*, vol. 20, no. 6, pp. 425–443, 2016.
- [84] E. S. Finn, X. Shen, D. Scheinost, M. D. Rosenberg, J. Huang, M. M. Chun, X. Papademetris, and R. T. Constable, "Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity," *Nature Neuroscience*, vol. 18, no. 11, p. 1664, 2015.
- [85] S. Mueller, D. Wang, M. D. Fox, B. T. Yeo, J. Sepulcre, M. R. Sabuncu, R. Shafee, J. Lu, and H. Liu, "Individual variability in functional connectivity architecture of the human brain," *Neuron*, vol. 77, no. 3, pp. 586–595, 2013.
- [86] A. Lancichinetti and S. Fortunato, "Consensus clustering in complex networks," *Scientific Reports*, vol. 2, p. 336, 2012.
- [87] X. Y. Stella and J. Shi, "Multiclass spectral clustering," in *Proc. of International Conference on Computer Vision*. IEEE, 2003, pp. 313–319.
- [88] M. W. Cole, T. Yarkoni, G. Repovš, A. Anticevic, and T. S. Braver, "Global connectivity of prefrontal cortex predicts cognitive control and intelligence," *Journal of Neuroscience*, vol. 32, no. 26, pp. 8988–8999, 2012.
- [89] E. Santarnecchi, G. Galli, N. R. Polizzotto, A. Rossi, and S. Rossi, "Efficiency of weak brain connections support general cognitive functioning," *Human Brain Mapping*, vol. 35, no. 9, pp. 4566–4582, 2014.
- [90] D. S. Bassett, B. G. Nelson, B. A. Mueller, J. Camchong, and K. O. Lim, "Altered resting state complexity in Schizophrenia," *NeuroImage*, vol. 59, no. 3, pp. 2196–2207, 2012.
- [91] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of statistical mechanics: theory and experiment*, vol. 2008, no. 10, p. P10008, 2008.
- [92] M. Girvan and M. E. Newman, "Community structure in social and biological networks," *Proceedings of the national academy of sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.
- [93] M. E. Newman, "Modularity and community structure in networks," *Proceedings of the National Academy of Sciences*, vol. 103, no. 23, pp. 8577–8582, 2006.
- [94] A. Clauset, M. E. Newman, and C. Moore, "Finding community structure in very large networks," *Physical Review E*, vol. 70, no. 6, p. 066111, 2004.
- [95] J. Reichardt and S. Bornholdt, "Statistical mechanics of community detection," *Physical Review E*, vol. 74, no. 1, p. 016110, 2006.
- [96] A. Lancichinetti and S. Fortunato, "Limits of modularity maximization in community detection," *Physical Review E*, vol. 84, no. 6, p. 066122, 2011.
- [97] D. A. Huffman, "A method for the construction of minimum-redundancy codes," *Proceedings of the IRE*, vol. 40, no. 9, pp. 1098–1101, 1952.
- [98] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 888–905, 2000.
- [99] T. Kawamoto and M. Rosvall, "Estimating the resolution limit of the map equation in community detection," *Physical Review E*, vol. 91, no. 1, p. 012809, 2015.
- [100] T. van Laarhoven and E. Marchiori, "Graph clustering with local search optimization: The resolution bias of the objective function matters most," *Physical Review E*, vol. 87, no. 1, p. 012812, 2013.

REFERENCES

- [101] E. M. Gordon, T. O. Laumann, B. Adeyemo, A. W. Gilmore, S. M. Nelson, N. U. Dosenbach, and S. E. Petersen, "Individual-specific features of brain systems identified with resting state functional correlations," *NeuroImage*, vol. 146, pp. 918–939, 2017.
- [102] C. D. Hacker, T. O. Laumann, and N. P. e. a. Szrama, "Resting state network estimation in individual subjects," *NeuroImage*, vol. 82, pp. 616–633, 2013.
- [103] B. Chen, T. Xu, C. Zhou, L. Wang, N. Yang, Z. Wang, H.-M. Dong, Z. Yang, Y.-F. Zang, X.-N. Zuo *et al.*, "Individual variability and test-retest reliability revealed by ten repeated resting-state brain scans over one month," *PLoS One*, vol. 10, no. 12, p. e0144963, 2015.
- [104] S. M. Smith, C. F. Beckmann, N. Ramnani, M. W. Woolrich, P. R. Bannister, M. Jenkinson, P. M. Matthews, and D. J. McGonigle, "Variability in fMRI: a re-examination of inter-session differences," *Human Brain Mapping*, vol. 24, no. 3, pp. 248–257, 2005.
- [105] M. Pannunzi, R. Hindriks, R. G. Bettinardi, E. Wenger, N. Lisofsky, J. Martensson, O. Butler, E. Filevich, M. Becker, M. Lochstet *et al.*, "Resting-state fMRI correlations: from link-wise unreliability to whole brain stability," *NeuroImage*, vol. 157, pp. 250–262, 2017.
- [106] N. X. Vinh, J. Epps, and J. Bailey, "Information theoretic measures for clusterings comparison: Variants, properties, normalization and correction for chance," *Journal of Machine Learning Research*, vol. 11, no. Oct, pp. 2837–2854, 2010.
- [107] A. Amelio and C. Pizzuti, "Is normalized mutual information a fair measure for comparing community detection methods?" in *Proceedings of the 2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2015*. ACM, 2015, pp. 1584–1585.
- [108] S. Romano, N. X. Vinh, J. Bailey, and K. Verspoor, "Adjusting for chance clustering comparison measures," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 4635–4666, 2016.
- [109] S. Gupta and J. C. Rajapakse, "Nodal degree distributions of resting-state functional brain modules," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018, pp. 1256–1259.
- [110] Y. Liu, Z. Li, H. Xiong, X. Gao, and J. Wu, "Understanding of internal clustering validation measures," in *2010 IEEE International Conference on Data Mining*. IEEE, 2010, pp. 911–916.
- [111] S. Sarkar and A. Dong, "Community detection in graphs using singular value decomposition," *Physical Review E*, vol. 83, no. 4, p. 046114, 2011.
- [112] V. Satopaa, J. Albrecht, D. Irwin, and B. Raghavan, "Finding a" kneedle" in a haystack: Detecting knee points in system behavior," in *2011 31st International Conference on Distributed Computing Systems Workshops*. IEEE, 2011, pp. 166–171.
- [113] G. Ojemann, J. Ojemann, E. Lettich, and M. Berger, "Cortical language localization in left, dominant hemisphere: an electrical stimulation mapping investigation in 117 patients," *Journal of Neurosurgery*, vol. 71, no. 3, pp. 316–326, 1989.
- [114] M. A. Frost and R. Goebel, "Measuring structural–functional correspondence: spatial variability of specialised brain regions after macro-anatomical alignment," *NeuroImage*, vol. 59, no. 2, pp. 1369–1381, 2012.
- [115] B. L. Finlay, R. B. Darlington, and N. Nicastro, "Developmental structure in brain evolution," *Behavioral and Brain Sciences*, vol. 24, no. 2, pp. 263–278, 2001.
- [116] W. de Haan, W. M. van der Flier, T. Koene, L. L. Smits, P. Scheltens, and C. J. Stam, "Disrupted modular brain dynamics reflect cognitive dysfunction in Alzheimer's disease," *NeuroImage*, vol. 59, no. 4, pp. 3085–3093, 2012.
- [117] D.-J. Kim, J. S. Kent, A. R. Bolbecker, O. Sporns, H. Cheng, S. D. Newman, A. Puce, B. F. O'donnell, and W. P. Hetrick, "Disrupted modular architecture of cerebellum in Schizophrenia: a graph theoretic analysis," *Schizophrenia Bulletin*, vol. 40, no. 6, pp. 1216–1226, 2014.
- [118] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, "Dynamic reconfiguration of human brain networks during learning," *Proceedings of the National Academy of Sciences*, vol. 108, no. 18, pp. 7641–7646, 2011.
- [119] L. Q. Uddin, K. Supekar, C. J. Lynch, A. Khouzam, J. Phillips, C. Feinstein, S. Ryali, and V. Menon, "Salience network–based classification and prediction of symptom severity in children with autism," *JAMA psychiatry*, vol. 70, no. 8, pp. 869–879, 2013.
- [120] X. Ma, Z. Zhuo, L. Wei, Z. Ma, Z. Li, H. Li, A. D. N. Initiative *et al.*, "Altered temporal organization of brief spontaneous brain activities in patients with Alzheimer's disease," *Neuroscience*, vol. 425, pp. 1–11, 2020.
- [121] Z. Dai, Q. Lin, T. Li, X. Wang, H. Yuan, X. Yu, Y. He, and H. Wang, "Disrupted structural and functional brain networks in Alzheimer's disease," *Neurobiology of aging*, vol. 75, pp. 71–82, 2019.

REFERENCES

- [122] Q. Ma, B. Huang, J. Wang, C. Seger, W. Yang, C. Li, J. Wang, J. Feng, L. Weng, W. Jiang *et al.*, “Altered modular organization of intrinsic brain functional networks in patients with parkinson’s disease,” *Brain imaging and behavior*, vol. 11, no. 2, pp. 430–443, 2017.
- [123] F. Chollet, *Deep Learning with Python and Keras: The practical guide from the developer of the Keras library*. MITP-Verlags GmbH & Co. KG, 2018.
- [124] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et al.*, “Tensorflow: A system for large-scale machine learning,” in *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, 2016, pp. 265–283.
- [125] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium *et al.*, “The WU-Minn human connectome project: an overview,” *NeuroImage*, vol. 80, pp. 62–79, 2013.
- [126] M. F. Glasser, S. M. Smith, D. S. Marcus, J. L. Andersson, E. J. Auerbach, T. E. Behrens, T. S. Coalson, M. P. Harms, M. Jenkinson, S. Moeller *et al.*, “The human connectome project’s neuroimaging approach,” *Nature Neuroscience*, vol. 19, no. 9, p. 1175, 2016.
- [127] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, Coalson *et al.*, “The minimal preprocessing pipelines for the human connectome project,” *NeuroImage*, vol. 80, pp. 105–124, 2013.
- [128] L. Heine, A. Soddu, F. Gómez, A. Vanhaudenhuyse, L. Tshibanda, M. Thonnard, V. Charland-Verville, M. Kirsch, S. Laureys, and A. Demertzi, “Resting state networks and consciousness,” *Frontiers in Psychology*, vol. 3, p. 295, 2012.
- [129] F. W. Sun, M. R. Stepanovic, J. Andreano, L. F. Barrett, A. Touroutoglou, and B. C. Dickerson, “Youthful brains in older adults: Preserved neuroanatomy in the default mode and salience networks contributes to youthful memory in superaging,” *Journal of Neuroscience*, vol. 36, no. 37, pp. 9659–9668, 2016.
- [130] J. R. Andrews-Hanna, J. S. Reidler, J. Sepulcre, R. Poulin, and R. L. Buckner, “Functional-anatomic fractionation of the brain’s default network,” *Neuron*, vol. 65, no. 4, pp. 550–562, 2010.
- [131] D. J. Felleman and D. E. Van, “Distributed hierarchical processing in the primate cerebral cortex,” *Cerebral Cortex (New York, NY: 1991)*, vol. 1, no. 1, pp. 1–47, 1991.
- [132] S. M. Smith, P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A. R. Laird *et al.*, “Correspondence of the brain’s functional architecture during activation and rest,” *Proceedings of the National Academy of Sciences*, vol. 106, no. 31, pp. 13 040–13 045, 2009.
- [133] M. Xia, J. Wang, and Y. He, “Brainnet viewer: a network visualization tool for human brain connectomics,” *PloS One*, vol. 8, no. 7, p. e68910, 2013.
- [134] C. Yan and Y. Zang, “DPARSF: a matlab toolbox for pipeline data analysis of resting-state fMRI,” *Frontiers in Systems Neuroscience*, vol. 4, p. 13, 2010.
- [135] J. D. Power, A. Mitra, T. O. Laumann, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, “Methods to detect, characterize, and remove motion artifact in resting state fMRI,” *NeuroImage*, vol. 84, pp. 320–341, 2014.
- [136] W. Cheng, E. T. Rolls, J. Qiu, W. Liu, Y. Tang, C.-C. Huang, X. Wang, J. Zhang, W. Lin, L. Zheng *et al.*, “Medial reward and lateral non-reward orbitofrontal cortex circuits change in opposite directions in depression,” *Brain*, vol. 139, no. 12, pp. 3296–3309, 2016.
- [137] M. H. Sneve, H. Grydeland, M. G. Rosa, T. Paus, T. Chaplin, K. Walhovd, and A. M. Fjell, “High-expanding regions in primate cortical brain evolution support supramodal cognitive flexibility,” *Cerebral Cortex*, vol. 29, no. 9, pp. 3891–3901, 2019.
- [138] J. Sepulcre, H. Liu, T. Talukdar, I. Martincorena, B. T. Yeo, and R. L. Buckner, “The organization of local and distant functional connectivity in the human brain,” *PLoS computational biology*, vol. 6, no. 6, p. e1000808, 2010.
- [139] L. K. Gallos, H. A. Makse, and M. Sigman, “A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks,” *Proceedings of the National Academy of Sciences*, vol. 109, no. 8, pp. 2825–2830, 2012.
- [140] A. Lancichinetti, S. Fortunato, and F. Radicchi, “Benchmark graphs for testing community detection algorithms,” *Physical Review E*, vol. 78, no. 4, p. 046110, 2008.
- [141] A. Lancichinetti and S. Fortunato, “Benchmarks for testing community detection algorithms on directed and weighted graphs with overlapping communities,” *Physical Review E*, vol. 80, no. 1, p. 016118, 2009.

REFERENCES

- [142] E. Yoneki, P. Hui, S. Chan, and J. Crowcroft, "A socio-aware overlay for publish/subscribe communication in delay tolerant networks," in *Proceedings of the 10th ACM Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*. ACM, 2007, pp. 225–234.
- [143] M. Van Den Heuvel, R. Mandl, and H. H. Pol, "Normalized cut group clustering of resting-state FMRI data," *PloS One*, vol. 3, no. 4, p. e2001, 2008.
- [144] J. C. Rajapakse, S. Gupta, and X. Sui, "Fitting networks models for functional brain connectivity," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, 2017, pp. 515–519.
- [145] P. Shaw, N. J. Kabani, J. P. Lerch, K. Eckstrand, R. Lenroot, N. Gogtay, D. Greenstein, L. Clasen, A. Evans, J. L. Rapoport *et al.*, "Neurodevelopmental trajectories of the human cerebral cortex," *Journal of Neuroscience*, vol. 28, no. 14, pp. 3586–3594, 2008.
- [146] Z. Petanjek, M. Judaš, G. Šimić, M. R. Rašin, H. B. Uylings, P. Rakic, and I. Kostović, "Extraordinary neoteny of synaptic spines in the human prefrontal cortex," *Proceedings of the National Academy of Sciences*, vol. 108, no. 32, pp. 13 281–13 286, 2011.
- [147] Y. Xia, Q. Chen, L. Shi, M. Li, W. Gong, H. Chen, and J. Qiu, "Tracking the dynamic functional connectivity structure of the human brain across the adult lifespan," *Human Brain Mapping*, vol. 40, no. 3, pp. 717–728, 2019.
- [148] J. A. Archer, A. Lee, A. Qiu, and S.-H. A. Chen, "A comprehensive analysis of connectivity and aging over the adult life span," *Brain connectivity*, vol. 6, no. 2, pp. 169–185, 2016.
- [149] M. G. Hart, S. J. Price, and J. Suckling, "Functional connectivity networks for preoperative brain mapping in neurosurgery," *Journal of Neurosurgery*, vol. 126, no. 6, pp. 1941–1950, 2016.
- [150] C.-G. Yan, R. C. Craddock, X.-N. Zuo, Y.-F. Zang, and M. P. Milham, "Standardizing the intrinsic brain: towards robust measurement of inter-individual variation in 1000 functional connectomes," *NeuroImage*, vol. 80, pp. 246–262, 2013.
- [151] X. Liao, M. Cao, M. Xia, and Y. He, "Individual differences and time-varying features of modular brain architecture," *NeuroImage*, vol. 152, pp. 94–107, 2017.
- [152] R. Kong, J. Li, C. Orban, M. R. Sabuncu, H. Liu, A. Schaefer, N. Sun, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff *et al.*, "Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion," *Cerebral cortex*, vol. 29, no. 6, pp. 2533–2551, 2019.
- [153] D. Wang, R. L. Buckner, M. D. Fox, D. J. Holt, A. J. Holmes, S. Stoecklein, G. Langs, R. Pan, T. Qian, K. Li *et al.*, "Parcellating cortical functional networks in individuals," *Nature Neuroscience*, vol. 18, no. 12, p. 1853, 2015.
- [154] S. Gupta, J. C. Rajapakse, R. E. Welsch, A. D. N. Initiative *et al.*, "Ambivert degree identifies crucial brain functional hubs and improves detection of Alzheimer's disease and Autism spectrum disorder," *NeuroImage: Clinical*, vol. 25, p. 102186, 2020.
- [155] J. D. Power and S. E. Petersen, "Control-related systems in the human brain," *Current opinion in neurobiology*, vol. 23, no. 2, pp. 223–228, 2013.
- [156] A. D. Friederici and S. M. Gierhan, "The language network," *Current Opinion in Neurobiology*, vol. 23, no. 2, pp. 250–254, 2013.
- [157] M. Behrmann and D. C. Plaut, "Distributed circuits, not circumscribed centers, mediate visual recognition," *Trends in Cognitive Sciences*, vol. 17, no. 5, pp. 210–219, 2013.
- [158] M. Senden, N. Reuter, M. P. van den Heuvel, R. Goebel, G. Deco, and M. Gilson, "Task-related effective connectivity reveals that the cortical rich club gates cortex-wide communication," *Human Brain Mapping*, vol. 39, no. 3, pp. 1246–1262, 2018.
- [159] E. M. Gordon, C. J. Lynch, C. Gratton, T. O. Laumann, A. W. Gilmore, D. J. Greene, M. Ortega, A. L. Nguyen, B. L. Schlaggar, S. E. Petersen *et al.*, "Three distinct sets of connector hubs integrate human brain function," *Cell Reports*, vol. 24, no. 7, pp. 1687–1695, 2018.
- [160] M. P. van den Heuvel and A. Fornito, "Brain networks in Schizophrenia," *Neuropsychology Review*, vol. 24, no. 1, pp. 32–48, 2014.
- [161] E. R. Vargas and L. M. Wahl, "The gateway coefficient: a novel metric for identifying critical connections in modular networks," *The European Physical Journal B*, vol. 87, no. 7, p. 161, 2014.
- [162] C. Gratton, E. M. Nomura, F. Pérez, and M. D'Esposito, "Focal brain lesions to critical locations cause widespread disruption of the modular organization of the brain," *Journal of Cognitive Neuroscience*, vol. 24, no. 6, pp. 1275–1285, 2012.

REFERENCES

- [163] A. Fornito, A. Zalesky, and M. Breakspear, "The connectomics of brain disorders," *Nature Reviews Neuroscience*, vol. 16, no. 3, pp. 159–172, 2015.
- [164] J. Guillon, Y. Attal, O. Colliot, V. La Corte, B. Dubois, D. Schwartz, M. Chavez, and F. D. V. Fallani, "Loss of brain inter-frequency hubs in Alzheimer's disease," *Scientific Reports*, vol. 7, no. 1, p. 10879, 2017.
- [165] M. Rubinov *et al.*, "Schizophrenia and abnormal brain network hubs," *Dialogues in Clinical Neuroscience*, vol. 15, no. 3, p. 339, 2013.
- [166] C. Bordier, C. Nicolini, G. Forcellini, and A. Bifone, "Disrupted modular organization of primary sensory brain areas in Schizophrenia," *NeuroImage: Clinical*, vol. 18, pp. 682–693, 2018.
- [167] T. Itahashi, T. Yamada, H. Watanabe, M. Nakamura, D. Jimbo, S. Shioda, K. Toriizuka, N. Kato, and R. Hashimoto, "Altered network topologies and hub organization in adults with autism: a resting-state fMRI study," *PloS One*, vol. 9, no. 4, p. e94115, 2014.
- [168] X. Liang, Q. Zou, Y. He, and Y. Yang, "Coupling of functional connectivity and regional cerebral blood flow reveals a physiological basis for network hubs of the human brain," *Proceedings of the National Academy of Sciences*, vol. 110, no. 5, pp. 1929–1934, 2013.
- [169] D. Tomasi, G.-J. Wang, and N. D. Volkow, "Energetic cost of brain functional connectivity," *Proceedings of the National Academy of Sciences*, vol. 110, no. 33, pp. 13 642–13 647, 2013.
- [170] S. Saxena and P. Caroni, "Selective neuronal vulnerability in neurodegenerative diseases: from stressor thresholds to degeneration," *Neuron*, vol. 71, no. 1, pp. 35–48, 2011.
- [171] J. V. Hull, L. B. Dokovna, Z. J. Jacokes, C. M. Torgerson, A. Irimia, and J. D. Van Horn, "Resting-state functional connectivity in autism spectrum disorders: a review," *Frontiers in Psychiatry*, vol. 7, p. 205, 2017.
- [172] W. de Haan, K. Mott, E. C. van Straaten, P. Scheltens, and C. J. Stam, "Activity dependent degeneration explains hub vulnerability in Alzheimer's disease," *PLoS computational biology*, vol. 8, no. 8, p. e1002582, 2012.
- [173] Y. I. Sheline and M. E. Raichle, "Resting state functional connectivity in preclinical Alzheimer's disease," *Biological Psychiatry*, vol. 74, no. 5, pp. 340–347, 2013.
- [174] M. Bertolero, B. Yeo, and M. D'esposito, "The diverse club," *Nature Communications*, vol. 8, no. 1, p. 1277, 2017.
- [175] S. Hayasaka and P. J. Laurienti, "Comparison of characteristics between region-and voxel-based network analyses in resting-state fMRI data," *NeuroImage*, vol. 50, no. 2, pp. 499–508, 2010.
- [176] D. Tomasi and N. D. Volkow, "Association between functional connectivity hubs and brain networks," *Cerebral Cortex*, vol. 21, no. 9, pp. 2003–2013, 2011.
- [177] A. Clauset, C. R. Shalizi, and M. E. Newman, "Power-law distributions in empirical data," *SIAM review*, vol. 51, no. 4, pp. 661–703, 2009.
- [178] T. Nishikawa, A. E. Motter, Y.-C. Lai, and F. C. Hoppensteadt, "Heterogeneity in oscillator networks: Are smaller worlds easier to synchronize?" *Physical Review Letters*, vol. 91, no. 1, p. 014101, 2003.
- [179] L. C. Freeman, "A set of measures of centrality based on betweenness," *Sociometry*, pp. 35–41, 1977.
- [180] R. Guimera and L. A. N. Amaral, "Functional cartography of complex metabolic networks," *Nature*, vol. 433, no. 7028, pp. 895–900, 2005.
- [181] M. A. Bertolero, B. T. Yeo, and M. D'Esposito, "The modular and integrative functional architecture of the human brain," *Proceedings of the National Academy of Sciences*, vol. 112, no. 49, pp. E6798–E6807, 2015.
- [182] C. Gratton, T. O. Laumann, E. M. Gordon, B. Adeyemo, and S. E. Petersen, "Evidence for two independent factors that modify brain networks to meet task goals," *Cell Reports*, vol. 17, no. 5, pp. 1276–1288, 2016.
- [183] O. Esteban, C. J. Markiewicz, R. W. Blair, C. A. Moodie, A. I. Isik, A. Erramuzpe, J. D. Kent, M. Goncalves, E. DuPre, M. Snyder *et al.*, "fmriprep: a robust preprocessing pipeline for functional mri," *Nature methods*, vol. 16, no. 1, pp. 111–116, 2019.
- [184] A. Di Martino, C.-G. Yan, Q. Li, E. Denio, F. X. Castellanos, K. Alaerts, J. S. Anderson, M. Assaf, S. Y. Bookheimer, M. Dapretto *et al.*, "The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism," *Molecular Psychiatry*, vol. 19, no. 6, p. 659, 2014.

REFERENCES

- [185] D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch *et al.*, “The human connectome project: a data acquisition perspective,” *NeuroImage*, vol. 62, no. 4, pp. 2222–2231, 2012.
- [186] J. Alstott, E. Bullmore, and D. Plenz, “powerlaw: a python package for analysis of heavy-tailed distributions,” *PloS One*, vol. 9, no. 1, p. e85777, 2014.
- [187] A. D. Milner and M. A. Goodale, “Two visual systems re-viewed,” *Neuropsychologia*, vol. 46, no. 3, pp. 774–785, 2008.
- [188] M. Benedek, E. Jauk, R. E. Beaty, A. Fink, K. Koschutnig, and A. C. Neubauer, “Brain mechanisms associated with internally directed attention and self-generated thought,” *Scientific Reports*, vol. 6, p. 22959, 2016.
- [189] M. Benedek, T. Schües, R. E. Beaty, E. Jauk, K. Koschutnig, A. Fink, and A. C. Neubauer, “To create or to recall original ideas: Brain processes associated with the imagination of novel object uses,” *Cortex*, vol. 99, pp. 93–102, 2018.
- [190] S. Gupta, Y. H. Chan, J. C. Rajapakse, A. D. N. Initiative *et al.*, “Decoding brain functional connectivity implicated in ad and mci,” in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2019, pp. 781–789.
- [191] J. Wang, X. Zuo, Z. Dai, M. Xia, Z. Zhao, X. Zhao, J. Jia, Y. Han, and Y. He, “Disrupted functional brain connectome in individuals at risk for Alzheimer’s disease,” *Biological Psychiatry*, vol. 73, no. 5, pp. 472–481, 2013.
- [192] R. J. Meszlényi, K. Buza, and Z. Vidnyánszky, “Resting state fMRI functional connectivity-based classification using a convolutional neural network architecture,” *Frontiers in Neuroinformatics*, vol. 11, p. 61, 2017.
- [193] Z. Dai, C. Yan, K. Li, Z. Wang, J. Wang, M. Cao, Q. Lin, N. Shu, M. Xia, Y. Bi *et al.*, “Identifying and mapping connectivity patterns of brain network hubs in Alzheimer’s disease,” *Cerebral Cortex*, vol. 25, no. 10, pp. 3723–3742, 2014.
- [194] J. Mutlu, B. Landeau, M. Gaubert, V. de La Sayette, B. Desgranges, and G. Chételat, “Distinct influence of specific versus global connectivity on the different Alzheimer’s disease biomarkers,” *Brain*, vol. 140, no. 12, pp. 3317–3328, 2017.
- [195] M. Weiler, A. Fukuda, L. HP Massabki, T. M Lopes, A. R Franco, B. P Damasceno, F. Cendes, and M. LF Balthazar, “Default mode, executive function, and language functional connectivity networks are compromised in mild alzheimer’s disease,” *Current Alzheimer Research*, vol. 11, no. 3, pp. 274–282, 2014.
- [196] M. R. Brier, J. B. Thomas, A. Z. Snyder, T. L. Benzinger, D. Zhang, M. E. Raichle, D. M. Holtzman, J. C. Morris, and B. M. Ances, “Loss of intranetwork and internetwork resting state functional connections with Alzheimer’s disease progression,” *Journal of Neuroscience*, vol. 32, no. 26, pp. 8890–8899, 2012.
- [197] B. C. Dickerson and R. A. Sperling, “Large-scale functional brain network abnormalities in Alzheimer’s disease: insights from functional neuroimaging,” *Behavioural neurology*, vol. 21, no. 1-2, pp. 63–75, 2009.
- [198] A. Padmanabhan, C. J. Lynch, M. Schaer, and V. Menon, “The default mode network in autism,” *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, vol. 2, no. 6, pp. 476–486, 2017.
- [199] B. A. Zielinski, J. S. Anderson, A. L. Froehlich, M. B. Prigge, J. A. Nielsen, J. R. Cooperrider, A. N. Cariello, P. T. Fletcher, A. L. Alexander, N. Lange *et al.*, “scmri reveals large-scale brain network abnormalities in autism,” *PloS One*, vol. 7, no. 11, p. e49172, 2012.
- [200] D. Tomasi and N. D. Volkow, “Aging and functional brain networks,” *Molecular psychiatry*, vol. 17, no. 5, p. 549, 2012.
- [201] A. C. Yang, S.-J. Tsai, M.-E. Liu, C.-C. Huang, and C.-P. Lin, “The association of aging with white matter integrity and functional connectivity hubs,” *Frontiers in aging neuroscience*, vol. 8, p. 143, 2016.
- [202] X. Hu, C. E. Teunissen, A. Spottke, M. T. Heneka, E. Düzel, O. Peters, S. Li, J. Priller, K. Buerger, S. Teipel *et al.*, “Smaller medial temporal lobe volumes in individuals with subjective cognitive decline and biomarker evidence of Alzheimer’s disease : Data from three memory clinic studies,” *Alzheimer’s & Dementia*, vol. 15, no. 2, pp. 185–193, 2019.
- [203] B. L. Klaassens, J. van Gerven, J. van der Grond, F. de Vos, C. Möller, and S. A. Rombouts, “Diminished posterior precuneus connectivity with the default mode network differentiates normal aging from Alzheimer’s disease,” *Frontiers in aging Neuroscience*, vol. 9, p. 97, 2017.
- [204] K.-i. Yamashita, T. Uehara, P. Prawiroharjo, K. Yamashita, O. Togao, A. Hiwatashi, Y. Taniwaki, H. Utsunomiya, T. Matsushita, R. Yamasaki *et al.*, “Functional connectivity change between posterior cingulate cortex and ventral attention network relates to the impairment of orientation for time in Alzheimer’s disease patients,” *Brain imaging and behavior*, vol. 13, no. 1, pp. 154–161, 2019.

REFERENCES

- [205] D. P. Kennedy and E. Courchesne, "The intrinsic functional organization of the brain is altered in autism," *NeuroImage*, vol. 39, no. 4, pp. 1877–1885, 2008.
- [206] N. de Lacy, D. Doherty, B. King, S. Rachakonda, and V. D. Calhoun, "Disruption to control network function correlates with altered dynamic connectivity in the wider autism spectrum," *NeuroImage: Clinical*, vol. 15, pp. 513–524, 2017.
- [207] T. Ikuta, K. M. Shafritz, J. Bregman, B. D. Peters, P. Gruner, A. K. Malhotra, and P. R. Szeszko, "Abnormal cingulum bundle development in autism: a probabilistic tractography study," *Psychiatry Research: Neuroimaging*, vol. 221, no. 1, pp. 63–68, 2014.
- [208] R. J. Jou, A. P. Jackowski, X. Papademetris, N. Rajeevan, L. H. Staib, and F. R. Volkmar, "Diffusion tensor imaging in autism spectrum disorders: preliminary evidence of abnormal neural connectivity," *Australian & New Zealand Journal of Psychiatry*, vol. 45, no. 2, pp. 153–162, 2011.
- [209] K. N. Thakkar, F. E. Polli, R. M. Joseph, D. S. Tuch, N. Hadjikhani, J. J. Barton, and D. S. Manoach, "Response monitoring, repetitive behaviour and anterior cingulate abnormalities in autism spectrum disorders (asd)," *Brain*, vol. 131, no. 9, pp. 2464–2478, 2008.
- [210] R. Ju, C. Hu, P. Zhou, and Q. Li, "Early diagnosis of alzheimer's disease based on resting-state brain networks and deep learning," *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, vol. 16, no. 1, pp. 244–257, 2019.
- [211] T. Iidaka, "Resting state functional magnetic resonance imaging and neural network classified autism and control," *Cortex*, vol. 63, pp. 55–67, 2015.
- [212] H. Wang, C. Chen, and H. Fushing, "Extracting multiscale pattern information of fMRI based functional brain connectivity with application on classification of autism spectrum disorders," *PLoS One*, vol. 7, no. 10, p. e45502, 2012.
- [213] A. Khazaei, A. Ebrahimzadeh, and A. Babajani-Feremi, "Application of advanced machine learning methods on resting-state fMRI network for identification of mild cognitive impairment and Alzheimer's disease," *Brain imaging and behavior*, vol. 10, no. 3, pp. 799–817, 2016.
- [214] Y. Li, Y. Qin, X. Chen, and W. Li, "Exploring the functional brain network of Alzheimer's disease: based on the computational experiment," *PLoS One*, vol. 8, no. 9, p. e73186, 2013.
- [215] D. A. Norman and T. Shallice, "Attention to action," in *Consciousness and Self-regulation*. Springer, 1986, pp. 1–18.
- [216] E. Düzel, R. Cabeza, T. W. Picton, A. P. Yonelinas, H. Scheich, H.-J. Heinze, and E. Tulving, "Task-related and item-related brain processes of memory retrieval," *Proceedings of the National Academy of Sciences*, vol. 96, no. 4, pp. 1794–1799, 1999.
- [217] L. Nyberg, A. R. McIntosh, R. Cabeza, R. Habib, S. Houle, and E. Tulving, "General and specific brain regions involved in encoding and retrieval of events: what, where, and when," *Proceedings of the National Academy of Sciences*, vol. 93, no. 20, pp. 11 280–11 285, 1996.
- [218] M. Meinzer, L. Seeds, T. Flaisch, and et al, "Impact of changed positive and negative task-related brain activity on word-retrieval in aging," *Neurobiology of Aging*, vol. 33, no. 4, pp. 656–669, 2012.
- [219] J. P. Hamilton, D. J. Furman, C. Chang, M. E. Thomason, E. Dennis, and I. H. Gotlib, "Default-mode and task-positive network activity in major depressive disorder: implications for adaptive and maladaptive rumination," *Biological Psychiatry*, vol. 70, no. 4, pp. 327–333, 2011.
- [220] G. G. Knyazev, A. N. Savostyanov, A. V. Bocharov *et al.*, "Task-positive and task-negative networks in major depressive disorder: a combined fMRI and EEG study," *Journal of affective disorders*, vol. 235, pp. 211–219, 2018.
- [221] B. D. Mills, O. Miranda-Dominguez, K. L. Mills *et al.*, "ADHD and attentional control: Impaired segregation of task positive and task negative brain networks," *Network Neuroscience*, vol. 2, no. 02, pp. 200–217, 2018.
- [222] R. J. Melrose, A. M. Jimenez, H. Riskin-Jones *et al.*, "Alterations to task positive and task negative networks during executive functioning in mild cognitive impairment," *NeuroImage: Clinical*, vol. 19, pp. 970–981, 2018.
- [223] B. Haaveit, J. Jensen, D. Alnæs *et al.*, "Reduced load-dependent default mode network deactivation across executive tasks in Schizophrenia spectrum disorders," *NeuroImage: Clinical*, vol. 12, pp. 389–396, 2016.
- [224] M. W. Cole, T. Ito, D. S. Bassett, and D. H. Schultz, "Activity flow over resting-state networks shapes cognitive task activations," *Nature Neuroscience*, vol. 19, no. 12, p. 1718, 2016.

REFERENCES

- [225] C. Gratton, T. O. Laumann, A. N. Nielsen, D. J. Greene, E. M. Gordon, A. W. Gilmore, S. M. Nelson, R. S. Coalson, A. Z. Snyder, B. L. Schlaggar *et al.*, “Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation,” *Neuron*, vol. 98, no. 2, pp. 439–452, 2018.
- [226] H. Xie, V. D. Calhoun, J. Gonzalez-Castillo, E. Damaraju, R. Miller, P. A. Bandettini, and S. Mitra, “Whole-brain connectivity dynamics reflect both task-specific and individual-specific modulation: A multitask study,” *NeuroImage*, vol. 180, pp. 495–504, 2018.
- [227] E. M. Gordon, T. O. Laumann, B. Adeyemo, and S. E. Petersen, “Individual variability of the system-level organization of the human brain,” *Cerebral cortex*, vol. 27, no. 1, pp. 386–399, 2017.
- [228] T. He, R. Kong, A. J. Holmes, M. Nguyen, M. R. Sabuncu, S. B. Eickhoff, D. Bzdok, J. Feng, and B. T. Yeo, “Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics,” *NeuroImage*, vol. 206, p. 116276, 2020.
- [229] H. Jang, S. M. Plis, V. D. Calhoun, and J.-H. Lee, “Task-specific feature extraction and classification of fMRI volumes using a deep neural network initialized with a deep belief network: Evaluation using sensorimotor tasks,” *NeuroImage*, vol. 145, pp. 314–328, 2017.
- [230] H. Li and Y. Fan, “Interpretable, highly accurate brain decoding of subtly distinct brain states from functional MRI using intrinsic functional networks and long short-term memory recurrent neural networks,” *NeuroImage*, vol. 202, p. 116059, 2019.
- [231] A. Floren, B. Naylor, R. Miikkulainen, and D. Ress, “Accurately decoding visual information from fMRI data obtained in a realistic virtual environment,” *Frontiers in Human Neuroscience*, vol. 9, p. 327, 2015.
- [232] A. Shrikumar, P. Greenside, and A. Kundaje, “Learning important features through propagating activation differences,” in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR.org, 2017, pp. 3145–3153.
- [233] M. Sundararajan, A. Taly, and Q. Yan, “Axiomatic attribution for deep networks,” in *Proceedings of the 34th International Conference on Machine Learning-Volume 70*. JMLR.org, 2017, pp. 3319–3328.
- [234] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Advances in Neural Information Processing Systems*, 2017, pp. 4765–4774.
- [235] A. Elton and W. Gao, “Task-positive functional connectivity of the default mode network transcends task domain,” *Journal of cognitive neuroscience*, vol. 27, no. 12, pp. 2369–2381, 2015.
- [236] J. R. Binder, W. L. Gross, J. B. Allendorfer, L. Bonilha, J. Chapin, J. C. Edwards, T. J. Grabowski, J. T. Langfitt, D. W. Loring, M. J. Lowe *et al.*, “Mapping anterior temporal lobe language areas with fMRI: a multicenter normative study,” *NeuroImage*, vol. 54, no. 2, pp. 1465–1475, 2011.
- [237] M. W. Cole, T. Ito, D. Schultz, R. Mill, R. Chen, and C. Cocuzza, “Task activations produce spurious but systematic inflation of task functional connectivity estimates,” *NeuroImage*, vol. 189, pp. 1–18, 2019.
- [238] C. J. Brown, J. Kawahara, and G. Hamarneh, “Connectome priors in deep neural networks to predict autism,” in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, 2018, pp. 110–113.
- [239] R. L. Buckner, F. M. Krienen, A. Castellanos, J. C. Diaz, and B. T. Yeo, “The organization of the human cerebellum estimated by intrinsic functional connectivity,” *Journal of neurophysiology*, vol. 106, no. 5, pp. 2322–2345, 2011.
- [240] R. M. Birn, E. K. Molloy, R. Patriat, T. Parker, T. B. Meier, G. R. Kirk, V. A. Nair, M. E. Meyerand, and V. Prabhakaran, “The effect of scan length on the reliability of resting-state fMRI connectivity estimates,” *NeuroImage*, vol. 83, pp. 550–558, 2013.
- [241] A. C. Kelly, L. Q. Uddin, B. B. Biswal, F. X. Castellanos, and M. P. Milham, “Competition between functional brain networks mediates behavioral variability,” *NeuroImage*, vol. 39, no. 1, pp. 527–537, 2008.
- [242] W. Gao, J. H. Gilmore, S. Alcauter, and W. Lin, “The dynamic reorganization of the default-mode network during a visual classification task,” *Frontiers in Systems Neuroscience*, vol. 7, p. 34, 2013.
- [243] S. Marek and N. U. Dosenbach, “The frontoparietal network: function, electrophysiology, and importance of individual precision mapping,” *Dialogues in clinical neuroscience*, vol. 20, no. 2, p. 133, 2018.
- [244] L. Cocchi, A. Zalesky, A. Fornito, and J. B. Mattingley, “Dynamic cooperation and competition between brain systems during cognitive control,” *Trends in Cognitive Sciences*, vol. 17, no. 10, pp. 493–501, 2013.

REFERENCES

- [245] W. W. Seeley, V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, A. L. Reiss, and M. D. Greicius, "Dissociable intrinsic connectivity networks for salience processing and executive control," *Journal of Neuroscience*, vol. 27, no. 9, pp. 2349–2356, 2007.
- [246] R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter, "The brain's default network," *Annals of the New York Academy of Sciences*, vol. 1124, no. 1, pp. 1–38, 2008.
- [247] M. W. Cole, G. Repovš, and A. Anticevic, "The frontoparietal control system: a central role in mental health," *The Neuroscientist*, vol. 20, no. 6, pp. 652–664, 2014.
- [248] M. Corbetta, J. M. Kincade, J. M. Ollinger, M. P. McAvoy, and G. L. Shulman, "Voluntary orienting is dissociated from target detection in human posterior parietal cortex," *Nature Neuroscience*, vol. 3, no. 3, pp. 292–297, 2000.
- [249] A. Flinker, E. Chang, N. Barbaro, M. Berger, and R. Knight, "Sub-centimeter language organization in the human temporal lobe," *Brain and language*, vol. 117, no. 3, pp. 103–109, 2011.
- [250] G. Meschyan and A. E. Hernandez, "Impact of language proficiency and orthographic transparency on bilingual word reading: An fMRI investigation," *NeuroImage*, vol. 29, no. 4, pp. 1135–1140, 2006.
- [251] J. T. Crinion, M. A. Lambon-Ralph, E. A. Warburton, D. Howard, and R. J. Wise, "Temporal lobe regions engaged during normal speech comprehension," *Brain*, vol. 126, no. 5, pp. 1193–1201, 2003.
- [252] S. Gupta, Y. H. Chan, and J. C. Rajapakse, "Obtaining leaner deep neural networks for decoding brain functional connectome in a single shot," *Neurocomputing (Accepted, In Press)*, 2020.
- [253] G. Hinton, L. Deng, D. Yu, G. Dahl, A.-r. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, B. Kingsbury *et al.*, "Deep neural networks for acoustic modeling in speech recognition," *IEEE Signal processing magazine*, vol. 29, 2012.
- [254] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [255] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [256] A. Weigend, "On overfitting and the effective number of hidden units," in *Proceedings of the 1993 connectionist models summer school*, vol. 1, 1994, pp. 335–342.
- [257] Y. S. Abu-Mostafa, M. Magdon-Ismail, and H.-T. Lin, *Learning from data*. AMLBook New York, NY, USA:, 2012, vol. 4.
- [258] S. Anwar, K. Hwang, and W. Sung, "Structured pruning of deep convolutional neural networks," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 13, no. 3, p. 32, 2017.
- [259] J. D. Power, K. A. Barnes, A. Z. Snyder, B. L. Schlaggar, and S. E. Petersen, "Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion," *NeuroImage*, vol. 59, no. 3, pp. 2142–2154, 2012.
- [260] S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. Müller, and W. Samek, "On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation," *PloS One*, vol. 10, no. 7, p. e0130140, 2015.
- [261] M. N. Hebart and C. I. Baker, "Deconstructing multivariate decoding for the study of brain function," *NeuroImage*, vol. 180, pp. 4–18, 2018.
- [262] M. J. Rosa, L. Portugal, T. Hahn, A. J. Fallgatter, M. I. Garrido, J. Shawe-Taylor, and J. Mourao-Miranda, "Sparse network-based models for patient classification using fMRI," *NeuroImage*, vol. 105, pp. 493–506, 2015.
- [263] T. N. Rubin, O. Koyejo, K. J. Gorgolewski, M. N. Jones, R. A. Poldrack, and T. Yarkoni, "Decoding brain activity using a large-scale probabilistic functional-anatomical atlas of human cognition," *PLoS computational biology*, vol. 13, no. 10, p. e1005649, 2017.
- [264] B. Sen, S.-H. Chu, and K. K. Parhi, "Ranking regions, edges and classifying tasks in functional brain graphs by sub-graph entropy," *Scientific Reports*, vol. 9, no. 1, p. 7628, 2019.
- [265] 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*, vol. 17, pp. 16–23, 2018.
- [266] S. Koyamada, Y. Shikauchi, K. Nakae, M. Koyama, and S. Ishii, "Deep learning of fMRI big data: a novel approach to subject-transfer decoding," *arXiv preprint arXiv:1502.00093*, 2015.

REFERENCES

- [267] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.
- [268] A. Hyvärinen and E. Oja, “Independent component analysis: algorithms and applications,” *Neural networks*, vol. 13, no. 4-5, pp. 411–430, 2000.
- [269] D. E. Hinkle, W. Wiersma, and S. G. Jurs, *Applied statistics for the behavioral sciences*. Houghton Mifflin College Division, 2003, vol. 663.
- [270] R. B. Lehoucq, D. C. Sorensen, and C. Yang, *ARPACK users’ guide: solution of large-scale eigenvalue problems with implicitly restarted Arnoldi methods*. Siam, 1998, vol. 6.
- [271] K. Yan and D. Zhang, “Feature selection and analysis on correlated gas sensor data with recursive feature elimination,” *Sensors and Actuators B: Chemical*, vol. 212, pp. 353–363, 2015.
- [272] L. Tološi and T. Lengauer, “Classification with correlated features: unreliability of feature ranking and solutions,” *Bioinformatics*, vol. 27, no. 14, pp. 1986–1994, 2011.
- [273] C. F. Dormann, J. Elith, S. Bacher, C. Buchmann, G. Carl, G. Carré, J. R. G. Marquéz, B. Gruber, B. Lafourcade, P. J. Leitão *et al.*, “Collinearity: a review of methods to deal with it and a simulation study evaluating their performance,” *Ecography*, vol. 36, no. 1, pp. 27–46, 2013.
- [274] P. Bellec, C. Chu, F. Chouinard-Decorte, Y. Benhajali, D. S. Margulies, and R. C. Craddock, “The neuro bureau ADHD-200 preprocessed repository,” *NeuroImage*, vol. 144, pp. 275–286, 2017.
- [275] M. Hamilton, “The hamilton depression scale—accelerator or break on antidepressant drug discovery,” *Psychiatry*, vol. 23, pp. 56–62, 1960.
- [276] A. T. Beck and A. Beamesderfer, “Assessment of depression: the depression inventory,” in *Psychological measurements in psychopharmacology*. Karger Publishers, 1974, vol. 7, pp. 151–169.
- [277] T. Poggio, K. Kawaguchi, Q. Liao, B. Miranda, L. Rosasco, X. Boix, J. Hidary, and H. Mhaskar, “Theory of deep learning iii: explaining the non-overfitting puzzle,” *arXiv preprint arXiv:1801.00173*, 2017.
- [278] G. Varoquaux, “Cross-validation failure: small sample sizes lead to large error bars,” *NeuroImage*, vol. 180, pp. 68–77, 2018.
- [279] J. Kim, V. D. Calhoun, E. Shim, and J.-H. Lee, “Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of Schizophrenia,” *NeuroImage*, vol. 124, pp. 127–146, 2016.
- [280] D. Chan, N. C. Fox, R. I. Scahill, W. R. Crum, J. L. Whitwell, G. Leschziner, A. M. Rossor, J. M. Stevens, L. Cipolotti, and M. N. Rossor, “Patterns of temporal lobe atrophy in semantic dementia and Alzheimer’s disease,” *Annals of neurology*, vol. 49, no. 4, pp. 433–442, 2001.
- [281] A. Convit, J. De Asis, M. De Leon, C. Tarshish, S. De Santi, and H. Rusinek, “Atrophy of the medial occipitotemporal, inferior, and middle temporal gyri in non-demented elderly predict decline to Alzheimer’s disease,” *Neurobiology of aging*, vol. 21, no. 1, pp. 19–26, 2000.
- [282] S. N. Thiyagesh, T. F. Farrow, R. W. Parks, H. Accosta-Mesa, C. Young, I. D. Wilkinson, M. D. Hunter, and P. W. Woodruff, “The neural basis of visuospatial perception in Alzheimer’s disease and healthy elderly comparison subjects: an fMRI study,” *Psychiatry Research: Neuroimaging*, vol. 172, no. 2, pp. 109–116, 2009.
- [283] L. Wang, Y. Zang, Y. He, M. Liang, X. Zhang, L. Tian, T. Wu, T. Jiang, and K. Li, “Changes in hippocampal connectivity in the early stages of Alzheimer’s disease: evidence from resting state fMRI,” *NeuroImage*, vol. 31, no. 2, pp. 496–504, 2006.
- [284] Z. Wang, M. Zhang, Y. Han, H. Song, R. Guo, and K. Li, “Differentially disrupted functional connectivity of the subregions of the amygdala in Alzheimer’s disease,” *Journal of X-ray Science and Technology*, vol. 24, no. 2, pp. 329–342, 2016.
- [285] Z. Zhao, J. Lu, X. Jia, W. Chao, Y. Han, J. Jia, and K. Li, “Selective changes of resting-state brain oscillations in aMCI: an fMRI study using ALFF,” *BioMed Research International*, vol. 2014, 2014.
- [286] R. H. Kaiser, J. R. Andrews-Hanna, J. M. Spielberg, S. L. Warren, B. P. Sutton, G. A. Miller, W. Heller, and M. T. Banich, “Distracted and down: neural mechanisms of affective interference in subclinical depression,” *Social cognitive and affective neuroscience*, vol. 10, no. 5, pp. 654–663, 2014.

REFERENCES

- [287] H. R. Snyder, "Major depressive disorder is associated with broad impairments on neuropsychological measures of executive function: a meta-analysis and review." *Psychological bulletin*, vol. 139, no. 1, p. 81, 2013.
- [288] L. Kong, K. Chen, Y. Tang, F. Wu, N. Driesen, F. Womer, G. Fan, L. Ren, W. Jiang, Y. Cao *et al.*, "Functional connectivity between the amygdala and prefrontal cortex in medication-naïve individuals with major depressive disorder," *Journal of psychiatry & neuroscience: JPN*, vol. 38, no. 6, p. 417, 2013.
- [289] R. Ramasubbu, N. Konduru, F. Cortese, S. Bray, I. Gaxiola, and B. Goodyear, "Reduced intrinsic connectivity of amygdala in adults with major depressive disorder," *Frontiers in psychiatry*, vol. 5, p. 17, 2014.
- [290] J. D. Townsend, N. K. Eberhart, S. Y. Bookheimer, N. I. Eisenberger, L. C. Foland-Ross, I. A. Cook, C. A. Sugar, and L. L. Altshuler, "fmri activation in the amygdala and the orbitofrontal cortex in unmedicated subjects with major depressive disorder," *Psychiatry Research: Neuroimaging*, vol. 183, no. 3, pp. 209–217, 2010.
- [291] Z. Yao, L. Wang, Q. Lu, H. Liu, and G. Teng, "Regional homogeneity in depression and its relationship with separate depressive symptom clusters: a resting-state fMRI study," *Journal of affective disorders*, vol. 115, no. 3, pp. 430–438, 2009.
- [292] B. Zhang, M. Li, W. Qin, L. R. Demenescu, C. D. Metzger, B. Bogerts, C. Yu, and M. Walter, "Altered functional connectivity density in major depressive disorder at rest," *European archives of psychiatry and clinical neuroscience*, vol. 266, no. 3, pp. 239–248, 2016.
- [293] S.-W. Xue, D. Wang, Z. Tan, Y. Wang, Z. Lian, Y. Sun, X. Hu, X. Wang, and X. Zhou, "Disrupted brain entropy and functional connectivity patterns of thalamic subregions in major depressive disorder," *Neuropsychiatric Disease and Treatment*, vol. 15, p. 2629, 2019.
- [294] A. B. Smith, E. Taylor, M. Brammer, R. Halari, and K. Rubia, "Reduced activation in right lateral prefrontal cortex and anterior cingulate gyrus in medication-naïve adolescents with attention deficit hyperactivity disorder during time discrimination," *Journal of Child Psychology and Psychiatry*, vol. 49, no. 9, pp. 977–985, 2008.
- [295] S. J. Suskauer, D. J. Simmonds, B. S. Caffo, M. B. Denckla, J. J. Pekar, and S. H. Mostofsky, "fMRI of intrasubject variability in adhd: anomalous premotor activity with prefrontal compensation," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 47, no. 10, pp. 1141–1150, 2008.
- [296] L. Tian, T. Jiang, Y. Wang, Y. Zang, Y. He, M. Liang, M. Sui, Q. Cao, S. Hu, M. Peng *et al.*, "Altered resting-state functional connectivity patterns of anterior cingulate cortex in adolescents with attention deficit hyperactivity disorder," *Neuroscience letters*, vol. 400, no. 1-2, pp. 39–43, 2006.
- [297] D. Tomasi and N. D. Volkow, "Abnormal functional connectivity in children with attention-deficit/hyperactivity disorder," *Biological psychiatry*, vol. 71, no. 5, pp. 443–450, 2012.
- [298] K. Rubia, R. Halari, A. Christakou, and E. Taylor, "Impulsiveness as a timing disturbance: neurocognitive abnormalities in attention-deficit hyperactivity disorder during temporal processes and normalization with methylphenidate," *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1525, pp. 1919–1931, 2009.
- [299] R. K. Kana, T. A. Keller, V. L. Cherkassky, N. J. Minshew, and M. A. Just, "Atypical frontal-posterior synchronization of theory of mind regions in autism during mental state attribution," *Social neuroscience*, vol. 4, no. 2, pp. 135–152, 2009.
- [300] A. Rausch, W. Zhang, K. V. Haak, M. Mennes, E. J. Hermans, E. van Oort, G. van Wingen, C. F. Beckmann, J. K. Buitelaar, and W. B. Groen, "Altered functional connectivity of the amygdaloid input nuclei in adolescents and young adults with autism spectrum disorder: a resting state fMRI study," *Molecular autism*, vol. 7, no. 1, p. 13, 2016.
- [301] W. Sato, Y. Kubota, T. Kochiyama, S. Uono, S. Yoshimura, R. Sawada, M. Sakihama, and M. Toichi, "Increased putamen volume in adults with autism spectrum disorder," *Frontiers in human neuroscience*, vol. 8, p. 957, 2014.
- [302] G. Allen and E. Courchesne, "Differential effects of developmental cerebellar abnormality on cognitive and motor functions in the cerebellum: an fMRI study of autism," *American Journal of Psychiatry*, vol. 160, no. 2, pp. 262–273, 2003.
- [303] T. J. Akiki and C. G. Abdallah, "Determining the hierarchical architecture of the human brain using subject-level clustering of functional networks," *Scientific reports*, vol. 9, no. 1, pp. 1–15, 2019.
- [304] P. M. Gleiser and V. I. Spoor, "Modelling hierarchical structure in functional brain networks," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 368, no. 1933, pp. 5633–5644, 2010.
- [305] L. G. Jeub, O. Sporns, and S. Fortunato, "Multiresolution consensus clustering in networks," *Scientific reports*, vol. 8, no. 1, pp. 1–16, 2018.

REFERENCES

- [306] L. Zheng, T. Li, and C. Ding, "Hierarchical ensemble clustering," in *2010 IEEE International Conference on Data Mining*. IEEE, 2010, pp. 1199–1204.
- [307] E. B. Fowlkes and C. L. Mallows, "A method for comparing two hierarchical clusterings," *Journal of the American statistical association*, vol. 78, no. 383, pp. 553–569, 1983.
- [308] Y. Milaneschi, F. Lamers, W. J. Peyrot, A. Abdellaoui, G. Willemsen, J. J. Hottenga, R. Jansen, H. Mbarek, A. Dehghan, C. Lu *et al.*, "Polygenic dissection of major depression clinical heterogeneity," *Molecular psychiatry*, vol. 21, no. 4, pp. 516–522, 2016.
- [309] W. Jones and A. Klin, "Heterogeneity and homogeneity across the autism spectrum: the role of development," *Journal of the American Academy of Child & Adolescent Psychiatry*, vol. 48, no. 5, p. 471, 2009.
- [310] X. Wang, Y. Ren, and W. Zhang, "Depression disorder classification of fmri data using sparse low-rank functional brain network and graph-based features," *Computational and mathematical methods in medicine*, vol. 2017, 2017.
- [311] J. D. Rudie, J. Brown, D. Beck-Pancer, L. Hernandez, E. Dennis, P. Thompson, S. Bookheimer, and M. Dapretto, "Altered functional and structural brain network organization in autism," *NeuroImage: clinical*, vol. 2, pp. 79–94, 2013.
- [312] S. Dey, A. R. Rao, and M. Shah, "Exploiting the brain's network structure in identifying adhd subjects," *Frontiers in systems neuroscience*, vol. 6, p. 75, 2012.
- [313] H. Li and Y. Fan, "Brain decoding from functional MRI using long short-term memory recurrent neural networks," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2018, pp. 320–328.
- [314] E. P. Ganella, C. F. Bartholomeusz, C. Seguin, S. Whittle, C. Bousman, C. Phassoulotis, I. Everall, C. Pantelis, and A. Zalesky, "Functional brain networks in treatment-resistant Schizophrenia," *Schizophrenia research*, vol. 184, pp. 73–81, 2017.
- [315] M. Jenkinson, P. Bannister, M. Brady, and S. Smith, "Improved optimization for the robust and accurate linear registration and motion correction of brain images," *NeuroImage*, vol. 17, no. 2, pp. 825–841, 2002.
- [316] R. W. Cox and J. S. Hyde, "Software tools for analysis and visualization of fMRI data," *NMR in Biomedicine*, vol. 10, no. 4-5, pp. 171–178, 1997.
- [317] C. Lanczos, "Evaluation of noisy data," *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis*, vol. 1, no. 1, pp. 76–85, 1964.
- [318] A. R. Hariri, A. Tessitore, V. S. Mattay, F. Fera, and D. R. Weinberger, "The amygdala response to emotional stimuli: a comparison of faces and scenes," *NeuroImage*, vol. 17, no. 1, pp. 317–323, 2002.
- [319] M. R. Delgado, L. E. Nystrom, C. Fissell, D. Noll, and J. A. Fiez, "Tracking the hemodynamic responses to reward and punishment in the striatum," *Journal of neurophysiology*, vol. 84, no. 6, pp. 3072–3077, 2000.
- [320] R. Smith, K. Keramatian, and K. Christoff, "Localizing the rostrolateral prefrontal cortex at the individual level," *NeuroImage*, vol. 36, no. 4, pp. 1387–1396, 2007.
- [321] F. Castelli, F. Happé, U. Frith, and C. Frith, "Movement and mind: a functional imaging study of perception and interpretation of complex intentional movement patterns," *NeuroImage*, vol. 12, no. 3, pp. 314–325, 2000.
- [322] A. Drobyshevsky, S. B. Baumann, and W. Schneider, "A rapid fMRI task battery for mapping of visual, motor, cognitive, and emotional function," *NeuroImage*, vol. 31, no. 2, pp. 732–744, 2006.