

Applications of Unsupervised Clustering in Functional Magnetic Resonance Imaging

by

Xinyu Zhao

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 6, 2017

Keywords: Functional MRI, Unsupervised Learning, Clustering, Functional Connectivity

Copyright 2016 by Xinyu Zhao

Approved by

Gopikrishna Deshpande, Chair, Associate Professor, Department of Electrical and Computer
Engineering

Thomas Denney, Director, Auburn University MRI Research Center

Stan Reeves, Professor Department of Electrical and Computer Engineering

Jennifer Robinson, Assistant Professor, Department of Psychology Cognitive Neuroscience

Abstract

Functional Magnetic resonance imaging (fMRI) is a noninvasive neuroimaging technique that measures brain activity by detecting changes associated with blood flow. This technique has been widely used in radiology, biomedical research, and clinic diagnostics. Clustering is one of the most popular techniques for fMRI data analysis. The goal of clustering is to group objects in a way such that objects in the same cluster are similar to each other whereas objects in the different clusters are dissimilar. The commonly used k-means clustering and semi-supervised clustering methods require the number of clusters to be predefined, which is difficult to be determined in the majority of real data. Thus, in this dissertation, three unsupervised clustering methods were specifically chosen, which did not require a priori specification of the number of clusters. We investigated the feasibility of these methods in three different fMRI studies.

In the first study, the selected unsupervised methods were adopted on resting-state functional magnetic resonance imaging connectivity measures to investigate whether the clinical diagnostic grouping of different disorders is grounded in underlying neurobiological and phenotypic clusters. A general analysis pipeline was derived along with three supplementary analyses, i.e., site-specific analysis, outlier subject elimination, and enrichment analysis. The effectiveness of proposed methods were verified on different disorders and the results suggest that

neurobiological and phenotypic biomarkers could potentially be used as an aid by the clinician, in addition to currently available clinical diagnostic standards, to improve diagnostic precision.

In the second study, we investigated the perforant pathway between entorhinal cortex and the hippocampus during encoding task by applying the selected clustering methods on the functional connectivity between cortical layer II of entorhinal cortex and hippocampus. The result showed that the functional connectivity between EC layer II and hippocampus parcellated the hippocampus into proximal and distal regions along perforant pathway. This parcellation was based on our observation of stronger connectivity between layer II of EC with hippocampal subfields such as DG/CA4/CA3/CA2 which are proximal to the EC along the perforant pathway, compared to subfields such as CA1/Subiculum which are distal. Further, this pattern was true more for the left, rather than the right, hippocampus. Our results provide the first direct non-invasive functional evidence for the perforant pathway in humans.

In the third study, we employed selected clustering methods on functional connectivity between the hippocampus and different layers of the dorsal attention network and the default mode network to investigate HERNET (hippocampal encoding/retrieval and network) model, which proposed an encoding/retrieval dichotomy with the anterior hippocampus more connected to the dorsal attention network during memory encoding, and the posterior portions more connected to the default mode network during retrieval. Our results support some predictions of the HERNET model including anterior-posterior gradient along the long axis of the hippocampus. While preferential relationships between the entire hippocampus and DAN/DMN during encoding/retrieval, respectively, were observed as predicted, anterior-posterior specificity in

these network relationships could not be confirmed. The strength and clarity of evidence for/against the HERNET model were superior with layer-specific data compared to conventional volume data.

Acknowledgments

First and foremost, I would like to express my deepest appreciation and gratitude to my advisor Dr. Gopikrishna Deshpande, who has been a tremendous mentor to me. I would like to thank him for encouraging and guiding my research, and for enabling me to grow as a research scientist. His priceless help and advices have been greatly helpful to both my graduate research and career development.

My gratitude also goes to Dr. Jennifer Robinson, Dr. Stan Reeves, and Dr. Thomas Denney, for serving on my advisory committee, and Dr. Jeffrey Katz for being my university reader. They have provided brilliant comments and suggestions for my dissertation, making the last part of my PhD live an enjoyable journey and the defense to be a delightful moment. Also I'd like to specially thank Dr. Jennifer Robinson for sharing valuable hippocampal data for my research.

I'm very grateful to have Dr. Mark Nelms being supportive for me. His encouragement and generous help allowed me to go through the darkest time in graduate school. This work would have been impossible without him.

I would also like to thank to my family. Words cannot express how grateful I am to my parents, and parents-in-law for all of the sacrifices that you've made to enable my success. I would also like to thank all of my friends, especially Shuwei Zhang, Jue Wang, Jiayi Xu, Liang Tang, Kang

Sun, Jie Zhang, and my sweet roomies Yawei Chen and Yusheng Ding. They have been great support and encouragement for me in working towards my degree and striving towards my goals.

Last but not least, I would like to express my gratitude to the cutest person in the world, my beloved husband Song Gao. He has always been with me through the tough and sweet moments of my life. I would like to thank him for always being patient with me, being there for me in many sleepless nights, and for providing various IT support, especially fixing my C++ compiler after upgrading OS X in 2012.

Table of Contents

Abstract.....	ii
Acknowledgments	v
List of Tables	xii
List of Illustrations.....	xv
List of Abbreviations	xxi
Chapter 1 Introduction.....	1
1.1 Overview of fMRI.....	1
1.1.1 Blood-oxygen-level Dependent and Hemodynamic Response.....	1
1.1.2 Spatial and Temporal Resolution	2
1.1.3 fMRI Analysis.....	4
1.1.4 Preprocessing fMRI Data	7
1.2 Overview of Clustering.....	12
1.2.1 Clustering Definition.....	13
1.2.2 Clustering Procedure.....	15
1.2.3 Distance (Similarity) Measures.....	17

1.2.4 Clustering Algorithms.....	19
1.2.5 Determination of Number of Clusters.....	22
1.2.6 Selection of Clustering Algorithms.....	24
1.2.7 Application of Clustering on fMRI Analysis.....	27
Time Series Analysis.....	27
Functional Parcellation of fMRI Data.....	27
Brain Network Analysis.....	28
1.3 Dissertation Structure.....	29
Chapter 2 Investigating the Correspondence of Clinical Diagnostic Grouping with Underlying Neurobiological and Phenotypic Clusters Using Unsupervised Learning.....	31
2.1 Introduction.....	32
2.2 Materials and Methods.....	37
2.2.1 Participants and Non-Imaging Measures.....	37
2.2.2 Data Acquisition.....	44
2.2.3 Preprocessing.....	47
2.2.4 Connectivity Measures.....	47
2.2.5 Clustering.....	49
2.2.6 Feature Selection and Cluster Identification.....	57
2.2.7 Site-specific Analysis.....	61
2.2.8 Elimination of Outlier Subjects.....	61
2.2.9 Functional Interpretation of Selected Connectivity Features—Enrichment Analysis	64

2.3 Results	66
2.4 Discussion	74
2.4.1 Connectivity Features Important for Clustering	75
2.4.2 Phenotypic Features Important for Clustering	90
2.4.3 Site-specific Analysis	93
2.4.4 Connectivity-based Reassignment of Diagnostic Labels	93
2.4.5 Outlier Subject Elimination	94
2.5 Conclusion	95
Chapter 3 Noninvasive Characterization of Layer-specific Perforant Pathway between Entorhinal Cortex and Hippocampus using Ultra High Field fMRI at 7T	97
3.1 Introduction	97
3.2 Materials and Methods	104
3.2.1 Data Acquisition	104
3.2.2 Preprocessing	105
3.2.3 Connectivity Measures	108
3.2.4 Layer Specific Clustering	108
3.2.5 Feature Selection and Cluster Identification	116
3.2.6. Volume Level Clustering	119
3.2.7. Validation of Identified Clusters	119
3.3 Results	121
3.4 Discussion	129
3.4.1 Layer-specific functional perforant pathway from EC to the hippocampus	130
3.4.2 Functional Hemispheric Specialization	131

3.4.3 High-resolution Functional Imaging	132
3.4.4 Limitations and Future Works	133
3.5 Conclusion.....	134
Chapter 4 Functional Parcellation of the Hippocampus based on its Connectivity with Default Mode and Dorsal Attention Networks.....	136
4.1 Introduction	137
4.2 Materials and Methods.....	143
4.2.1 Data Acquisition.....	143
4.2.2 Preprocessing	144
4.2.3 Connectivity Measures.....	146
4.2.4 Layer Specific Clustering.....	147
4.2.5 Feature Selection and Cluster Identification	157
4.2.6. Volume Level Clustering	159
4.2.7 Comparison with Anatomical Anterior-Posterior Segmentation	160
4.3 Results	162
4.4 Discussion	169
4.4.1 Anterior-Posterior Functional Differentiation of the Hippocampus	170
4.4.2 Layer-specific Connectivity Between the DAN/DMN and Hippocampus	171
4.4.3 Layer-specific Functional Pathway Between the DAN/DMN and Hippocampus..	172
4.4.4 High-resolution Functional Imaging	173
4.4.5 Long-axis Differentiation of the Hippocampus	174
4.5 Limitation and Future Works	175
4.6 Conclusion.....	177

Chapter 5 Conclusion and Future Work	179
Bibliography	183
Appendix A	234
Appendix B	238

List of Tables

Table 1.1 Comparison of different clustering algorithms. Parameters: k = number of clusters, N = sample size, d = feature dimensions, t = number of iterations. User-specified input parameters include: 1. number of clusters, 2. density threshold, 3. radius distance, 4. stop criterion, 5. threshold cutoff, 6. distribution parameters, 7. training data as background knowledge, 8. degree, 9. matrix inflation, 10. minimum distance with higher density, 11. local density.....	26
Table 2.1 Phenotypic variables selected by GA with different clustering methods (ADHD).....	39
Table 2.2 Phenotypic variables selected by GA with different clustering methods (AD).	40
Table 2.3 Phenotypic variables selected by GA with different clustering methods (ASD).	42
Table 2.4 Phenotypic variables selected by GA with different clustering methods (PTSD).	46
Table 2.5 Estimated optimal values of each input parameter in clustering for Clinical vs. Connectivity comparison.	65
Table 2.6 Estimated optimal values of each input parameter in clustering for Phenotypic vs. Connectivity comparison.	66
Table 2.7 Peak similarity (highlighted), corresponding number of features, and number of clusters obtained using SFR, SFS, and GA with different clustering methods for ADHD dataset.	69
Table 2.8 Peak similarity (highlighted), corresponding number of features, and number of clusters obtained using SFR, SFS, and GA with different clustering methods for AD dataset.....	69
Table 2.9 Peak similarity (highlighted), corresponding number of features, and number of clusters obtained using SFR, SFS, and GA with different clustering methods for ASD dataset.....	69

Table 2.10 Peak similarity (highlighted), corresponding number of features, and number of clusters obtained using SFR, SFS, and GA with different clustering methods for PTSD dataset.....	70
Table 2.11 Number of subjects provided by each site in the ASD sample.	70
Table 2.12 Number of subjects provided by each site in the ADHD sample.....	71
Table 2.13 Similarity achieved using data from individual sites and for the whole dataset using features commonly selected by NYU and Peking.	71
Table 2.14 Comparison of peak similarity obtained with and without elimination of outlier subjects. p: number of outliers, k: number of clusters, sim: clustering similarity. In AD dataset, the number of clusters identified by hierarchical clustering was changed from 5 to 4 with outlier elimination (highlighted), which matched with the grouping obtained using clinical diagnosis.	74
Table 2.15 Network-to-network interactions selected by enrichment analysis for ADHD dataset.	81
Table 2.16 Network-to-network interactions selected by enrichment analysis for AD dataset. ...	83
Table 2.17 Network-to-network interactions selected by enrichment analysis for ASD dataset.	86
Table 2.18 Network-to-network interactions selected by enrichment analysis for PTSD dataset.....	89
Table 3.1 Estimated optimal values of each input parameter in clustering.	122
Table 3.2 Cluster similarity between different clustering methods for layer-specific analysis. For the same input data, the similarity of clusters obtained using different clustering methods were measured using Torres' method [188].	123
Table 3.3 Quantitative characterization of the overlap between functional clusters (obtained from layer-specific analysis) and anatomical subfields. The percentage of overlap between each pair of corresponding clusters, the mean correlation and corresponding p-values obtained within each cluster using different clustering methods are shown. Functional cluster I corresponds to anatomical subfields CA2/CA3/CA4/DG and functional cluster II corresponds to anatomical subfields CA1/subiculum.....	126
Table 3.4 Cluster similarity between different clustering methods for volume level analysis. For the same input data, the similarity of clusters obtained using different clustering methods was measured using Torres' method [188].	127
Table 3.5 Quantitative characterization of the overlap between functional clusters obtained from volume-level analysis and anatomical subfields. The percentage of overlap between each pair of corresponding clusters, and the mean correlation obtained within each cluster using different clustering methods are shown. Functional cluster I corresponds	

to anatomical subfields CA2/CA3/CA4/DG and functional cluster II corresponds to anatomical subfields CA1/subiculum.	129
Table 4.1 Estimated optimal values of each input parameter in layer-specific clustering. <i>h</i> : cutting height, and <i>s</i> : reachability threshold.	163
Table 4.2 Estimated optimal values of each input parameter in volume-level clustering. <i>h</i> : cutting height, and <i>s</i> : reachability threshold.	163
Table 4.3 Cluster similarity between functional and anatomical anterior-posterior parcellations using different clustering methods and mean correlation obtained within each cluster on the left side of the hippocampus.	165
Table 4.4 Comparison of the absolute correlations obtained between the DAN with the hippocampus and the DMN with the hippocampus during the encoding task by conducting one-tailed two-sample t-test (H_0 : correlation of DAN and the hippocampus \leq correlation of the DMN and the hippocampus).	166
Table 4.5 Comparison of the absolute correlations obtained between the DMN with the hippocampus and the DAN with the hippocampus during the retrieval task by conducting one-tailed two-sample t-test (H_0 : correlation of the DMN with the hippocampus \leq correlation of the DAN and the hippocampus).	167

List of Illustrations

Figure 1.1 An example of the hemodynamic response.	3
Figure 1.2 Processing stream of fMRI analysis using different software packages.....	8
Figure 1.3 A depiction of interleaved MRI acquisition. The slices are acquired in the order 1-7-2-8-3-9-4-10-5-11-6.	9
Figure 1.4 A depiction of slice timing correction. The blue line corresponds to the original time series from a single voxel in the slice acquired at the beginning of each volume acquisition. The red line reflects the interpolated timecourse that would be obtained to correct this slice to match the center slice.....	10
Figure 1.5 An example of filters applied to fMRI image. The low-pass filter blurs the image whereas a high-pass filter enhances the edges in the image.	12
Figure 1.6 The effect of spatial smoothing on fMRI image with different width of Gaussian kernel.....	14
Figure 1.7 An example of clustering. Given some data points (left), the target of clustering is to group them into different clusters (right). In this case, three clusters and noise are marked using different colors.	15
Figure 1.8 An example of determination of number of clusters using elbow criterion. The elbow is indicated by the red circle, which means in this case the optimal number of clusters is 5.	24
Figure 2.1 Illustration of proposed pipeline for identifying different brain-based disorders using unsupervised clustering methods. The main pipeline is depicted in purple color along with two supplementary analyses, site-specific analysis in salmon color and outlier subject elimination in green color.	38
Figure 2.2 Dendrogram derived from hierarchical clustering. The final clustering result is obtained by cutting the tree at defined level. The two clusters are shown by different colors (red and cyan).....	50

Figure 2.3 Illustration of OPTICS. (a) Original simulated dataset, (b) reachability plot obtained from OPTICS, and (c) clustering result. Each cluster corresponds to one valley in the reachability plot.....52

Figure 2.4 Illustration of DPC method. Plot of δ as a function of ρ for each object. Objects with larger ρ and δ are cluster centers and objects with smaller ρ , and larger δ are outliers.54

Figure 2.5 Flowchart of GA for feature selection. In the M_0 -by- N_0 matrix, each row represents a candidate solution, describing a subset of selected features. Each of the N_0 bits in a row represents whether a feature is selected (1) or discarded (0).59

Figure 2.6 Illustration of outlier subject elimination process (using upper triangular connected pairs). x_1, x_2, \dots, x_6 are six representative subjects used in the illustration.62

Figure 2.7 Similarity between connectivity and phenotypic variables obtained from different iterations using hierarchical clustering and different feature selection methods: (a) SFR, (b) SFS, and (c) GA.68

Figure 2.8 Statistical significance (p-value) of selected phenotypic/genotypic variables with both the old clinical diagnostic grouping and the new connectivity grouping. The results are shown here for the ADHD dataset. Logarithmic scale is used for the y-axis of p-values.72

Figure 2.9 Statistical significance (p-value) of selected phenotypic/genotypic variables with both the traditional clinical diagnostic grouping and the new connectivity-based grouping. The results are shown here for the AD dataset. Logarithmic scale is used for the y-axis of p-values.73

Figure 2.10 SFC, vDFC, SEC, and vDEC features selected by GA and hierarchical (ADHD). Selected features were split into two groups, i.e., 1) control > disease (ADHD-C and ADHD-I) and 2) disease > control. DMN: Default mode network, VN: Visual network, BGN: Basal ganglia network, SMN: Sensory motor network, SCAN: Semantic cognition and attention network.....80

Figure 2.11 SFC and vDFC features selected by GA and hierarchical (AD). Selected features were split into two groups, i.e., 1) control > disease (AD, LMCI and EMCI) and 2) disease > control. DMN: Default mode network, VN: Visual network, BGN: Basal ganglia network, SMN: Sensory motor network, SCAN: Semantic cognition and attention network.....83

Figure 2.12 SFC, vDFC, SEC, and vDEC features selected by GA and hierarchical (ASD). Selected features were split into two groups, i.e., 1) control > disease (autism and asperger's) and 2) disease > control. DMN: Default mode network, VN: Visual network, BGN: Basal ganglia network, SMN: Sensory motor network, SCAN: Semantic cognition and attention network.....85

Figure 2.13 SFC, vDFC, SEC, and vDEC features selected by GA and hierarchical (PTSD/PCS). Selected features were split into two groups, i.e., 1) control > disease (PTSD and PTSD/PCS) and 2) disease > control. DMN: Default mode network, VN: Visual network, BGN: Basal ganglia network, SMN: Sensory motor network, SCAN: Semantic cognition and attention network.	89
Figure 3.1 Perforant pathway from the entorhinal cortex and the hippocampus.	99
Figure 3.2 Illustration of proposed analysis pipeline for investigating the perforant pathway between layer II of EC and subfields of the hippocampus.	103
Figure 3.3 Hippocampus and EC ROIs used in this study.	105
Figure 3.4 Cortical layer reconstruction with FreeSurfer. The surface reconstructions, i.e, the white-gray interface (black) and gray-pial interface (orange), and cortical layers reconstruction. Six cortical layers were reconstructed within the cortical gray matter at fixed relative distances between the white and pial surfaces determined from the cortical gray matter.	107
Figure 3.5 Illustration of hierarchical clustering. a) Original simulated dataset, b) Dendrogram derived from hierarchical clustering, and c) Clustering results obtained with a specific cutting height. Two clusters that were identified are marked with different colors. ..	109
Figure 3.6 Illustration of OPTICS clustering. a) Original simulated dataset, b) reachability plot obtained from OPTICS, and c) clustering results. Two clusters were identified corresponding to valleys in the reachability plot.	112
Figure 3.7 Illustration of DPC clustering. a) Original simulated dataset, b) Plot of δ as a function of ρ for each object. Objects with larger ρ and δ are cluster centers and objects with smaller ρ , and larger δ are outliers. c) Clustering results. Two clusters were identified corresponding to two cluster centers in the decision graph.	114
Figure 3.8 Flowchart of GA for feature selection. In the m-by-d matrix, each row represents a candidate solution, describing a subset of selected features. Each of the d bits in a row represents whether a feature is selected (1) or discarded (0).	117
Figure 3.9 3D visualization of the entire hippocampal atlas in sagittal view. Each subfield is represented by a different color.	120
Figure 3.10 Anatomical parcellation of the hippocampus used in this study. (Coordinates are in MNI space).....	122
Figure 3.11 Clusters of hippocampal voxels determined (using the DPC method) based on their functional connectivity with layer II of the entorhinal cortex during the encoding task. (Coordinates are in MNI space).....	124

Figure 3.12	Overlap between anatomical subfields of the hippocampus and functional parcels obtained from clustering (using DPC method) of hippocampal voxels based on their functional connectivity with layer II of EC during an encoding task.	125
Figure 3.13	Clusters of hippocampal voxels determined (using DPC method) based on their functional connectivity with voxels in the entorhinal cortex volume during the encoding task. (Coordinates are in MNI space).....	128
Figure 3.14	Overlap between anatomical subfields of the hippocampus and functional parcels obtained from clustering (using DPC method) of hippocampal voxels based on their functional connectivity with voxels in EC volume during an encoding task.....	128
Figure 4.1	Illustration of our hypotheses based on predictions from the HERNET model as well as known anatomical pathway between different layers of DAN/DMN and hippocampus: a) anatomical pathway between hippocampus and different layers of DAN/DMN during encoding/retrieval processes, b) illustration of proposed first hypothesis, c) illustration of proposed second hypothesis, and d) illustration of proposed third hypothesis.	141
Figure 4.2	Illustration of proposed analysis pipeline for investigating hippocampal parcellation based on its layer-specific connectivity with DAN/DMN ROIs. The same process was repeated for encoding and retrieval tasks, separately, as well as with conventional volume data (as opposed to layer-specific data).	142
Figure 4.3	Hippocampal ROI used in this study. Only the left hippocampus was considered in this study.	144
Figure 4.4	Cortical layer reconstruction with FreeSurfer. Six cortical layers were reconstructed within the cortical gray matter at fixed relative distances between the white and pial surfaces.....	147
Figure 4.5	Illustration of hierarchical clustering. a) Original simulated dataset, b) Dendrogram derived from hierarchical clustering, and c) Clustering results obtained with a specific cutting height. Two clusters that were identified are marked with different colors. ...	149
Figure 4.6	Illustration of OPTICS clustering. a) Original simulated dataset, b) reachability plot obtained from OPTICS, and c) clustering results. Two clusters were identified corresponding to valleys in the reachability plot.	150
Figure 4.7	Illustration of DPC clustering. a) Original simulated dataset, b) Plot of δ as a function of ρ for each object. Objects with larger ρ and δ are cluster centers and objects with smaller ρ , and larger δ are outliers. c) Clustering results. Two clusters were identified corresponding to two cluster centers in the decision graph.	154
Figure 4.8	Illustration of dependency of clustering results on input parameters. Take hierarchical clustering as an example. With a relatively high cutting height (a), two clusters (red and blue) were identified (b). As cutting height was reduced (c), one big cluster (red) was separated into two smaller clusters (d).	156

Figure 4.9 Flowchart of GA for feature selection. In the M-by-d matrix, each row represents a candidate solution, describing a subset of selected features. Each of the d bits in a row represents whether a feature is selected (1) or discarded (0).	158
Figure 4.10 Anatomical anterior-posterior segmentation used in this study. (Coordinates are in MNI space).....	161
Figure 4.11 Clusters of hippocampal voxels determined (using the DPC method) based on their functional connectivity with 1) layer V of the DMN/DAN during the encoding task, and 2) layer II of the DMN/DAN during the retrieval task. (Coordinates are in MNI space)	164
Figure 4.12 Clusters of hippocampal voxels determined (using the DPC method) based on their functional connectivity DMN/DAN volume during encoding and retrieval tasks. (Coordinates are in MNI space).....	168
Figure A.1 Clusters of hippocampal voxels determined (using the hierarchical clustering method) based on their functional connectivity with layer II of the entorhinal cortex during the encoding task. (Coordinates are in MNI space).....	234
Figure A.2 Clusters of hippocampal voxels determined (using the OPTICS method) based on their functional connectivity with layer II of the entorhinal cortex during the encoding task. (Coordinates are in MNI space).....	235
Figure A.3 Clusters of hippocampal voxels determined (using the hierarchical clustering method) based on their functional connectivity with voxels in the entorhinal cortex volume during the encoding task. (Coordinates are in MNI space).....	236
Figure A.4 Clusters of hippocampal voxels determined (using the hierarchical clustering method) based on their functional connectivity with voxels in the entorhinal cortex volume during the encoding task. (Coordinates are in MNI space).....	237
Figure B.1 Clusters of hippocampal voxels determined (using the hierarchical method) based on their functional connectivity with 1) layer V of the DMN/DAN during the encoding task, and 2) layer II of the DMN/DAN during the retrieval task. (Coordinates are in MNI space).....	238
Figure B.2 Clusters of hippocampal voxels determined (using the OPTICS method) based on their functional connectivity with 1) layer V of the DMN/DAN during the encoding task, and 2) layer II of the DMN/DAN during the retrieval task. (Coordinates are in MNI space).....	239
Figure B.3 Clusters of hippocampal voxels determined (using the DPC method) based on their functional connectivity DMN/DAN volume during encoding and retrieval tasks. (Coordinates are in MNI space).....	240

Figure B.4 Clusters of hippocampal voxels determined (using the OPTICS method) based on their functional connectivity DMN/DAN volume during encoding and retrieval tasks. (Coordinates are in MNI space).....241

List of Abbreviations

fMRI	Functional Magnetic Resonance Imaging
PET	Positron Emission Tomography
BOLD	Blood-Oxygen-Level Dependent
HDR	Hemodynamic Response
ROIs	Region of Interests
SPM	Statistical Parametric Mapping
FSL	FMRIB Software Library
AFNI	Analysis of Functional NeuroImages
ADHD	Attention Deficit Hyperactivity Disorder
AD	Alzheimer's Disease
ASD	Autism Spectrum Disorder
PTSD	Post-Traumatic Stress Disorder
PCS	Post-Concussion Syndrome
ADHD	FMRIB Software Library
OPTICS	Ordering Points to Identify the Clustering Structure
DPC	Density Peak Clustering
SFS	Sequential Forward Searching
GA	Genetic Algorithm

SFR	Sequential Feature Ranking
SFC	Statistic Functional Connectivity
vDFC	Variance of Dynamic Functional Connectivity
SEC	Statistic Effective Connectivity
vDEC	Variance of Dynamic Effective Connectivity
FC	Functional Connectivity
EC	Effective Connectivity
CH	Calinski-Harabasz
ICNs	Intrinsic Connectivity Networks
DMN	Default Mode Network
VN	Visual Network
BGN	Basal Ganglia Network
SMN	Sensory Motor Network
SCAN	Semantic Cognition and Attention
PCC	Posterior Cingulate Cortex
mPFC	Medial Prefrontal Cortex
MTL	Medial Temporal Lobe
ACC	Anterior Cingulate Cortex
IPL	Inferior Parietal Lobe
MOG	Medial Orbital Gyrus
ITG	Inferior Temporal Gyrus
SMA	Sensory Motor Cortex
MTG	Middle Temporal Gyrus

IFG	Inferior Frontal Gyrus
STG	Supperior Temporal Gyrus
DAN	Dorsal Attention Network
VAN	Ventral Attention Network
N2N	Network-To-Network
EC	Entorhinal Cortex
CA	Cornu Ammonis
DG	Dentate Gyrus
SR	Stratum Radiatum
SL	Stratum Lacunosum
SM	Stratum Moleculare
SNR	Signal To Noise Ratio
HERNET	Hippocampal Encoding/Retrieval and network
HIPER	Hippocampal Encoding/Retrieval
FEF	Frontal Eye Field
SPL	Superior Parietal Lobe

Chapter 1

Introduction

1.1 Overview of fMRI

Functional magnetic resonance imaging (fMRI) is a noninvasive neuroimaging technique that measures brain activity by detecting changes associated with blood flow [1] [2]. This technology has been widely used in radiology [3], biomedical research [4] [5], and clinic diagnostics [6] [7]. In the early 1990s, fMRI was first introduced to the world [1] and has rapidly become one of the most valuable tools in brain mapping research. The reason for this explosion is that fMRI provides an unprecedented ability to safely and noninvasively detect brain activity, e.g., it does not require people to undergo shots, surgery, or to ingest substances, or to be exposed to ionizing radiation, etc. Also, it provides superb spatial resolution and relatively good temporal resolution compared to previous methods such as positron emission tomography (PET).

1.1.1 Blood-oxygen-level Dependent and Hemodynamic Response

The primary form of fMRI uses the blood-oxygen-level dependent (BOLD) contrast, discovered by Seiji Ogawa [2]. It works by detecting the changes in blood oxygenation and flow that occur in response to neural activity [8].

Although brain does not store glucose, it is considered as one of the most critical sources of brain's energy. When neurons become active getting them back to their original state requires energy, which is produced from glucose, to actively pumping ions back and forth across the neuronal cell membranes. More blood flows in to transport more glucose. It also brings in more

oxygen in the form of oxygenated hemoglobin (Hb) molecules in red blood cells. This is from both a higher rate of blood flow and an expansion of blood vessels. The blood-flow change is localized to within 2 or 3 mm of where the neural activity is. Usually, the brought-in oxygen is more than the oxygen consumed in burning glucose, and this causes a decrease in deoxygenated hemoglobin (dHb). Since dHb is more magnetic than Hb, this difference leads to an improved magnetic resonance (MR) signal, which can be used to determine which neurons are active at a time. To summarize, neuronal activity and cerebral blood flow are coupled in such a way that when neuron fires, it consumes more oxygen and to meet this demand blood flow increases to the active neuron. Since the amount of blood that is sent to the neuron is more than the amount that is required, it leads to a relative surplus in local blood oxygen. The MR signal measured in fMRI depends on this change in oxygenation which is referred to as BOLD signal [2].

The changes in MR signal from the neuronal activity is called the hemodynamic response (HDR), which is the increase in blood flow that follows a brief period of neuronal activity [2]. Figure 1.1 depicts an example of HDR, which shows two obvious behaviors. First, there is a latency (about 1-2 sec) between HDR and neuronal events, since it takes that long for the vascular system to respond to the brain's need for glucose. Second, after rises to a peak (about 5 sec), it drops down below the baseline (which is called undershoot) and last for 15-20 sec to return to the baseline. There is some evidence that continuous metabolic requirements in a brain region contribute to this undershoot.

1.1.2 Spatial and Temporal Resolution

The spatial resolution of an fMRI image refers to how well it discriminates between nearby locations. It is measured by the size of voxels, which is a three-dimensional rectangular cuboid,

whose dimensions are determined by several factors, e.g., the slice thickness, the area of a slice, etc. The spatial resolution usually ranges from 1 to 5mm. A full brain scan uses larger voxels while a region of interest (ROI) study within the brain typically uses smaller voxels. Novel techniques have been developed to improve the spatial resolution. Ultra-high-resolution MRI or MR spectroscopy works at a resolution of tens of micrometers [9]. It uses 7T fields, small-bore scanners that can fit small animals and external contrast agents. Parallel imaging is another technique to improve spatial resolution [10]. It uses multiple coils for excitation and reception. Spatial resolution improves as the square root of the number of coils used.

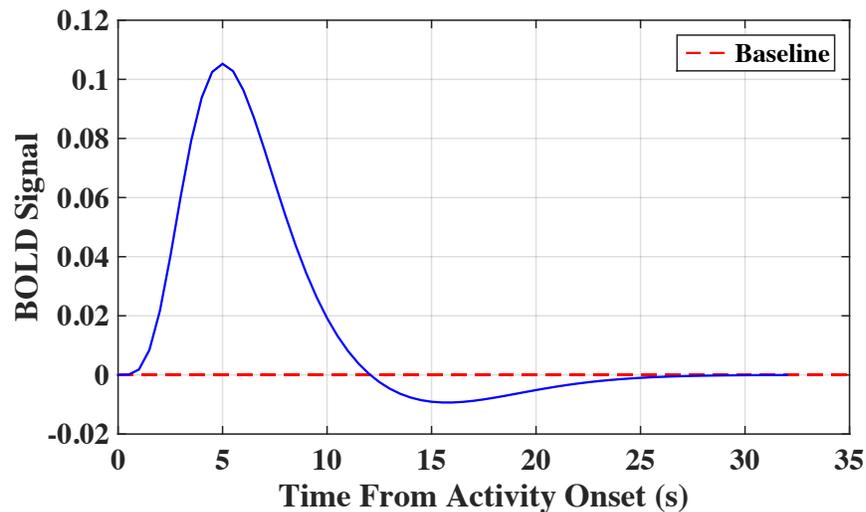


Figure 1.1 An example of the hemodynamic response.

Temporal resolution is the smallest time period of neural activity reliably separated out by fMRI. One element deciding this is the sampling time (TR). For every TR, a 3D volume image is generated by the scanner. All 3D images generated during the scan can be put together to form a 4D volume image. Nowadays, the temporal resolution can reach to as small as 200ms. The temporal resolution of fMRI is limited by multiple factors, e.g., recovering time for net magnetization, a long latency for blood-flow operation, etc. Many techniques have been

proposed to address these issues. One method is to use multiple coils to speed up acquisition time [11]. Another technique is to decide which parts of the signal matter less and drop those [12].

1.1.3 fMRI Analysis

The analysis of fMRI data is made complex by a number of factors. First, the data are affected by many artifacts, e.g., head movement, respiration, etc. Second, there are a variety of sources of variability in the data, including variability between individuals and variability across time within individuals. Third, the dimensionality of the data is very large, which causes a number of challenges in comparison to the small datasets that many scientists are accustomed to working with.

The main components of fMRI analysis are described below.

- Quality control: ensure the data is not corrupted by artifacts.
- Distortion correction: the correction of spatial distortions.
- Motion correction: the realignment of scans across time to correct head motion.
- Slice timing correction: the correction of differences in timing across different slices in the image.
- Spatial normalization: the alignment of data from individuals into a common spatial framework so that they can be compared in a group analysis.
- Temporal filtering: the filtering of the data in time to remove high-frequency noise.
- Statistical inference: the estimation of statistical significance of the results, correcting for a large number of statistical tests performed on the brain.

- Clustering and classification: the partition of the brain or a group of subjects based on homogeneous and heterogeneous characteristics.
- Visualization: Visualization of the results and estimation of effect sizes.

In the early days of fMRI, there is no open source fMRI software, which can be used to analyze fMRI data systematically. Nowadays, there are many comprehensive software packages for fMRI analysis, each of which has its pros and cons.

Statistical Parametric Mapping (SPM) was the first widely used and openly distributed software package for fMRI analysis, which was developed by Karl Friston and colleagues in the lab at University College London [13]. SPM started in the early 1990s as a program for analysis of PET data and was then adapted in the mid-1990s for analysis of fMRI data. It remains the most popular software package for fMRI analysis. SPM was built in MATLAB, which makes it accessible for a very broad range of computer platforms. In addition, MATLAB code is relatively readable, which makes it easy to check the code and understand thoroughly what has been done by the programs. Even if one does not use SPM as a primary analysis package, many of the MATLAB functions in the SPM package are useful for processing data, reading and writing data files, etc. SPM is also extensible through its toolbox functionality, and a large number of extensions are available via the SPM Web site. One unique feature of SPM is its connectivity modeling tools, including psychophysiological interaction and dynamic causal modeling. The visualization tools available with SPM are relatively limited, and many users take advantage of other packages for visualization.

FMRIB Software Library (FSL) is one of the most popular software package in recent years, which was developed by Stephen Smith and colleagues at Oxford University in 2000 [14]. It has

several desired characteristics. First, FSL has been at the front edge of statistical modeling for fMRI data, developing and implementing a number of novel modeling, estimation, and inference techniques that are implemented in their FEAT, FLAME, and RANDOMISE modules. Second, FSL includes a robust toolbox for independent components analysis (ICA), which has become very popular both for artifact detection and for modeling of resting-state fMRI data. Third, FSL includes a sophisticated set of tools for analysis of diffusion tensor imaging data, which is used to analyze the structure of white matter. FSL includes an increasingly powerful visualization tool called FSLView, which includes the ability to overlay a number of probabilistic atlases and to view time series as a movie. Another major advantage of FSL is its integration with grid computing, which allows for the use of computing clusters to greatly speed the analysis of very large datasets.

Analysis of Functional NeuroImages (AFNI) was created by Robert Cox and his colleagues, first at the Medical College of Wisconsin and then at the National Institutes of Mental Health [15]. Its primary strength is in its very powerful and flexible visualization abilities, including the capacity to integrate visualization of volumes and cortical surfaces using the SUMA toolbox. AFNI's statistical modeling and inference tools have historically been less sophisticated than those available in SPM and FSL. However, recent work has integrated AFNI with the R statistical package, which allows a use of more sophisticated modeling techniques available within R.

Although there are a variety of software packages that can be adopted for fMRI analysis, the overall processing stream looks similar. Figure 1.2 depicts the general processing stream for SPM, FSL and AFNI in fMRI data analysis. From the figure, it can be seen that SPM spatial

normalization is applied prior to statistical analysis, whereas in FSL and AFNI it is applied to the output of the statistical analysis. However, the major procedures are the same between these packages.

1.1.4 Preprocessing fMRI Data

The 4D volume fMRI data acquired from the scanner cannot be used directly for brain functionality analysis. A preprocessing pipeline needs to be applied to detect and repair potential artifacts in the data, or prepare the data for later processing. Although preprocessing of fMRI data varies between using different software packages, there is a standard set of methods to be chosen from.

Distortion Correction

Due to the inhomogeneity of the main magnetic field caused by the air-tissue interfaces, the acquisition of fMRI suffers from two forms of artifacts, which are referred to as dropout and geometric distortion. Dropout is seen as reduced signal in the brain areas adjacent to air-tissue interfaces. Once the data has been acquired, there is no way to recover data from dropout area. Thus, it is necessary to develop a method to reduce dropout. In addition to the signal loss, fMRI images can also be spatially distorted in the same regions. When gradients are applied to encode spatial information in the MRI image, the inhomogeneity in the magnetic field results in errors in the location of structures in the resulting images. These distortions make it difficult to align fMRI data with structural images.

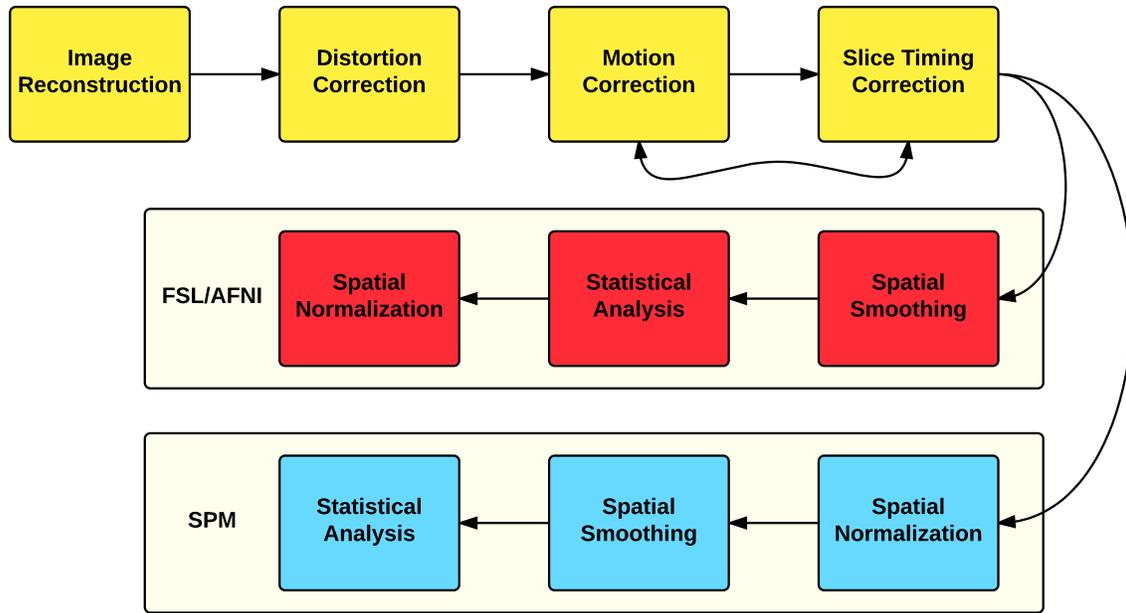


Figure 1.2 Processing stream of fMRI analysis using different software packages.

One of the most effective ways to correct inhomogeneity of magnetic field is to apply a field map [16]. The field map can be computed by comparing the phase difference between two fMRI images obtained at two echo times. By inverting this map, the original location of the data in each voxel can be easily determined. Pass-band balanced-steady-state free procession (b-SSFP) fMRI is another developed method to minimize dropout and spatial distortion [17]. It utilizes rapid radio frequency excitation pulses combined with fully balanced gradient pulses during each excitation repetition interval.

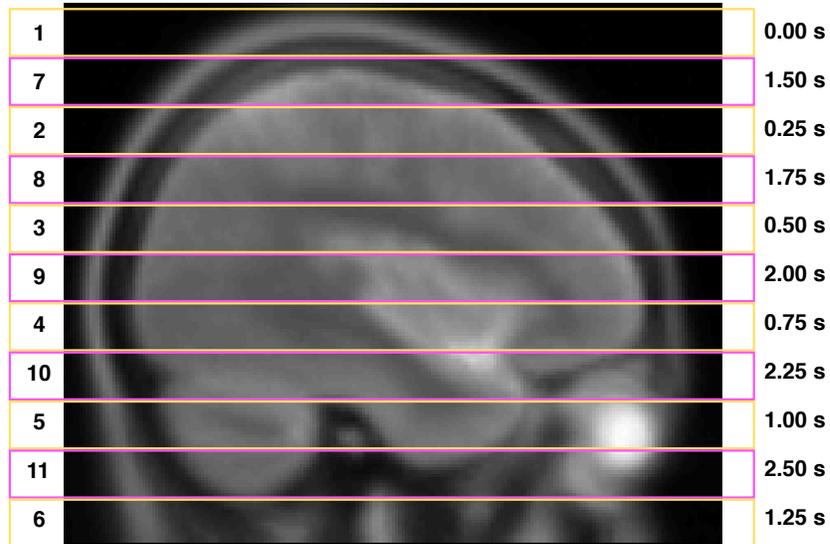


Figure 1.3 A depiction of interleaved MRI acquisition. The slices are acquired in the order 1-7-2-8-3-9-4-10-5-11-6.

Slice Timing Correction

Most of the fMRI data are collected using two-dimensional MRI acquisition, in which the data are acquired one slice at a time either in ascending/descending order or by acquiring the odd slices followed by the even slices (referred to as interleaved acquisition; see Figure 1.3). Thus, data in different part of the image are acquired at different times. These differences in the acquisition time of different voxels can cause huge problems for fMRI analysis. For example, using these data to create a statistical model can lead to a mismatch between the model and the data that varies across the brain. To avoid this issue, slice timing correction was developed [18] by interpolating the data in all slices to match the timing of a reference slice (Figure 1.4). Different data interpolation methods have been proposed for slice timing correction including linear, sinc and cubic spline interpolation.

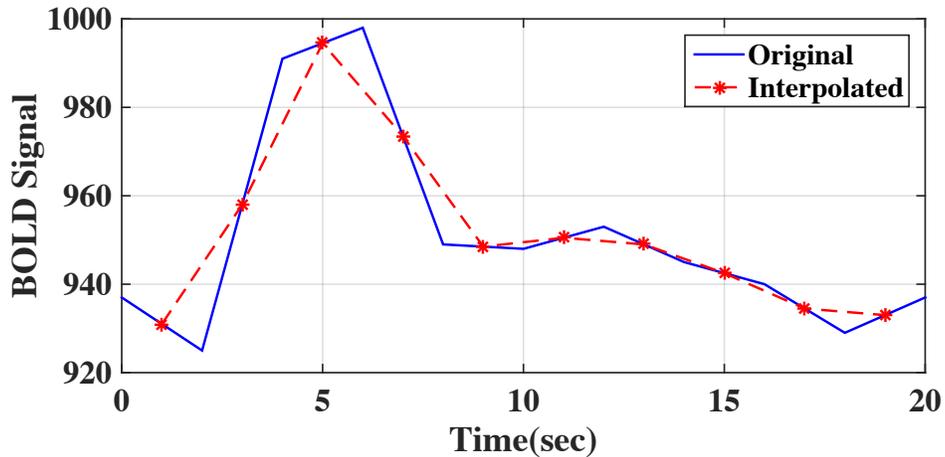


Figure 1.4 A depiction of slice timing correction. The blue line corresponds to the original time series from a single voxel in the slice acquired at the beginning of each volume acquisition. The red line reflects the interpolated timecourse that would be obtained to correct this slice to match the center slice.

Motion Correction

Motion correction of fMRI data is a commonly used procedure prior to data analysis. There are two major effects of head motion: it leads to a mismatch of the location of subsequent images obtained at different times, and it can result in disruption of the MRI signal. These effects can be corrected by using a rigid-body registration, i.e., translations in three directions and rotations in three directions. Each image in the fMRI time series is aligned to a reference image, and the images are then resliced to create realigned versions of the original data. A lot of registration algorithms have been developed, e.g., Roger Woods' automated image registration (AIR) method [19] [20], Jiang's Prospective Acquisition CorrEction (PACE) method [21], Ostuni's 3D sinc-interpolation method [22], etc.

Coregistration and Normalization

Most fMRI studies try to reach a generic conclusion that is applicable to not only the study itself, but also other studies. This requires coregistration and normalization to integrated fMRI images across individuals. Coregistration refers to the alignment between functional and structural images from the same subject, which is often done by mapping functional images into anatomical spaces. On the other hand, normalization, which refers to the alignment of image to a standard template, is to eliminate the variability of brain shape from different. Coregistration and normalization differ in the transformation algorithm. Coregistration usually adopts rigid-body transformation with six parameters to register images, whereas normalization uses a nine-parameter transformation with another three additional parameters accounting for scaling differences on x , y , and z axis. Although the algorithms employed in these two procedures may differ, the basic idea is similar, i.e., defining a cost function with the goal to minimize the differences on image parameters among images. The cost function can be the mutual information, entropy correlation coefficient, or sum of the squared difference plus membrane energy of the deformation field. Different templates can be selected as a target for individual images alignment. One of the commonly used templates for normalization is the Montreal Neurological Institute (MNI) template, which was developed to provide an MRI-based template that would allow automated registration rather than landmark-based registration [23].

Temporal Filtering

Temporal filtering aims to remove or attenuate noise and artifacts within the raw signal, which may come from thermal noise, physiological noise from heart beat and breathing, or magnetic field shifting, etc. The idea is to decide which frequencies band are noise so that they can be

removed by using a filter (low-pass, high-pass, or band-pass; see figure 1.5) corresponding to that band.

Spatial Smoothing

Spatial smoothing is the idea of averaging the intensities of nearby voxels to produce a smooth spatial map of intensity across the image. This equals to apply a low-pass filter on the image so that high frequencies of the signal are removed while enhancing low frequencies. This procedure leads to a blurred image and spatial correlation within the image is more pronounced. The standard procedure of spatial smoothing is employed by convolving the fMRI image with a Gaussian function of a specific width, which is called Gaussian kernel. The width of the kernel determines how much the data is smoothed (see Figure 1.6), which is usually expressed as Full Width at Half Maximum (FWHM).

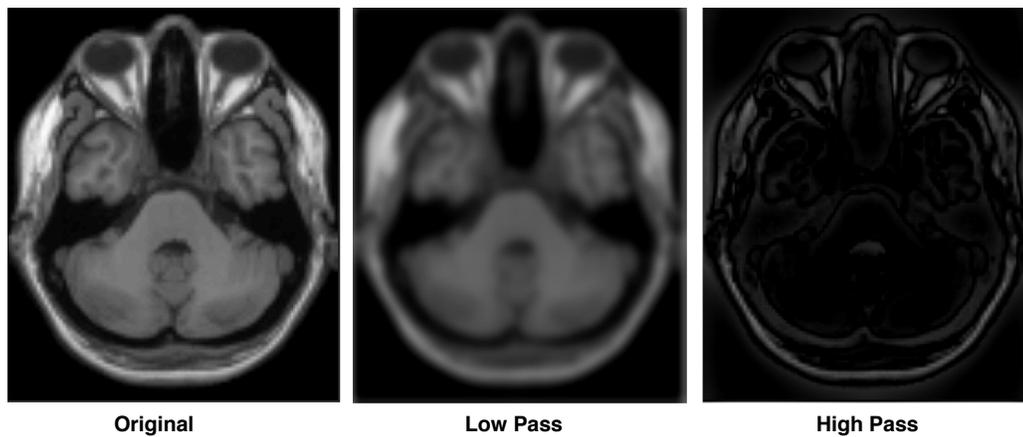


Figure 1.5 An example of filters applied to fMRI image. The low-pass filter blurs the image whereas a high-pass filter enhances the edges in the image.

1.2 Overview of Clustering

Machine learning is a subfield of computer science, which gives the computers the ability to learn without being explicitly programmed [24]. It explores statistical models and construction of algorithms that can derive from and make predictions on data. In the past decade, machine learning has been employed in various areas, e.g., self-driving car, practical speech recognition, effective web search, gene expression, etc. Machine learning tasks are typically classified into three broad categories--supervised learning, reinforcement learning, and unsupervised learning--depending on the nature of the learning data available to a learning system. In supervised learning, the training data given to the computer corresponds to the desired label. A supervised learning algorithm analyzes the training data and produces an inferred function, which can be used for mapping new examples. In reinforcement learning, the machine interacts with its environment by producing actions. These actions affect the state of the environment, which in turn results in the machine receiving some scalar rewards. The goal of the machine is to learn to act in a way that maximizes the feature rewards it receives over its lifetime. In unsupervised learning, a machine is trained to discover hidden structure from data. In contrast to supervised learning and reinforcement learning, the data given to the learner are unlabeled, i.e., there is no error or reward signal to evaluate a potential solution. The most common unsupervised learning method is clustering, which is used for exploring the unknown nature of data through the separation of a finite dataset, with little or no ground truth, into a finite and discrete set of “natural” hidden data structures [25] [26].

1.2.1 Clustering Definition

Given a set of data, the target of clustering is to group them into a certain number of clusters so that data in the same cluster are similar to each other while data in different clusters are

dissimilar (Figure 1.7). Due to the inherent subjectivity of clustering, there is no universally agreed-upon definition of the term cluster. Thus, no criterion has been established to evaluate the effectiveness of all clustering techniques [27]. Depending on the problems and data properties, the definition of cluster differs from one to another. Despite this difficulty, several definitions are still valid:

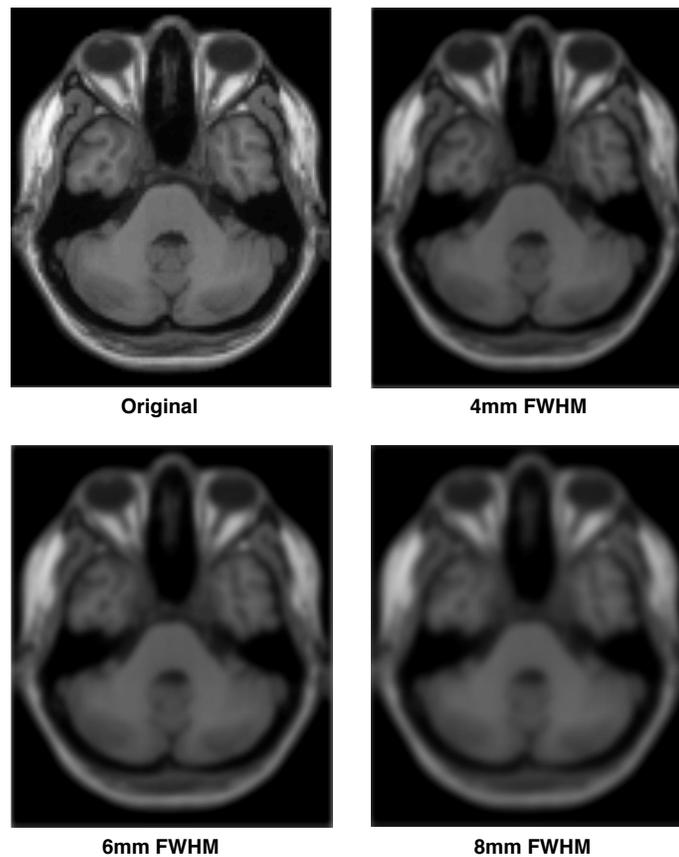


Figure 1.6 The effect of spatial smoothing on fMRI image with different width of Gaussian kernel.

- A cluster is a set of data objects that are similar to each other while data objects in different clusters are different from one another.

- A cluster is a set of data objects such that the distance between an object in a cluster and the centroid of the cluster is less than the distance between this object to the centroids of any other clusters.
- A cluster is a set of data objects such that the distance between any two objects in the cluster is less than the distance between any object in the cluster and any object not in it.
- A cluster is a continuous region of data objects with a relatively high density, which is separated from other such dense regions by low-density regions.

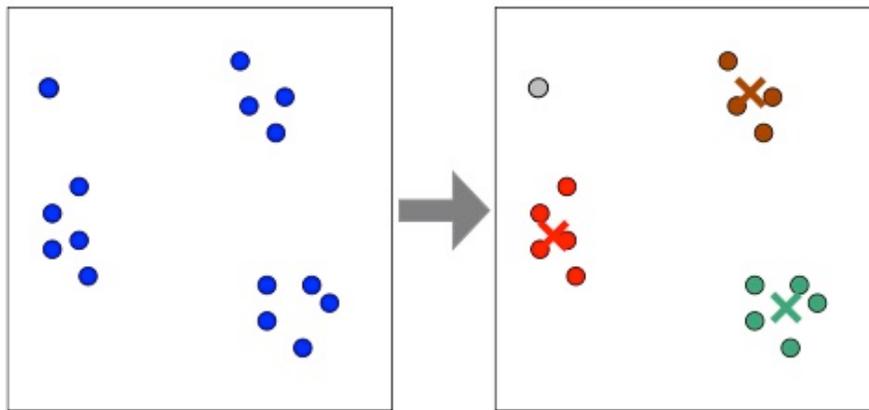


Figure 1.7 An example of clustering. Given some data points (left), the target of clustering is to group them into different clusters (right). In this case, three clusters and noise are marked using different colors.

From the above definitions, it can be seen that the internal homogeneity and external separation are the two emphasized conditions in the clustering [28].

1.2.2 Clustering Procedure

In general, clustering analysis consists of the following four basic steps, i.e., feature selection, clustering algorithm selection, validation, result interpretation, which are equally important and closely related.

Feature selection or extraction. Feature selection picks distinguishing features from a set of candidates, whereas feature extraction uses some transformations to create new features from the original ones [29] [30]. Feature extraction is potentially capable of generating features that could be of better use in separating dataset. However, those features may not be physically interpretable and therefore are difficult to understand. In contrast, feature selection assures the retention of the original physical meaning of the selected features.

- 1) Clustering algorithm selection. The selection of the clustering algorithm is usually accompanied with two questions: what type of distance measure should be used and what kind of optimization function should be selected. Intuitively, data objects are grouped into different clusters according to whether they are similar with each other or not in terms of distance. Almost all clustering algorithms are explicitly or implicitly related to some specific definition of distance measures. Some algorithms even work directly on the distance matrix. Once the distance measure is determined, clustering analysis could be considered as an optimization problem with a specific criterion function.
- 2) Validation. Given an input dataset, the clustering algorithm always produces a separation even if there is no structure in the data. Moreover, different clustering algorithms usually lead to different clustering results. Even for the same algorithm, the selection of a parameter or the presentation order of input data may affect the final results. Therefore, effective evaluation criteria are critically important to the entire clustering process.

- 3) Result interpretation. This step can be easily ignored, although it is as important as the other three steps as the ultimate goal of clustering is to provide a better understanding of the data.

1.2.3 Distance (Similarity) Measures

Clustering algorithms are derived based on the distance measures of data object, each described by a set of features, denoted as a multidimensional vector. The features can be quantitative or qualitative, continuous or discrete, which leads to different measure mechanisms.

Given two data objects \mathbf{Z}_i and \mathbf{Z}_j , $\mathbf{Z}_i = (Z_{i1}, Z_{i2}, \dots, Z_{id}) \in \mathbb{R}^d$, where d equals to the number of features. The commonly used distance measures are described below.

Distance Measures of Numeric Data

When the dimension of data object is numeric, the general distance measure is called the Minkowski (or L_p norm) distance [31], which is defined as,

$$D(\mathbf{Z}_i, \mathbf{Z}_j) = (\sum_{l=1}^d |Z_{il} - Z_{jl}|^p)^{1/p} \quad (1.1)$$

When $p = 1$, this is equivalent to Manhattan distance, whereas when $p = 2$, this is equivalent to Euclidean distance. This measure is independent of the underlying data distribution, thus it has been widely used in different clustering algorithms, e.g., fuzzy c-means, k-means, etc [32] [33] [34].

Minkowski distance measure, although taking into consideration the distribution of each feature dimension, assumes that the feature dimensions are independent of each other. However, if feature dimensions have some correlations with each other, this measure might affect

clustering result. To consider the correlations among different feature dimensions, Mahalanobis distance measures [35] can be used, which is defined as,

$$D(\mathbf{Z}_i, \mathbf{Z}_j) = (\mathbf{Z}_i, \mathbf{Z}_j)^T \mathbf{S}^{-1}(\mathbf{Z}_i, \mathbf{Z}_j) \quad (1.2)$$

Where \mathbf{S} represents the within-cluster covariance matrix. Mahalanobis distance measure has been used ellipsoidal ART clustering and hyper-ellipsoidal clustering algorithms [36] [37]. When features are not correlated, this measure is equivalent to squared Euclidean distance measure.

Cosine similarity [38] is another commonly used distance measure when the direction of the data is mainly considered rather than the magnitude of the data. The definition of Cosine similarity between two data objects is shown below:

$$D(\mathbf{Z}_i, \mathbf{Z}_j) = \frac{\mathbf{z}_i^T \mathbf{z}_j}{\|\mathbf{z}_i\| \|\mathbf{z}_j\|} \quad (1.3)$$

Cosine similarity measures normalized inner product between data objects, which is commonly used in document clustering [39].

Distance Measures of Categorical data

There is no ordering between categorical values, thus only the similarity can be measured, i.e., the overlap between two categorical data objects. Hamming distance [40] is one of the most popular distance measures of categorical data, which counts the number of attributes that need to be changed in order to match one another. The ratio of number of changes to number of all attributes can then be used to determine the similarity between two data objects.

Jaccard similarity [41] is another commonly used measure, which is defined as

$$D(\mathbf{Z}_i, \mathbf{Z}_j) = \frac{\text{Size}(\mathbf{Z}_i \cap \mathbf{Z}_j)}{\text{Size}(\mathbf{Z}_i \cup \mathbf{Z}_j)} \quad (1.4)$$

Where $\mathbf{Z}_i \cap \mathbf{Z}_j$ computes the intersection between two objects, whereas $\mathbf{Z}_i \cup \mathbf{Z}_j$ computes the union between two objects.

1.2.4 Clustering Algorithms

Different clustering algorithms are derived based on different cluster models, which include connectivity model, centroid model, distribution model, density model, subspace model, group model, and graph-based model. The most appropriate clustering algorithm for a particular problem often needs to be chosen experimentally. Also, an algorithm that is designed for one particular model has no chance to work on a data set that contains a radically different kind of model [42].

Connectivity-based Clustering

Connectivity based clustering, also known as hierarchical clustering [43] [44], organizes objects into a hierarchical tree structure, where each node in the tree represents a cluster. It can be performed either from singleton clusters to a cluster including all objects or vice versa. The former is known as agglomerative hierarchical clustering, and the latter is called divisive hierarchical clustering. Compared with agglomerative clustering, divisive clustering is more computationally expensive. Thus, agglomerative clustering is often preferred [45]. However, divisive clustering provides clearer insights into the main structure of the data. This is because that the larger clusters are generated at the early stage of the clustering process and are less likely to suffer from the accumulated erroneous decisions, which cannot be corrected by the successive process [46].

In both agglomerative and divisive algorithm, the distance between two clusters is determined by linkage criterion. There are three commonly used linkage criterion--single linkage, complete linkage, and average linkage [47]. The single linkage defines the distance between two clusters as the distance between their two closest objects. It may cause the chaining problem, which forces clusters together due to single objects being close to each other. Complete linkage uses the maximum distance between two clusters as their distance. It is useful if objects are in high-dimensional space. However, it is unsuitable for dataset with a lot of noise, since outliers are given more weight in the cluster decision. Average linkage takes the mean distance between all pairs of objects in two clusters. It is more computationally intensive than other two linkages, but it can avoid the chaining problem without giving larger weight to outliers.

Hierarchical clustering is one of the most popular clustering methods due to the fact that clusters can be navigated at different levels. However, the time complexity and space complexity of this algorithm are very high. Also, errors in merging clusters cannot be undone and will affect the result.

Centroid-based Clustering

In centroid-based clustering, clusters are characterized by a center vector. With the fixed number of clusters (k), each object is assigned to one specific cluster. Centroid-based clustering can be constructed as an optimization problem with a pre-specified cost function.

The k-means algorithm is one of the best-known centroid-based clustering algorithms [48] [49]. It is based on an iterative optimization procedure to seek an optimal k partition of data with k predefined. K-means works well for many practical problems, particularly when the resulting clusters are compact and hyperspherical in shape. Meanwhile, the time complexity of k-means is

relatively low compared to hierarchical clustering, which is a critical factor in large scale data clustering. However, as k-means is based on a hill-climbing method for optimization, the inherent limitations cause it to suffer from several major drawbacks, such as easy trapping in local minima and sensitivity to the initial setting of centroids. Another major drawback lies in the requirement of determining the parameter k , which is usually seen as part of the unknown natural structure in the data and has to be estimated empirically or using cluster analysis.

Distribution-based Clustering

Distribution-based clustering directly relates to the use of distribution models in statistics [50] [51]. Clusters are defined based on how likely the objects included are likely to belong to the same distribution. Although the theoretical foundation of this kind of clustering is valid, it suffers from overfitting problem unless constraints are put on the complexity of the model. A more complex model will usually be able to explain the data better, which makes choosing the appropriate model complexity inherently difficult. One prominent method is known as expectation-maximization algorithm [52] [53]. The dataset is modeled by a fixed number of Gaussian distributions that are initialized randomly, and the parameters of these Gaussian distributions are iteratively tuned to fit better to the dataset.

Distribution-based clustering can provide information beyond the cluster assignments of objects, such as correlation and dependence between attributes. However, this algorithm does not work if there is no mathematical model inherent in the dataset for the model to optimize. Moreover, assuming that data adhere to Gaussian distribution models can cause fatal problems.

Density-based Clustering

In density-based clustering [54] [55], clusters are defined as areas of higher density than the remainder of the data. Objects in the sparse areas are usually considered to be noise and border points, which are used to separate clusters. The most popular density based clustering method is DBSCAN [56]. Given a radius, each object of a cluster has to contain at least a minimum number of neighborhoods within the radius. Objects not assigned to any cluster are considered as noise. DBSCAN can discover clusters of arbitrary shapes and is resistant to noise. However, the clustering results highly depend on the user-specified parameters: radius and a minimum number of neighborhoods. OPTICS [57] is a generalization of DBSCAN that removes the need to choose an appropriate value for the radius and produces a hierarchical result related to that of hierarchical clustering.

Advantages of density-based clustering include the ability of find clusters with arbitrary shapes and time efficiency. Also, it does not require a user-specified number of clusters. However, this method requires some kind of density drop to detect cluster borders which may not be the case in real problems. Also, it cannot detect intrinsic cluster structures which are prevalent in the majority of real life-data.

1.2.5 Determination of Number of Clusters

The number of clusters (k) is often required as a predefined parameter or many clustering algorithms, e.g., k-means, expectation-maximization, etc. Other algorithms such as hierarchical, DBSCAN, and OPTICS do not require the specification of this parameter. However, other different parameters need to be determined in these algorithms, e.g., cutting level for hierarchical clustering, radius and a minimum number of neighborhoods for DBSCAN and OPTICS, which

can result in a different number of clusters. Thus, it is critical to define a criterion to determine the optimal values for these parameters.

One of the commonly used methods is called elbow criterion [58]. It looks at the percentage of variance explained as a function of the number of clusters and select a number of clusters as the optimal one so that adding another cluster does not provide a much larger percentage of variance explained. As shown in Figure 1.8, the percentage of variance explained by the clusters is plotted as a function of the number of clusters. As the number of clusters increases from 1 to 4, the curve increases dramatically. However, after the number of clusters increases beyond 4, the increase rate drops down, which shows an angle in the figure. The optimal number of clusters is then chosen at this point. The main drawback of this method is that “elbow” point cannot always be unambiguously identified [59], e.g., the elbow may not be pronounced, or the graph may have multiple elbows.

The average silhouette of the clustering is another useful criterion for assessing the natural number of clusters [60]. The silhouette of a clustering result is a measure of how closely it is matched to objects within its cluster and how loosely it is matched to objects of the neighboring cluster, i.e., the cluster whose average distance from the object is lowest. The range of silhouette varies between -1 and 1. If a silhouette is close to 1, it implies that objects are assigned to appropriate clusters. If a silhouette is close to -1, it implies that most of the objects are assigned to wrong clusters. A silhouette close to zero means that the objects are one the border of two clusters.

The Calinski-Harabasz criterion (CH) is defined as a ratio between within-cluster variance and between-cluster variance [61]. A well-defined clustering result has a large between-cluster

variance and a small within-cluster variance. Thus, the optimal number of clusters is indicated by the peak of the CH.

Finally, a priori theory can be used as a non-statistical tool for determining the number of clusters [62]. Although a priori theory is not central to exploratory research, it does provide a benchmark for assessing the results of the theory-testing inquiry. For example, comparison of emergent clusters with a theory-based typology can provide evidence regarding the typology's descriptive validity [63].

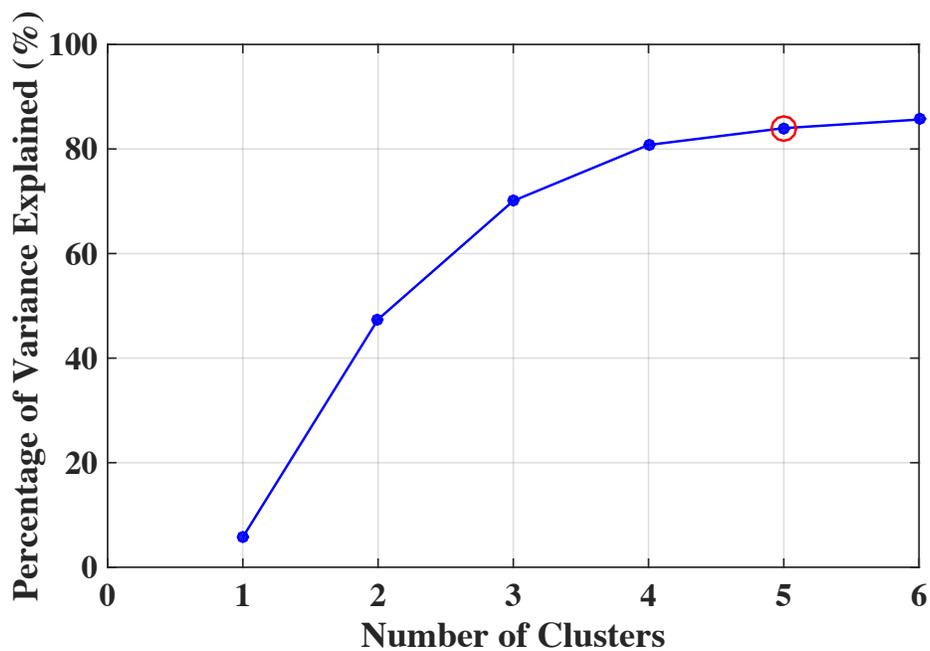


Figure 1.8 An example of determination of number of clusters using elbow criterion. The elbow is indicated by the red circle, which means in this case the optimal number of clusters is 5.

1.2.6 Selection of Clustering Algorithms

There are different kinds of clustering algorithms, and each of them has various features. To evaluate a clustering algorithm's suitability for a problem, a general set of desirable features is applied [64].

- Scalability: the runtime and memory requirement.
- Robustness: ability to detect noise, which is defined as objects that are distant from rest of data.
- Order independence: the key to ensure reproducibility of results.
- Minimum user-specified input: parameters, such as the number of clusters, can significantly affect the result.
- Arbitrary-shaped clusters: ability to find arbitrary shaped clusters.

Regarding to these evaluation criteria, the commonly used clustering methods (see Table 1.1) are compared.

Algorithm	Time Complexity	Robustness	Order Independency	Input Parameters	Arbitrary Shape
K-means	$O(tkN)$	No	No	1, 4	No
K-medoids	$O(tkN)$	Yes	No	1	No
Fuzz k-means	$O(tkN)$	No	No	1	Yes
K-modes	$O(tkN)$	No	No	1	No
CLARANS	$O(N^2)$	Yes	Yes	1	Yes
Hierarchical	$O(N^2 \log N)$	Yes	Yes	5	Yes
BIRCH	$O(N)$	Yes	Yes	-	No
Spectral	$O(N)$	No	Yes	5	No
HIERDENC	$O(N)$	Yes	Yes	-	No
MULIC	$O(N^2)$	Yes	No	-	No

DBSCAN	$O(N \log N)$	Yes	Yes	2, 3	Yes
OPTICS	$O(N \log N)$	Yes	Yes	2, 5	Yes
DENCLUE	$O(N^2)$	Yes	No	3	Yes
DPC	$O(N^2)$	Yes	Yes	10, 11	Yes
EM	$O(N \log N)$	Yes	No	1, 6	No
SOMs	$O(N^2)$	No	No	1, 5, 7	Yes
COBWEB	$O(Nd^2)$	Yes	No	-	No
BILCOM	$O(N^2)$	Yes	No	5	No
MCODE	$O(Nd^3)$	No	Yes	8	No
MCL	$O(N^3)$	Yes	Yes	9	No

Table 1.1 Comparison of different clustering algorithms. Parameters: k = number of clusters, N = sample size, d = feature dimensions, t = number of iterations. User-specified input parameters include: 1. number of clusters, 2. density threshold, 3. radius distance, 4. stop criterion, 5. threshold cutoff, 6. distribution parameters, 7. training data as background knowledge, 8. degree, 9. matrix inflation, 10. minimum distance with higher density, 11. local density.

In this proposal, we want to apply clustering methods on fMRI data where the feature dimension is large and the data structures are unknown. Thus, the selected clustering algorithms must satisfy following requirements: 1) The number of clusters does not need to be pre-defined, 2) the ability to identify arbitrary-shaped clusters, and 3) high efficiency including the order independence and low complexity. Based on these requirements, three clustering methods are selected: hierarchical, OPTICS, and DPC. Although DBSCAN is qualified, since it cannot identify clusters with various densities (which has been resolved in OPTICS), it was dropped out from the selected methods.

1.2.7 Application of Clustering on fMRI Analysis

Application of clustering algorithms in fMRI data analysis is ubiquitous, with typical examples including time series analysis, brain parcellation, and brain network analysis.

Time Series Analysis

Clustering is one the most useful method to analyze fMRI time series. One may want to group a collection of voxels or region of interests (ROIs) based upon similarities in BOLD time series using some distance measures. Clustering is applied to either raw time series or to a function of the time series. For example, Baumgartner [65] demonstrated that fuzzy clustering was a robust, model-independent method to extract functional information in time and space. By applying fuzzy clustering on raw fMRI time series, he showed that intertrial reproducibility of cortical activation is significantly improved. Goutte [66] used two clustering methods, k-means and hierarchical clustering, to detect similarities in activation between voxels. Rather than using the raw fMRI time series, a cross-correlation function was applied as a feature space. The results showed that clustering can effectively identify regions of similar activations. Mezer [67] applied k-means clustering on the cortex, sub-cortical regions and white matter using both raw time series of resting state fMRI (rs-fMRI) signal as well as its Fourier transform.

Functional Parcellation of fMRI Data

Analysis and interpretation of fMRI data often require one to divide the whole brain into a number of regions with homogeneous characteristics. While predefined brain atlases do not adapt to the signal in the individual subject images, parcellation approaches use clustering

methods to divide the brain into a set of non-overlapping regions with homogeneous functionality.

K-means is one of the most popular clustering methods that has been used in different brain parcellation studies [68] [69] [70]. Most of these studies applied k-means clustering on fMRI voxels without considering their spatial coordination. An alternative approach is using hierarchical clustering [71] [72]. Among different hierarchical agglomerative algorithms, the ward's hierarchical clustering [73] was demonstrated with minimum variances. Spectral clustering [74] [75] is often done by performing k-means clustering on a representation of the data that preserves the spatial coordination and the similarity of functional features. This representation is typically obtained by using the first eigenvectors of the Laplacian matrix of the graph that encodes the spatial relationships weighted by the functional features similarity between adjacent locations. A model-based dense clustering [76] method was also developed based on the assumption that activation in brain maps is fundamentally based on spatial continuity and spatial variance.

Brain Network Analysis

The brain is a network, which consists of spatially distributed but functionally linked regions that continuously share information with each other [77]. In the past decade, a lot of neuroimaging studies have started to explore functional brain connections of specific brain regions and overall organization of functional communication in the whole-brain networks. A variety of methods have been proposed to analyze these connections, e.g., seed methods [78], principal component analysis [79], independent component analysis [80], etc., and clustering is one the most popular techniques.

Several studies have been proposed that use clustering on rs-fMRI data to exam the existence and extent of functional connections between brain regions. For example, Cordes [81] used a hierarchical clustering method on rs-fMRI data to detect similarities of low-frequency fluctuations. Thirion [74] proposed a generic framework based on clustering to study resting-state activity networks obtained in fMRI systematically. In addition, model-free methods based clustering have been developed, which can be used to explore connectivity patterns of whole-brain without the need for defining a seed region. In contrast to seed-based methods, model-free methods are designed to look for general patterns of unique connectivity across brain regions. Several clustering-based methods have been successfully applied to fMRI data. For example, Salvador [82] adopted multivariate analysis by hierarchical clustering to investigate large-scale systems organization of the whole human brain using fMRI data. Heuvel [77] proposed a voxel-based model-free normalized cut graph clustering approach with whole-brain coverage for group analysis of rs-fMRI data, and found seven resting state networks that showed a large overlap with others previous results.

1.3 Dissertation Structure

The rest of dissertation is organized as follows. Chapter 2 introduces work that proposes a general pipeline to identify different brain-based disorders along with several supplementary analyses, e.g., site-specific analysis, elimination of outlier subjects, etc. The effectiveness of proposed pipeline is verified on five different disorders: ADHD, AD, ASD, PTSD and PCS. Chapter 3 and Chapter 4 discuss works that relate to hippocampal clustering. Specifically, Chapter 3 validates the major pathway, i.e., perforant pathway, between entorhinal cortex and hippocampus using the connectivity between layer II of entorhinal cortex with different subfields

of the hippocampus during an encoding task. Chapter 4 follows with work that validates the HERNET model by clustering the connectivity between different layers of default mode network and dorsal attention network, and different regions of the hippocampus during encoding and retrieval tasks. Finally, Chapter 5 concludes with works completed so far and outlines future works.

Chapter 2

Investigating the Correspondence of Clinical Diagnostic Grouping with Underlying Neurobiological and Phenotypic Clusters Using Unsupervised Learning

Many brain-based disorders are traditionally diagnosed based on clinical interviews and behavioral assessments, which are recognized to be largely imperfect. Therefore, it is necessary to establish neuroimaging-based biomarkers to improve diagnostic precision. Resting-state functional magnetic resonance imaging (rs-fMRI) is a promising technique for the characterization and classification of varying disorders. However, most of these classification methods are supervised, i.e., they require *a priori* clinical labels to guide classification. In this study, we adopted various unsupervised clustering methods using static and dynamic rs-fMRI connectivity measures to investigate whether the clinical diagnostic grouping of different disorders is grounded in underlying neurobiological and phenotypic clusters. In order to do so, we derived a general analysis pipeline for identifying different brain-based disorders using genetic algorithm-based feature selection, and unsupervised clustering methods using four different datasets; three of them – ADNI, ADHD-200, and ABIDE – which are publicly available, and a fourth one – PTSD and PCS – which was acquired in-house. Using these datasets, the effectiveness of the proposed pipeline was verified on different disorders: Attention Deficit Hyperactivity Disorder (ADHD), Alzheimer’s Disease (AD), Autism Spectrum Disorder (ASD), Post-Traumatic Stress Disorder (PTSD), and Post-Concussion Syndrome (PCS). For ADHD and AD, highest similarity was achieved between connectivity and phenotypic clusters, whereas for ASD and PTSD/PCS, highest similarity was achieved between connectivity and clinical diagnostic clusters. For multi-site data (ABIDE and ADHD-200), we report site-specific

results, as well as the effect of elimination of outlier subjects. Overall, our results suggest that neurobiological and phenotypic biomarkers could potentially be used as an aid by the clinician, in addition to currently available clinical diagnostic standards, to improve diagnostic precision.

2.1 Introduction

A brain-based or neuropsychiatric disorder is a condition of affect, cognition, and behavior with underlying cerebral dysfunction. There are many factors that contribute to these disorders, e.g., genes, family history, substance abuse, traumatic brain injury, life experience, etc. Conventional diagnosis mainly consists of clinical interviews and standardized testing, which are recognized to be largely imperfect [83] [84] [85]. Because neuropsychiatric pathologies are complex, which can lead to inconsistencies between clinicians' diagnoses, there is increasing interest in identifying non-invasive neuroimaging biomarkers. The most commonly used approach for achieving this is by employing *supervised* learning models such as support vector machines [86] [87], artificial neural networks [88], and decision trees [89], wherein the model learns the associations between patterns in the data and diagnostic labels using a training data set. This model can then be tested on an independent validation data set. However, the problem with this approach is that the model itself is based on clinical labels, and hence, it cannot be used to uncover novel structures and groupings from the data. This can be achieved by employing *unsupervised* models. *Unsupervised* models have been used to uncover clusters of subjects with similar patterns of imaging data, with applications in identifying disease clusters [24] [90] as well as sub-clusters [91] within a disease. Most of these studies use k-means or hierarchical clustering [92] [93]. However, these approaches are besieged with methodological issues such as (i) *a priori* choice of clusters needed in k-means, (ii) a stopping criterion needed in hierarchical

clustering, and (iii) the large dimensionality of imaging data necessitates some type of dimensionality reduction for clustering to work properly, and this step is either not carried out [94] or carried out by preselecting features not from the structure in the data, but by some external considerations such as previous findings in a given disorder [95] [96]. Such approaches rob the method of its advantages of being truly data-driven in that (iv) the clusters obtained from imaging data are seldom compared to data obtained from clinical diagnostic criteria and related behavioral phenotypes. This is important because disease clusters obtained from any method, be it imaging or another diagnostic tool, should be linked with the behavioral phenotype. In this study, we address the above shortcomings using resting state functional magnetic resonance imaging (rs-fMRI) data obtained from five different neuropsychiatric disorders: Attention Deficit Hyperactivity Disorder (ADHD), Alzheimer's Disease (AD), Autism Spectrum Disorder (ASD), Post-Concussion Syndrome (PCS) and Post-Traumatic Stress Disorder (PTSD).

Attention Deficit Hyperactivity Disorder

ADHD is a psychiatric disorder characterized by impulsiveness, inattention, and hyperactivity. This condition affects about 5% of children and adolescents worldwide [97]. Symptoms include difficulty staying focused and paying attention, difficulty controlling behavior, and hyperactivity. ADHD has three subtypes: ADHD hyperactive-impulsive (ADHD-H), ADHD inattentive (ADHD-I), and ADHD combined hyperactive-impulsive and inattentive subtype (ADHD-C). Because symptoms vary from person to person, ADHD can be difficult to identify. Also, there has been a debate that ADHD is over-diagnosed in children and adolescents by current clinical criterion [98].

Alzheimer's disease

AD is the most commonly diagnosed type of dementia in elderly patients [99], which is characterized by memory dysfunction, cognitive decline, etc. Before the onset of dementia, patients may develop an intermediate stage of dysfunction known as mild cognitive impairment (MCI). Patients with MCI have a higher risk of progressing to AD [100]. There are a variety of clinical diagnoses of AD, each with varying degrees of accuracy. According to results from the Honolulu-Asia Aging Study [101], as many as one-third of all Alzheimer's diagnoses may actually be false positive. In addition, the diagnostic boundary between AD and MCI is not well established.

Autism Spectrum Disorder

ASD is a pervasive developmental disorder clinically characterized by social and communication impairments as well as restricted interests and repetitive behaviors [83]. While the boundaries between ASD, its comorbidities, and neurotypicals with sub-clinical ASD-like traits are blurred, several diagnostic subcategories within ASD were defined: autism, Asperger's disorder, and pervasive developmental disorder-not otherwise specified (PDD-NOS). It has been often argued that the Asperger's disorder criteria is problematic [102] [103]. In the latest DSM-V classification, Asperger's and PDD-NOS were eliminated, in favor of the so called "dimensional assessment" of the autism spectrum [104]. This highlights the confusion in the field due to lack of objective biomarkers based on underlying neurobiology.

Post-Traumatic Stress Disorder

PTSD is a disabling condition in individuals exposed to a traumatic event, such as war, violent crime, and motor vehicle accidents [105]. PTSD is characterized by intrusive, avoidance, negative alterations in cognitions and mood, and alterations in arousal and reactivity [106].

PTSD is associated with compromised functioning of the amygdala, hippocampus, insula, and regions of the prefrontal cortex such as the ventromedial PFC [107] [108] [109]. Although cognitive decrements are associated with PTSD, there is evidence that they are mediated by symptoms of the disorder (e.g., depression and anxiety) [110].

Post-Concussion Syndrome

PCS is a syndrome following a mild traumatic brain injury or concussion that includes a constellation of symptoms which can be categorized as vestibular, cognitive, affective, and somatosensory [111] [112] [113]. Symptoms can last for weeks and even months. In military service members, diagnosis can be more complex since there is a high co-morbidity with PTSD and homogeneity of symptomatology between the two disorders [114].

In summary, although the neuropsychiatric disorders delineated have well established diagnostic criteria, there still exists overlapping symptoms as well as some potential commonalities in implicated neuroanatomy. Also, diagnostic categories often do not adequately capture the spectrum of symptom and impairments ranging from mild to severe. Further, categorization of subgroups within many disorders have yet to be fully characterized. Thus, neuroimaging-based biomarkers can improve our understanding of subgroups within a specific neuropsychiatric disorder and eventually improve diagnostic precision.

Resting-state functional magnetic resonance imaging (rs-fMRI) is a promising technique for identification of different neuropsychiatric disorders [115] [116] [117] [118] [119] [120]. It measures spontaneous fluctuations in blood oxygen level-dependent (BOLD) signal without performing any explicit task [121] [122]. A common methodology is to apply supervised classification methods using functional connectivity. For example, some studies [88] used

support vector machine (SVM) and artificial neural network (ANN) on different brain connectivity measures to identify ADHD. Khazaei et al [85] combined a graph theoretical approach with SVM to classify patients with AD and mild cognitive impairment from healthy controls. Plitt and colleagues [84] applied different classification methods, e.g., K-Nearest Neighbor (KNN), Linear Support Vector Machines (L-SVM), Gaussian Kernel support vector machines (rbf-SVM), L1-regularized logistic regression, on rs-fMRI connectivity measures to establish biomarkers for Autism spectrum disorders (ASD). However, these methods all require *a priori* clinical diagnoses to guide classification. Also, the majority only targeted diagnosing one specific illness [94]. There are very few studies using unsupervised clustering on rs-fMRI to identify different neuropsychiatric disorders [83].

In this work, we attempt to address the four challenges in *supervised* models of MRI data enumerated above by deriving a general analysis pipeline for identifying different neuropsychiatric disorders using unsupervised clustering methods. The main idea of clustering is to group objects in such a way that objects in the same group are more similar to each other than to those in other groups. Three clustering methods were specifically chosen, i.e., hierarchical clustering [123], ordering points to identify the clustering structure (OPTICS) [57], and density peak clustering (DPC) [124], since they did not require *a priori* specification of the number of clusters. The commonly used k-means clustering [125] [126] was not considered in this study due to the uncertainty of the number of clusters and sensitivity to outliers. Since clustering accuracy is often lower in high dimensional feature space, feature selection methods were applied. Most existing feature selection algorithms in the machine learning literature focus on heuristic search such as sequential forward searching (SFS) [127], non-linear optimization [128], genetic algorithm (GA) [127], etc. Bradley et al. proposed a non-linear optimization using a

nonlinear kernel support vector machine. Although this method provides high accuracy, it can only be used in the supervised learning context. SFS was proposed based on a greedy algorithm, which follows the problem-solving heuristic of making the locally optimal decision at each step. Similar to SFS, here we propose a sequential feature ranking (SFR) method by applying ANOVA test among different groups (e.g., control group, disease subgroups) and then sequentially selecting features from the original dataset based on the p-value of each feature. Although SFS and SFR can be applied in unsupervised learning, they do not guarantee an optimal solution. Therefore, we propose GA as a robust feature selection method for unsupervised learning approaches to identification of disease clusters from resting state functional connectivity by maximizing the similarity between connectivity and clinical diagnosis, and between connectivity and behavioral phenotypes, respectively. The identified clusters are then compared with those obtained from clinical diagnostic criteria and behavioral phenotypes.

2.2 Materials and Methods

In this work a general pipeline has been derived (Figure 2.1) for identifying different brain-based disorders using unsupervised clustering methods along with several supplementary analyses, i.e., site-specific analysis for multi-site data, elimination of outlier subjects, and enrichment analysis. The details of each step in the pipeline are described below.

2.2.1 Participants and Non-Imaging Measures

a) ADHD: 487 subjects with complete phenotypic data were selected from ADHD-200 sample (http://fcon_1000.projects.nitrc.org/indi/adhd200/), including 272 healthy controls (HC), 118 ADHD-C, and 97 ADHD-I. The total number of subjects for ADHD-H was too small, and

therefore ADHD-H was not considered in this work. The subjects were scanned at one of these three different sites: Peking University, Kennedy Krieger Institute (KKI), and New York University Child Study Center (NYU).

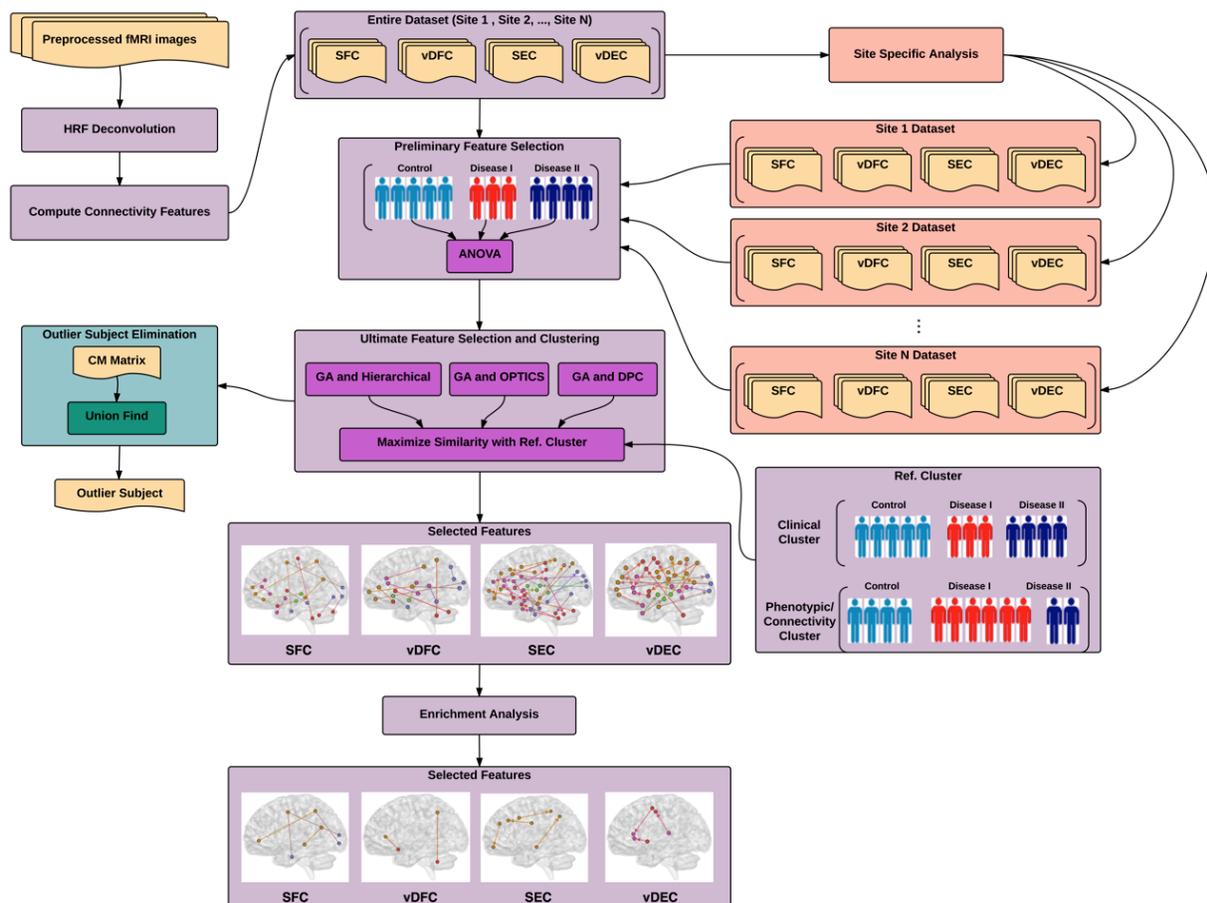


Figure 2.1 Illustration of proposed pipeline for identifying different brain-based disorders using unsupervised clustering methods. The main pipeline is depicted in purple color along with two supplementary analyses, site-specific analysis in salmon color and outlier subject elimination in green color.

Peking University subjects with diagnosis of ADHD were initially identified using the Computerized Diagnostic Interview Schedule IV [C-DIS-IV] [129]. All participants (ADHD and HC) were evaluated with the Schedule of Affective Disorders and Schizophrenia for Children—

Present and Lifetime Version [KSADS-PL] [130] with one parent for the establishment of the diagnosis for study inclusion. The ADHD Rating Scale [ADHD-RS-IV] [131] [132] was employed to provide dimensional measures of ADHD symptoms. Intelligence was evaluated with the Wechsler Intelligence Scale for Chinese Children-Revised [WISCC-R] [133].

In the KKI sample, psychiatric diagnoses were based on evaluations with the Diagnostic Interview for Children and Adolescents, Fourth Edition [DICA-IV] [134], a structured parent interview based on DSM-IV criteria; the Conners' Parent Rating Scale-Revised, Long Form [CPRS-R] [135], and ADHD-RS-IV. Intelligence was evaluated with the Wechsler Intelligence Scale for Children-Fourth Edition [WISC-IV] [136] and academic achievement was assessed with the Wechsler Individual Achievement Test-II [137].

In the NYU sample, psychiatric diagnoses were based on evaluations with KSADS-PL administered to parents and children and CPRS-R. Intelligence was evaluated with the Wechsler Abbreviated Scale of Intelligence [WASI] [138].

Phenotypic/Genetic Variables	Selected Variables		
	DPC	OPTICS	Hierarchical
ADHD index score	✓		
Inattentive score	✓	✓	✓
Hyper/Impulsive score	✓	✓	✓
VIQ		✓	✓
PIQ			✓
FIQ	✓		✓

Table 2.1 Phenotypic variables selected by GA with different clustering methods (ADHD).

Six Phenotypic variables were measured for all sites (Table 2.1), i.e., three ADHD measures including ADHD index score, Inattentive score, and Hyper/Impulsive score [139], and three IQ measures including Verbal IQ [VIQ], Performance IQ [PIQ], and Full Scale IQ [FIQ] [140].

b) AD: Rs-fMRI data from Alzheimer’s disease neuroimaging initiative (ADNI) database (<http://adni.loni.ucla.edu>) was analyzed in this study. The sample consisted of subjects with three progressive stages of cognitive impairment – early MCI [EMCI] (n = 23), late MCI [LMCI] (n = 29), and AD (n = 13) – along with matched HC (n = 31).

The patients with AD had a Mini-Mental State Examination [MMSE] [141] score of 14-26, a Clinical Dementia Rating [CDR] [142] of 0.5 or 1.0 and met the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer’s disease and Related Disorders Association [NINCDS/ADRDA] criteria [143] for probable AD. The patients with MCI had MMSE scores between 24 and 30, a memory complaint, objective memory loss measured by education adjusted scores on Wechsler Memory Scale Logical Memory II, a CDR of 0.5, absence of significant levels of impairment in other cognitive domains, essentially preserved activities of daily living and an absence of dementia [85].

Phenotypic/Genetic Variables	Selected Variables		
	DPC	OPTICS	Hierarchical
APOE A1 and A2		✓	✓
NPI			
GDS			
MMSE	✓	✓	✓
CDR	✓	✓	✓
FAQ	✓	✓	✓

Table 2.2 Phenotypic variables selected by GA with different clustering methods (AD).

Eight phenotypic variables, i.e., neuropsychiatric inventory [NPI] score [144], geriatric depression scale [GDS] [142], MMSE, CDR and functional assessment questionnaire [FAQ] [145], and one genetic variable i.e., apolipoprotein [APOE] A1 and A2 genotypes [146], were measured (Table 2.2). Except for AD dataset, all other three datasets (ADHD, ASD, and PTSD) only have phenotypic variables. Thus, we just refer to these variables as phenotypic variables henceforth.

c) ASD: 454 subjects with complete phenotypic data were selected from the Autism Brain Imaging Data Exchange (ABIDE) database (http://fcon_1000.projects.nitrc.org/indi/abide/index.html). The sample consisted of 256 HC, 166 Autism, and 32 Asperger's. Including PDD-NOS and "Asperger's or PDD-NOS" would make the whole dataset a lot more unbalanced, and therefore these two subgroups were not considered in this study. Each of subjects was scanned at one of the following seven different sites: California Institute of Technology (Caltech), Carnegie Mellon University (CMU), NYU Langone Medical Center (NYU), University of Pittsburgh School of Medicine (Pitt), San Diego State University (SDSU), Trinity Center for Health Sciences (Trinity), and University of California Los Angeles (UCLA).

For most of sites, diagnosis of ASD was consistent with Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision [DSM-IV-TR] criteria [147], and classification of either autism or Asperger's was made by a clinician based on the Autism Diagnostic Observation Schedule [ADOS] [148] and Autism Diagnostic interview-Revised [ADI-R] [149]. HC subjects were screened through a self-report history questionnaire to rule out neurological disorders, such as ASD, ADHD, or Tourette's Disorder.

Ten phenotypic variables were measured (Table 2.3) for all sites including three IQ measures, i.e., FIQ, VIQ, PIQ, four ADI_R measures, i.e., Reciprocal Social Interaction Subscore [ADI_R_SOCIAL], Abnormalities in Communication Subscore [ADI_R_VERBAL], Restricted, Repetitive, and Stereotyped Patterns of Behavior Subscore [ADI_RRB], Abnormality of Development Evident at or Before 36 Months Subscore [ADI_R_ONSET], and three ADOS measures, i.e., Classic Total ADOS Score [ADOS_TOTAL], Communication Total Subscore of the Classic ADOS [ADOS_COMM], and Social Total Subscore of the Classic ADOS [ADOS_SOCIAL].

Phenotypic/Genetic Variables	Selected Variables		
	DPC	OPTICS	Hierarchical
FIQ	✓	✓	✓
VIQ	✓		
PIQ	✓		
ADI_R_SOCIAL	✓	✓	
ADI_R_VERBAL		✓	✓
ADI_RRB		✓	
ADI_R_ONSET			
ADOS_TOTAL			✓
ADOS_COMM		✓	✓
ADOS_SOCIAL	✓		✓

Table 2.3 Phenotypic variables selected by GA with different clustering methods (ASD).

d) PTSD and PCS: Eighty-seven active-duty male U.S. Army Soldiers were recruited from Fort Benning, GA, USA and Fort Rucker, AL, USA to participate voluntarily in the current study. The subjects were grouped into 17 having PTSD, 42 having both PCS and PTSD (PCS+PTSD) and 28 combat controls (all groups matched in age, race, deployment history and education), all having combat experience in Iraq (Operation Iraqi Freedom [OIF]) and/or

Afghanistan (Operation Enduring Freedom [OEF]). The study protocol and procedures were approved by the Auburn University Institutional Review Board (IRB) and the Headquarters U.S. Army Medical Research and Materiel Command, IRB (HQ USAMRMC IRB).

Subjects were grouped using symptom severity in PTSD using the ‘PTSD Checklist-5’ [PCL5] score [150], symptom severity in PCS using the ‘Neurobehavioral Symptom Inventory’ [NSI] score [151] and medical history. i) Subjects with no history of mTBI in the last five years, a total score ≥ 38 on PCL5 and < 26 on NSI were grouped as PTSD. ii) Subjects with a history of medically documented mTBI, post-concussive symptoms, and scores ≥ 38 on PCL5 and ≥ 26 on NSI were grouped as the comorbid PCS+PTSD. iii) Subjects with a score < 38 on PCL5 and < 26 on NSI, no DSM-IV-TR or DSM-V diagnosis of a psychotic disorder (e.g. schizophrenia), no mTBI within the last 5 years, and no history of a moderate-to-severe TBI were grouped as combat controls. The post-concussive symptom (NSI) scores were significantly different between the PCS+PTSD group and the PTSD and control groups combined ($p = 1.32 \times 10^{-29}$). Also the PCL5 scores were significantly different between the control group and the PTSD and PCS+PTSD groups combined ($p = 3.64 \times 10^{-44}$).

Scores from 32 measures were included in the analysis as phenotypic variables. They consisted of 10 primary Neurocognitive measures CNS-Vital Signs® [CNS-VS] measures, 7 derived CNS-VS domain scores, 8 self-report psychological health measures, and 7 neurocognitive measures from a second battery, the Automated Neuropsychological Assessment Metric (ANAM 4.0) (Table 2.4). CNS-VS [152] is a computerized neurocognitive assessment battery. The 10 primary CNS-VS measures were Symbol Digit Coding [SDC; correct responses], Stroop Test [ST] (simple and complex), Shifting Attention Test (SAT), Continuous Performance

Test [CPT; correct responses and reaction time, RT], Dual-Task Test [DTT; correct responses and RT], and Digit Span Test [DST]. The following domain scores were calculated from the subtests to derive domain scores: verbal memory [VM], complex attention [CA], reaction time [RT], processing speed [PS], cognitive flexibility [CF], and executive functioning [EF]. Domain scores were standardized to have a mean of 100 and standard deviation of 15. In addition, domain scores were averaged to form a single score called neurocognitive composite index [NCI]. Additionally, data from the ANAM seven subtests were included - Coded Digit Substitution [CDS], Coded Digit Substitution-Delayed [CDD], Matching to Sample [MTS], Mathematical Processing [MP], Procedural Reaction Time [PRT], Simple Reaction Time [SRT], and Simple Reaction Time-Delayed [SRT2]. Effort was also assessed to improve the validity of our assessment data. To this end, the Test of Memory Malingering (TOMM; Tombaugh, 2003) was administered, which consists of two learning trials and a retention trial that uses pictures of common, everyday objects (e.g., chair, pencil). A cut-off score (<45 correct) for the first two learning trials was used to determine eligibility for participation in the study.

Psychological health was assessed using five self-report measures - Perceived Stress Scale [PPS], Pittsburgh Sleep Quality Index [PSQI], Epworth Sleepiness Scale [ESS], Zung Anxiety Scale [ZAS] and Zung Depression Scale [ZDS]; and three exposure/injury descriptive measures - Combat Exposure Scale [CES], lifetime concussions [LC], and Life Events Checklist [LEC].

2.2.2 Data Acquisition

a) ADHD: Subjects were scanned on different MRI scanners (KKI: Philips 3T, Peking: SIEMENS MAGNETOM Trio 3T, NYU: SIEMENS MAGNETOM Allegra 3T) using standard T2* weighted echo-planar imaging sequence with the following parameters at resting state: For

KKI, TR=2500ms, TE=30ms, FA=75°, slice thickness=3mm, number of slices=47. For Peking, TR=2000ms, TE=30ms, FA=90°, slice thickness=3mm, number of slices=33. For NYU, TR=2000ms, TE=15ms, FA=90°, slice thickness=4mm, number of slices=33.

Phenotypic/Genetic Variables		Selected Variables		
		DPC	OPTICS	Hierarchical
Primary CNS- VS Measures	SDC Correct		✓	✓
	ST Simple			
	ST Complex			
	SAT Correct			
	SAT RT			
	CPT Correct			
	CPT RT			
	DTT Percent Box			
	DTT Correct			
	DST			
Derived CNS- VS Measures	NCI			
	RT			
	VM			
	CA			
	CF			
	EF			
	PS			
Self-report Measures	LC			
	PSS			
	PSQI			
	ESS			
	ZDS	✓	✓	✓
	ZAS	✓		
	CES	✓	✓	✓
	LEC		✓	✓
	CDD_SS			✓
	CDS_SS			

Other Psychometric Measures	MTS			
	MP			
	PRT			
	SRT2			
	SRT			

Table 2.4 Phenotypic variables selected by GA with different clustering methods (PTSD).

b) AD: Data were acquired using a Philips 3T MRI scanner. A total of 140 volumes (TR=3000ms, TE=30ms, FA=80°, and slice thickness=3.3mm, number of slices=48) were obtained. For each subject, the first 7 volumes were discarded for signal equilibrium and to allow the participant's adaptation to the circumstances.

c) ASD: For Caltech, data were acquired using a 3T MAGNETOM Trio scanner with TR=2000ms, TE=30ms, FA=75°, slice thickness=3.5mm, number of slices=34. For CMU, data were acquired using a 3T MAGNETOM Verio scanner with TR=2000ms, TE=30ms, FA=73°, slice thickness=3.0mm, number of slices=34. For NYU, data were acquired using a 3T MAGNETOM allegra scanner with TR=2000ms, TE=15ms, FA=90°, slice thickness=4.0mm, number of slices=33. For Pitt, data were acquired using a 3T MAGNETOM allegra scanner with TR=1500ms, TE=25ms, FA=70°, slice thickness=4.0mm, number of slices=29. For SDSU, data were acquired using a GE 3T MR750 scanner. A total of 180 volumes (TR=2000ms, TE=30ms, FA=90°, slice thickness=3.4mm) were obtained. For Trinity, data were acquired using a Philips 3T Achieva MRI scanner with TR=2000ms, TE=28ms, FA=90°, slice thickness=3.5mm, number of slices=38. For UCLA, data were acquired using a 3T MAGNETOM Trio scanner with TR=3000ms, TE=28ms, FA=90°, slice thickness=4.0mm, number of slices=34.

d) PTSD/PCS: Participants were scanned in a 3T MAGNETOM Verio scanner (Siemens Healthcare, Erlangen, Germany) at Auburn University using T2* weighted multiband echo-planar imaging (EPI) sequence in resting state (the participants were asked to keep their eyes open and fixated on a white cross displayed with dark background on the screen using an Avotec projection system and not think of anything specific), with TR=600ms, TE=30ms, FA=55°, multiband factor=2, voxel size=3×3×5 mm³, 1000 volumes and two sessions (cerebellum was excluded).

For all sites, the data acquisition protocol was approved by local Institutional Review Boards (IRBs), the experiments were performed in compliance with internationally accepted ethical standards and all subjects provided informed consent.

2.2.3 Preprocessing

Standard pre-processing of fMRI data was performed in different datasets, including slice timing correction, motion correction, realignment, normalization to MNI space, and regressing out nuisance variance (six head motion parameters, white matter signal, cerebrospinal fluid signal and global mean signal regression). Below, we elaborate on pre-processing steps specific to each dataset. For ADHD dataset, preprocessing was performed by Athena pipeline using AFNI [15] and FSL [153] while DPARSF toolbox [154] and SPM8 package (<http://www.fil.ion.ucl.ac.uk/spm>) was used to preprocess other datasets.

2.2.4 Connectivity Measures

Given the high dimensionality of whole-brain data, each rs-fMRI image was partitioned into 200 (for ADHD, AD, and ASD) or 125 (for PTSD/PCS) functionally homogenous regions of

interests (ROIs) using normalized cut spectral clustering [cc200 template] [155]. Even though the same parcellation was used on all data, we ended up with only 125 regions for the PTSD/PCS dataset because we had limited coverage. The mean time series for each ROI was subsequently extracted and deconvolved using the method proposed by Wu et al [156] to obtain hidden neuronal time series [157] [158] [159] [160]. Deconvolution was performed because inter-subject and spatial variability of the HRF could potentially give rise to false connectivity estimates [161] [86] [162]. Then, four connectivity matrices—statistic functional connectivity (SFC), variance of dynamic functional connectivity (vDFC) [163], statistic effective connectivity (SEC) [164] [165] [166] [167], and variance of dynamic effective connectivity (vDEC) [168] [169] [170]—were computed using the latent neuronal time series.

Functional connectivity (FC) measures the functional interrelationship between pairs of brain regions. We evaluated SFC using Pearson’s correlation, which gives a single measure of connection strength between two time series. Most studies investigate SFC, assuming connectivity to be temporally stationary. Dynamic fluctuations of connectivity are not captured when using static connectivity. It has been shown that dynamic changes in FC are relevant to neuropathology [171] as well as behavioral performance in different cognitive domains in healthy individuals [172]. For a comprehensive overview of DFC of resting state fMRI see Hutchison et al. [173]. DFC was evaluated using a sliding windowed Pearson’s correlation with variable window length determined adaptively by timeseries stationarity assessed through the augmented Dickey-Fuller test (ADF test), as in our earlier study [172]. This procedure searches for the optimal window length within a specified range using stationarity of the signal as the criteria for optimization. We have used a liberal range of 20 to 140 data points. The justification for using this range for resting state fMRI data is provided in Jia et al. [172].

While FC is a non-directional quantity, another approach to brain connectivity modeling is effective connectivity (EC), which characterizes directional interactions in the brain. It gives characteristically different information from FC, i.e. the former characterizes causal influences while the latter captures synchronization, both of which have been acknowledged to be modes of communication in the brain. We evaluated SEC using Granger causality [174] [86] [175], which quantifies the directional influence of one region over the other. We also evaluated its time-varying version, called DEC [176] [177] [178] [179], using time-varying Granger causality evaluated in a dynamic Kalman filter framework [180] [181] [182] [183].

SFC, DFC, SEC and DEC values were obtained between all pairs of brain regions. Variance of DFC and DEC were computed to obtain vDFC and vDEC. Significant group differences were obtained in four matrices using one way ANOVA, and only the top significant features ($p < 0.01$) were used in further clustering analysis.

2.2.5 Clustering

In order to verify whether clinical diagnostic grouping grounded in underlying neurobiological and phenotypic clusters, the three clustering methods were applied on three types of features: (i) connectivity-based features: SFC, SEC, vDFC, and vDEC, (ii) clinical diagnostic labels, and (iii) phenotypic and genetic (when available) variables. Let $\mathbf{Z} = \{\mathbf{Z}_1, \dots, \mathbf{Z}_j, \dots, \mathbf{Z}_N\}$ represent a set of N objects, i.e., number of subjects in the dataset. $\mathbf{Z}_j = (Z_{j1}, Z_{j2}, \dots, Z_{jp}) \in \mathbb{R}^p$, where p equals to the number of selected features. Assume the data is partitioned into k clusters. Each cluster is a set of indexes from $\{1, \dots, N\}$, and each object \mathbf{Z}_j belongs to exactly one cluster.

a) **Hierarchical Clustering (HC; Agglomerative):** The main idea of hierarchical clustering [184] [185] [123] is that objects that are more related to nearby objects than to objects farther away in terms of Euclidean distance are clustered together. A brief description of the procedure is as follows:

- 1) Assign each object Z_j in a cluster of its own.
- 2) Calculate the distance between any two clusters and merge the closest pair of clusters.
- 3) Repeat step 2-3 until all Z_j are in one big cluster.

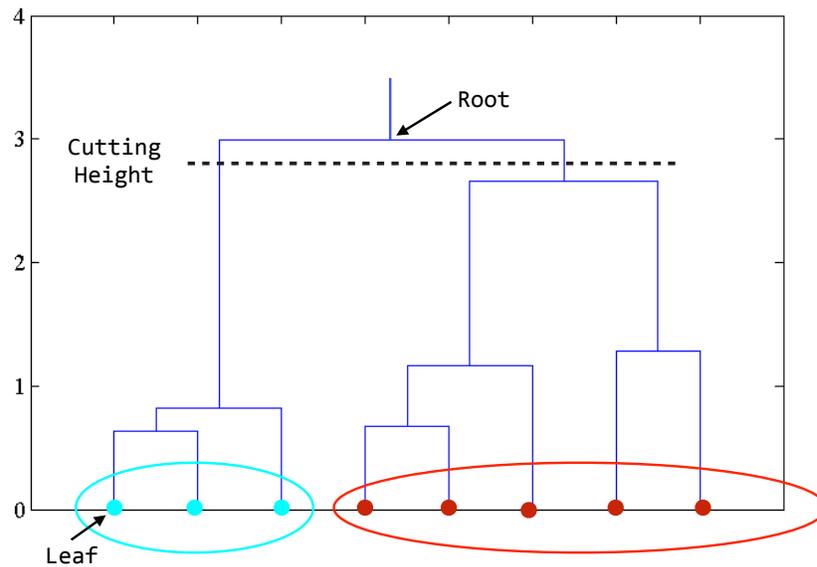


Figure 2.2 Dendrogram derived from hierarchical clustering. The final clustering result is obtained by cutting the tree at defined level. The two clusters are shown by different colors (red and cyan)

The results of hierarchical clustering are usually depicted by a tree-like structure, the dendrogram (Figure 2.2). The root of the dendrogram represents the entire data, each leaf represents one object, and the height of the dendrogram represents the distance between each

pair of clusters. Different data partitions can be obtained by cutting the dendrogram at different levels. Note that the distance between two clusters can be measured in a variety of ways, referred to as linkage methods. The single linkage [186] calculates the shortest distance between two clusters, the complete linkage [186] calculates the longest distance, and the average linkage [186] calculates the mean distance. The single linkage method can handle non-elliptical shape of clusters, but can be affected by noise and outliers. The complete linkage method is less sensitive to noise and outliers but tends to break large clusters. The average linkage is a compromise between single-linkage and complete linkage methods. Thus, the average linkage method was employed in this work.

b) Ordering Points to Identify the Clustering Structure (OPTICS): OPTICS [57] is one of the most popular density-based clustering methods [55]. Similar to hierarchical clustering, it is based on grouping objects within certain distance thresholds (ϵ). However, in OPTICS each cluster has to contain at least a minimum number of objects (*MinPts*), which is not the case in hierarchical clustering. OPTICS can discover clusters with arbitrary shapes and allow “noise” objects that do not belong to any of the clusters.

There are two critical values for each object: core-distance and reachability-distance. Let $N_\epsilon(\mathbf{Z}_j)$ represent the number of neighborhood objects within ϵ (called $\epsilon - neighborhood$), and $MinPts - distance(\mathbf{Z}_j)$ represent the distance from \mathbf{Z}_j to its *MinPts*' neighbor. An object \mathbf{Z}_j is a core object if at least *MinPts* objects are found with its $\epsilon - neighborhood$. The core distance of \mathbf{Z}_j is defined as:

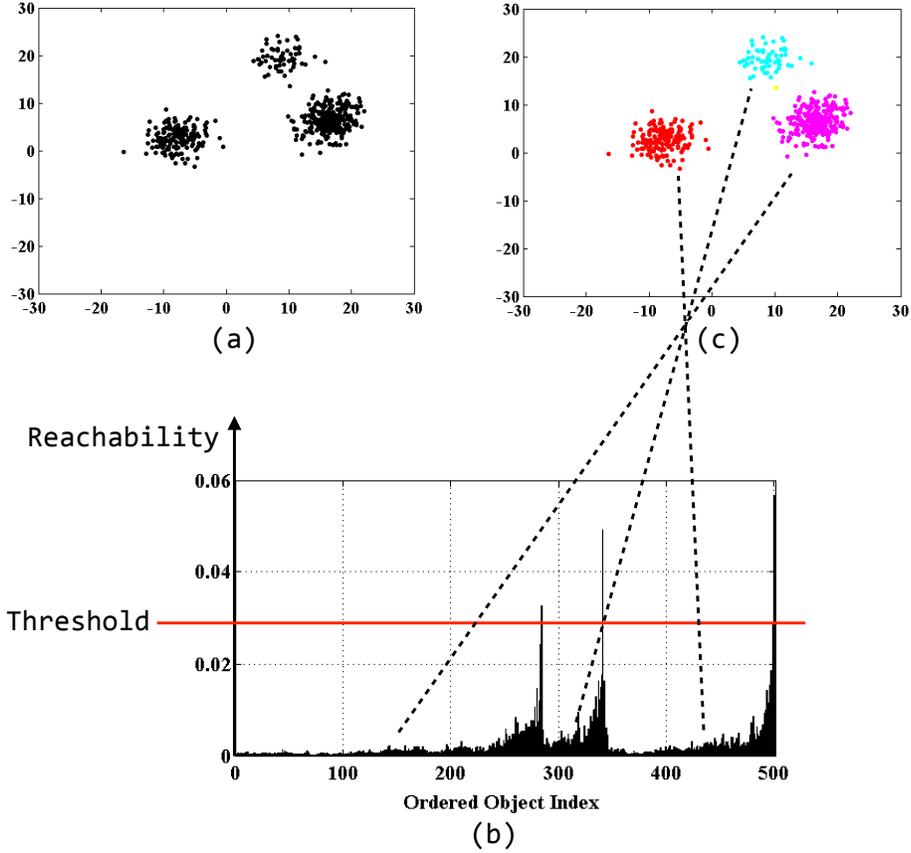


Figure 2.3 Illustration of OPTICS. (a) Original simulated dataset, (b) reachability plot obtained from OPTICS, and (c) clustering result. Each cluster corresponds to one valley in the reachability plot.

$$core - dist_{\varepsilon, MinPts}(\mathbf{Z}_j) = \begin{cases} Undefined & \text{if } N_{\varepsilon}(\mathbf{Z}_j) < MinPts \\ MinPts - distance(\mathbf{Z}_j) & \text{otherwise} \end{cases} \quad (2.1)$$

Which is the smallest distance between \mathbf{Z}_j and an object in its $\varepsilon - neighborhood$ such that \mathbf{Z}_j would be a core object. The reachability-distance of object \mathbf{Z}_i with respect to object \mathbf{Z}_j is defined as:

$$\begin{aligned}
& \text{reachability} - \text{dist}_{\varepsilon, \text{MinPts}}(\mathbf{Z}_i, \mathbf{Z}_j) \\
& = \begin{cases} \text{Undefined} & \text{if } N_{\varepsilon}(\mathbf{Z}_j) < \text{MinPts} \\ \max(\text{core} - \text{distance}(\mathbf{Z}_j), \text{dist}(\mathbf{Z}_i, \mathbf{Z}_j)) & \text{otherwise} \end{cases} \quad (2.2)
\end{aligned}$$

Where $\text{dist}(\mathbf{Z}_i, \mathbf{Z}_j)$ is the distance between \mathbf{Z}_i and \mathbf{Z}_j . The complete procedure of OPTICS is described below:

- 1) Choose one object \mathbf{Z}_j arbitrarily.
- 2) Retrieve the ε - *neighborhood* of \mathbf{Z}_j , determine the core-distance, and set the reachability-distance to undefined.
- 3) If \mathbf{Z}_j is not a core object, go to step 5. Otherwise, go to step 4.
- 4) For each object \mathbf{Z}_i in the ε - *neighborhood* of \mathbf{Z}_j , update its reachability-distance from \mathbf{Z}_j and insert \mathbf{Z}_i into an OrderSeeds list if it has not been processed yet.
- 5) If the input dataset is fully consumed and the OrderSeeds list is empty, go to step 6. Otherwise, move on to the next object in the OrderSeeds list (or the input list, if the OrderSeeds list is empty) and go to step 2.
- 6) Output core-distance, reachability-distance of each object, and processed order.

The data objects are plotted in the processed order together with their respective reachability-distance (called reachability plot) depicting the hierarchical structure of the clusters. Since objects belonging to a cluster have a low reachability-distance to their nearest neighbor, the clusters show up as valleys in the reachability plot (see Figure 2.3). The ultimate data partition can be obtained by using a threshold on the reachability plot.

c) **Density Peak Clustering (DPC):** DPC method [124] is a relatively new method which was proposed based on the idea that cluster centers can be characterized by two criteria: it has a higher density than its neighbors and a relatively large distance from objects with higher densities. Like OPTICS, this method can detect clusters with arbitrary shapes and spot outliers. Also, DPC outperforms commonly used clustering methods, e.g., k-means and hierarchical clustering, when the dataset contains complicated features such as narrow bridges between clusters, uneven-sized clusters, clusters with high overlap, etc.

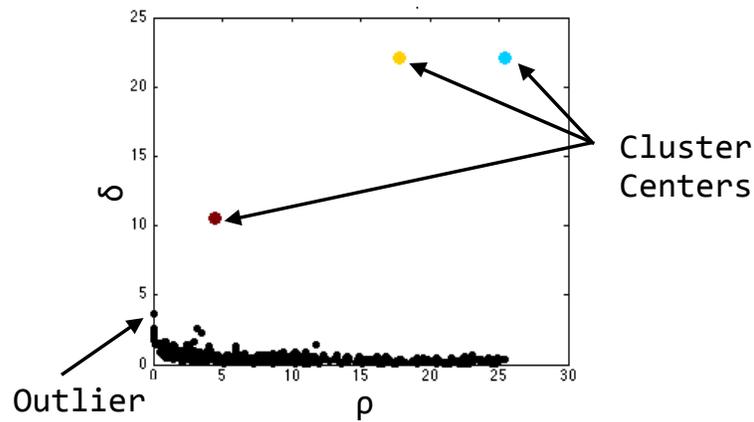


Figure 2.4 Illustration of DPC method. Plot of δ as a function of ρ for each object. Objects with larger ρ and δ are cluster centers and objects with smaller ρ , and larger δ are outliers.

For each object \mathbf{Z}_j , two quantities are computed: local density $\rho(\mathbf{Z}_j)$ and minimum distance with higher density $\delta(\mathbf{Z}_j)$. $\rho(\mathbf{Z}_j)$ is defined as:

$$\rho(\mathbf{Z}_j) = \sum_i \chi(\text{dist}(\mathbf{Z}_i, \mathbf{Z}_j) - d_c) \quad (2.3)$$

Where d_c is a cutoff distance, and $\chi(x)$ can be computed by,

$$\chi(x) = \begin{cases} 1 & \text{if } x < 0 \\ 0 & \text{otherwise} \end{cases} \quad (2.4)$$

From Equation (2.3) and Equation (2.4), it can be seen that $\rho(\mathbf{Z}_j)$ equals to the number of objects within d_c with respect to object \mathbf{Z}_j . $\delta(\mathbf{Z}_j)$ is measured by,

$$\delta(\mathbf{Z}_j) = \min_{i:\rho(\mathbf{Z}_i) > \rho(\mathbf{Z}_j)} \text{dist}(\mathbf{Z}_i, \mathbf{Z}_j) \quad (2.5)$$

For the object with highest density, $\delta(\mathbf{Z}_j)$ is conventionally set to,

$$\delta(\mathbf{Z}_j) = \max_i \text{dist}(\mathbf{Z}_i, \mathbf{Z}_j) \quad (2.6)$$

Note that if \mathbf{Z}_j is local or global maxima in the density, $\delta(\mathbf{Z}_j)$ will be much larger than its typical nearest neighbor. Thus, objects with larger ρ and δ are considered as cluster centers. Objects with smaller ρ and larger δ are considered as outliers. Other objects are assigned to the same cluster as their nearest neighbor of higher density (see Figure 2.4).

d) Input Parameter Optimization: In each clustering method, there are several user-specified input parameters, which can affect clustering results significantly. For hierarchical method, the cutting height of the dendrogram needs to be specified. For OPTICS, ϵ can simply be set to the maximum possible value, and Ankerst and colleagues [57] showed that for *MinPts* using values between 10 and 20 would always lead to good results. However, the threshold for the reachability plot, which is used to extract clusters, still needs to be properly determined. For DPC, d_c can be chosen based on the rule that the average number of neighbors is around 1-2% of the total number of objects in the data set [124]. A threshold for ρ and δ needs to be defined to distinguish cluster centers and outliers. To find the optimal value of these parameters, the

Calinski-Harabasz (CH) index [61] was applied in this work. Given a clustering result with identified K clusters. Let C_k represents the center of cluster k , where $1 \leq k \leq K$, and let C represent the center of entire data set, then the CH index is defined as:

$$CH = \frac{B/(K - 1)}{W/(N - K)} \quad (2.7)$$

Where the between-cluster variation B is computed by,

$$B = \sum_{k=1}^K \|C_k - C\|^2 \quad (2.8)$$

And the within-cluster variation W is computed by:

$$W = \sum_{k=1}^K \sum_{z_j \in \text{cluster}-k} \|z_j - C_k\|^2 \quad (2.9)$$

Based on the definition of clustering, we want to minimize W and maximize B . Thus, the optimal parameters are determined by maximizing the CH index. The optimal number of clusters can be identified, simultaneously. Specifically, for hierarchical clustering, we started with a relatively high cutting height for the dendrogram. In each iteration, the cutting height was reduced by a small amount and the CH index was computed and recorded based on the current data partition. The iteration continued until the cutting height was smaller than a specified baseline (e.g., the average height of the dendrogram). The optimal height was determined as the one with the largest CH index. The same iterative procedure was applied to OPTICS to determine optimal threshold of reachability plot, and to DPC to determine the optimal threshold

of ρ and δ . This iterative method is referred to as “grid search” [187], which has been commonly used for hyper-parameter optimization.

2.2.6 Feature Selection and Cluster Identification

The clustering accuracy is often lower in high dimensional feature space, because most of the features may be irrelevant, redundant, or sometimes may even misguide results. Moreover, a large number of features make the clustering results difficult to interpret. Therefore, it is necessary to select a minimum subset of relevant features to achieve a better clustering result. For supervised learning, feature selection can be trivial, i.e., only the features that are related to the given cluster labels are maintained. Nevertheless, for unsupervised learning, the cluster labels are unknown. Thus, finding the relevant subset of features and clustering the subset of the data must be accomplished simultaneously. To evaluate the result for both feature selection and clustering, a “similarity criteria” was also applied by computing the similarities of diagnostic clusters obtained by the following pairs of features using Torres’ method [188]: (i) clinic diagnostic labels and connectivity, and (ii) phenotypic variables and connectivity.

Let $C = \{C_1, C_2, \dots, C_m\}$ and $D = \{D_1, D_2, \dots, D_n\}$ represent two clustering results. The similarity matrix for C and D is an $m \times n$ matrix defined as:

$$S_{C,D} = \begin{bmatrix} S_{11} & \dots & S_{1j} & \dots & S_{1n} \\ \vdots & & \vdots & & \vdots \\ S_{i1} & \dots & S_{ij} & \dots & S_{in} \\ \vdots & & \vdots & & \vdots \\ S_{m1} & \dots & S_{mj} & \dots & S_{mn} \end{bmatrix} \quad (2.10)$$

Where $S_{ij} = i/u$, which is Jaccard's Similarity Coefficient with i being the size of intersection and u being the size of the union of cluster sets C_i and D_j . The similarity of clustering C and D is then defined as:

$$Sim(C, D) = \frac{\sum_{i \leq m, j \leq n} S_{ij}}{\max(m, n)} \quad (2.11)$$

From Equation (2.10) and Equation (2.11), it can be seen that $0 \leq Sim(C, D) \leq 1$, and $Sim(C, D) = 1$ when two clustering results are identical.

Assuming d to be the initial number of features, an exhaustive search of 2^d possible subsets is computationally intractable. Thus, three alternative methods are proposed in our framework, i.e., SFR, SFS, and GA, to find the optimal subset of features.

a) Sequential Feature Ranking (SFR): We propose a new feature selection method by ranking all features based on statistical significance in descending order. Each time, a feature was added to the subset from the sorted list, and the corresponding value of similarity criterion was computed. The iteration stopped when all features were added to the subset. The optimal subset of features was determined to be the one resulting in largest similarity between (i) clinic diagnostic labels and connectivity, and (ii) phenotypic variables and connectivity, respectively. SFR is derived based on an intuitive thinking that statistically significant features can lead to better clustering results. However, as the statistical significance of individual feature does not necessarily guarantee cluster separation when they are combined, SFR does not guarantee a global optimum.

b) Sequential Forward Searching of Greedy Algorithm (SFS): SFS is one of commonly used heuristic methods for finding an optimal solution [127]. Specifically, when used for feature selection, the method started with zero features in the “optimal subset” and sequentially added one feature at a time. The feature added was the one that provided the largest similarity criterion value when used in combination with the features chosen previously. The iteration continued until adding more features did not increase the similarity criterion value. Although SFS converges faster than other heuristic methods, it does not guarantee a global optimum. We used this method in comparison with GA.

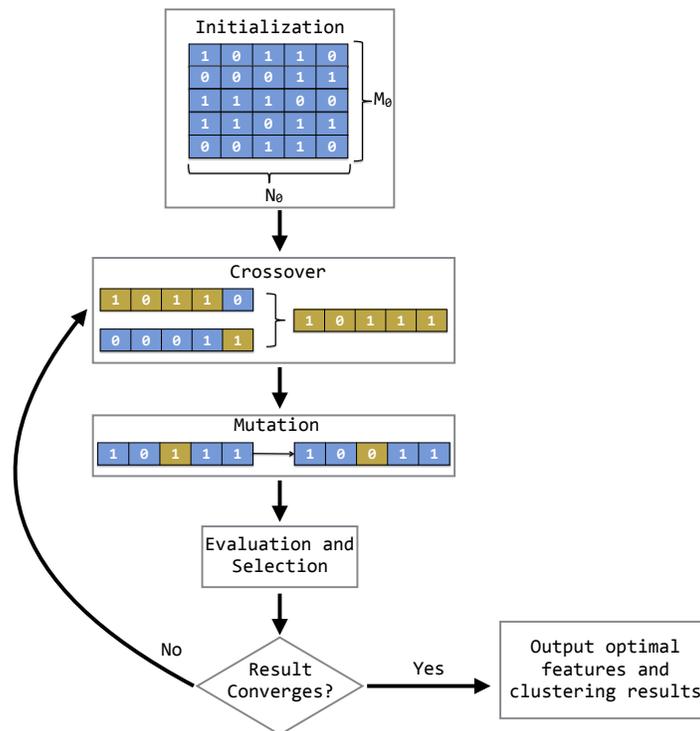


Figure 2.5 Flowchart of GA for feature selection. In the M_0 -by- N_0 matrix, each row represents a candidate solution, describing a subset of selected features. Each of the N_0 bits in a row represents whether a feature is selected (1) or discarded (0).

c) Genetic Algorithm (GA): GA is also a search heuristic method inspired by stochastic evolution theory that is routinely used in generating solutions to optimization and searching problems [189] [190]. GA maintains a set of candidate solutions (M_0) called “population”. It is an iterative process of selecting “survival solutions” for the next iteration. The survival solutions were those with larger values of similarity criteria, which were generated from the crossover and mutation operations. Crossover was applied to produce new solutions by randomly combining two current solutions, and mutation was applied on newly generated solutions with a small probability. The iteration continued until the maximum similarity among candidate solutions did not increase any further.

In this study, an array of N_0 bits was used to represent the selected subset of features, where M_0 is the total number of significant features used in clustering analysis (see Connectivity Measure). Each bit in the array indicates the activation status of one specific feature: 1 indicates selected and 0 indicated discarded. The complete procedure of GA is described below (Figure 2.5):

- 1) Initialization: 400 candidate solutions were generated by randomly setting 1 or 0 for each bit in vectors.
- 2) Crossover: two candidate solutions A and B were randomly selected from the current population. A value v between 1 and N_0 was randomly selected. Then a new solution was formed by combining the feature bits 1 to v from A and feature bits $v + 1$ to N_0 from B .
- 3) Mutation: for each new generated new solution, a mutation was applied by reversing bits in the vector with a probability of 0.1.
- 4) Evaluation: the value of similarity criteria for all candidate solutions was computed.

- 5) Selection: 280 solutions with higher similarity were selected and 120 solutions from the rest of the solutions were randomly selected to increase the diversity of the solution.
- 6) If the largest similarity did not converge, we iterated back to step 2. Otherwise, the clustering results with the largest similarity and the corresponding selected subset of features were saved as outputs.

2.2.7 Site-specific Analysis

As discussed in section 2.2.2, ADHD and ASD datasets were obtained from different sites using different scanners, which might introduce inter-site variance and affect the clustering accuracy. To eliminate this variance, site-specific feature selection and clustering were individually applied on data acquired at each site. Let $S_1 = \{F_1, F_7, \dots, F_m\}$ and $S_2 = \{F_3, F_5, \dots, F_n\}$ represent connectivity features selected by the proposed feature selection and clustering framework from site 1 and site 2, respectively. The intersection between S_1 and S_2 was then used as the new “selected features” for the whole dataset.

2.2.8 Elimination of Outlier Subjects

Real-world data always suffer from different sources of noise, which can introduce outliers in the feature space. The accuracy of clustering depends vitally on the quality of the input data. Accordingly, the most feasible and direct way to improve the effectiveness is to eliminate outlier subjects from the data.

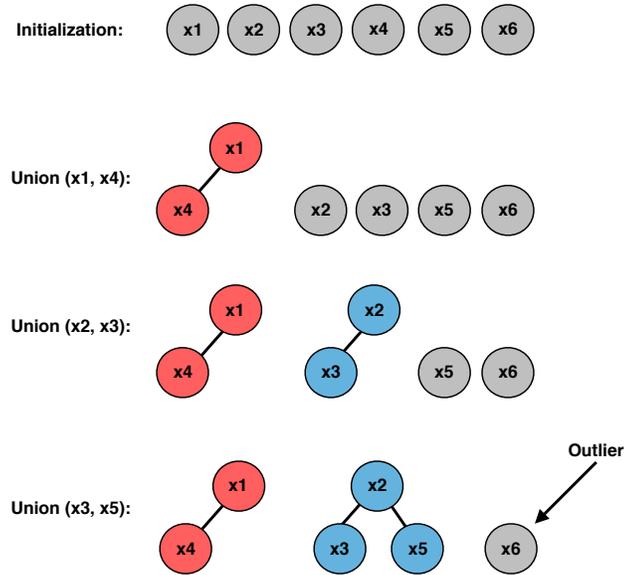
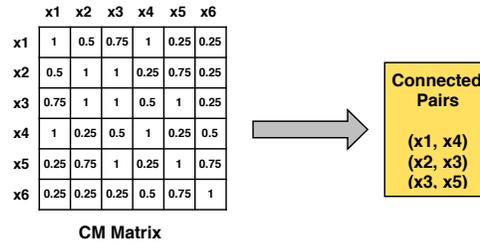


Figure 2.6 Illustration of outlier subject elimination process (using upper triangular connected pairs). x_1, x_2, \dots, x_6 are six representative subjects used in the illustration.

In this study, three different clustering methods based on three distinct principles were used for revealing hidden structures in the data. For the same input data, different clustering methods will, in general, result in different partitions in terms of the number of clusters, or the membership of clusters, and it is difficult to find a single clustering method that can handle different datasets. However, it has been demonstrated that by combining results from different clustering methods into a “co-association” matrix (CM) [191] [192], true underlying data membership can be identified. Inspired by this theory, we propose a new outlier subject elimination method by

applying the union-find algorithm [193] on the co-association matrix so that isolated outlier subjects can be identified, considered as noise in the dataset and eliminated from the analysis.

Given M different partitions for a given dataset with N subjects. The $N \times N$ co-association (CM) matrix is then defined as:

$$CM = \begin{bmatrix} CM_{11} & \dots & CM_{1j} & \dots & CM_{1N} \\ & & \vdots & & \\ CM_{i1} & \dots & CM_{ij} & \dots & CM_{iN} \\ & & \vdots & & \\ CM_{N1} & \dots & CM_{Nj} & \dots & CM_{NN} \end{bmatrix} \quad (2.12)$$

Each element in CM matrix is computed by:

$$CM_{ij} = \frac{m_{ij}}{M} \quad (2.13)$$

where CM_{ij} is the number of times subjects i and j are assigned to the same cluster among the M partitions.

With CM matrix we define subjects i and j as a connected pair with condition $CM_{ij} = 1$, which indicates that subjects i and j are always grouped together among the M partitions. Note that CM is a symmetric matrix, thus only upper triangular (or lower triangular) connected pairs need to be considered. A union-find algorithm is then applied so that connected subjects are merged together. Given N subjects and its corresponding CM matrix, the union-find algorithm is illustrated below (Figure 2.6):

- 1) Initially, each subject was considered as a singleton tree with only itself in it.
- 2) By looking up CM matrix, the connected pairs were identified.

- 3) For one connected pair a and b , where a belongs to tree A and b belongs to tree B , a “find” operation was applied to find the root of tree A and root of tree B , respectively.
- 4) A union operation was then applied to merge trees A and B .
- 5) Step 3 and 4 were repeated for all connected pairs.

The output of union-find algorithm was a set of trees, and those trees with only one node in it were considered as outlier subjects.

2.2.9 Functional Interpretation of Selected Connectivity Features—Enrichment Analysis

Functional interpretation of large-scale neuroimaging finds, e.g., brain connectivity analysis, is often done by associating identified regions or connections to previous studies. Such an approach is developed based on a subjective visual inspection or on percent of overlap with existing maps without any statistical justification. Therefore, it has potential risk to discover false positive results and overlooking additional results. In this study, to avoid these shortcomings, a novel functional interpretation method—enrichment analysis [194]—was employed, which provides a quantitative statistical measure on the association between selected connectivity features and pre-defined functional brain networks.

Let us define the following: 1) a background set S with m predefined ROIs, i.e., 200 ROIs (for ADHD, AD, and ASD) or 125 ROIs (for PTSD/PCS), and 2) a group of n selected connectivity features $A = \{(p_1, q_1), (p_2, q_2), \dots, (p_n, q_n)\}$, where each p_i and q_i represents ROIs. Two disjoint subsets of S , C and D (with size m_c and m_d), were generated by enrichment analysis, each of which constitutes a known brain network identified in previous studies. A group B was then generated with all possible ROI pairs (i.e., connectivity features) between C and D . The size of B was determined by $K = m_c \times m_d$. Let x represent the intersection between A and B . The

significance of x is the probability of having x or more elements in the intersection, which can be calculated by,

$$p = F(x|M, n, K) = \sum_{i=x}^{\min(n, K)} \frac{\binom{K}{i} \binom{M-K}{n-i}}{\binom{M}{n}} \quad (2.14)$$

where $M = \frac{m(m-1)}{2}$ is the total pairs of ROIs in the background set S . Equation (2.14) is the so called the hypergeometric (HG) cumulative distribution, which is equivalent to a one-tailed Fisher's exact test. The underlying null hypothesis of this test is that A was randomly selected from the set of all groups of ROI pairs with the same number of connectivities n over the same set of ROIs. By using this method, the statistically significant brain network-to-network (N2N) connections can be verified and quantified with corresponding p -values.

The entire pipeline for identifying different brain-based disorders along with several supplementary analyses, i.e., site-specific analysis, elimination of outlier subjects, and enrichment analysis, is illustrated in Figure 2.1.

Disease Name	DPC		OPTICS	Hierarchical
	ρ	δ	Reachability Threshold	Cutting Height
ADHD	46.93	1.22	1.03	1.15
AD	6.37	1.17	1.11	1.17
ASD	22.65	1.06	1.18	1.16
PTSD/PCS	12.05	1.17	0.42	1.16

Table 2.5 Estimated optimal values of each input parameter in clustering for Clinical vs. Connectivity comparison.

Disease Name	DPC		OPTICS	Hierarchical
	ρ	δ	Reachability Threshold	Cutting Height
ADHD	37.64	1.19	1.21	1.14
AD	6.00	1.18	1.18	1.16
ASD	26.54	1.02	1.11	1.16
PTSD/PCS	14.89	1.11	0.40	1.08

Table 2.6 Estimated optimal values of each input parameter in clustering for Phenotypic vs. Connectivity comparison.

2.3 Results

The optimal values of each input parameter determined for the three clustering methods are presented in Table 2.5 and Table 2.6. Different feature selection methods were compared in terms of peak similarity obtained for the different neuropsychiatric disorders. From Table 2.7-Table 2.10, it can be seen that the minimum subset of features selected by GA consistently resulted in highest similarity between clusters obtained from the clinical diagnoses, fMRI-based connectivity and phenotypic variables. Using GA, the average and maximum similarities between connectivity and clinical diagnosis were 80.59% and 100% respectively. The average and maximum similarities between connectivity and phenotypic variables were 76.72% and 80.38%, respectively. The average and maximum similarities between clinical diagnosis and phenotypic variables were 73.06% and 76.62% respectively. SFS was less reliable than GA in that the average and maximum similarities achieved between connectivity and clinical diagnosis were 72.20 and 100%, respectively; and the average and maximum similarities between connectivity and phenotypic variables were 66.95% and 72.22%, respectively. For similarity, the number of features determined by SFS was larger than that selected by GA. For instance, in the

PTSD/PCS dataset, although the peak similarities obtained by using SFS and GA with OPTICS were similar, the number of features selected by these two methods were 84 and 15, respectively. The similarities obtained by SFR were much lower than that obtained by SFS and GA, and the number of clusters determined using SFR was different from that using SFS and GA in all datasets. The convergence of SFR, SFS, and GA were also compared. In Figure 2.7, the similarity between connectivity and phenotypic variables obtained using hierarchical clustering and different feature selection methods was plotted as a function of the number of iterations in ADHD dataset. The shape of the curve looks comparable between connectivity and clinical diagnosis for the different clustering methods, but the amplitude may be different. With GA and SFS, a clearly step-wise convergence was observed. Although SFS converged faster than GA, a lower similarity was achieved after the curve became stable. With SFR, no clear convergence was observed (i.e., the curve oscillated dramatically).

The performance of the different clustering methods varied across the datasets. Hierarchical clustering gave higher similarity in ADHD (Table 2.7) and ASD (Table 2.8) datasets. OPTICS performed better in AD (Table 2.9) and PTSD/PCS (Table 2.10) datasets. DPC also resulted in a higher similarity in PTSD/PCS. The computation time of DPC was longer than hierarchical and OPTICS. For example, using 2.3 GHz Intel Core i7 processor, the computing time for one iteration using the PTSD dataset were as follows: hierarchical clustering took 0.27 sec, OPTICS took 0.42 sec, and DPC took 5.22 sec, due to the fact that more input parameters (ρ and δ) were required to be optimized in DPC than that in hierarchical (cutting height) and OPTICS (threshold of reachability plot). More parameters result in larger searching space.

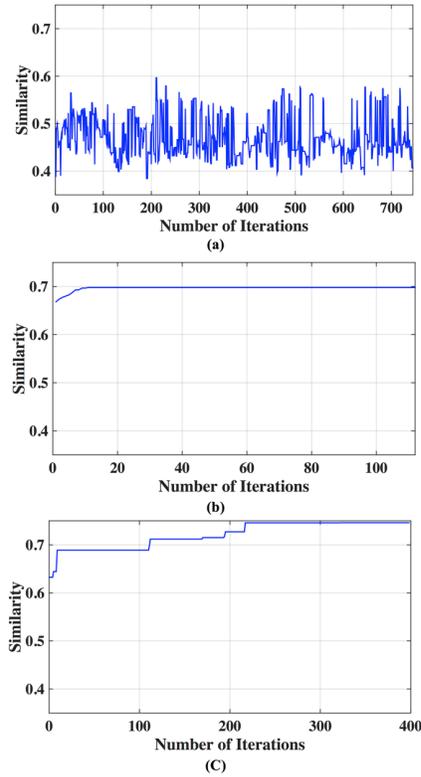


Figure 2.7 Similarity between connectivity and phenotypic variables obtained from different iterations using hierarchical clustering and different feature selection methods: (a) SFR, (b) SFS, and (c) GA.

Feature Selection Method	Clustering Method	Clinical v.s. Phenotypic			Clinical v.s. Connectivity			Phenotypic v.s. Connectivity		
		Peak Sim. (%)	# of Features	# of Clusters	Peak Sim. (%)	# of Features	# of Clusters	Peak Sim. (%)	# of Features	# of Clusters
SFR	DPC	66.19	2	3	48.91	3	2	49.94	3	2
	OPTICS	64.19	2	2	56.10	4	3	58.22	4	3
	Hierarchical	61.41	3	2	51.38	195	2	51.84	18	2
SFS	DPC	66.19	2	3	58.36	37	3	68.30	17	3
	OPTICS	64.19	2	2	58.40	13	3	66.46	26	3
	Hierarchical	67.23	3	3	59.32	54	3	69.79	51	3
GA	DPC	73.51	4	3	62.48	72	3	63.83	63	3
	OPTICS	70.52	3	3	64.44	122	3	64.42	25	3
	Hierarchical	69.29	5	3	69.34	83	3	74.61	121	3

Table 2.7 Peak similarity (highlighted), corresponding number of features, and number of clusters obtained using SFR, SFS, and GA with different clustering methods for ADHD dataset.

Feature Selection Method	Clustering Method	Clinical v.s. Phenotypic			Clinical v.s. Connectivity			Phenotypic v.s. Connectivity		
		Peak Sim. (%)	# of Features	# of Clusters	Peak Sim. (%)	# of Features	# of Clusters	Peak Sim. (%)	# of Features	# of Clusters
SFR	DPC	64.46	2	4	56.68	61	4	57.99	175	3
	OPTICS	64.16	2	3	66.35	84	4	60.48	84	4
	Hierarchical	64.46	2	4	67.63	241	4	57.20	241	4
SFS	DPC	57.05	3	4	59.26	2	4	56.30	4	4
	OPTICS	65.93	4	4	64.03	20	4	57.58	12	4
	Hierarchical	65.14	4	5	64.82	155	4	60.26	193	5
GA	DPC	62.53	3	4	63.52	44	4	74.55	97	4
	OPTICS	65.93	4	4	64.23	116	4	76.75	58	4
	Hierarchical	65.14	4	5	66.65	74	4	73.28	124	5

Table 2.8 Peak similarity (highlighted), corresponding number of features, and number of clusters obtained using SFR, SFS, and GA with different clustering methods for AD dataset.

Feature Selection Method	Clustering Method	Clinical v.s. Phenotypic			Clinical v.s. Connectivity			Phenotypic v.s. Connectivity		
		Peak Sim. (%)	# of Features	# of Clusters	Peak Sim. (%)	# of Features	# of Clusters	Peak Sim. (%)	# of Features	# of Clusters
SFR	DPC	66.67	9	2	58.91	5	3	61.23	5	3
	OPTICS	74.75	10	3	62.14	10	3	64.53	99	3
	Hierarchical	75.81	9	3	61.04	681	2	60.64	336	2
SFS	DPC	66.20	2	2	64.41	33	3	63.79	42	3
	OPTICS	76.00	8	3	64.50	144	3	61.75	2	3
	Hierarchical	76.00	5	3	64.79	85	3	65.52	143	3
GA	DPC	75.85	5	3	72.45	60	3	68.98	50	3
	OPTICS	76.46	5	3	79.18	103	3	74.47	70	3
	Hierarchical	76.63	5	3	89.20	101	3	75.66	54	3

Table 2.9 Peak similarity (highlighted), corresponding number of features, and number of clusters obtained using SFR, SFS, and GA with different clustering methods for ASD dataset.

Feature Selection Method	Clustering Method	Clinical v.s. Phenotypic			Clinical v.s. Connectivity			Phenotypic v.s. Connectivity		
		Peak Sim. (%)	# of Features	# of Clusters	Peak Sim. (%)	# of Features	# of Clusters	Peak Sim. (%)	# of Features	# of Clusters
SFR	DPC	63.82	6	2	66.67	9	2	61.38	9	2
	OPTICS	69.20	4	3	66.67	16	2	76.15	18	3
	Hierarchical	61.89	6	2	66.67	6	2	57.31	4	2
SFS	DPC	74.25	3	3	73.01	2	3	70.40	4	3
	OPTICS	69.76	7	3	100	84	3	66.39	2	3
	Hierarchical	69.35	3	3	75.30	2	3	72.22	2	3
GA	DPC	74.25	3	3	100	40	3	77.25	25	3
	OPTICS	76.15	4	3	100	15	3	80.38	8	3
	Hierarchical	70.15	5	3	77.65	2	3	64.70	1	3

Table 2.10 Peak similarity (highlighted), corresponding number of features, and number of clusters obtained using SFR, SFS, and GA with different clustering methods for PTSD dataset.

Site Name	HC	Autism	Asperger's
Caltech	19	13	0
CMU	13	14	0
Pitt	27	30	0
UCLA	45	44	0
SDSU	23	2	5
Trinity	24	10	6
NYU	105	53	21

Table 2.11 Number of subjects provided by each site in the ASD sample.

Site-specific analysis was applied on the ADHD dataset. We could not apply this analysis on ASD dataset since there was only one site that had enough samples for HC and disease subgroups (Table 2.11) whereas AD and PTSD datasets were obtained on the same scanner. For ADHD dataset, NYU and Peking had more than 30 samples for control, ADHD-C, and ADHD-I (Table 2.12). Thus, a site-specific analysis was applied on these two sites, individually.

The peak similarity obtained between clinical diagnostic and phenotypic clusters, between clinical diagnostic and connectivity clusters, and between phenotypic and connectivity clusters for site-specific analysis are shown in Table 2.13. Compared with previous results presented in Table 2.7 using feature selection and clustering on the entire dataset across different sites, the similarity was increased by applying site-specific analysis for Peking and NYU, individually. The similarity was reduced by applying clustering on the whole datasets with commonly selected features from these two sites.

Site Name	HC	ADHD-C	ADHD-I
NYU	98	73	43
Peking	116	29	49
KKI	58	16	5

Table 2.12 Number of subjects provided by each site in the ADHD sample.

Site Name	Clinical v.s. Phenotypic			Clinical v.s. Connectivity			Phenotypic v.s. Connectivity		
	Hier.	OPTICS	DPC	Hier.	OPTICS	DPC	Hier.	OPTICS	DPC
NYU	74.32	75.87	77.02	100	100	79.38	87.02	87.02	74.55
Peking	78.99	66.67	82.18	74.66	89.02	88.83	82.32	89.84	88.42
Whole Dataset	63.27	65.14	66.95	62.92	58.37	54.69	71.54	59.38	54.79

Table 2.13 Similarity achieved using data from individual sites and for the whole dataset using features commonly selected by NYU and Peking.

For ADHD and AD, highest similarity was achieved between connectivity and phenotypic clusters and the corresponding similarity between clinical diagnostic and phenotypic clusters was lower. On the other hand, for ASD and PTSD/PCS, highest similarity was achieved between connectivity and clinical diagnostic clusters. This suggests that diagnostic criteria for ASD and

PTSD/PCS are mapped well onto underlying neurobiological clusters, while that was not the case for ADHD and AD. Consequently, for ADHD and AD, we reassigned diagnostic labels based on those generated by connectivity clusters to form new neurobiologically-informed groups. In order to verify whether this new grouping is valid, we estimated the statistical separation of phenotypic variables based on the traditional diagnostic grouping as well as with the new neurobiologically-informed groups. The results shown in Figure 2.8 and Figure 2.9 indicate that almost all p-values were smaller with the new grouping by applying 2-sample t-test. This suggests that when traditional diagnostic groups do not map well onto underlying neurobiological clusters, connectivity can be used to regroup the subjects so that they map better onto the behavioral phenotypes.

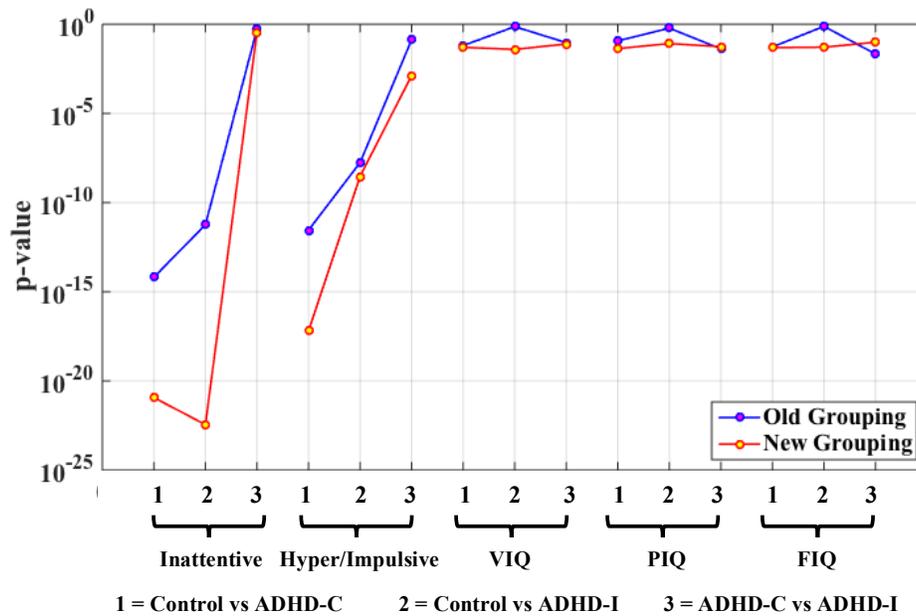


Figure 2.8 Statistical significance (p-value) of selected phenotypic/genotypic variables with both the old clinical diagnostic grouping and the new connectivity grouping. The results are shown here for the ADHD dataset. Logarithmic scale is used for the y-axis of p-values.

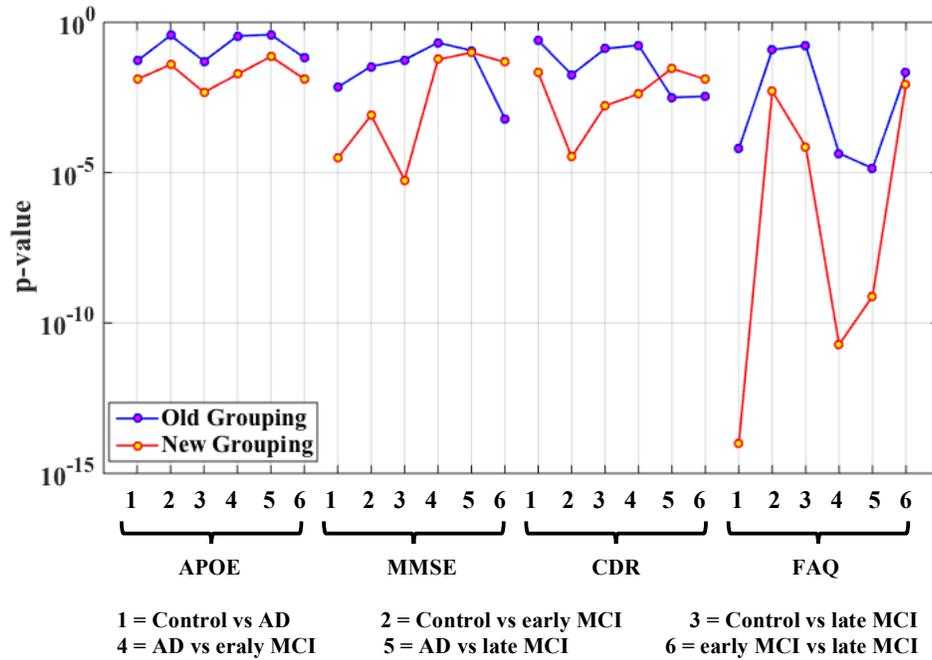


Figure 2.9 Statistical significance (p-value) of selected phenotypic/genotypic variables with both the traditional clinical diagnostic grouping and the new connectivity-based grouping. The results are shown here for the AD dataset. Logarithmic scale is used for the y-axis of p-values.

The peak similarity obtained with and without outlier subject elimination was compared and is shown in Table 2.14. Consistently higher similarity was achieved by removing the identified outlier subjects from the dataset. Moreover, in AD dataset, the number of clusters identified by hierarchical clustering was changed from 5 to 4 with outlier elimination (highlighted in Table 2.14), which matched with the grouping obtained using clinical diagnosis. The data in ADHD and ASD datasets comprised of data acquired at different sites using different scanners, which might explain the fact that the number of outliers identified in ADHD and ASD were generally greater than the other two datasets.

Disease Name	Clustering Method	With Outlier Subjects Removed								Including All Subjects			
		Clinical v.s. Connectivity				Phenotypic v.s. Connectivity				Clinical v.s. Connectivity		Phenotypic v.s. Connectivity	
		p	p (%)	sim	k	p	p (%)	sim	k	sim	k	sim	k
ADHD	DPC	54	11.09	69.72	3	38	7.80	64.86	3	62.48	3	63.83	3
	OPTICS			69.79	3			68.92	3	64.44	3	64.42	3
	Hierarchical			69.62	3			75.66	3	69.34	3	74.61	3
AD	DPC	14	14.58	65.33	4	9	9.38	76.70	4	63.52	4	74.55	4
	OPTICS			65.88	4			77.75	4	64.23	4	76.75	4
	Hierarchical			69.68	4			78.38	4	66.65	5	73.28	5
ASD	DPC	15	3.30	72.65	3	49	10.79	71.95	3	72.45	3	68.98	3
	OPTICS			79.83	3			73.63	3	79.18	3	74.47	3
	Hierarchical			89.84	3			76.05	3	89.20	3	75.66	3
PTSD	DPC	3	3.45	100	3	5	5.57	77.57	3	100	3	77.25	3
	OPTICS			100	3			81.54	3	100	3	80.38	3
	Hierarchical			80.60	3			72.54	3	77.65	3	64.70	3

Table 2.14 Comparison of peak similarity obtained with and without elimination of outlier subjects. p: number of outliers, k: number of clusters, sim: clustering similarity. In AD dataset, the number of clusters identified by hierarchical clustering was changed from 5 to 4 with outlier elimination (highlighted), which matched with the grouping obtained using clinical diagnosis.

2.4 Discussion

In this work, we have proposed a general analysis pipeline for characterizing different neuropsychiatric disorders using unsupervised learning methods. First, we discuss the selected brain connectivity features and phenotypic variables for each disorder and compare our results with previous studies. Second, elaborate on the implications of results obtained within specific sites in comparison to those obtained from the entire ADHD dataset. Third, we discuss the

reassignment of diagnostic labels based on those generated by connectivity clusters. Finally, delineate the role of outlier subject elimination in unsupervised learning methods as applied to neuroimaging.

2.4.1 Connectivity Features Important for Clustering

After applying clustering, the selected connectivity features were split into two networks, i.e., 1) a network in which functional/effective connectivities and temporal variability of constituent paths were significantly ($p < 0.05$, FDR corrected) larger in the control group, and 2) a network in which functional/effective connectivities and temporal variability of the constituent paths were significantly ($p < 0.05$, FDR corrected) larger in the disease group. Here, “disease group” refers to all pathological subgroups combined. This was done since all disease groups have two or more pathological sub-groups and it becomes increasingly complex to interpret all pairwise differences. Then, these two networks were mapped back to the image space and overlaid on an anatomical glass brain [using BrainNet Viewer [195]] for the visualization, respectively. The identified brain networks were then qualitatively interpreted and compared with previous studies using enrichment analysis [194].

Intrinsic connectivity networks (ICNs) denote groups of brain regions that show correlated spontaneous activities at “resting” state [196]. It has been shown that ICNs reflect strong coupling of spontaneous fluctuations in ongoing activity and remain robust under different mental states, e.g., sleep, loss of consciousness, etc. [197]. ICNs provide a common neurofunctional framework for investigating cognitive dysfunction in different neuropsychiatric disorders. There are many stable ICNs that have been identified in the human brain so far. Five of them—default mode network (DMN), visual network (VN), basal ganglia network (BGN),

sensory motor network (SMN), and the semantic cognition and attention (SCAN)—have been demonstrated to be particularly important for understanding higher cognitive function and dysfunction, and provide useful models for identifying rs-fMRI connectivity patterns. Below, we discuss the significance of each of these networks to provide a context for presenting alterations in the interactions within and between these networks observed in neuropsychiatric disorders.

DMN is one of the most well-known ICNs, which is a distributed network anchored in the posterior cingulate cortex [PCC], the medial prefrontal cortex [mPFC], the medial temporal lobe [MTL], the precuneus, the anterior cingulate cortex [ACC], the inferior parietal lobe [IPL], and the medial orbital gyrus [MOG] [198]. PCC, hippocampus, and angular gyrus are typically associated with episodic memory retrieval [199] [200], autobiographical memory [201], and semantic memory related to internal thought [202]. mPFC has been demonstrated to be associated with self-related and social cognitive processes [203], value-based decision making [204], and emotion regulation [205]. Together, the entire DMN comprises an integrated system involving episodic memory, autobiographical memory, and self-related mental processes.

VN [206] involves the occipital and bilateral temporal regions including the middle occipital gyrus, the inferior temporal gyrus [ITG], the fusiform gyrus, the cuneus, etc., which is involved in visual processing and mental imagery [207] [208]. The middle occipital gyrus, ITG, and the fusiform gyrus are primarily involved in the higher functions of vision processing, e.g., distinguishing objects among different categories, face recognition, visual words recognition, representation of complex object features, etc. [209] [210]. The cuneus has been demonstrated to be involved in basic visual processing, which receives visual information from retina [211].

BGN is predominantly located in the basal ganglia including the striatum (which is subdivided into the caudate nucleus and putamen), the globus pallidus or pallidum, the substantia nigra and the thalamus [212]. BGN is associated with a variety of functions including control of voluntary motor movements [213], procedural learning, eye movements [214], cognition [215], emotion [216], etc.

SMN involves the precentral gyrus, the postcentral gyrus, cerebellum, the posterior insula, and part of the frontal gyrus corresponding to the primary sensory motor cortex and the supplementary motor area [SMA] [217] [218]. Studies have indicated that this network is processing somatosensory stimuli, executing motor movements and sensorimotor integration [219] [220].

SCAN is defined as regions associated with semantic cognition network and attention network, which is a network of lateral structures in frontal and parietal cortices, as well as some temporal regions. The semantic cognition network is primarily made up of three regions, Broca's area, Wernicke's area, as well as parts of the middle temporal gyrus [MTG] [221] [222]. Broca's area is generally defined as comprising Brodmann areas 44 and 45. Area 44 (the posterior part of the inferior frontal gyrus [IFG]) is involved in phonological processing and language production whereas area 45 (the anterior part of the IFG) engages in the semantic aspects of language. Together, Broca's area plays an important role in processing of verbal information [223]. Wernicke's area is traditionally thought to be located in the posterior part of the superior temporal gyrus [STG], which is involved in the comprehension or understanding of written and spoken language [224]. Some studies have showed that MTG is involved in the retrieval of lexical syntactic information [225]. The attention network is commonly segregated into two

distinct networks: a bilateral dorsal attention network (DAN), which includes the dorsal frontal and parietal cortices, and the ventral attention network (VAN), largely right-lateralized, which includes the ventral frontal and parietal cortices [199] [226]. DAN has been associated with goal-directed, top-down attention processes as inhibitory control, working memory and response selection, whereas the VAN is related with salience processing and mediates stimulus-driven, bottom-up attention processes [227]. Moreover, it is relevant to note that dorsal and ventral systems appear to interact not only during cognitive tasks [228] [229] but also during spontaneous activity [230]. Previous literature suggests that semantic cognition and attention are intimately related. This is also borne out by the fact that many disorders such as ADHD and ASD have simultaneous deficits in semantic cognition and attention. Therefore, we considered this as one network.

A qualitative as well as quantitative interpretation of alterations of these INCs and other related brain regions in different neuropsychiatric disorders are discussed below. For each pathology, we chose the features that gave us highest similarity between clusters obtained from clinical labels, connectivity features and phenotypic features. For ADHD and AD, highest similarity was obtained between connectivity and phenotypic clusters while for ASD and PTSD/PCS, highest similarity was obtained between clinical labels and connectivity clusters. Therefore, the features obtained in these two different scenarios have different implications. For ADHD and AD data sets, it suggests that traditional clinical diagnostic grouping may not neatly map onto neurobiological and neurobehavioral clusters. This may be because of uncertainty in clearly identifying differences between disease sub-groups in ADHD (ADHD-C and ADHD-I) and AD (EMCI, LMCI and AD). Contrarily, for ASD and PTSD/PCS data sets, it suggests that traditional clinical diagnostic grouping may in fact map well onto at least neurobiological clusters.

- DMN
- VN
- BGN
- SMN
- SCAN

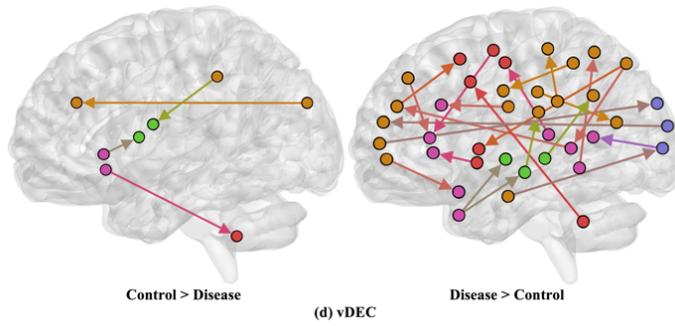
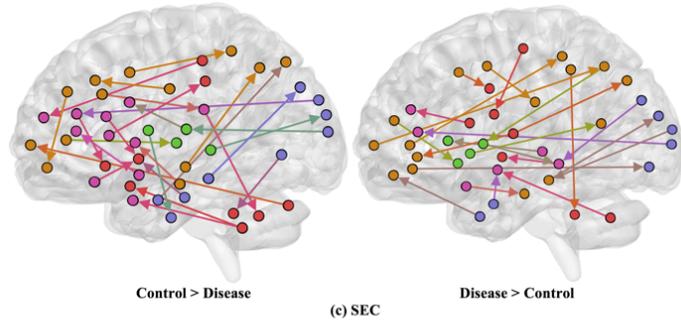
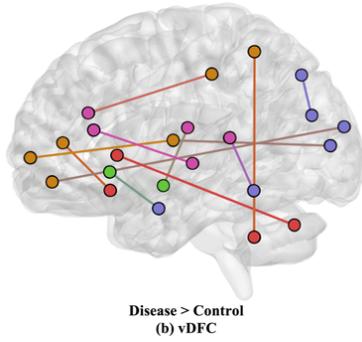
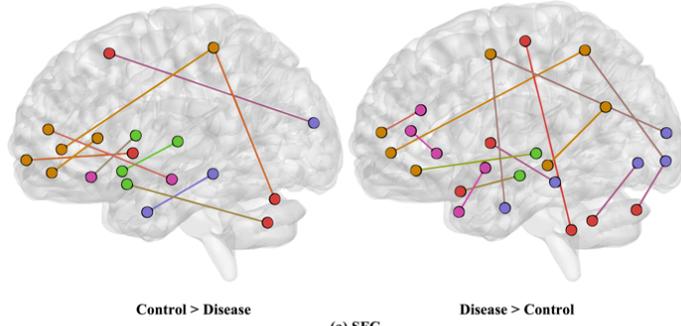


Figure 2.10 SFC, vDFC, SEC, and vDEC features selected by GA and hierarchical (ADHD). Selected features were split into two groups, i.e., 1) control > disease (ADHD-C and ADHD-I) and 2) disease > control. DMN: Default mode network, VN: Visual network, BGN: Basal ganglia network, SMN: Sensory motor network, SCAN: Semantic cognition and attention network.

These facts are borne out by computing the purity of clusters obtained from connectivity features for disease sub-groups within each data set. To measure cluster purity, the clusters obtained using connectivity features were regrouped using the diagnostic label, and each object was assigned to majority class in the current cluster. Then the accuracy was measured by counting the number of correctly assigned objects within each cluster and took the average. The cluster purity for ADHD, AD, ASD, and PTSD/PCS were 0.73, 0.75, 0.94, and 1.00, respectively. It can be seen that ASD and PTSD/PCS data sets had high purity while for ADHD and AD, the purity of clusters for disease subgroups was comparatively lower.

a) ADHD: 121 relevant connectivity features were selected by GA and hierarchical clustering (since this combination gave highest similarity between connectivity and phenotypic features), which were 26 SFC, 14 vDFC, 53 SEC, and 28 vDEC. These features include connections in all lobes of the brain (Figure 2.10). With enrichment analysis, two N2N interactions were selected for SFC, i.e the interactions within BGN and the interaction between VN and SMN, including connections between cerebellum and occipital lobe, between insula and fusiform, and between caudate and thalamus. In addition, two N2N interactions were selected for SEC, i.e. from BGN to VN, and from SCAN to SMN, including connections from caudate to occipital lobe and ITG,

from IFG and middle frontal gyrus [MFG] to posterior insula, from IFG to postcentral gyrus, and from STG to cerebellum (Table 2.15).

Feature Type	Comparison	P-Value	Selected Features
SFC	Control > Disease	0.006	BGN—BGN
SFC	Disease > Control	0.04	VN—SMN
SEC	Control > Disease	0.03	BGN→VN
SEC	Control > Disease	0.007	SCAN→SMN

Table 2.15 Network-to-network interactions selected by enrichment analysis for ADHD dataset.

Most of the rs-fMRI studies have demonstrated atypical functional activations in frontal, temporal, parietal lobes, and cerebellar regions [231] [232] [233] in ADHD. Multiple studies have found aberrant functional connectivity among the brain regions of the DMN, SCAN, and BGN [234] [235] [236] [237]. Abnormal functional activations in the orbitofrontal cortex [OFC] have been suggested to influence behavioral inhibition in children with ADHD [238]. Resting-state fMRI studies have frequently reported disrupted functional connectivity between ACC and PCC in ADHD [239] [236]. Significantly decreased activations have been reported in PFC, SPL, and IFG in ADHD, during multiple cognitive performance tasks and in resting-state [240] [241] [242]. One fMRI study conducted in adults with childhood ADHD showed reduced activations in bilateral IFG, left parietal lobe, caudate, and thalamus [243]. Another study found reduced functional connectivity between thalamus and other BGN areas (e.g., putamen, caudate) with ADHD [237]. Some studies have also identified reduced activations in IFG [232] and STG [244] in ADHD patients. Kessler and colleagues [245] observed reduced connectivity between SCAN and SMN and increased connectivity within VN by applying joint independent component analysis on the ADHD-200 sample. On the other hand, increased functional connectivity in

DMN, BGN, SMN, and VN has been observed in some studies [246] [240]. Significantly increased functional connectivity between ACC and the thalamus, cerebellum, and insula have been shown during resting-state in children with ADHD, compared to controls [247] [236] [248]. Li and colleagues [249] found increased connectivity between the right pulvinar and occipital regions, during a visual sustained attention task-based fMRI study. Hale and colleagues [250] also observed reduced activations in VN and DMN, during letter and location judgment tasks. The features selected by GA for maximizing the similarity between connectivity and phenotypic clusters and the subset of significant N2N interactions determined by enrichment analysis are in agreement with previous literature implicating the very same regions and connections in ADHD.

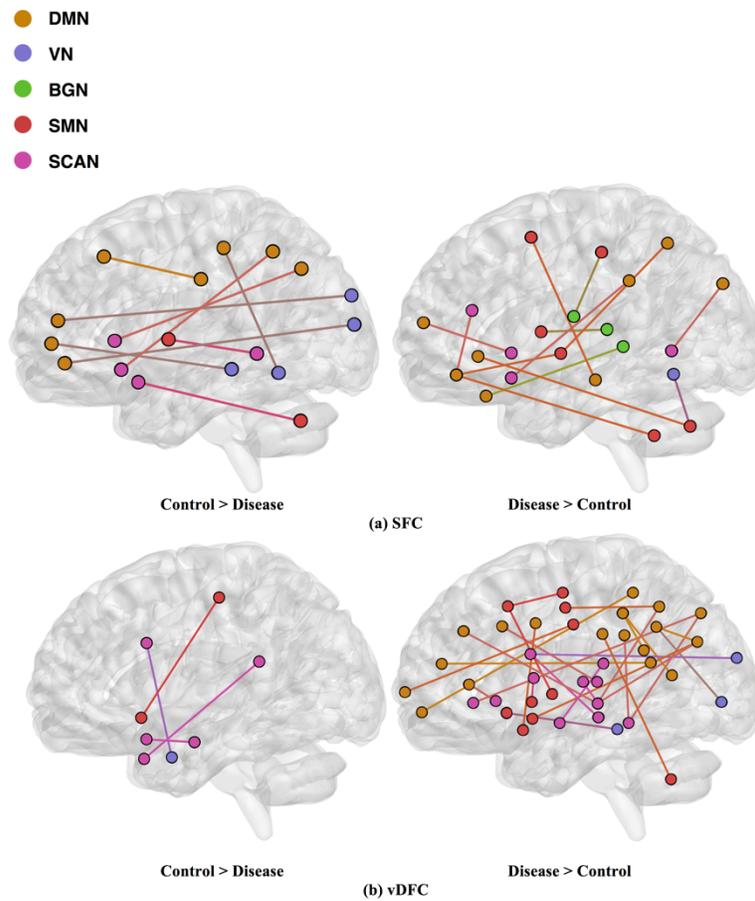


Figure 2.11 SFC and vDFC features selected by GA and hierarchical (AD). Selected features were split into two groups, i.e., 1) control > disease (AD, LMCI and EMCI) and 2) disease > control. DMN: Default mode network, VN: Visual network, BGN: Basal ganglia network, SMN: Sensory motor network, SCAN: Semantic cognition and attention network.

b) AD: 58 features were selected by GA and OPTICS (since this combination gave highest similarity between connectivity and phenotypic features), including 32 vDFC features and 26 SFC features. Most of features were related to DMN, VN, SMN, and SCAN (Figure 2.11). With enrichment analysis, two N2N interactions were selected for SFC, i.e., the interaction between DMN and SMN, and that between DMN and VN, including connections between ACC and middle occipital gyrus, between PFC and fusiform, between IPL and ITG, between SFG and insula, between hippocampus and SMA, between cerebellum and SFG, and between cerebellum and PFC. In addition, two N2N interactions were selected for vDFC, i.e., the interactions within SCAN, and between DMN and SCAN, including connections between MTG and STG, between PFC and IFG, between precuneus and IFG, between precuneus and MTG, between PFC and STG, and between MTG and IPL (Table 2.16).

Feature Type	Comparison	P-Value	Selected Features
SFC	Disease > Control	0.036	DMN—SMN
SFC	Control > Disease	0.01	DMN—VN
vDFC	Control > Disease	0.01	SCAN—SCAN
vDFC	Disease > Control	0.04	DMN—SCAN

Table 2.16 Network-to-network interactions selected by enrichment analysis for AD dataset.

Several previous studies have indicated dysfunctions in different regions of DMN, VN, SMN and SCAN in the AD and MCI populations [251] [252]. Some studies have observed decreased

connectivity in DMN coupled with an increased connectivity within prefrontal regions [253] [254] [255]. Significant alterations of connectivity in MTG, the PCC, hippocampus, and the angular gyrus, have been observed in AD [256], [200]. The dysfunction in the MTG, which is referred to as a central hub of SCAN [257], is suggested as an early feature of AD [258]. A lesser degree of MTG activation has been observed in MCI [259] [260] compared to controls. The medial parietal cortex, including the PCC and precuneus, are selectively vulnerable to amyloid deposition in AD [257], and studies of cortical metabolism using positron emission tomography and single photon emission computed tomography in AD suggest that abnormalities in the PCC and precuneus are early features of AD [258]. A voxel-based study showed that AD patients had both decreased activity of the right MFG and an increased activity of the right parietal cortex [261]. Reduced connectivity in the temporal lobe was also observed in different rs-fMRI studies [252] [254]. Multiple studies have suggested that the insula is involved in AD [262] [263] [264] and some of the behavioral abnormalities in AD may reflect insular pathology. Brier and colleagues [116] observed reduced anti-correlations between DMN and SMN, and between DMN and SCAN, during a resting state fMRI study. Li and colleagues [265] also found aberrant connectivity between DMN and SCAN, as well as between DMN and SMN. These previous studies seem to support our findings regarding features which are important for unsupervised clustering of control, EMCI, LMCI and AD groups.

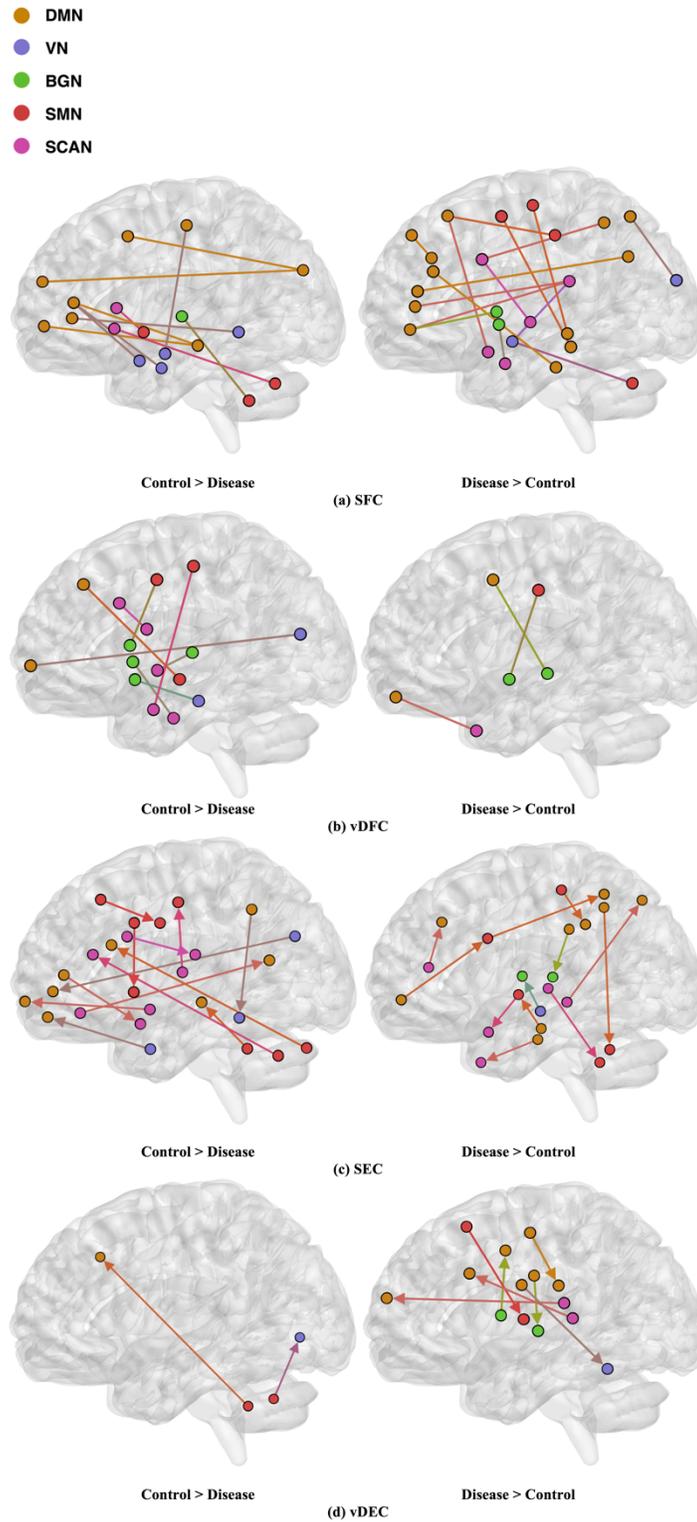


Figure 2.12 SFC, vDFC, SEC, and vDEC features selected by GA and hierarchical (ASD). Selected features were split into two groups, i.e., 1) control > disease (autism and asperger's) and 85

2) disease > control. DMN: Default mode network, VN: Visual network, BGN: Basal ganglia network, SMN: Sensory motor network, SCAN: Semantic cognition and attention network.

Feature Type	Comparison	P-Value	Selected Features
SFC	Control > Disease	0.033	DMN—DMN
SFC	Control > Disease	0.015	DMN—VN
vDFC	Control > Disease	0.04	BGN—SCAN
SEC	Disease > Control	0.025	DMN→SMN

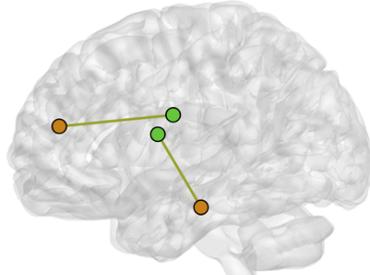
Table 2.17 Network-to-network interactions selected by enrichment analysis for ASD dataset.

c) ASD: 76 features were selected using GA and hierarchical (since this combination gave highest similarity between connectivity features and clinical diagnosis)—30 SFC, 11 vDFC, 27 SEC, and 9 vDEC—involving frontal, parietal, temporal lobes, and cerebellar regions (Figure 2.12). With enrichment analysis, two N2N interactions were selected for SFC, i.e., the interaction within DMN, and between DMN and VN, including connections between PFC and angular gyrus, between SFG and angular gyrus, between ACC and parahippocampal gyrus, between MOG and parahippocampal gyrus, between ACC and fusiform, between SFG and ITG, and between ACC and ITG. One N2N interaction was selected for vDFC, i.e., the interaction between BGN and SCAN, including connections between caudate and MTG, and between thalamus and STG. In addition, one N2N interaction was selected for SEC, i.e., from DMN to SMN, including connections from MOG to precentral, from hippocampus to posterior insula, and from precuneus to cerebellum (Table 2.17).

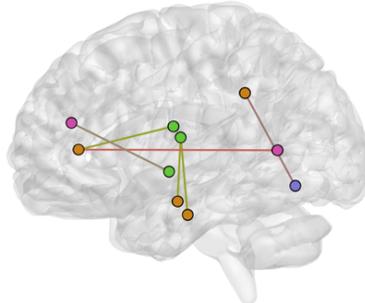
Several recent studies have observed abnormal connectivity in DMN, SCAN, SMN, BGN and VN in the pathophysiology of ASD [266] [267]. A recent meta-analysis showed alterations in MTG, hippocampus, as well as the posterior medial cortex in ASD [268], which were suggested

to be related to deficits in social information processing. It has been shown that PCC and mPFC in ASD are hypoactive compared with healthy controls [269]. Decreased connectivity between the PCC and the SFG, the PCC and the temporal lobes, as well as the PCC and the parahippocampal gyri were observed, which were associated with poor social skills [270]. Dysfunction in the SCAN has been shown to be related to deficits in language and communication in individuals with ASD. Reduced activation and functional connectivity in frontal-temporal SCAN were observed by Mody et al [271]. A recent rs-fMRI study found a marked loss of functional connectivity between the right cerebellar region and regions in SCAN [272]. Weaker connection between the SMA and ventral premotor cortex was found in the ASD group compared with controls, which has been hypothesized to underlie the initiation of speech motor actions [273]. Decreased connectivity between BGN and the occipital region and prefrontal cortical regions was also found by Prat et al. [274]. A meta-analysis identified posterior insula as a consistent locus of hypoactivity in ASD [269]. Other fMRI studies have also suggested that insula is one possible key dysfunctional area in ASD [275]. In contrast, a recent rs-fMRI study [276] observed stronger functional connectivity within several large-scale brain networks in children with ASD compared with controls, including DMN, SCAN, SMN, BGN, and VN. It has been suggested that developmental trajectories in ASD can be both heterogeneous and aberrant compared to neurotypicals and hyper- or hypo-connectivity is observed depending on when the data is acquired during development [276]. Our results are in broad agreement with previous fMRI literature in ASD discussed above.

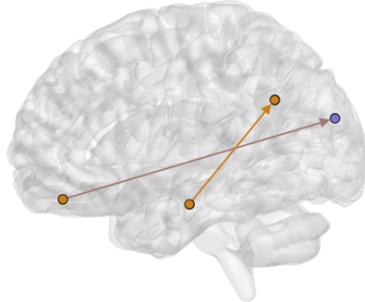
- DMN
- VN
- BGN
- SMN
- SCAN



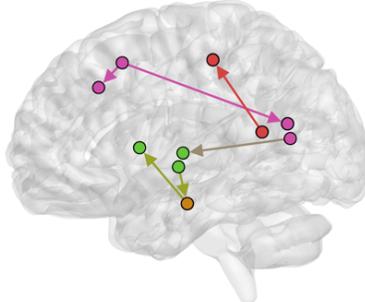
Disease > Control
(a) SFC



Control > Disease
(b) vDFC



Disease > Control
(c) SEC



Control > Disease
(d) vDEC

Figure 2.13 SFC, vDFC, SEC, and vDEC features selected by GA and hierarchical (PTSD/PCS). Selected features were split into two groups, i.e., 1) control > disease (PTSD and PTSD/PCS) and 2) disease > control. DMN: Default mode network, VN: Visual network, BGN: Basal ganglia network, SMN: Sensory motor network, SCAN: Semantic cognition and attention network.

d) PTSD/PCS: 15 features were selected by GA and OPTICS (since this combination gave highest similarity between connectivity features and clinical diagnosis); 2 SFC, 5 vDFC, 2 SEC, and 6 vDEC. These features were mainly located in DMN, BGN, and SCAN (Figure 2.13). With enrichment analysis, one N2N interaction between DMN and BGN was selected for both SFC and vDFC. This involved connections between ACC and caudate, and between parahippocampal gyrus and caudate. In addition, one N2N interaction from DMN to BGN was selected for vDEC, which included the connection from parahippocampal gyrus to caudate (Table 2.18).

Feature Type	Comparison	P-Value	Selected Features
SFC	Disease > Control	0.001	DMN—BGN
vDFC	Control > Disease	< 0.001	DMN—BGN
vDEC	Control > Disease	0.016	DMN→BGN and DMN←BGN

Table 2.18 Network-to-network interactions selected by enrichment analysis for PTSD dataset.

Several resting-state studies of PTSD have showed aberrant connectivity within brain structures associated with DMN [277] [278] [279]. The parahippocampal gyri and hippocampus are critical structures in the DMN, which have been shown to be essential for memory functions, especially memorizing facts and events, and memory consolidation [280]. A previous rs-fMRI study found decreased functional connectivity in the hippocampal regions in PTSD patients

[281]. BGN has also been reported to be associated with PTSD [282] [283] [284]. PTSD has been linked with abnormal activation of different BGN regions, brain stem, and limbic regions compared with control group [107] [285] [286]. The connectivity between DMN and BGN and between DMN and VN have been observed to be impacted in PTSD in many functional connectivity studies [287] [288]. Lanius and colleagues [287] found increased connectivity between ACC and caudate, PCC, the right parietal lobe, and the right occipital lobe, during a resting-state fMRI study using subjects with PTSD. Stark and colleagues found changes in connectivity between DMN and BGN, e.g., connections between ACC and caudate, between parahippocampal gyri and caudate, etc., by applying a systematic, quantitative meta-data analysis on previous studies. SCAN has also been demonstrated to be linked to PTSD. Reduced connectivity was observed in MTG, MFG and several BGN regions in the PTSD group, compared with controls [289]. Yin and colleagues [290] also found reduced connectivity in MTG and lingual gyrus, during a resting-state fMRI study. It is interesting to note that increased static connectivity and reduced variability of dynamic connectivity between the hippocampal formation and BGN regions such as caudate has been recently reported in PTSD and PCS [291] and our results seem to confirm these findings and show that those aberrations are important for unsupervised clustering of subjects into these groups.

From above discussions, it can be seen that for each individual neuropsychiatric disorder, connectivity features selected by GA with optimal clustering method are consistent with previous studies, which suggest the effectiveness of our general pipeline for identifying different brain-based disorders using unsupervised learning.

2.4.2 Phenotypic Features Important for Clustering

The phenotypic variables important for clustering were selected for each psychiatric disease. Below, we discuss the relevance of these variables in the context of existing literature on those measures.

a) ADHD: Four phenotypic variables were selected by GA and DPC including ADHD index score, Inattentive score, Hyper/Impulsive score (all are subscales in ADHD-RS), and FIQ in intelligence scale. ADHD-RS has been considered as an effective clinical diagnostic tool for assessing the severity of ADHD in children and adolescents [292] [293]. It gathers information on the severity and frequency of symptoms, the establishment of childhood onset of symptoms, the chronicity and pervasiveness of symptoms, and the impact of symptoms on major life activities. Intelligence scale has been demonstrated to be helpful in predicting symptomatology and outcome in children with ADHD [294]. A meta-analysis showed that FIQ was lower in adults with ADHD compared to HC [140].

b) AD: Three phenotypic variables, i.e., MSE, CDR, and FAQ, and one genotypic variable, i.e., APOE were selected by GA and OPTICS. APOE is considered as the major genetic risk factor for AD [146]. Although the presence of APOE does not necessarily entail the development of AD, this genetic isoform probably accelerates the rate of AD conversion and progression [295]. The MMSE is the most commonly used instrument for screening memory problems and other deficits related to cognitive aging. It has been widely used to screen for dementia [141]. CDR is a global scale developed to clinically denote the presence of AD and stage its severity [296]. Several methods have been derived based on CDR to identify AD accurately [297] [298]. FAQ is a standardized assessment of instrumental activities of daily living, which delineates the clinical distinction between MCI and AD [299].

c) ASD: Five phenotypic variables were selected including ADOS_TOTAL, ADOS_COMM and ADOS_SOCIAL (all makeup ADOS test), FIQ in intelligence scale, and ADI-R_VERBAL in ADI_R test. ADOS has been extensively used in the clinic for diagnosing ASD [300] [301]. It consists of a series of structured and semi-structured presses for an interaction of specific target behaviors associated with particular tasks and by general ratings of the quality of behaviors. Further, several studies have observed higher VIQ and FIQ in ASD compared to neurotypicals [302] [303]. ADI-R is a structured interview conducted with the parents of the referred individual and covers the subject's full developmental history [304]. The communication and language score, as one of the three content areas in ADI-R, is useful in assessing the presence and severity of delay or total lack of language.

d) PTSD/PCS: Four phenotypic variables—SDC correct, ZDS, CES, and LEC—were selected by GA and OPTICS. SDC is a test of psychomotor performance, visual-motor coordination, sustained attention, and motor and mental speed, which has been shown to be related to PTSD [305]. ZDS is a short self-administered survey to quantify the depressed status of a patient. Burriss and colleagues (2008) showed that PTSD was associated with general learning and memory impairments, and depression was considered as a mediator of these deficits. [306] In addition, Dretsch and colleagues (2012) revealed that depressive symptoms in individuals with PTSD account for working memory impairments. CES was constructed to measure the subjective report of wartime stressors experienced by combatants [307]. It has been demonstrated that CES is a useful tool for diagnosing PTSD [308]. LEC, a measure of exposure to potentially traumatic events, was developed for diagnosing PTSD as well. In a clinical sample of combat veterans, a significantly correlated relationship between LEC and PTSD symptoms was observed [309].

It can be seen from the discussion above that the phenotypic (and genotypic in case of AD) variables selected by GA for maximizing the similarity of clusters obtained from them and from connectivity features indicates that they are clinically meaningful and relevant to the behavioral deficits observed in each disorder.

2.4.3 Site-specific Analysis

Modern machine learning systems often integrate data from several different sources. Usually, these sources provide data of a similar type but collected under different circumstances. For example, the ADHD dataset used in this study was collected from different sites. Although fMRI images provided by these sites had similar qualities, these images were obtained from different scanners with different scanning parameters. The accuracy of machine learning algorithms can be affected by the heterogeneity of input data. To address this issue, we performed a site-specific analysis. By applying feature selection and clustering on data obtained from each individual site, the cluster similarity was increased considerably (see Table 2.7 and Table 2.13). However, as when we applied clustering on whole dataset with commonly selected features from individual sites, the similarity was reduced. Due to inter-site variance, it is difficult for us to translate high accuracy obtained for individual site into the whole dataset. It also affects the diagnostic precision obtained from brain connectivity measures. This calls for data acquisition standards and homogenization of data acquired from different scanners.

2.4.4 Connectivity-based Reassignment of Diagnostic Labels

Many brain-based disorders are highly heterogeneous, and categorization of subgroups within many disorders is yet to be completely established. Traditionally, brain-based disorders are diagnosed by clinical interviews associated with different behavioral assessments. However, it is

widely acknowledged that current clinical criteria are insufficient to clearly identify most of the brain-based disorders, separate them from healthy subjects and identify sub-groups within them. Therefore, it is necessary to develop brain imaging based models for understanding how, precisely, neural circuits generate flexible behaviors and their impairments give rise to psychiatric symptoms [91]. In this study, we used unsupervised learning algorithms to discover brain connectivity-based clusters, which were not limited to existing diagnostic criteria. Instead, it focused on separating subjects into isolated clusters with maximized inter-cluster variance and minimized intra-cluster variance. After clustering, we reassigned diagnostic labels based on those generated by connectivity clusters. Compared with clinical diagnostic groups, the neurobiologically-informed groups provided better mapping from subjects to the behavioral phenotypes. This result indicates that it might be possible to view brain-based disorders from the perspective of brain connectivity measures, establishing neuroimaging-based biomarkers for different neuropsychiatric disorders.

2.4.5 Outlier Subject Elimination

The overarching aim of healthcare is personalized medicine. However, basing individualized treatments on brain imaging characteristics is in the nascent stages, i.e., some subjects will deviate considerably from the normative population distribution and it becomes easier to assess population level characteristics when such subjects are eliminated from the analysis. As shown in this study, with the proposed subject outlier elimination process, the precision of clustering was improved. Note that, the inter-individual variability may be introduced not just by the variability in the underlying neuropathology, but also by non-neural sources of variance such as different scanners and/or different scanning parameters. Until a standard data acquisition process is

established, outlier subject elimination will serve to homogenize the data and make better inferences at the population level.

2.5 Conclusion

Many neuropsychiatric disorders are conventionally diagnosed based on clinical interviews and behavioral assessments. Inherent limitations of specific measures and clinical judgement contribute to a far from perfect process. Therefore, it is necessary to establish neuroimaging-based biomarkers to improve diagnostic precision and accuracy. Rs-fMRI has been used as a promising technique for characterization and classification of different disorders. However, these approaches are besieged with methodological issues such as (i) *a priori* choice of clusters needed in k-means, (ii) a stopping criterion needed in hierarchical clustering, (iii) the large dimensionality of imaging data necessitates some type of dimensionality reduction for clustering to work properly and this step is either not carried out, or carried out by preselecting features not from the structure in the data, but by some external considerations such as previous findings in a given disorder, and (iv) the clusters obtained from imaging data are seldom compared by those obtained from clinical diagnostic criteria or behavioral phenotypes.

To address these four issues, a general pipeline was derived on identifying different brain-based disorders using unsupervised clustering methods. In addition, site-specific analysis and elimination of outlier subjects were also applied to improve clustering accuracy. Three selected clustering methods were adopted on three types of features: 1) fMRI connectivity measures, 2) clinical diagnostic labels, and 3) phenotypic variables. GA based feature selection method was also applied to improve clustering accuracy. The accuracy of the clustering and feature selection was assessed by computing the similarity of clusters between all three types of features. The

effectiveness of proposed pipeline was verified on five different disorders: ADHD, AD, ASD, PTSD and PCS. For ADHD and AD, highest similarity was achieved between connectivity and phenotypic clusters, whereas for ASD and PTSD/PCS, highest similarity was achieved between connectivity and clinical diagnostic clusters. The result suggests that neurobiological and phenotypic biomarkers could potentially be used as an aid by the clinician, in addition to currently available subjective markers, to improve diagnostic precision.

Chapter 3

Noninvasive Characterization of Layer-specific Perforant Pathway between Entorhinal Cortex and Hippocampus using Ultra High Field fMRI at 7T

The medial temporal lobe houses several brain structures necessary for learning and memory. Two of these structures, the hippocampus and the entorhinal cortex (EC), are connected via the perforant pathway. Specifically, the signal from the superficial layers of the EC project to DG/CA4 and CA3, and from there to CA1 and the subiculum. In this study, we investigated this pathway non-invasively in humans using ultra-high-field functional magnetic resonance imaging (fMRI) data acquired during memory encoding. We found that the functional connectivity between EC layer II and hippocampus, when clustered in an unsupervised way, parcellated the hippocampus into proximal and distal regions along perforant pathway. This parcellation was based on our observation of stronger connectivity between layer II of EC with hippocampal subfields such as DG/CA4/CA3/CA2 which are proximal to the EC along the perforant pathway, compared to subfields such as CA1/Subiculum which are distal. Further, this pattern was true more for the left, rather than the right, hippocampus. Our results provide the first direct non-invasive functional evidence for the perforant pathway in humans.

3.1 Introduction

The medial temporal lobe is an important brain region consisting of the hippocampus and the entorhinal cortex (EC) among several other regions, and plays important roles in various brain functions including episodic memory and spatial navigation [310] [311] [312]. The medial temporal lobe has been implicated in many neuropsychiatric and neurological disorders,

including Alzheimer's disease [313] [314] [315], major depression [316] [317] [318] [319], post-traumatic stress disorder [320] [321], and schizophrenia [322] [323] [324]. Therefore, a better understanding of the functional microcircuits connecting the hippocampus and EC within the medial temporal lobe has the potential to lead to transformative advances in our understanding of these disorders.

Like the cerebral cortex, the hippocampus is a paired structure, with mirror-image halves in the left and right sides of the brain. The hippocampus consists of many subfields including the cornu ammonis (CA) and the subiculum. The CA is composed of thin stacked layers that are distinguishable from each other histologically by their differing cell types and cell densities [325] [326]. These layers are the stratum pyramidale, which is further sub-divided into CA1, CA2, CA3, and DG (or CA4), stratum radiatum (SR), stratum lacunosum (SL), and stratum moleculare (SM) [325]. Note that, CA4 is in fact the deep, polymorphic layer of the DG [327]. Thus, in this study, DG and CA4 are considered as one subfield in the hippocampus.

The entorhinal cortex (EC), also located in the medial temporal lobe, constitutes the major gateway between the hippocampus and neocortex [328]. As part of the hippocampal memory system, the EC plays an important role in declarative memories and spatial memories [329]. Similar to other cortical regions, the EC is organized into cytoarchitectonically distinct layers [328]. The superficial layer (layer I) is relatively free of neurons, which contains a dense band of transversely oriented fibers. The outermost cell layer (layer II) mainly contains so-called "stellate" or "modified pyramidal cells", which are fairly large, making them distinctly different from the adjacent cortical regions. Layer III is a wide layer of loosely arranged, large to medium sized cells that are predominantly of the pyramidal type. The layer V is clearly stratified and

sometimes subdivided into a superficial layer of large to medium-sized, darkly stained pyramidal cells, which is referred to as layer IV in primates. Subsequent deeper portions of layer V have an overall stratified appearance and mainly consist of rather small pyramidal cells with moderately dense packing. The deepest layer, i.e., layer VI, is delineated by the white matter [328].

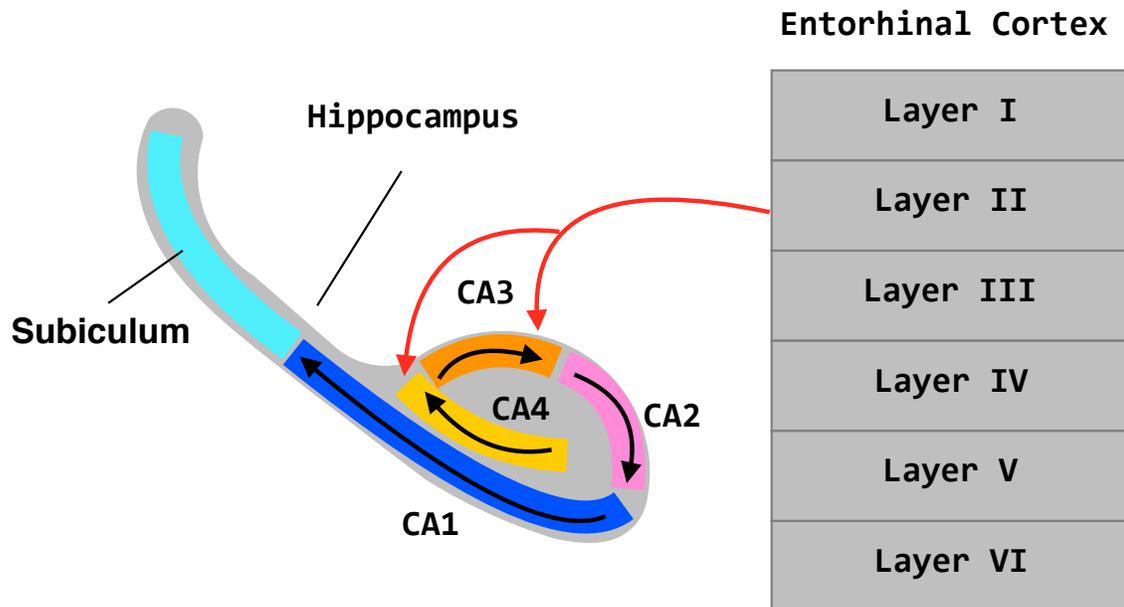


Figure 3.1 Perforant pathway from the entorhinal cortex and the hippocampus.

Subfields of the hippocampus are connected to different cortical layers of EC through anatomically well-characterized circuits [330]. During an encoding task, the major pathway from the EC to the hippocampus is the perforant path, wherein layer II of EC projects to DG (also referred to as CA4) and CA3 and from there to CA1 and subiculum [331] (Figure 3.1). While other pathways between EC and hippocampus do exist (such as the “indirect” pathway from layer III of EC to CA1 and Subiculum) [332], the perforant pathway is the most dominant one, and hence the one likely to be detectable with fMRI. The involvement of EC and hippocampus in the perforant pathway has been supported by many previous studies [330] [331] [333] [334]. Suh

et al. [333] applied spatial working-memory tasks on a transgenic mouse and showed that the projection to CA3 and DG in mice is primarily from layer II of EC. Witter et al. [334] studied the anatomical organization of the hippocampal memory system in rats and confirmed the perforant pathway. Nevertheless, our understanding of the EC-hippocampal structural wiring as well as functional layer-specific microcircuits underlying memory encoding in the perforant pathway has primarily relied on invasive techniques such as anatomical tract tracing and single unit recordings in animals, respectively. While such approaches are clearly useful, being able to carry out such investigations using non-invasive methods such as functional magnetic resonance imaging (fMRI) allows us to characterize EC-hippocampal functional microcircuitry in both healthy and clinical human populations. However, such attempts have been extremely scarce in the literature given that it is feasible only using very high spatial resolution images acquired at ultra-high fields (≥ 7 Tesla) coupled with an appropriate analysis strategy. Recently, Maass et al. [331] measured encoding-related activation in layers of the EC and sub-regions of the hippocampus using high resolution fMRI at 7T. They provided evidence that activation in superficial layers (layer II and probably parts of layer III) of EC was specifically correlated with activity in DG/CA2-3, as well as with the SR/SL/SM of CA1, the latter being compatible with input from the EC to CA1 via the perforant pathway. Furthermore, they demonstrated that the perforant pathway input from EC to hippocampus was associated with “novelty response” (stronger fMRI activation for novel as compared with familiar stimuli) during encoding process.

Notwithstanding the study discussed above, non-invasive modalities including fMRI have typically lacked the spatial resolution to resolve layers at conventionally used field strengths (3 Tesla). Here, we overcome this issue by acquiring fMRI data with very high resolution at ultra-high fields (7 Tesla). This technique provides several advantages over conventional field

strengths, such as improvements in spatial resolution from few millimeters to sub-millimeter, increases signal-to-noise ratios (SNR) [335] [336], increase in T_2^* contrast [337] [338], etc. Importantly, sub-millimeter voxels make it feasible to reasonably resolve different cortical layers, albeit with some partial voluming effects. Several fMRI studies on laminar specific activation have been reported, both in animals [339] [340] [341] and in humans [342] [343] [344] [338] [345]. In contrast to investigating the BOLD response (aka activation) in specific layers, our focus is on the layer specific connectional architecture of the EC-hippocampal microcircuits.

The present study sought to investigate the functional perforant pathway using functional connectivity between superficial layers of EC with different subfields of the hippocampus during an encoding task. Specifically, we hypothesized that functional connectivity between layer II of EC and the hippocampus, when clustered in an unsupervised way, must parcellate the hippocampus into proximal and distal regions along the perforant pathway, with stronger connectivity between layer II of EC with hippocampal subfields such as DG/CA4/CA3/CA2 (which are proximal to the EC along the perforant pathway) compared to subfields such as CA1/Subiculum (which are distal to the EC along the perforant pathway). In order to test these hypotheses, fMRI data was acquired from healthy subjects performing a memory encoding task involving objects, faces and words in the 7T scanner. Ultra-high resolution anatomical data was also acquired to resolve layers in the EC. Functional connectivity between layer II of EC and hippocampal voxels was computed. In order to unravel connectivity patterns between EC layers and hippocampal subfields in a data driven way, different unsupervised clustering methods were applied on functional connectivity estimates. The main idea of clustering is to group objects in such a way that objects in the same group are more similar to each other than to those in other

groups. Three clustering methods were specifically chosen, i.e., hierarchical clustering [123], ordering points to identify the clustering structure (OPTICS) [57], and density peak clustering (DPC) [124]. These methods were specifically chosen since they did not require *a priori* specification of the number of clusters. The commonly used k-means clustering [125] [126] method was not considered in this study due to the uncertainty of the number of clusters (which needs to be pre-specified) and sensitivity of the method to outliers. Since clustering accuracy is often lower in high dimensional feature space, feature selection methods were employed. Most existing feature selection algorithms in the machine learning literature focus on heuristic search such as sequential forward searching (SFS) [127], non-linear optimization [128], genetic algorithm (GA) [127], etc. Bradley et al proposed non-linear optimization using a nonlinear kernel support vector machine. Although this method provides high accuracy, it can only be used in the supervised learning context. SFS was proposed based on a greedy algorithm, which follows the problem-solving heuristic of making the locally optimal decision at each step. Although SFS can be applied in unsupervised learning, it does not guarantee a globally optimal solution. Therefore, we propose GA as a robust feature selection method for selecting specific functional connections between EC layers and hippocampal voxels which can then be clustered by unsupervised learning approaches.

It is noteworthy that our task did not involve spatial memory typically used in navigation, which preferentially recruits the right hippocampus [346]. Rather, our task involved factual memories of objects, pictures of scenes and words, which likely preferentially recruits the left hippocampus [346] [347] [348]. Therefore, we predicted that our hypothesis will hold true more for the left, rather than the right, hippocampus during an encoding task.

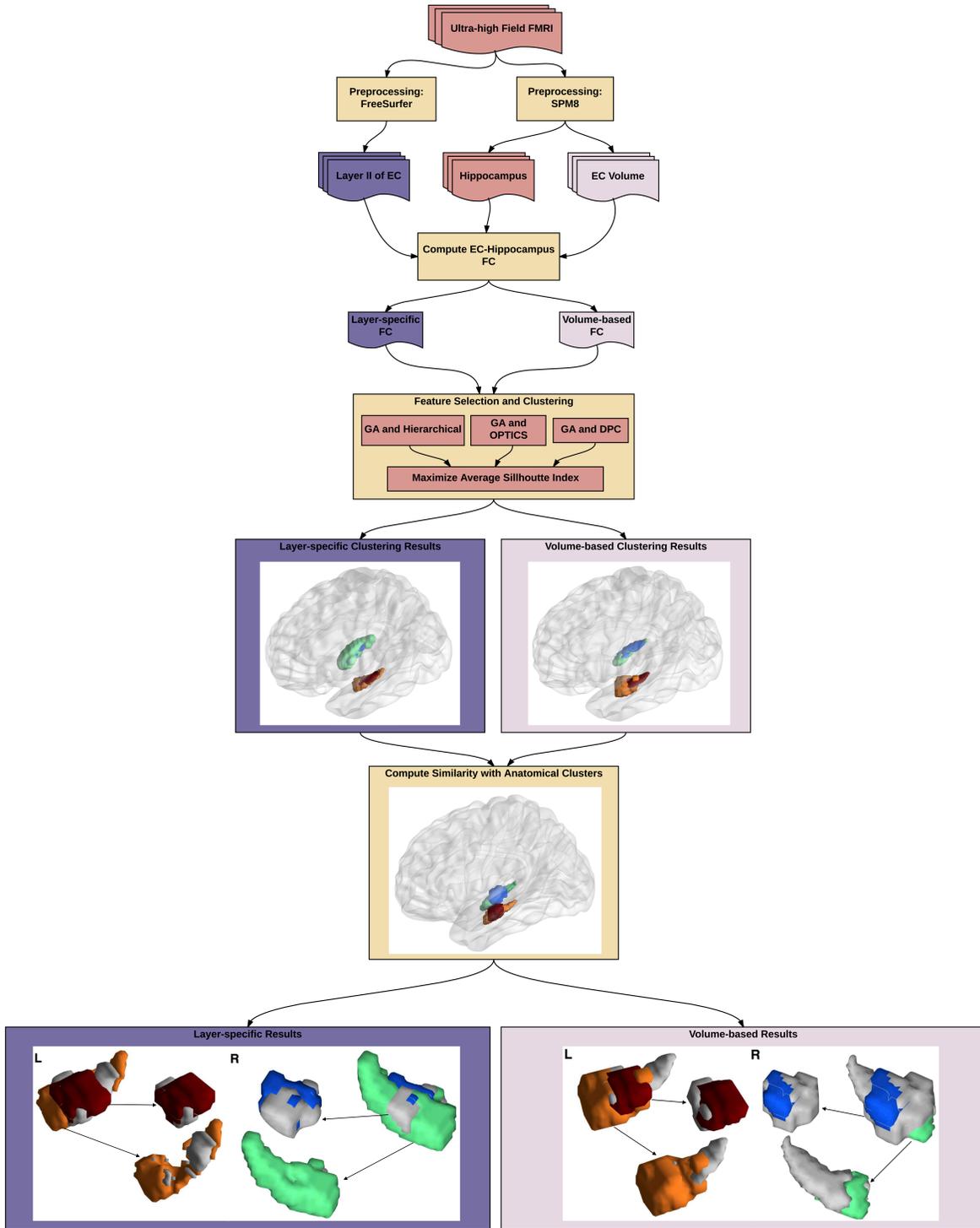


Figure 3.2 Illustration of proposed analysis pipeline for investigating the perforant pathway between layer II of EC and subfields of the hippocampus.

3.2 Materials and Methods

In this work the functional perforant pathway was investigated by applying unsupervised clustering methods on the functional connectivity between layer II of the EC with different subfields of the hippocampus during an encoding task. The identified functional clusters of the hippocampus were then compared with well-known anatomical parcellation of the hippocampus. The entire analysis pipeline is illustrated in Figure 3.2 and will be elaborated below.

3.2.1 Data Acquisition

Thirty-one healthy individuals (26 right-handed, 12 males, 19 females, age = 21.13 ± 1.43) were recruited for the study. The Internal Review Board (IRB) at Auburn University approved the study, subjects provided informed consent and the experimental procedures were performed in accordance with internationally accepted ethical standards. Echo-planar imaging (EPI) data were acquired on the Auburn University MRI Research Center (AUMRIRC) Siemens 7T MAGNETOM scanner outfitted with a 32-channel head coil by Nova Medical (Wilmington, MA). The sequence was optimized for the hippocampus (37 slices acquired parallel to the AC-PC line, $0.85\text{mm} \times 0.85\text{mm} \times 1.4\text{mm}$ voxels, TR/TE: 3000/28ms, 70° flip angle, base/phase resolution: 234/100, A→P phase encode direction, iPAT GRAPPA acceleration factor = 3, interleaved acquisition, 123 time points, total acquisition time of 6 minutes). During encoding task, the participants were asked to view a series of faces, pictures of scenes and words. Each trial lasted for 30 seconds in which the subjects were presented 10 images of the same category for 3 seconds each. These trials were interspersed with a 6 second inter-trial interval. A whole-brain high-resolution 3D MPRAGE sequence was used to acquire anatomical data (256 slices, $0.63\text{mm} \times 0.63\text{mm} \times 0.60\text{mm}$, TR/TE: 2200/2.8, 7° flip angle, base/phase resolution 384/100%,

collected in an ascending fashion, acquisition time = 14:06) for extracting different cortical layers in the entorhinal cortex and for registration purposes.

3.2.2 Preprocessing

a) Hippocampal data: Standard pre-processing steps were carried out using SPM8 [349] including brain extraction, slice timing correction, temporal band-pass filtering (0.01 to 0.10), regression of motion and physiological artifacts (using CompCor), registration to anatomical space, and normalization to MNI standard space.

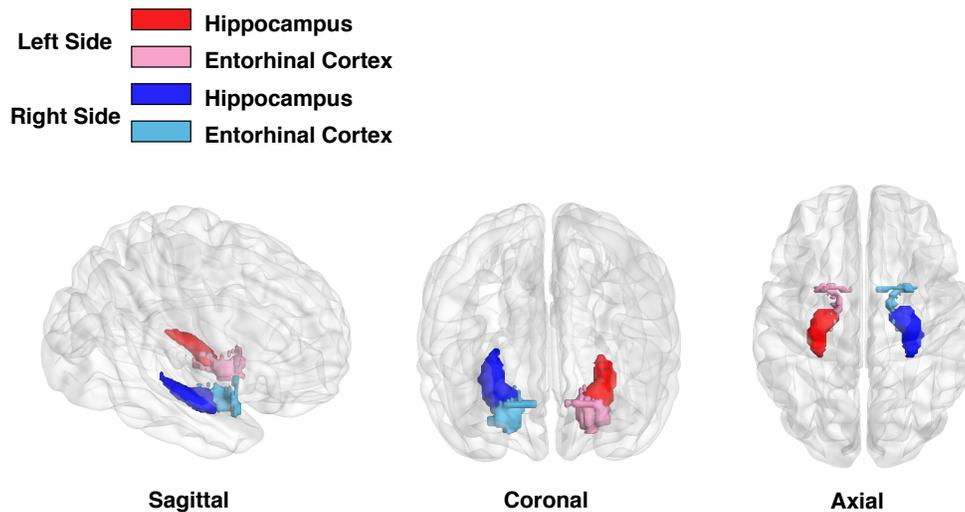


Figure 3.3 Hippocampus and EC ROIs used in this study.

The left and right hippocampal ROIs were defined using the Harvard-Oxford Structural Probability Atlas distributed with the FSL neuroimaging analysis software package (<http://www.fmrib.ox.ac.uk/fsl/fslview/atlas-descriptions.html#ho>). Each ROI was thresholded at 75%, which yields an accurate anatomical representation. The mean probability for the voxels in the left ($M \pm SD$: $86.41\% \pm 7.10\%$) and right hippocampal ROIs ($87.75\% \pm 7.39\%$), which belong to the hippocampus, was over 87%. The probability that the voxel centroid for each ROI

belonged to the left (MNI coordinates [-26, -18.8, -17.2]) and right (MNI coordinates [27.52, -18.2, -16.8]) hippocampus was 97% and 97.3%, respectively. The total volume for the left and right hippocampus was 1880 mm³ and 2072 mm³, respectively (Figure 3.3).

b) Cortical layer-specific EC data: Data extracted from the whole brain was first preprocessed using SPM8. To preserve high spatial resolution, no spatial filtering was applied. After intensity normalization, registration to MNI space, and skull stripping, the interface between the cortical gray matter and the underlying white matter (white-gray interface) and the interface between the cortical gray matter and the pial surface (gray-pial interface) were automatically generated from the anatomical image using FreeSurfer [350]. In order to reconstruct cortical layers, cortical thickness maps need to be derived. The cortical thickness was calculated as the average of the distance from the white-gray interface to the closest possible point on the gray-pial interface, then from that point back to the closest point on the white-gray interface again. To improve accuracy, surface smoothing and automatic topology correction were also applied [351] [352] [353] [354]. Six cortical layers were then reconstructed within the cortical gray matter at fixed relative distances between the white and pial surfaces determined from the cortical thickness, i.e., the first layer was located at 96% of the cortical thickness away from the white matter, the second layer at 80%, the third layer at 64%, the fourth layer at 48%, the fifth layer at 32%, and the sixth layer at 16% (Figure 3.4). The laminar layers were derived from the anatomical image, so it was necessary to align the EPI volume to these layers. Since the layer specific analysis of EC is the primary goal in this study, a boundary-based registration method [355] was employed, which aligns the EPI image to the anatomical image by maximizing the intensity gradient across interfaces (i.e., white-gray interface, gray-pial interface). The entire cortical surface was automatically divided into 34 cortical ROIs in each of the individual

hemispheres based on the Desikan-Killany atlas in Freesurfer [356]. Since EC is one of those 34 regions in the atlas, vertices in layer II of the entorhinal cortex were then identified. It is noteworthy that voxels in the volume become vertices on surfaces.

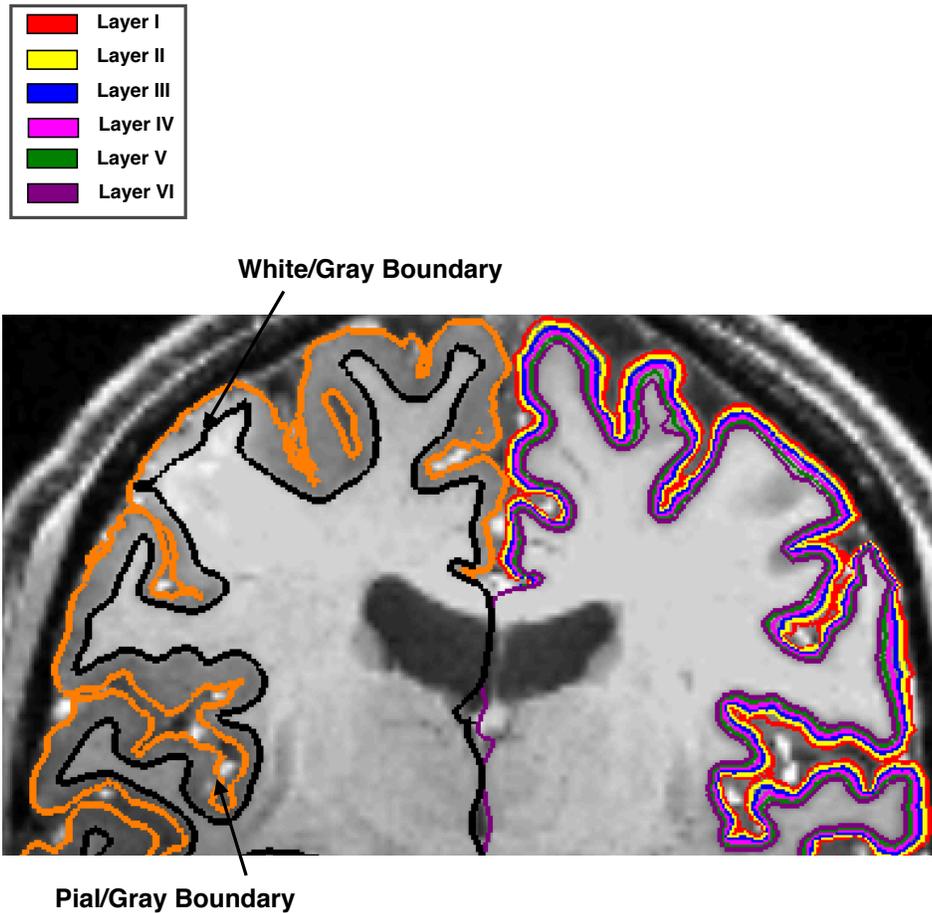


Figure 3.4 Cortical layer reconstruction with FreeSurfer. The surface reconstructions, i.e, the white-gray interface (black) and gray-pial interface (orange), and cortical layers reconstruction. Six cortical layers were reconstructed within the cortical gray matter at fixed relative distances between the white and pial surfaces determined from the cortical gray matter.

3.2.3 Connectivity Measures

Functional connectivity (FC) was estimated between the hippocampal voxels and vertices in layer II of EC using time series extracted from them in individual subjects. FC measures the functional interrelationship between pairs of brain regions by estimating Pearson's correlation.

3.2.4 Layer Specific Clustering

Three clustering methods were applied on the FC computed between the hippocampal voxels and vertices in layer II of EC. The same process was repeated on the left and right hemisphere, separately. Let $\mathbf{X} = \{\mathbf{X}_1, \dots, \mathbf{X}_j, \dots, \mathbf{X}_N\}$ represent a set of N objects, i.e., number of hippocampal voxels. $\mathbf{X}_j = (X_{j1}, X_{j2}, \dots, X_{jd}) \in \mathbb{R}^d$, where d equals to the number of FC features. Let us assume that the N objects are to be partitioned into k clusters. Each cluster is a set of indexes from $\{1, \dots, N\}$, and each object \mathbf{X}_j belongs to exactly one cluster. In this formulation, we sought to parcellate the hippocampus based on the functional connectivity between vertices in layer II of EC and hippocampal voxels.

a) Hierarchical Clustering (Agglomerative): The main idea of hierarchical clustering [184] [185] [123] is to build a binary tree structure that iteratively merges two closest clusters (in terms of Euclidean distance) together. A brief description of the procedure is as follows:

- 1) Initially, each object \mathbf{X}_j is assigned to a cluster with only itself in it.
- 2) Euclidean distance between any two clusters is estimated and the closest pair of clusters are merged.
- 3) Step 2-3 are repeated until all \mathbf{X}_j are in one big cluster.

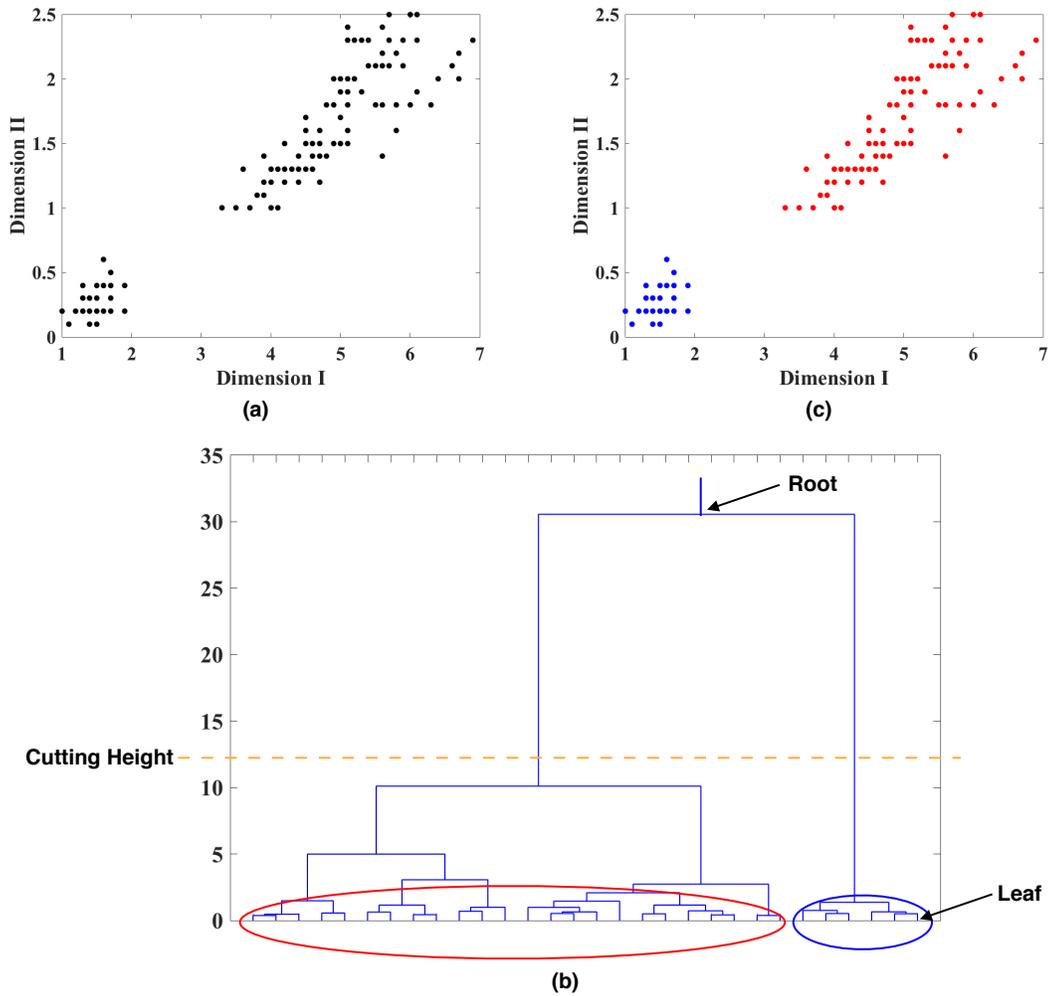


Figure 3.5 Illustration of hierarchical clustering. a) Original simulated dataset, b) Dendrogram derived from hierarchical clustering, and c) Clustering results obtained with a specific cutting height. Two clusters that were identified are marked with different colors.

The resulting tree structure is usually referred to as the dendrogram (Figure 3.5). The root of the dendrogram represents the entire data, each leaf represents one object, and the height of the dendrogram represents the distance between each pair of clusters. Different data partitions can be obtained by cutting the dendrogram at different levels. Note that the distance between two clusters can be measured in a variety of ways, referred to as linkage methods. The single linkage

[186] calculates the shortest distance between two clusters, the complete linkage [186] calculates the longest distance, and the average linkage [186] calculates the mean distance. The single linkage method can handle non-elliptical shape of clusters, but can be affected by noise and outliers. The complete linkage method is less sensitive to noise and outliers but tends to break large clusters. The average linkage is a compromise between single-linkage and complete linkage methods. Thus, the average linkage method was employed in this work.

b) Ordering Points to Identify the Clustering Structure: OPTICS [57] is one of the most popular density-based clustering methods [55]. Given a distance threshold (ϵ) and the minimum number of objects required to form a cluster (*MinPts*), objects in high-density regions are grouped together, whereas objects in low-density regions are marked as outliers. OPTICS can discover clusters with arbitrary shapes and has the ability to identify outlier objects that do not belong to any of the clusters.

In OPTICS, two variables are computed for each object in the dataset: *core-distance* and *reachability-distance*. Let $N_\epsilon(\mathbf{X}_j)$ represent the number of nearby objects within ϵ (called ϵ - *neighborhood*), and $MinPts - distance(\mathbf{X}_j)$ represent the distance from \mathbf{X}_j to its *MinPts*' neighbor. An object \mathbf{X}_j is a core object if at least *MinPts* objects are found with its ϵ - *neighborhood*. The *core-distance* of \mathbf{X}_j is defined as:

$$core - dist_{\epsilon, MinPts}(\mathbf{X}_j) = \begin{cases} Undefined & \text{if } N_\epsilon(\mathbf{X}_j) < MinPts \\ MinPts - distance(\mathbf{X}_j) & \text{otherwise} \end{cases} \quad (3.1)$$

which is the smallest distance between \mathbf{X}_j and an object in its ϵ - *neighborhood* such that \mathbf{X}_j would be a core object.

The reachability-distance of object X_i with respect to object X_j is defined as:

$$\begin{aligned}
 & \text{reachability} - \text{dist}_{\epsilon, \text{MinPts}}(X_i, X_j) \\
 &= \begin{cases} \text{Undefined} & \text{if } N_{\epsilon}(X_j) < \text{MinPts} \\ \max(\text{core} - \text{distance}(X_j), \text{dist}(X_i, X_j)) & \text{otherwise} \end{cases} \quad (3.2)
 \end{aligned}$$

Where $\text{dist}(X_i, X_j)$ is the distance between X_i and X_j . The complete procedure of OPTICS is described below:

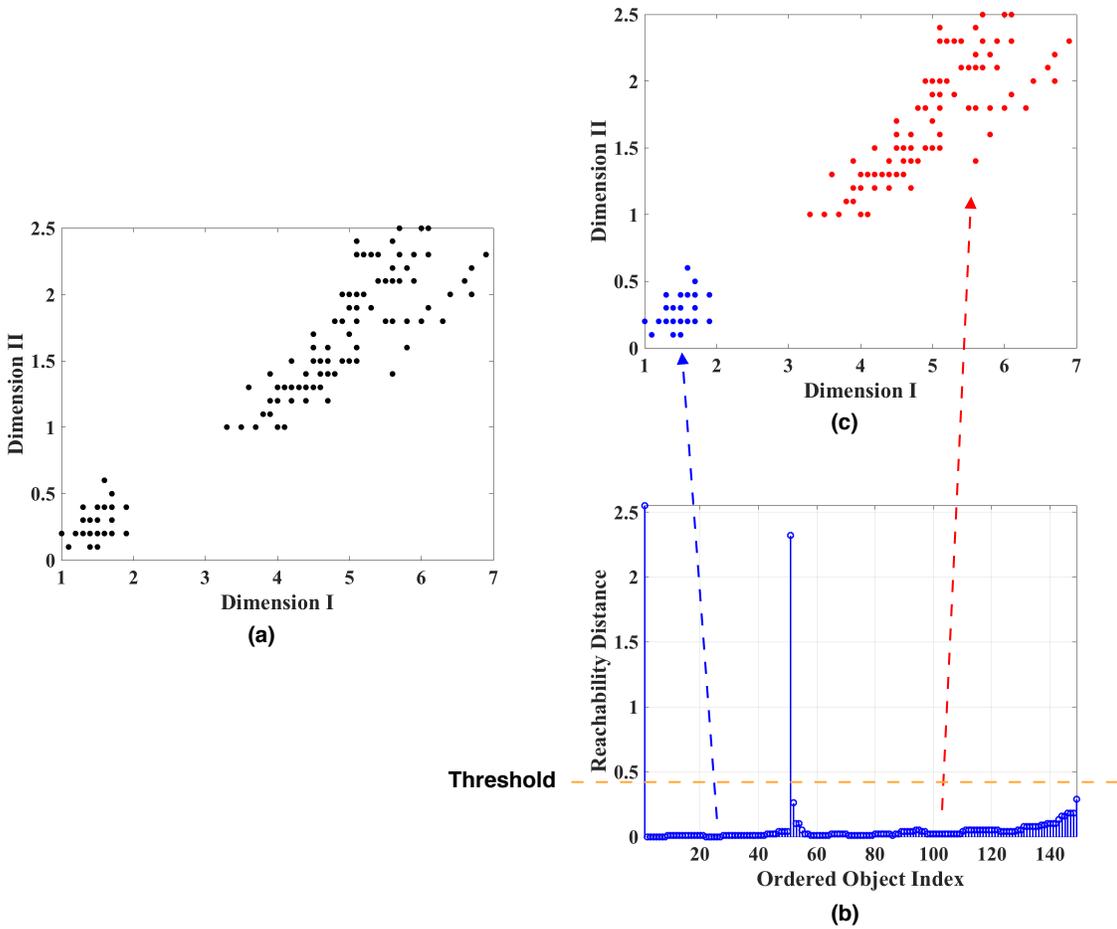


Figure 3.6 Illustration of OPTICS clustering. a) Original simulated dataset, b) reachability plot obtained from OPTICS, and c) clustering results. Two clusters were identified corresponding to valleys in the reachability plot.

- 1) Choose one object X_j arbitrarily.
- 2) Retrieve the ε – neighborhood of X_j , determine the core-distance, and set the reachability-distance to undefined.
- 3) If X_j is not a core object, go to step 5. Otherwise, go to step 4.
- 4) For each object X_i in the ε – neighborhood of X_j , update its reachability-distance from X_j and insert X_i into an OrderSeeds list if it has not been processed yet.
- 5) If the input dataset is fully consumed and the OrderSeeds list is empty, go to step 6. Otherwise, move on to the next object in the OrderSeeds list (or the input list, if the OrderSeeds list is empty) and go to step 2.
- 6) Output core-distance, reachability-distance of each object, and processed order.

The data objects are plotted in the processed order together with their respective reachability-distance (called reachability plot) depicting the hierarchical structure of the clusters. Since objects belonging to a cluster have a low reachability-distance to their nearest neighbor, the clusters show up as valleys in the reachability plot (see Figure 3.6). The final data partition can be obtained by using a threshold on the reachability plot.

c) Density Peak Clustering: DPC method [124] is a novel density-based clustering method. In DPC, cluster centers are characterized by two criteria: it has a higher density than its neighbors and a relatively large distance from objects with higher densities. Like other density-based clustering methods, e.g., OPTICS, it has ability to detect arbitrarily shaped clusters and

spot outlier objects. Moreover, DPC outperforms commonly used clustering methods, e.g., k-means and hierarchical clustering, when the dataset contains complicated features such as narrow bridges between clusters, uneven-sized clusters, clusters with high overlap, etc.

For each object \mathbf{X}_j , two quantities are computed: local density $\rho(\mathbf{X}_j)$ and minimum distance with higher density $\delta(\mathbf{X}_j)$. $\rho(\mathbf{X}_j)$ is defined as:

$$\rho(\mathbf{X}_j) = \sum_i \chi(\text{dist}(\mathbf{X}_i, \mathbf{X}_j) - d_c) \quad (3.3)$$

where d_c is a cutoff distance, and $\chi(x)$ can be computed by,

$$\chi(x) = \begin{cases} 1 & \text{if } x < 0 \\ 0 & \text{otherwise} \end{cases} \quad (3.4)$$

From Equation (3.3) and (3.4), it can be seen that $\rho(\mathbf{X}_j)$ equals to the number of objects within d_c with respect to object \mathbf{X}_j . $\delta(\mathbf{X}_j)$ is measured by,

$$\delta(\mathbf{Z}_j) = \min_{i:\rho(\mathbf{Z}_i) > \rho(\mathbf{Z}_j)} \text{dist}(\mathbf{X}_i, \mathbf{X}_j) \quad (3.5)$$

For the object with highest density, $\delta(\mathbf{X}_j)$ is conventionally set to,

$$\delta(\mathbf{X}_j) = \max_i \text{dist}(\mathbf{X}_i, \mathbf{X}_j) \quad (3.6)$$

Note that if \mathbf{X}_j is local or global maxima in the density, $\delta(\mathbf{X}_j)$ will be much larger than its typical nearest neighbor. Thus, objects with larger ρ and δ are considered as cluster centers, whereas objects with smaller ρ and larger δ are considered as outliers. Other objects are assigned to the same cluster as their nearest neighbor of higher density (see Figure 3.7).

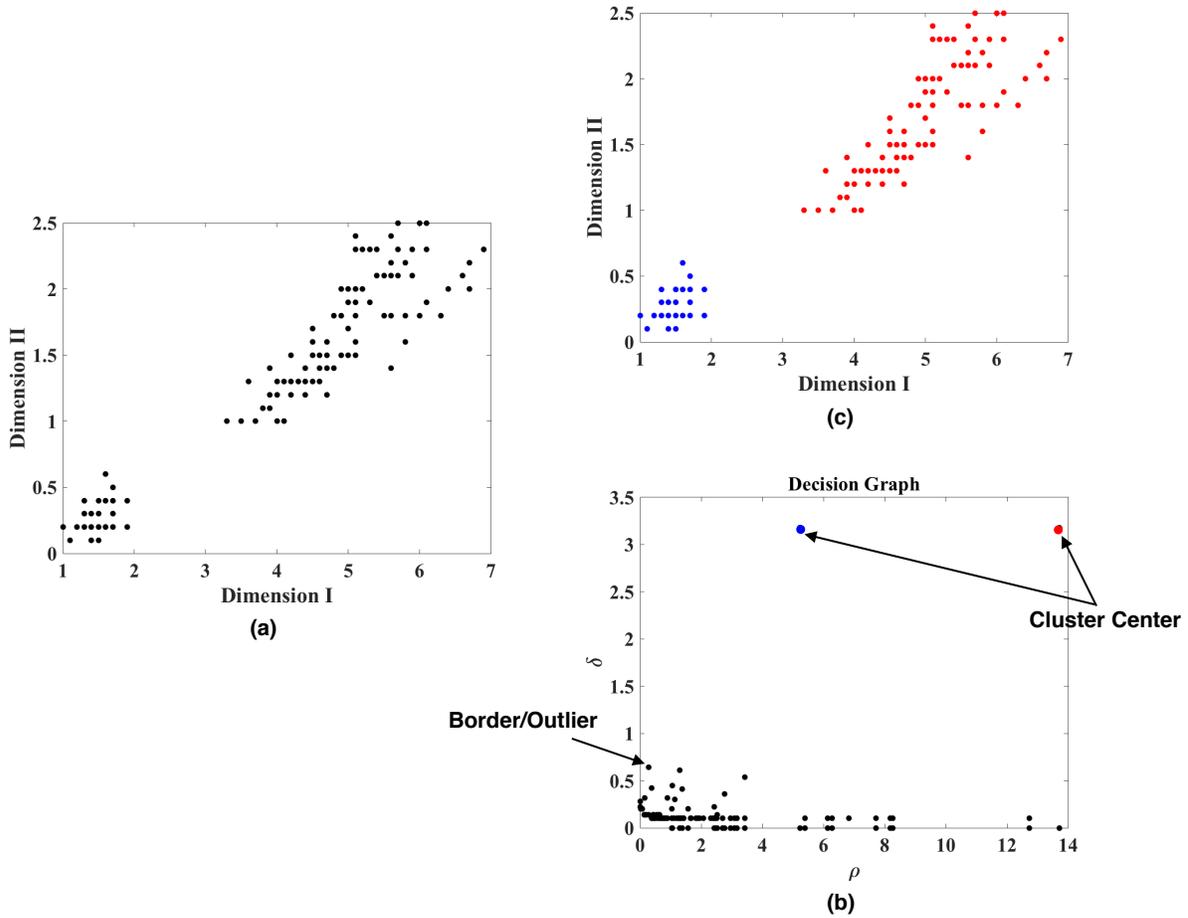


Figure 3.7 Illustration of DPC clustering. a) Original simulated dataset, b) Plot of δ as a function of ρ for each object. Objects with larger ρ and δ are cluster centers and objects with smaller ρ , and larger δ are outliers. c) Clustering results. Two clusters were identified corresponding to two cluster centers in the decision graph.

d) Input Parameter Optimization: In each clustering method, there are several user-specified input parameters, which can affect clustering results significantly. For hierarchical method, the cutting height of the dendrogram needs to be specified and the number of clusters varies with different cutting heights. For OPTICS, ϵ can simply be set to the maximum possible value, and Ankerst and colleagues [57] showed that for *MinPts* using values between 10 and 20 would

always lead to good results. However, the threshold for the reachability plot, which is used to extract clusters, still needs to be properly determined. For DPC, d_c can be chosen based on the rule that the average number of neighbors is around 1-2% of the total number of objects in the data set [124]. A threshold for ρ and δ needs to be defined to distinguish cluster centers, borders, and outliers. To find the optimal values of these parameters, the silhouette index [60] was applied in this work. Assume the data have been clustered via any clustering algorithm, such as hierarchical clustering, into K clusters. For each object \mathbf{X}_i , let $a(\mathbf{X}_i)$ represent the average distance of \mathbf{X}_i with all other object in the same cluster, and $b(\mathbf{X}_i)$ represent the smallest average distance of \mathbf{X}_i to any other cluster, of which \mathbf{X}_i does not belong to. Then the silhouette index of \mathbf{X}_i is defined as:

$$s(\mathbf{X}_i) = \frac{b(\mathbf{X}_i) - a(\mathbf{X}_i)}{\max \{a(\mathbf{X}_i), b(\mathbf{X}_i)\}} \quad (3.7)$$

From Equation (3.7) it is clear that $s(\mathbf{X}_i)$ is bounded between -1 and 1. More specifically, if \mathbf{X}_i has been assigned to an appropriate cluster, i.e., $a(\mathbf{X}_i) \ll b(\mathbf{X}_i)$, $s(\mathbf{X}_i)$ will be close to 1. On the contrary, $s(\mathbf{X}_i)$ is close to -1, if $s(\mathbf{X}_i)$ has been assigned to a “wrong” cluster. A $s(\mathbf{X}_i)$ near zero indicate \mathbf{X}_i is located on the border of two natural clusters. The average $s(\mathbf{X}_i)$ over all objects in the entire dataset measures how appropriately the dataset has been clustered. Thus, we want to maximize this average value identifying dense and well-separated clusters.

Based on the above definition, the optimal parameters can be determined by maximizing the average $s(\mathbf{X}_i)$ over the entire dataset. The optimal number of clusters can be identified, simultaneously. For example, for hierarchical clustering, we started with a relatively high cutting height for the dendrogram. In each iteration, the cutting height was reduced by a small amount

and the average $s(\mathbf{X}_i)$ was computed and recorded based on the current data partition. The iteration continued until the cutting height was smaller than a specified baseline (e.g., the average height of the dendrogram). The optimal height was determined as the one with the largest average $s(\mathbf{X}_i)$. The same iterative procedure was applied to OPTICS to determine optimal threshold of reachability plot, and to DPC to determine the optimal threshold of ρ and δ . This iterative method is referred to as “grid search” [187], which has been commonly used for hyper-parameter optimization.

3.2.5 Feature Selection and Cluster Identification

The clustering accuracy is often lower in high dimensional feature space, which is referred to as the “dimensional curse”. This is due to the fact that most of features in the dataset may be irrelevant, redundant, or sometimes may even misguide results. Moreover, a large number of features make the clustering results difficult to interpret. Therefore, a feature selection method is required to improve the clustering accuracy. For supervised learning, feature selection can be trivial, i.e., only the features that are related to the given cluster labels are maintained. Nevertheless, for unsupervised learning, the cluster labels are unknown. Thus, finding the relevant subset of features and clustering the subset of the data must be accomplished simultaneously. To evaluate the result for feature selection and clustering, the same optimization criterion, i.e., the average silhouette index, was applied.

Assuming d to be the initial number of features, an exhaustive search of 2^d possible subsets is computationally intractable. Thus, an alternative GA method was employed in our framework to find the optimal subset of features.

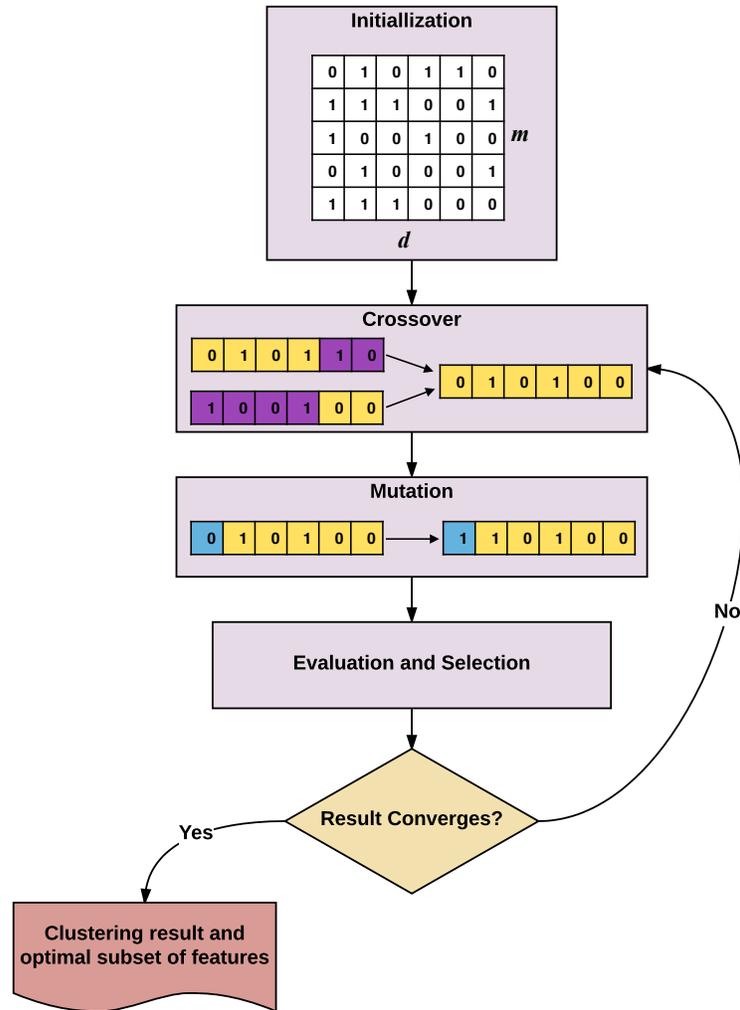


Figure 3.8 Flowchart of GA for feature selection. In the m -by- d matrix, each row represents a candidate solution, describing a subset of selected features. Each of the d bits in a row represents whether a feature is selected (1) or discarded (0).

GA is a search heuristic method inspired by stochastic evolution theory that is routinely used in generating solutions to optimization and searching problems [189] [190]. Initially, a set of candidate solutions were randomly generated and maintained during the entire process. This solution set is referred to as “population” in GA. In each iteration, “survivor solutions” with larger values of similarity criteria, i.e., the average silhouette index, were selected for the next

iteration. The survival solutions were generated from the crossover, which produced new solutions by randomly combining two current solutions, and mutations resulting in randomly changing new solutions with a small probability, or from the initial population. The iteration continued until the maximum similarity among candidate solutions did not increase any further.

In this study, an array of d bits was used to represent the selected subset of features and the population size is represented using m . Each bit in the array indicates the activation status of one specific feature: 1 indicates selected and 0 indicated discarded. The complete procedure of GA is described below (Figure 3.8):

- 1) Initialization: 400 candidate solutions were generated by randomly setting 1 or 0 for each bit in vectors.
- 2) Crossover: two candidate solutions A and B were randomly selected from the current population. A value v between 1 and d was randomly selected. Then a new solution was formed by combining the feature bits 1 to v from A and feature bits $v + 1$ to d from B.
- 3) Mutation: for each new generated new solution, a mutation was applied by reversing bits in the vector with a probability of 0.1.
- 4) Evaluation: the clustering methods were applied on each candidate solution (i.e., a subset of features) and the average silhouette index was computed on each data partition.
- 5) Selection: 280 solutions with higher average silhouette index were selected and 120 solutions from the rest of the solutions were randomly selected to increase the diversity of the solution.

- 6) If the highest average silhouette index did not converge, we iterated back to step 2. Otherwise, the clustering results with the highest average silhouette index and the corresponding selected subset of features were saved as outputs.

3.2.6. Volume Level Clustering

In order to determine whether characterizing layer-specific microcircuits using ultra-high field fMRI provides any advantages over conventionally computed voxel-level connectivity, we performed the clustering procedure enumerated above using FC computed between hippocampal and EC voxels. Specifically, let $\mathbf{X} = \{\mathbf{X}_1, \dots, \mathbf{X}_j, \dots, \mathbf{X}_N\}$ represent a set of N objects, i.e., number of hippocampal voxels. $\mathbf{X}_j = (X_{j1}, X_{j2}, \dots, X_{jd}) \in \mathbb{R}^d$, where d equals to the number of FC features computed between hippocampal voxels and voxels in EC. Subsequently, the same clustering and feature selection process was repeated on the left and right hemisphere, separately.

3.2.7. Validation of Identified Clusters

In this study, clusters were identified from EC-hippocampal functional connectivity during an encoding task. This provided a functional parcellation of the hippocampus which were compared with the well-known anatomical parcellation of the hippocampus. Specifically, the percentage of overlap between functional and anatomical parcels were computed, which provides quantitative insights into the layer-specific connectional architecture of the EC-hippocampus microcircuits.

Neuroanatomically, the hippocampus is subdivided into several subfields with intricate morphologies and complex three-dimensional relationships. Recent studies have utilized high-field MRI techniques to identify the hippocampal subfields [357] [358]. However, most of them

provided a segmentation protocol that only included a part of the hippocampus rather than the complete anterior-posterior hippocampal axis.

In this study, a recent hippocampal atlas proposed by Winterburn et al was utilized [359]. This atlas provides five clearly identified subfields, i.e., DG/CA4, CA1, CA2/CA3, subiculum, and SR/SL/SM (Figure 3.9). The SR, SL, and SM cannot be distinguished from each other, and they were combined into one region. CA2 and CA3 were also combined into one to minimize anatomical inaccuracies. This atlas was downloaded from the following website: <http://cobralab.ca/atlasses/Hippocampus-subfields/>, and normalized into MNI space.

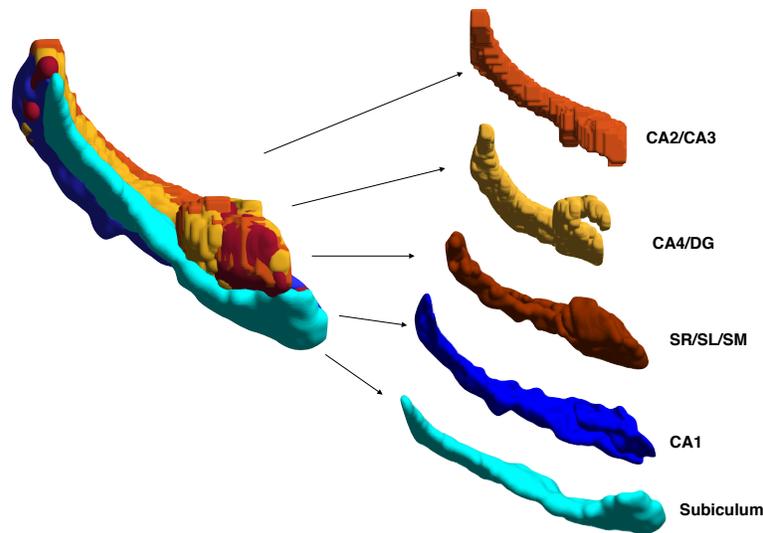


Figure 3.9 3D visualization of the entire hippocampal atlas in sagittal view. Each subfield is represented by a different color.

Let $A = \{A_1, A_2, \dots, A_M\}$ represent M anatomical parcels, and $F = \{F_1, F_2, \dots, F_M\}$ denote functional parcels identified by applying clustering methods on FC features. The percentage of overlap between any pair of anatomical and functional parcels, A_i and F_i , was computed by,

$$\text{overlap}(A_i, F_i) = \frac{\text{inter}(A_i, F_i)}{m_i} \times 100 \quad (3.8)$$

where $\text{inter}(A_i, F_i)$ is the intersection size between A_i and F_i , and m_i is the total number of voxels in A_i .

The proposed analysis pipeline for investigating perforant pathway, including feature selection, layer-specific clustering, volume level clustering, and validation of identified clusters, are illustrated in Figure 3.2.

3.3 Results

The optimal values of each input parameter determined for the three clustering methods are presented Table 3.1. Using each clustering method, the hippocampal voxels were clustered into two different functional parcels based on their functional connectivity with layer II of the EC during an encoding task. This was true across methods and for both left and right hippocampi. The obtained clusters were then mapped back to the image space and the resulting hippocampal parcels were overlaid on the anatomical image for the visualization.

These functional clusters had highest overlap with the following two anatomical parcels: CA2/CA3 and CA4/DG together (which belong to perforant pathway), and CA1 and subiculum together. Therefore, while comparing functional and anatomical hippocampal parcels, we considered CA2/CA3 and CA4/DG together as one anatomical parcel and CA1 and subiculum together as another anatomical parcel as shown in Figure 3.10.

Method Name	Parameter	Optimal Value			
		Left		Right	
		Layer II	Volume	Layer II	Volume
DPC	ρ	7.54	8.92	7.56	8.45
	δ	0.21	3.01	0.06	1.87
Hierarchical	Cutting Height	1.15	1.16	1.15	1.16
OPTICS	Reachability Threshold	0.01	0.20	0.04	0.02

Table 3.1 Estimated optimal values of each input parameter in clustering.

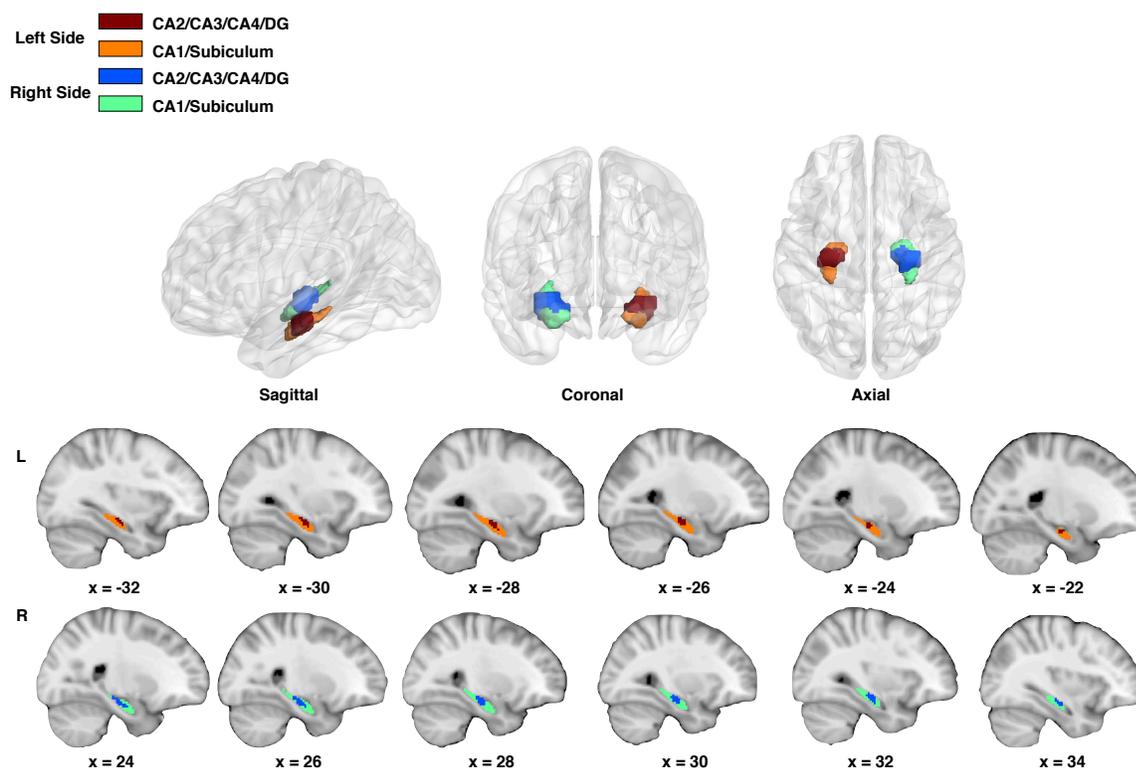


Figure 3.10 Anatomical parcellation of the hippocampus used in this study. (Coordinates are in MNI space)

Comparison	Side	Similarity
DPC v.s. OPTICS	L	0.93
	R	0.78
DPC v.s. Hierarchical	L	0.94
	R	0.84
OPTICS v.s. Hierarchical	L	0.98
	R	0.81

Table 3.2 Cluster similarity between different clustering methods for layer-specific analysis. For the same input data, the similarity of clusters obtained using different clustering methods were measured using Torres' method [188].

Similar hippocampal parcellations were discovered using their functional connectivity with EC layer II using different clustering methods and the corresponding cluster similarity is quantified as shown in Table 3.2. For illustration, the clustering results obtained using the DPC method are shown in Figure 11. The clustering results obtained using hierarchical clustering and OPTICS are shown in Supplementary Information Figure A.1 and Figure A.2. From Figure 3.11, it can be seen that the parcellation of the left hippocampus was consistent with the gradient shown in the anatomical parcellation (Figure 3.10). However, the right hippocampus did not show this pattern, as evidenced by much lower % of overlap with the DG/CA4/CA3/CA2 cluster. Figure 3.12 visually depicts the overlap between anatomical subfields of the hippocampus and functional parcels obtained from clustering (using DPC method) of hippocampal voxels based on their functional connectivity with layer II of EC during an encoding task.

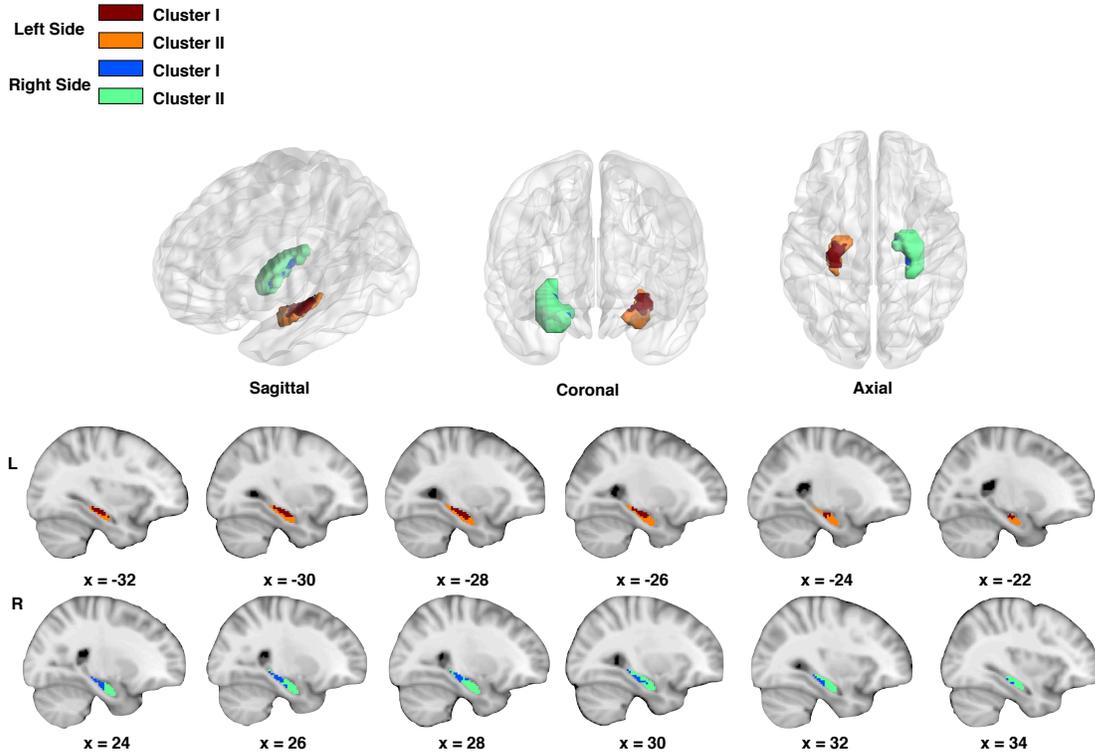


Figure 3.11 Clusters of hippocampal voxels determined (using the DPC method) based on their functional connectivity with layer II of the entorhinal cortex during the encoding task. (Coordinates are in MNI space)

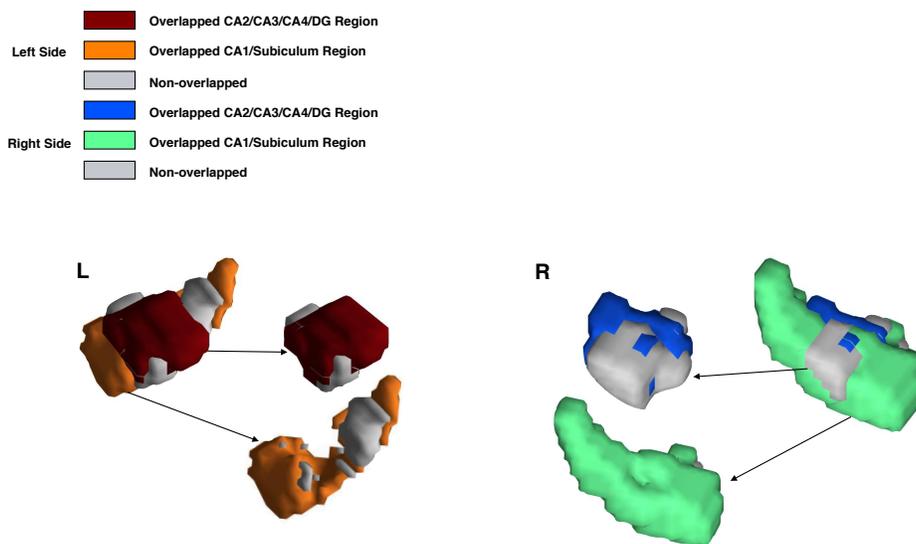


Figure 3.12 Overlap between anatomical subfields of the hippocampus and functional parcels obtained from clustering (using DPC method) of hippocampal voxels based on their functional connectivity with layer II of EC during an encoding task.

To quantitatively characterize the identified clusters, the overlap between our functionally obtained hippocampal parcels and anatomically defined parcels (CA2/CA3 and CA4/DG together as one anatomical parcel and CA1 and subiculum together as another anatomical parcel as shown in Figure 3.10) was computed. The mean correlation between hippocampal voxels and selected vertices (using GA-based feature selection method) in layer II of EC was also computed within each cluster along with p-values indicating the statistical significance of those correlations (Table 3.3).

For the left hippocampus, the percentage of overlap was consistently high with different clustering methods. The average overlap between DG/CA4/CA3/CA2 and functional cluster I was 83.61%, whereas the average overlap between CA1/subiculum and functional cluster II was 72.99%, respectively. In addition, a significant positive mean correlation (p -value < 0.001) was consistently observed within cluster I, whereas a significant negative correlation (p -value < 0.001) was obtained within cluster II.

For the right hippocampus, most of voxels were clustered into functional cluster II. Thus, high overlap was obtained between CA1/subiculum and functional cluster II, but the overlap between DG/CA4/CA3/CA2 and functional cluster I was quite low. The average overlap between DG/CA4/CA3/CA2 and functional cluster I was 30.69%, whereas the average overlap between CA1/subiculum and functional cluster II was 76.90%. Compared to the left hippocampus, mean

correlation within cluster I and cluster II on the right hippocampus was lower, i.e., the corresponding p-value was greater than 0.05.

Method	Side	Percentage of Overlap (%)		Mean Correlation		P-Value	
		CA2/CA3/CA4	CA1/Subiculum	Cluster I	Cluster II	Cluster I	Cluster II
DPC	L	83.61	74.07	0.457	-0.301	< 0.001	< 0.001
	R	38.61	71.29	-0.090	0.090	0.121	0.118
Hierarchical	L	83.61	73.15	0.400	-0.325	< 0.001	< 0.001
	R	32.67	77.23	-0.120	0.064	0.062	0.244
OPTICS	L	83.61	71.76	0.349	-0.338	< 0.001	< 0.001
	R	20.79	82.18	-0.201	0.054	0.072	0.316

Table 3.3 Quantitative characterization of the overlap between functional clusters (obtained from layer-specific analysis) and anatomical subfields. The percentage of overlap between each pair of corresponding clusters, the mean correlation and corresponding p-values obtained within each cluster using different clustering methods are shown. Functional cluster I corresponds to anatomical subfields CA2/CA3/CA4/DG and functional cluster II corresponds to anatomical subfields CA1/subiculum.

For volume level analysis, i.e. clustering of hippocampal voxels based on their functional connectivity with voxels in EC volume (as opposed layer II of EC as before), similar structures were discovered using different clustering methods on the left hippocampus, whereas the pattern identified on the right hippocampus using DPC method was different from that obtained using hierarchical clustering and OPTICS. The cluster similarity was quantified as shown in Table 3.4. For illustration, the clustering results using DPC was shown in Figure 3.13, and the clustering results using hierarchical clustering and OPTICS are shown in supplementary information Figure A.3 and Figure A.4.

Comparison	Side	Similarity
DPC v.s. OPTICS	L	0.95
	R	0.68
DPC v.s. Hierarchical	L	0.96
	R	0.67
OPTICS v.s. Hierarchical	L	0.99
	R	0.96

Table 3.4 Cluster similarity between different clustering methods for volume level analysis. For the same input data, the similarity of clusters obtained using different clustering methods was measured using Torres' method [188].

■ Cluster I
■ Cluster II
 Left Side
■ Cluster I
■ Cluster II
 Right Side

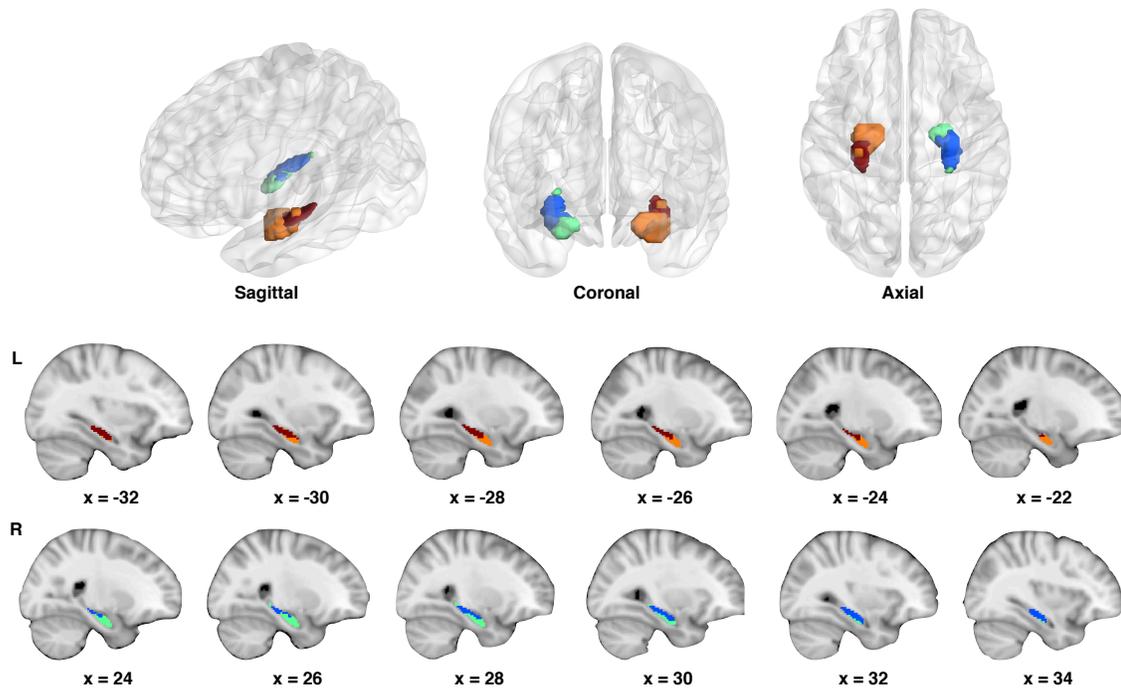


Figure 3.13 Clusters of hippocampal voxels determined (using DPC method) based on their functional connectivity with voxels in the entorhinal cortex volume during the encoding task.

(Coordinates are in MNI space)

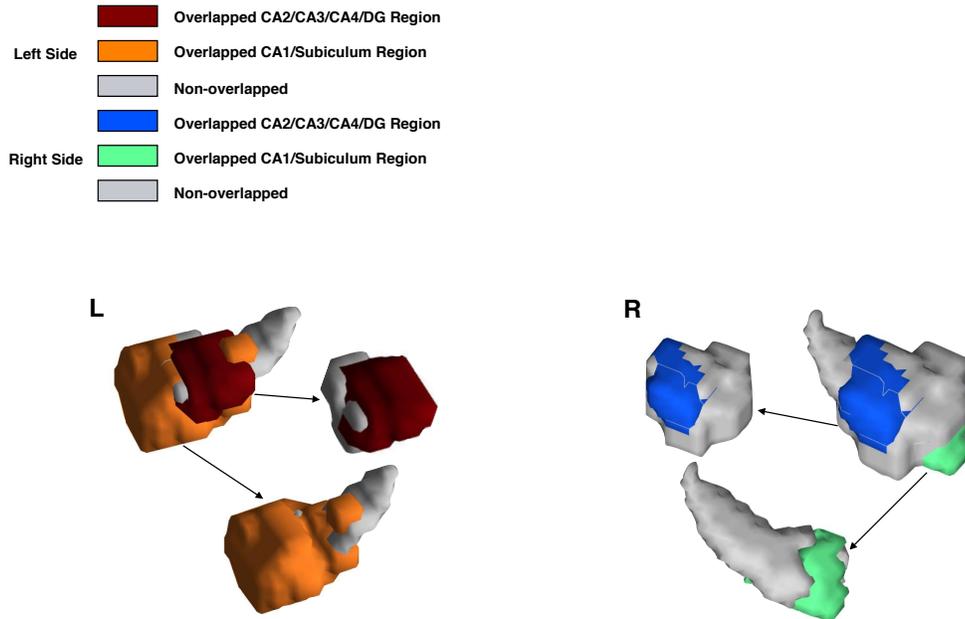


Figure 3.14 Overlap between anatomical subfields of the hippocampus and functional parcels obtained from clustering (using DPC method) of hippocampal voxels based on their functional connectivity with voxels in EC volume during an encoding task.

As shown in Figure 3.13, we obtained two clusters on each side of the hippocampus as well. The structures obtained on the left and right hippocampus were different from the patterns obtained using layer-specific data (Figure 3.10). The percentage of overlap (also see Figure 3.14 for visualization of overlap), mean correlation and corresponding p-value were also computed for volume level clustering. As shown in Table 3.5, although high overlap was obtained between DG/CA4/CA3/CA2 and functional cluster I in the left hippocampus, the overlap between CA1/subiculum and functional cluster II was notably low. The right hippocampus displayed

inconsistency across methods. For DPC, the overlap between both functional clusters with proximal and distal anatomical clusters were qualitatively low. On the other hand, for Hierarchical and OPTICS clustering methods, the functional clusters within the right hippocampus showed qualitatively larger overlap with the proximal cluster (DG/CA4/CA3/CA2) with notably low overlap with the distal cluster (CA1/subiculum) The functional connectivity between EC voxels and hippocampal voxels within both functional clusters I and II approached zero and the corresponding p-value was greater than 0.05.

Method	Side	Percentage of Overlap (%)		Mean Correlation		P-Value	
		CA2/CA3/CA4	CA1/Subiculum	Cluster I	Cluster II	Cluster I	Cluster II
DPC	L	81.97	58.33	-0.014	0.050	0.593	0.498
	R	58.91	61.39	0.019	0.013	0.398	0.629
Hierarchical	L	86.89	56.94	-0.025	0.045	0.604	0.467
	R	83.17	21.29	0.118	-0.047	0.457	0.666
OPTICS	L	88.52	56.48	-0.006	0.048	0.591	0.489
	R	82.18	23.27	0.106	-0.055	0.488	0.627

Table 3.5 Quantitative characterization of the overlap between functional clusters obtained from volume-level analysis and anatomical subfields. The percentage of overlap between each pair of corresponding clusters, and the mean correlation obtained within each cluster using different clustering methods are shown. Functional cluster I corresponds to anatomical subfields CA2/CA3/CA4/DG and functional cluster II corresponds to anatomical subfields CA1/subiculum.

3.4 Discussion

In this work, we investigated functional connectivity between the hippocampus and the entorhinal cortex during an encoding task by characterizing layer-specific microcircuits between

layer II of EC and the hippocampus using ultra high field fMRI. Given known perforant anatomical pathways between layer II of EC and the hippocampus, we hypothesized that functional connectivity between them, when clustered in an unsupervised way, must parcellate the hippocampus into proximal and distal regions along the perforant pathway, with stronger connectivity between layer II of EC with proximal hippocampal subfields such as DG/CA4/CA3/CA2 compared to distal subfields such as CA1/Subiculum. Further, given that our task involved factual memory encoding of faces, objects and words which preferentially recruits the left hippocampus, rather than spatial memory typically used in navigation which preferentially recruits the right hippocampus, we predicted that our hypothesis will hold true more for the left, rather than the right, hippocampus during an encoding task. We found support for our hypotheses stated above and we discuss the same below. The discussion of results is organized as follows. First, we discuss the results obtained by the characterization of layer-specific functional microcircuits between the EC and hippocampus during an encoding task. Second, we discuss the functional hemispheric specialization of the hippocampus. Finally, we discuss the importance of ultra-high field functional neuroimaging in developing accurate and robust models of functional connectivity in layer-specific microcircuits of the human brain.

3.4.1 Layer-specific functional perforant pathway from EC to the hippocampus

The topologically arranged circuitry between the EC and hippocampus has been historically studied by many researchers [360] [361] [362]. During an encoding task, the more superficial layers (layer II and III) of EC project to DG/CA4, CA3, CA2, and from there to CA1 and subiculum. This organization is referred to as the “perforant pathway”, which is considered as the major pathway from the EC to the hippocampus during memory encoding. Maass et al. [331]

used 7T high resolution fMRI to show layer-specific activation during memory encoding as predicted by the perforant pathway. However, the perforant pathway has never been directly investigated using connectivity based non-invasive methods in humans, although invasive studies in animals have demonstrated the existence of this pathway [363].

In this study, we investigated the functional perforant pathway using functional connectivity between layer II of EC with voxels in different subfields of the hippocampus during a memory encoding task. For the left hippocampus, our results yielded a parcellation which consistently overlapped with two anatomical clusters, i.e., DG/CA4/CA3/CA2 and CA1/subiculum, to a great extent. Meanwhile, the identified proximal subfields (DG/CA4/CA3/CA2) showed a significant positive correlation with layer II of EC, whereas the distal subfields (CA1/subiculum) exhibited a significant negative correlation with EC. Together, these findings demonstrate that during a memory encoding task, layer II of EC is more correlated with DG/CA4/CA3/CA2, than with CA1/subiculum. This result makes sense because the direction of information flow in the perforant pathway fits the following pattern: EC layer II→DG→CA4→CA3→CA2→CA1→subiculum. Consequently, layer II in EC must be more correlated with proximal subfields compared to distal subfields of the hippocampus.

3.4.2 Functional Hemispheric Specialization

Recent neuroimaging results have been suggested that there is distinct structural and functional specialization between the left and right hippocampus. Shipton et al. [364] conducted experiment on mice and found evidence for hemispheric dissociation of CA3 of the hippocampus. Papanicolaou et al. [365] and Burgess et al. [346] also found functional differentiation between the left and right hippocampus in humans as well. More specifically, the left hippocampus is

associated with more verbal, factual and procedural memory processes, whereas the right hippocampus is involved in more spatially-dependent memory processes typically used in navigation [346] [355]. Therefore, hemispheric specialization in the hippocampus seems to be an evolutionarily preserved feature. In this study, we found evidence for hemispheric specialization with a memory encoding task involving faces, objects and words. For the left hippocampus, the functional parcellation was consistent with the anatomical parcellation, i.e., DG/CA4/CA3/CA2 and CA1/subiculum. However, we did not find a similar pattern in the right hippocampus. These results indicate that the left and right hippocampus have substantially different functional roles.

3.4.3 High-resolution Functional Imaging

Recent advances in ultra-high field fMRI have provided a non-invasive way of investigating cortical columns. This technique provides several advantages over conventional field strengths, e.g., improved spatial resolution, increased signal to noise ratio, etc. More importantly, this technique makes it feasible to examine layer-specific brain activation across different brain areas. Several recent studies have showed that investigating changes in fMRI activation as a function of laminar depth can lead to more precise result [342] [343] [344]. In this study, we estimated EC-hippocampus functional connectivity during a memory encoding task using time series from layer II of EC as well as those from voxels in EC (which corresponds to conventional volume level analysis). As we expected, the hippocampal parcellation obtained using layer-specific connectivity data were more closely related to the anatomical parcellation of hippocampal subfields, which was not the case using volume level data. Also, the functional connectivity obtained between EC and hippocampal voxels did not conform to our prior expectations about this circuitry during memory encoding. Therefore, it is important to note the

relevance of high-resolution functional neuroimaging as the field progresses toward developing accurate and robust models of functional connectivity.

3.4.4 Limitations and Future Works

We note a few limitations of the present study which may be addressed in future research. First, before performing layer-specific fMRI connectivity analysis, different cortical layers need to be precisely delineated. Existing methods include the equidistant model and the Laplace model. The former one, which was used in our study, constructs laminar profiles that preserve a relatively fixed distance to the cortical boundaries [366]. The latter one solves the Laplace equation within the boundaries of the cortex and generates cortical profiles along the gradient of the solution [367]. Previous studies indicate that both these two models may fail to precisely follow anatomical layers all along the cortex [368]. To solve this issue, a novel equivolume model was proposed [368], which was derived based on the theory that the cortical segments preserve their volume, while layer thickness changes to compensate cortical folding. Future work should investigate the feasibility of this novel equivolume model and assess how well the cortical profiles can be constructed by using different methods. If found to be superior, our findings need to be replicated with more accurate reconstructions of cortical layers.

Second, our results indicate that the functional parcellation in the left hippocampus is in line with the anatomical parcellation. Our results do make sense given that our parcellation was based on functional connectivity in the perforant pathway during a memory encoding task. However, other studies have shown an anterior-posterior gradient along the long axis in the hippocampus [369] [370] [371], with the head and body of the hippocampus mainly constituting the anterior parcel and the tail of the hippocampus representing the posterior parcel. This indicates that gradients

within the hippocampus are dependent on the neurocognitive context in which they are investigated. This factor must be kept in mind while interpreting our results and likely forms a basis for future research as well.

Third, the present study investigated the dominant perforant pathway between the hippocampus and layer II in EC during memory encoding. According to previous studies, there are some other pathways between different layers of EC and subfields of the hippocampus during encoding and retrieval processes [330] [331]. For example, a laminar specific activity analysis [331] indicates that the retrieval process is dependent on activation of deeper layers of EC and CA1 of the hippocampus. Our study demonstrates the possibility to investigate layer-specific microcircuits in both healthy and clinical human populations, using non-invasive ultra-high fields fMRI. Thus, further fine-grained investigation can be applied in the future research.

3.5 Conclusion

As main component in the medial temporal lobe, the hippocampus and the EC are involved in encoding and retrieving processes. There is anatomical and electrophysiological evidence that during an encoding task, superficial layers of EC appear to be connected to different subfields of the hippocampus through anatomically characterized circuits so called perforant pathway. This pattern has been supported by previous studies. However, a comprehensive understanding of circuit-level mechanisms underlying encoding task in the perforant pathway has not been resolved.

In this study, we investigated the functional perforant pathway using the connectivity between layer II of EC with different subfields of the hippocampus during an encoding task. We found that when the functional connectivity between EC layer II and hippocampus, when clustered in

an unsupervised way, parcellated the hippocampus into proximal and distal regions along perforant pathway. This parcellation was based on our observation of stronger connectivity between layer II of EC with hippocampal subfields such as DG/CA4/CA3/CA2 which are proximal to the EC along the perforant pathway, compared to subfields such as CA1/Subiculum which are distal.

In addition, we used a verbal memory task and only the left hippocampus showed this dorsal and ventral pattern. This result suggests that there is a functional hemispheric specialization between the left and right hippocampus, and the left hippocampus is associated with more verbal memory task.

Finally, we applied the same clustering and feature selection process on volume level data in comparison with layer-specific data. As we expected the volume level data failed to reveal a dorsal-ventral pattern, and the identified structures did not match well with anatomical parcellation. This result suggests that examining functional connectivity as a function of laminar depth can lead to more precise and robust result.

Chapter 4

Functional Parcellation of the Hippocampus based on its Connectivity with Default Mode and Dorsal Attention Networks

Recent neuroimaging evidence suggests that there might be an anterior-posterior functional differentiation of the hippocampus along the long-axis. The HERNET (hippocampal encoding/retrieval and network) model proposed an encoding/retrieval dichotomy with the anterior hippocampus more connected to the dorsal attention network (DAN) during memory encoding, and the posterior portions more connected to the default mode network (DMN) during retrieval. Evidence both for and against the HERNET model has been reported. In this study, we test the validity of the HERNET model non-invasively in humans by computing functional connectivity (FC) in layer-specific cortico-hippocampal microcircuits. This was achieved by acquiring sub-millimeter functional magnetic resonance imaging (fMRI) data during encoding/retrieval tasks at 7T. Specifically, FC between infra-granular output layers of DAN with hippocampus during encoding and FC between the hippocampus and supra-granular input layers of DMN with during retrieval were computed to test the predictions of the HERNET model. Our results support some predictions of the HERNET model including anterior-posterior gradient along the long axis of the hippocampus. While preferential relationships between the entire hippocampus and DAN/DMN during encoding/retrieval, respectively, were observed as predicted, anterior-posterior specificity in these network relationships could not be confirmed. The strength and clarity of evidence for/against the HERNET model were superior with layer-specific data compared to conventional volume data.

4.1 Introduction

The hippocampus, located in the medial temporal lobe (MTL), plays important roles in many brain functions including episodic memory and spatial navigation [310] [311] [312]. Investigation of functional specialization within the hippocampus has received increased attention in neuroimaging. The abnormalities of the hippocampus have been identified in many neuropsychiatric disorders, including Alzheimer's disease [313] [314] [315], major depression [316] [317] [318] [319], post-traumatic stress disorder [320] [321], and schizophrenia [322] [323] [324]. Therefore a better understanding of the functional specialization within the hippocampus has the potential to lead to a better understanding of these disorders.

Recently, many studies have posited that there may be a functional differentiation along the long-axis of the hippocampus [372] [373]. Lepage et al. [374] performed a meta-analysis of positron emission tomography (PET) of episodic memory and discovered an orderly functional anatomic pattern in the hippocampus. More specifically, the anterior portions were primarily activated with episodic memory encoding, whereas the posterior portions were primarily associated with episodic memory retrieval. This model is referred to as HIPER (hippocampus encoding/retrieval) model.

The HIPER model has received support from many recent studies. Spaniol et al. [375] conducted meta-analyses of event-related fMRI studies of episodic memory and revealed an anterior-posterior gradient in the hippocampal activations associated with encoding and retrieval. Nadel [376] and his colleagues compared the anterior and posterior hippocampal activations during retrieval of different types of spatial information. They found that there is a functional differentiation along the longitudinal axis of the hippocampus with the posterior hippocampus

being crucial for precise spatial behavior, and the anterior hippocampus being involved in context coding. Baumann and Mattingley [377] examined retrieval-related activity in the taxi drivers' hippocampus using fMRI data and found that taxi drivers with a small anterior hippocampus had difficulty encoding new spatial associations. Moreover, Kim [378] conducted a meta-analysis and revealed that the encoding of sensory input involved mainly the anterior hippocampus and the external attention network, whereas retrieval engaged mainly the posterior hippocampus and the internal attention network. This model was referred to as the HERNET (hippocampal encoding/retrieval and network) model.

Memory encoding is inherently linked with external attention, whereas retrieval is intrinsically related to internal attention [379] [380]. Many studies have identified two brain networks, i.e., the dorsal attention network (DAN) and default mode network (DMN), that are closely associated with external and internal attention, respectively [198] [381]. Therefore, the HERNET model predicts that the anterior hippocampus and regions of the DAN co-activate during encoding while the posterior hippocampus and regions of the DMN co-activate during encoding. In fact Kim's [378] meta-analysis confirms this prediction. However, evidence conflicting the HERNET model also exists. This includes meta-analyses of imaging studies which contradict Kim's findings [382], suggestions that the anterior hippocampus is activated by novelty (which is purportedly mistaken for encoding since encoding tasks typically use novel stimuli) [383] [384] [385] [386] [387] and alternative models of functional specialization which attribute "hot" processing (emotion/motivation) to anterior hippocampus and "cold" processing (cognition) to the posterior part [388] [370] [371]. Given this state of affairs, we set out to directly test the HERNET model using functional connectivity between the hippocampus and DAN/DMN regions during memory encoding and retrieval tasks. Unlike previous studies which employed

voxel-level analysis from data obtained at conventional field strengths ($\leq 3T$), we employed fMRI data with ultra-high spatial resolution (sub-millimeter) obtained at 7T. This allowed us to investigate layer-specific microcircuits between DMN/DAN regions and the hippocampus with the hypothesis that they would provide a finer grained characterization of the connectivity between them, which in turn may provide more definitive evidence for or against the HERNET model.

Previous invasive studies in animals have shown that the connections between the hippocampus and DAN/DMN primarily involve cortical layers II and V, with layer V of the higher order cortex (frontal and parietal cortices) projecting to the hippocampus, whereas layer II of higher order cortex receiving the signal back from the hippocampus [360] [389]. This layer-specific pathway between the hippocampus and DAN/DMN is not exclusive since pathways which originate/terminate in other layers of the cortex may also contribute to the hippocampal input or output. This is not surprising given the highly complex underlying microcircuitry and given that signals between any two brain regions can relay via multiple structures including the thalamus. However, the pathways between the hippocampus and layers II and V of the DMN/DAN seem to be the dominant ones based on prior invasive animal literature [390] [391].

The present study sought to investigate the HIPER/HERNET model using the functional connectivity between the hippocampus and 1) deeper layers of DAN/DMN during an encoding task, and 2) superficial layers of DAN/DMN during a retrieval task. Specifically, we hypothesized that during a memory encoding task, clustering hippocampal voxels based on their functional connectivity with layer V of the DAN must parcellate the hippocampus in an anterior-posterior gradient along the long axis. Similarly, during a memory retrieval task, clustering

hippocampal voxels based on their functional connectivity with layer II of the DMN must also show an anterior-posterior segmentation. Second, during an encoding task, the hippocampal voxels must have stronger connectivity with layer V of the DAN than with layer V of the DMN. Contrarily, during a retrieval task, the hippocampal voxels must have stronger connectivity with layer II of the DMN than with layer II of the DAN. Third, considering the directionality of signal projection, during an encoding task, layer V of the DAN must show stronger correlation with anterior hippocampal regions than with posterior hippocampal regions, whereas during retrieval task, layer II of the DMN must exhibit stronger correlation with posterior hippocampal regions than with anterior hippocampal regions. Finally, we predicted that using layer-specific data would lead to more definitive results than using conventional volume-level data while investigating the connection between the hippocampus and the DAN/DMN. All four hypotheses, based on predictions from the HERNET model, are illustrated in Figure 4.1. Here it is noteworthy that although we intend to test these hypotheses using fMRI data extracted from layers II and V of DAN/DMN regions, some partial volume effects are expected for the spatial resolution of our data. Therefore, signal from layer II broadly represents those from supra-granular layers with dominant contribution from layer II. Likewise, signal from layer V broadly represents those from infra-granular layers with dominant contribution from layer V.

In order to test these hypotheses, fMRI data was acquired from healthy subjects performing memory encoding and retrieval tasks with faces, pictures of scenes and words in the 7T scanner. Ultra-high resolution anatomical data was also acquired to resolve layers in the DAN/DMN. Unsupervised clustering methods were applied, which groups objects in a data driven way without using any labels to guide the results, to parcellate the hippocampal voxels based on their connectivity with different layers in the DAN/DMN. Three clustering methods were specifically

chosen, i.e., hierarchical clustering [123], ordering points to identify the clustering structure (OPTICS) [57], and density peak clustering (DPC) [124], since they did not require *a priori* specification of the number of clusters. Since clustering accuracy is often lower in high dimensional feature space, a genetic algorithm (GA) based feature selection method was also employed, which is less prone to local optimum [127], comparing with other existing feature selection methods, e.g., sequential forward searching [127], non-linear optimization [128], etc.

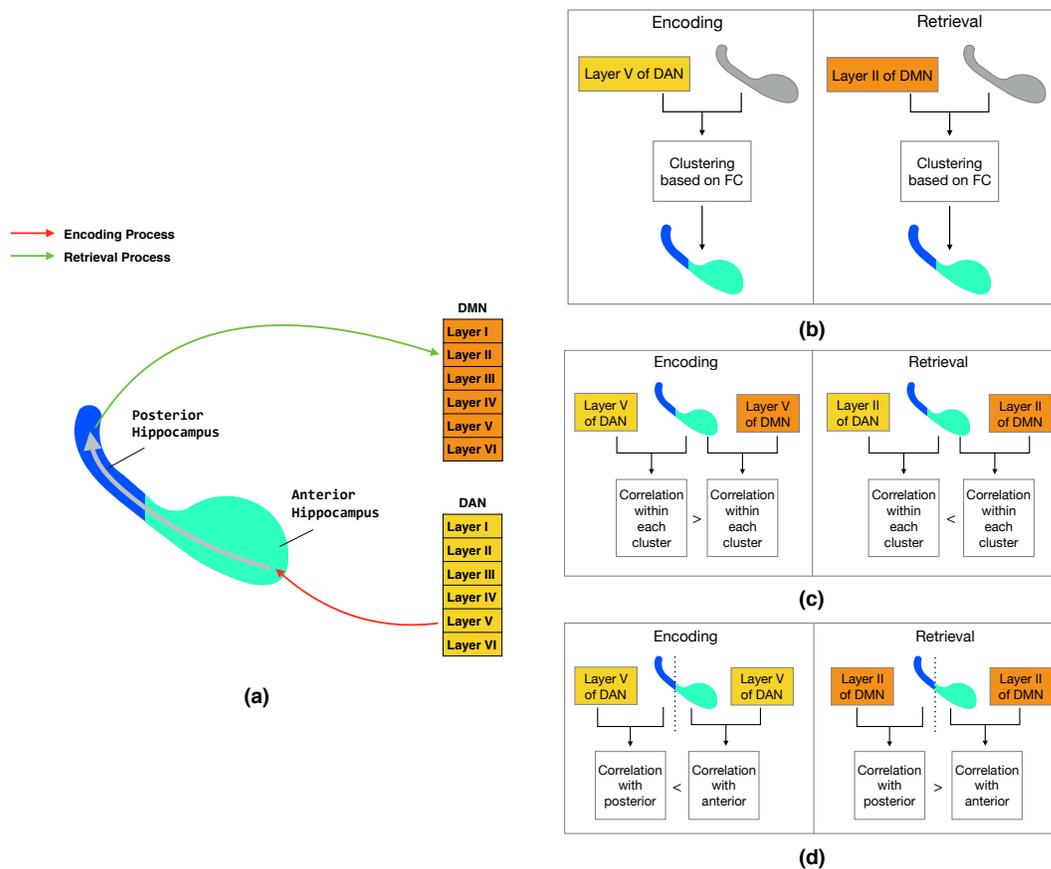


Figure 4.1 Illustration of our hypotheses based on predictions from the HERNET model as well as known anatomical pathway between different layers of DAN/DMN and hippocampus: a) anatomical pathway between hippocampus and different layers of DAN/DMN during

encoding/retrieval processes, b) illustration of proposed first hypothesis, c) illustration of proposed second hypothesis, and d) illustration of proposed third hypothesis.

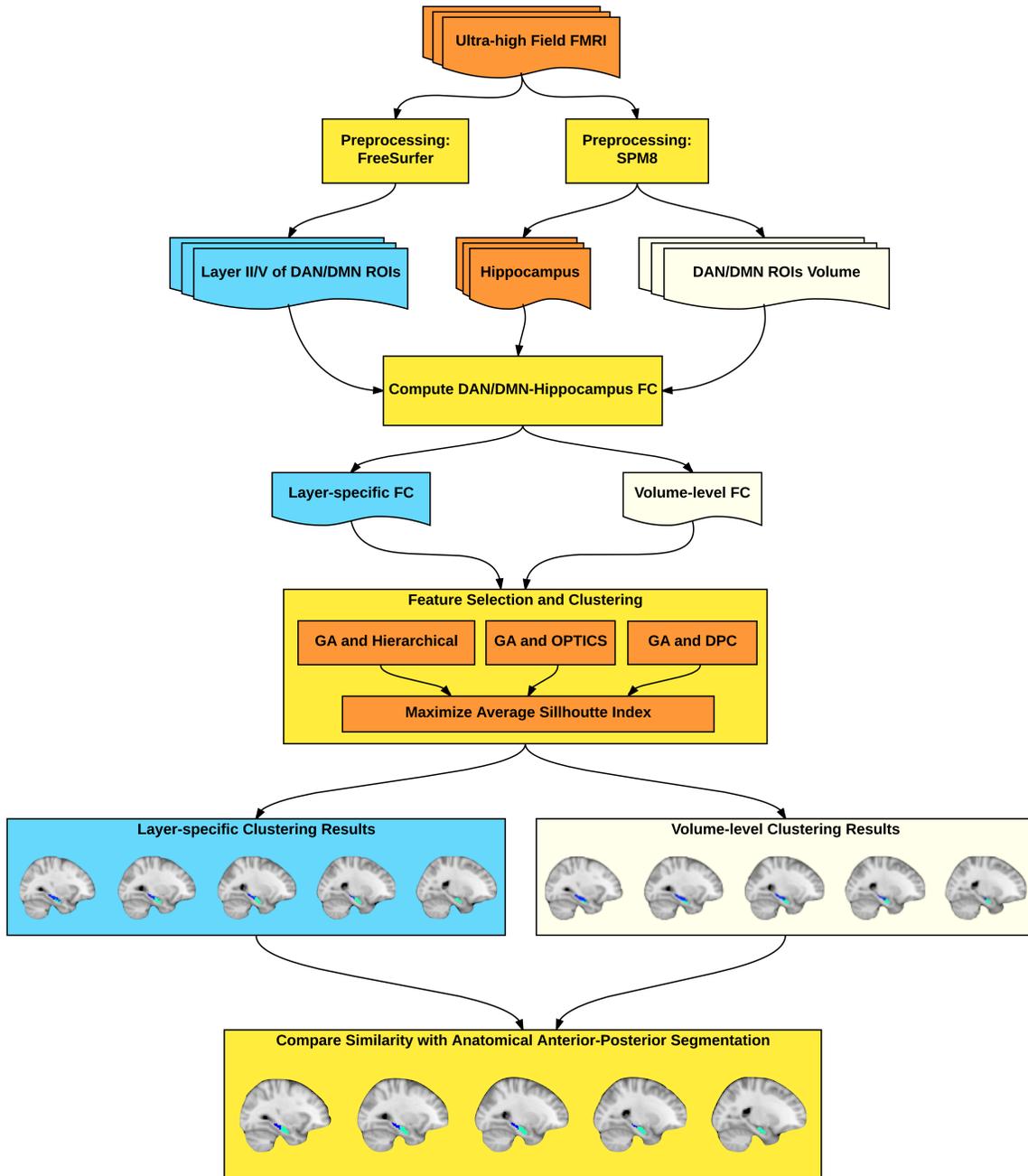


Figure 4.2 Illustration of proposed analysis pipeline for investigating hippocampal parcellation based on its layer-specific connectivity with DAN/DMN ROIs. The same process was repeated

for encoding and retrieval tasks, separately, as well as with conventional volume data (as opposed to layer-specific data).

4.2 Materials and Methods

In this work the connectivity between the hippocampus and the DAN/DMN were investigated by clustering hippocampal voxels, in an unsupervised way, based on their connectivity with 1) layer V of DAN/DMN during an encoding task, and 2) layer II of DAN/DMN during a retrieval task. The identified functional clusters of the hippocampus were then compared with an anatomical anterior-posterior segmentation [369]. The entire analysis pipeline is illustrated in Figure 4.2 and will be elaborated below.

4.2.1 Data Acquisition

Thirty-one healthy individuals (26 right-handed, 12 males, 19 females, age =21.1±1.4) were recruited for the study. The Internal Review Board (IRB) at Auburn University approved the study, subjects provided informed consent and the experimental procedures were performed in accordance with internationally accepted ethical standards. Echo-planar imaging (EPI) data were acquired on the Auburn University MRI Research Center (AUMRIRC) Siemens 7T MAGNETOM scanner outfitted with a 32-channel head coil by Nova Medical (Wilmington, MA). The sequence was optimized for the hippocampus (37 slices acquired parallel to the AC-PC line, 0.85mm×0.85mm×1.4mm voxels, TR/TE: 3000/28ms, 70° flip angle, base/phase resolution: 234/100, A→P phase encode direction, iPAT GRAPPA acceleration factor = 3, interleaved acquisition, 123 time points, total acquisition time 6 minutes). During encoding task, the participants were asked to view a series of faces, pictures of scenes and words. Each trial lasted for 30 seconds in which the subjects were presented 10 images of the same category for 3

seconds each. These trials were interspersed with a 6 second inter-trial interval. The paradigm for the retrieval task was identical to the encoding task with the exception that the subjects were provided with an MR-compatible button box to indicate, via button presses, whether they recognize the image as having seen during the encoding task or not. A whole-brain high-resolution 3D MPRAGE sequence was used to acquire anatomical data (256 slices, 0.63mm x 0.63mm x 0.60mm, TR/TE: 2200/2.8, 7° flip angle, base/phase resolution 384/100%, collected in an ascending fashion, acquisition time = 14:06) for extracting different cortical layers in the DAN/DMN ROIs and for registration purposes.

4.2.2 Preprocessing

a) Hippocampal data: Standard pre-processing steps were carried out using SPM8 [349] including brain extraction, slice timing correction, temporal band-pass filtering (0.01 to 0.10), regression of motion and physiological artifacts (using CompCor [392]), registration to anatomical space, and normalization to MNI standard space.

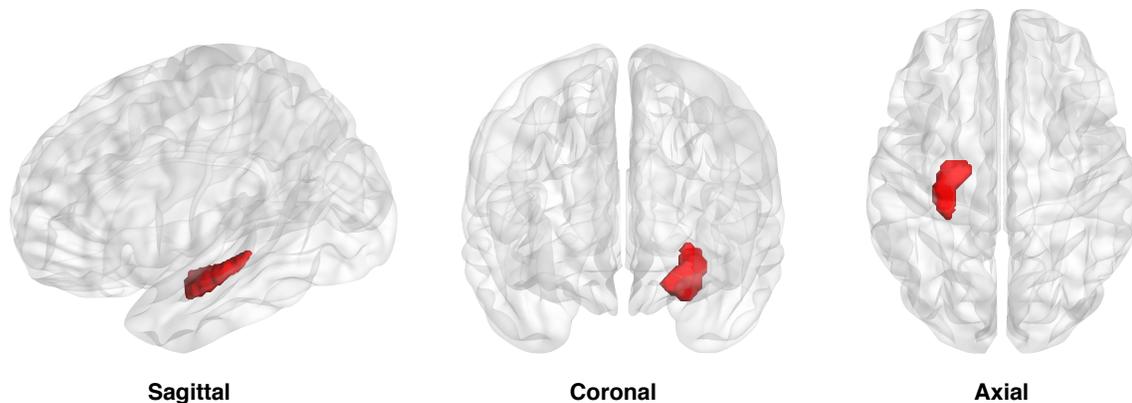


Figure 4.3 Hippocampal ROI used in this study. Only the left hippocampus was considered in this study.

The Harvard-Oxford Structural Probability Atlas distributed with the FSL neuroimaging analysis software package (<http://www.fmrib.ox.ac.uk/fsl/fslview/atlas-descriptions.html#ho>) was used to define hippocampal ROIs. Our task did not explicitly involve spatial memory typically used in navigation, which is primarily associated with the right hippocampus [346]. Therefore, only the left hippocampus was considered in this study. In order to determine a conservative anatomical representation, the hippocampal ROI was thresholded at 75%. The mean probability for the voxels in the hippocampal ROIs belonging to the hippocampus ($M \pm SD$: $86.41\% \pm 7.10\%$) and the voxel centroid (MNI coordinates $[-26, -18.8, -17.2]$) belonging to the hippocampus was $> 86\%$ and 97.1% , respectively. The identified hippocampal ROIs are illustrated in Figure 4.3 with a total volume of 1880 mm^3 .

b) Cortical layer-specific data of the DAN/DMN: Data extracted from the whole brain was first preprocessed using SPM8. To preserve high spatial resolution, no spatial filtering was applied. The cortical layers were then reconstructed using FreeSurfer [350]. Specifically, two interfaces, i.e., the cortical gray matter and the underlying white matter (white-gray interface) and the interface between the cortical gray matter and the pial surface (gray-pial interface), were automatically reconstructed from the anatomical image. Then, the cortical thickness was calculated as the average of the distance from the white-gray interface to the closest possible point on the gray-pial interface, then from that point back to the closest point on the white-gray interface again. To improve accuracy, surface smoothing and automatic topology correction were also applied [351] [352] [353] [354]. From the cortical thickness map, six cortical layers were then reconstructed within the cortical gray matter at fixed relative distance between the white and pial surfaces determined from the cortical thickness, i.e., the first layer was located at 96% of the cortical thickness away from the white matter, the second layer at 80%, the third layer at 64%,

the fourth layer at 48%, the fifth layer at 32%, and the six layer at 16% (Figure 4.4). The laminar layers were derived from the anatomical image, so it was necessary to align the EPI data to these layers. A boundary-based registration method [355] was employed, which aligns the EPI image to the anatomical image by maximizing the intensity gradient across white-gray interface and gray-pial interface. The entire cortical surface was automatically divided into 34 cortical ROIs in each of the individual hemispheres based on the Desikan-Killany atlas in Freesurfer [356]. The major DAN regions, i.e., frontal eye field [FEF], inferior temporal cortex [ITC], inferior frontal gyrus [IFG], and superior parietal lobe [SPL], and major DMN regions, i.e., anterior cingulate cortex [ACC], medial prefrontal cortex [mPFC], inferior parietal lobe [IPL], posterior cingulate cortex [PCC], and precuneus, were identified from those 34 ROIs, and vertices (voxels in the volume become vertices on surfaces) in layer II and layer V of these regions were then identified.

4.2.3 Connectivity Measures

FC measures the functional interrelationship between pairs of brain regions by estimating Pearson's correlation between time series representing those brain regions. Functional connectivity (FC) was estimated between the hippocampal voxels and 1) vertices in layer V of the DAN/DMN during the encoding task, and 2) vertices in layer II of the DAN/DMN during the retrieval task.

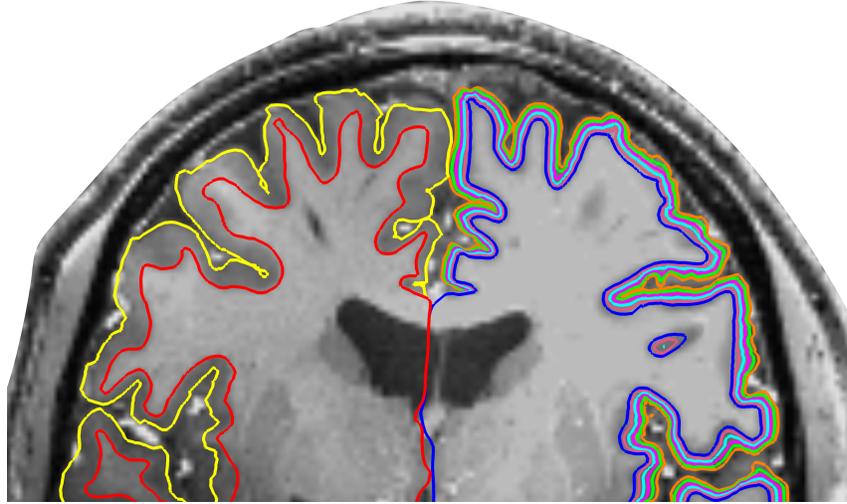
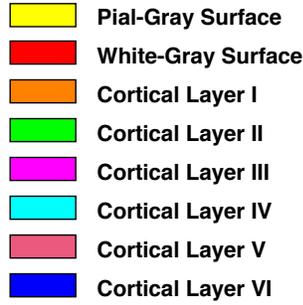


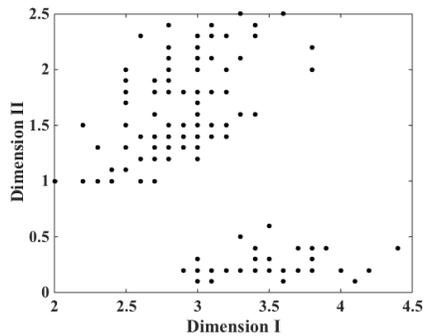
Figure 4.4 Cortical layer reconstruction with FreeSurfer. Six cortical layers were reconstructed within the cortical gray matter at fixed relative distances between the white and pial surfaces.

4.2.4 Layer Specific Clustering

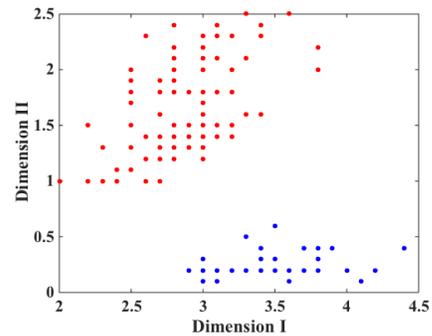
The hippocampal voxels were parcellated using three clustering methods based on their FC with 1) vertices in layer V of the DAN/DMN during the encoding task, and 2) vertices in layer II of the DAN/DMN during the retrieval task. The same process was repeated on DAN and DMN ROIs, separately. Let $\mathbf{Y} = \{\mathbf{Y}_1, \dots, \mathbf{Y}_i, \dots, \mathbf{Y}_N\}$ represent a set of N objects, i.e., number of hippocampal voxels. $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \dots, Y_{id}) \in \mathbb{R}^d$, where d equals to the number of FC features. Assume the N objects are separated into k clusters. Each cluster is a set of indexes from $\{1, \dots, N\}$, and each object \mathbf{Y}_i belongs to exactly one cluster.

a) Hierarchical Clustering (Agglomerative): As one of the most commonly used connectivity-based clustering method, the hierarchical clustering [184] [185] [123] groups objects into different clusters by building a hierarchical tree structure. The procedure of this method is illustrated below:

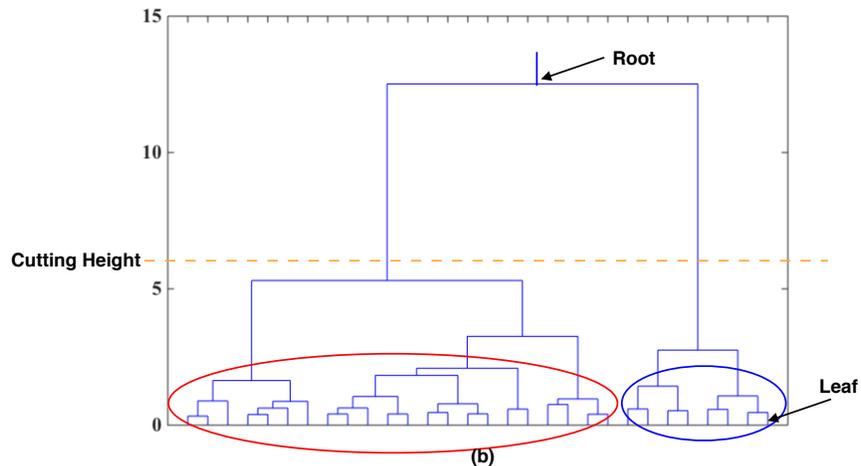
- 1) Initially, each object Y_i is assigned to a cluster with only itself in it.
- 2) Distance between any two clusters is measured. Then, the closest pair of clusters are merged.
- 3) Step 2-3 are repeated until all Y_i are in one big cluster.



(a)



(c)



(b)

Figure 4.5 Illustration of hierarchical clustering. a) Original simulated dataset, b) Dendrogram derived from hierarchical clustering, and c) Clustering results obtained with a specific cutting height. Two clusters that were identified are marked with different colors.

The resulting tree structure is usually referred to as the dendrogram (Figure 4.5). The root of the dendrogram represents the entire data, each leaf node represents one single object, and the height of the dendrogram represents the distance between each pair of clusters. Different data partitions can be obtained by cutting the dendrogram at different heights. Note that there are three linkage criteria, single-linkage, complete linkage, and average-linkage, which have been widely used in measuring distance between two clusters. The single linkage [186] calculates the shortest distance between two clusters, the complete linkage [186] calculates the longest distance, and the average linkage [186] calculates the mean distance. The single linkage method can handle non-elliptical shape of clusters, but can be affected by noise and outliers. The complete linkage method is less sensitive to noise and outliers but tends to break large clusters. The average linkage is a compromise between single-linkage and complete linkage methods. Thus, the average linkage method was employed in this work.

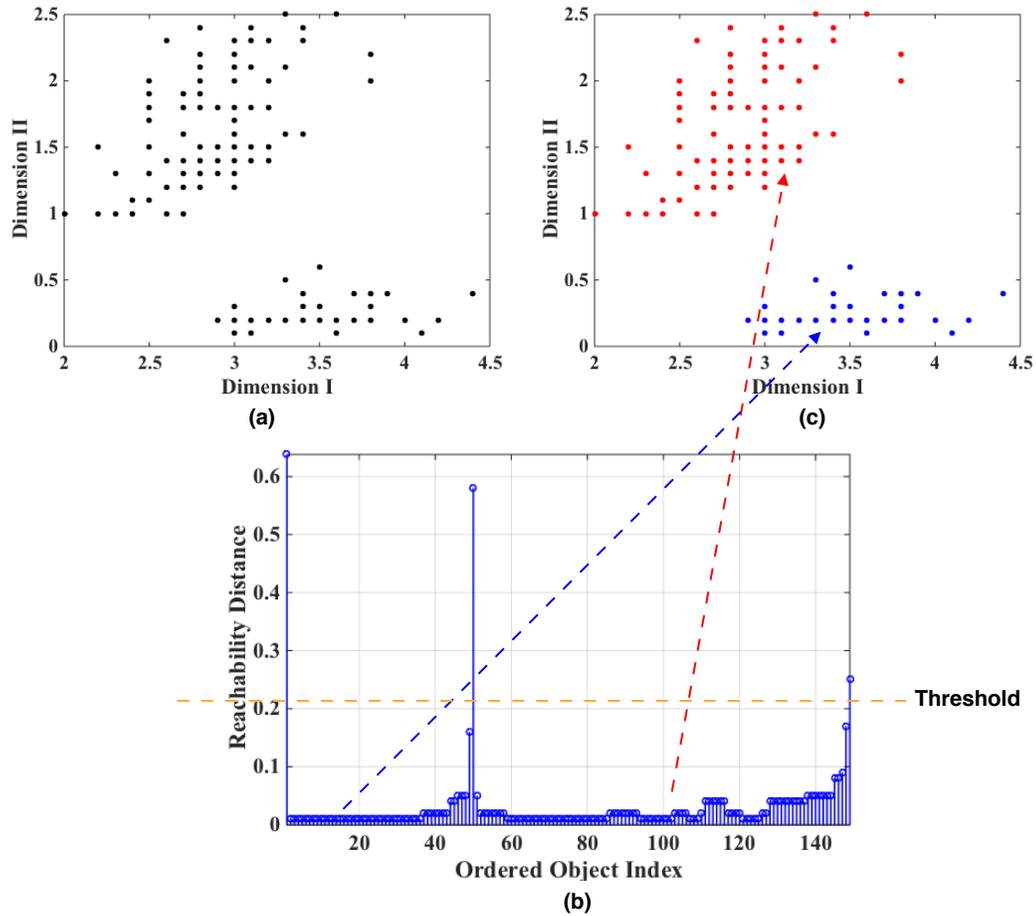


Figure 4.6 Illustration of OPTICS clustering. a) Original simulated dataset, b) reachability plot obtained from OPTICS, and c) clustering results. Two clusters were identified corresponding to valleys in the reachability plot.

b) Ordering Points to Identify the Clustering Structure: OPTICS [57] is one of the most popular density-based clustering methods [55]. Given a distance threshold (ϵ) and the minimum number of objects required to form a cluster (*MinPts*), objects in high-density areas are grouped together, whereas objects in sparse areas, which are required to separate clusters, are usually considered to be noise or outliers. OPTICS can discover clusters with arbitrary shapes and has the ability to identify outlier objects that do not belong to any of the clusters.

In OPTICS, two variables are computed for each object in the dataset: *core-distance* and *reachability-distance*. Let $N_\varepsilon(\mathbf{Y}_i)$ represent the number of nearby objects within ε (called ε - *neighborhood*), and $MinPts - distance(\mathbf{Y}_i)$ represent the distance from \mathbf{Y}_i to its *MinPts*' neighbor. An object \mathbf{Y}_i is a core object if at least *MinPts* objects are found with its ε - *neighborhood*. The *core-distance* of \mathbf{Y}_i is defined as:

$$core - dist_{\varepsilon, MinPts}(\mathbf{Y}_i) = \begin{cases} Undefined & \text{if } N_\varepsilon(\mathbf{Y}_i) < MinPts \\ MinPts - distance(\mathbf{Y}_i) & \text{otherwise} \end{cases} \quad (4.1)$$

,which is the smallest distance for \mathbf{Y}_i to have *MinPts* in its ε - *neighborhood*.

The reachability-distance of object \mathbf{Y}_j with respect to object \mathbf{Y}_i is defined as:

$$\begin{aligned} reachability - dist_{\varepsilon, MinPts}(\mathbf{Y}_j, \mathbf{Y}_i) \\ = \begin{cases} Undefined & \text{if } N_\varepsilon(\mathbf{Y}_i) < MinPts \\ \max(core - distance(\mathbf{Y}_i), dist(\mathbf{Y}_j, \mathbf{Y}_i)) & \text{otherwise} \end{cases} \end{aligned} \quad (4.2)$$

Where $dist(\mathbf{Y}_j, \mathbf{Y}_i)$ is the distance measure (e.g., Euclidean distance) between \mathbf{Y}_j and \mathbf{Y}_i . The complete procedure of OPTICS is described below:

- 1) Choose one object \mathbf{Y}_i arbitrarily.
- 2) Retrieve the ε - *neighborhood* of \mathbf{Y}_i , determine the core-distance of \mathbf{Y}_i , and set the reachability-distance of each object \mathbf{Y}_j in the ε - *neighborhood* of \mathbf{Y}_i to undefined.
- 3) If \mathbf{Y}_i is not a core object, go to step 5. Otherwise, go to step 4.
- 4) For each object \mathbf{Y}_j in the ε - *neighborhood* of \mathbf{Y}_i , update its reachability-distance from \mathbf{Y}_i and insert \mathbf{Y}_j into an OrderSeeds list if it has not been processed yet.

- 5) If the input dataset is fully consumed and the OrderSeeds list is empty, go to step 6. Otherwise, move on to the next object in the OrderSeeds list (or the input list, if the OrderSeeds list is empty) and go to step 2.
- 6) Output core-distance, reachability-distance of each object, and processed order.

The data objects are plotted in the processed order together with their respective reachability-distance (called reachability plot) depicting the hierarchical structure of the clusters. Since objects belonging to a cluster have a low reachability-distance to their nearest neighbor, the clusters show up as valleys in the reachability plot (see Figure 4.6). The final data partition can be obtained by using a threshold on the reachability plot.

c) Density Peak Clustering: recently Rodriguez and Laio [124] proposed a novel density-based clustering method (referred to as DPC) based on the idea that the cluster centers are characterized by a higher density than their neighbors and by a relatively large distance from objects with higher densities. Like other density-based clustering methods, e.g., OPTICS, it has ability to detect arbitrarily shaped clusters and spot outlier objects. Moreover, DPC outperforms commonly used clustering methods, e.g., k-means and hierarchical clustering, when the dataset contains complicated features such as narrow bridges between clusters, uneven-sized clusters, clusters with high overlap, etc.

For each object \mathbf{Y}_i , two quantities are computed: local density $\rho(\mathbf{Y}_i)$ and minimum distance with higher density $\delta(\mathbf{Y}_i)$. $\rho(\mathbf{Y}_i)$ is defined as:

$$\rho(\mathbf{Y}_i) = \sum_i \chi(\text{dist}(\mathbf{Y}_i, \mathbf{Y}_j) - d_c) \quad (4.3)$$

where d_c is a cutoff distance, and $\chi(y)$ can be computed by,

$$\chi(y) = \begin{cases} 1 & \text{if } y < 0 \\ 0 & \text{otherwise} \end{cases} \quad (4.4)$$

From Equation (4.3) and (4.4), it can be seen that $\rho(\mathbf{Y}_i)$ equals to the number of objects within d_c with respect to object \mathbf{Y}_i . $\delta(\mathbf{Y}_i)$ is measured by,

$$\delta(\mathbf{Y}_i) = \min_{i:\rho(\mathbf{Y}_j) > \rho(\mathbf{Y}_i)} \text{dist}(\mathbf{Y}_i, \mathbf{Y}_j) \quad (4.5)$$

For the object with highest density, $\delta(\mathbf{Y}_i)$ is conventionally set to,

$$\delta(\mathbf{Y}_i) = \max_i \text{dist}(\mathbf{Y}_i, \mathbf{Y}_j) \quad (4.6)$$

Note that if \mathbf{Y}_i is local or global maxima in the density, $\delta(\mathbf{Y}_i)$ will be much larger than its typical nearest neighbor. Thus, objects with larger ρ and δ are considered as cluster centers, whereas objects with smaller ρ and larger δ are considered as outliers. Other objects are assigned to the same cluster as their nearest neighbor of higher density (see Figure 4.7).

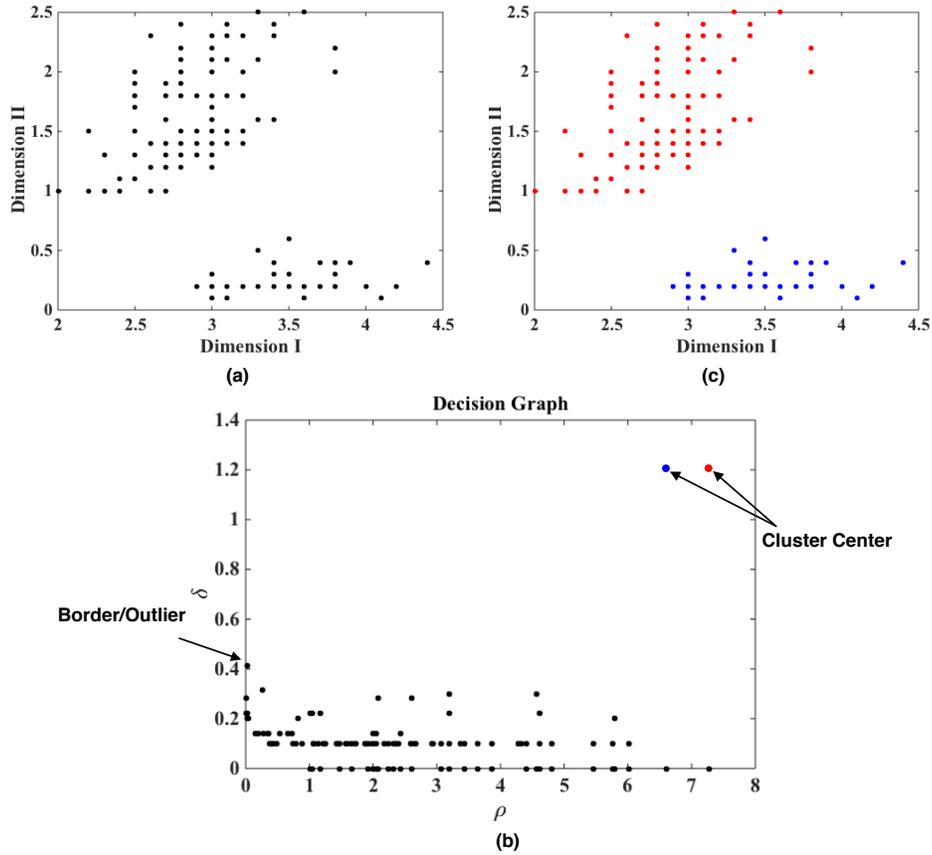


Figure 4.7 Illustration of DPC clustering. a) Original simulated dataset, b) Plot of δ as a function of ρ for each object. Objects with larger ρ and δ are cluster centers and objects with smaller ρ , and larger δ are outliers. c) Clustering results. Two clusters were identified corresponding to two cluster centers in the decision graph.

d) Input Parameter Optimization: In each clustering method, there are several user-specified input parameters, which can significantly affect the shape of the cluster and the number of the cluster (Figure 4.8). For hierarchical method, the cutting height of the dendrogram needs to be specified and the number of clusters varies with different cutting heights. For OPTICS, ε can simply be set to the maximum possible value, and Ankerst and colleagues [57] showed that for *MinPts* using values between 10 and 20 would always lead to good results. However, the

threshold for the reachability plot, which is used to extract clusters, still needs to be properly determined. For DPC, d_c can be chosen based on the rule that the average number of neighbors is around 1-2% of the total number of objects in the data set [124]. A threshold for ρ and δ needs to be defined to distinguish cluster centers, borders, and outliers. In this study, the average silhouette index [60] was employed to determine the optimal values of these parameters. Assume the data have been clustered via any clustering algorithm, such as OPTICS, into K clusters. For each object Y_i , let $a(Y_i)$ represent the average distance of Y_i with all other object in the same cluster, and $b(Y_i)$ represent the smallest average distance of Y_i to any other cluster, of which Y_i does not belong to. Then the silhouette index of Y_i is defined as:

$$s(Y_i) = \frac{b(Y_i) - a(Y_i)}{\max \{a(Y_i), b(Y_i)\}} \quad (4.7)$$

From Equation (4.7), it can be seen that $s(Y_i)$ is bounded between -1 and 1. If Y_i has been assigned to a “correct” cluster, $s(Y_i)$ will be close to 1. Contrarily, if Y_i has been assigned to a “wrong” cluster, $s(Y_i)$ will be close to -1. $s(Y_i)$ will be close to 0, if Y_i is located on the border of two natural clusters. By computing the average $s(Y_i)$ over all objects in the entire dataset, the accuracy of the clustering results can be quantified.

With the average silhouette index as the optimization criterion, a “grid search” [187] method was applied on determining the optimal value of each parameter. The optimal number of clusters can be determined simultaneously. For example, in OPTICS, we started with a relatively high threshold for the reachability-plot. In each iteration, the threshold was reduced by a small amount and the average $s(Y_i)$ was computed and recorded based on the current partition. The iteration continues until the threshold was smaller than a specified baseline, e.g., the average reachability-

distance of the reachability-plot. The optimal threshold was then determined as the one with the largest average $s(Y_i)$. The same iterative procedure was applied to hierarchical clustering to determine optimal cutting height of dendrogram, and to DPC to determine the optimal threshold of ρ and δ .

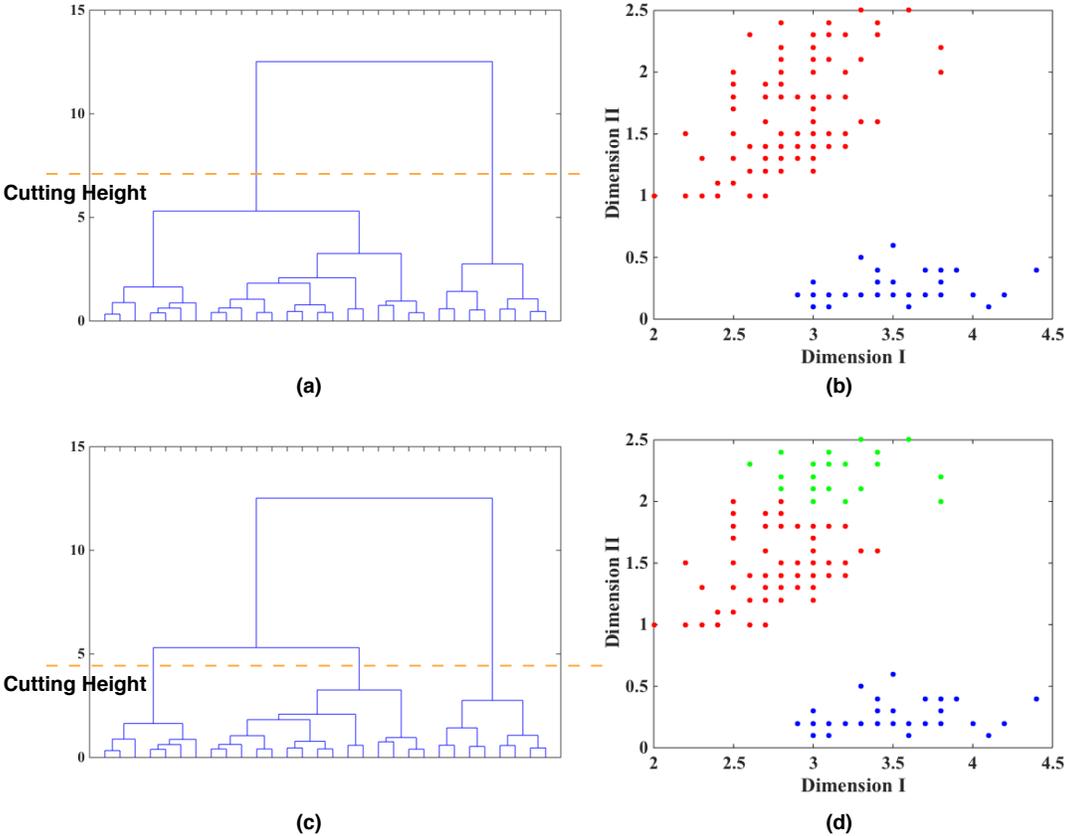


Figure 4.8 Illustration of dependency of clustering results on input parameters. Take hierarchical clustering as an example. With a relatively high cutting height (a), two clusters (red and blue) were identified (b). As cutting height was reduced (c), one big cluster (red) was separated into two smaller clusters (d).

4.2.5 Feature Selection and Cluster Identification

The clustering accuracy is often lower in high dimensional feature space, which may due to the fact that most of features in the dataset may be irrelevant, redundant, or sometimes may even misguide results. Moreover, a large number of features make the clustering results difficult to interpret. Therefore, a feature selection method is required to improve the clustering accuracy. For supervised learning, feature selection can be trivial, i.e., only the features that are related to the given cluster labels are maintained. Nevertheless, for unsupervised learning, the cluster labels are unknown. Thus, finding the relevant subset of features and clustering the subset of the data must be accomplished simultaneously.

Assuming d to be the initial number of features, an exhaustive search of 2^d possible subsets need to be examined, which is computationally expensive. Therefore, in this study an alternative GA based method was applied to determine the optimal subset of features, as well as the optimal clustering results. The average silhouette index was used as the optimization criteria.

As one of the most popular search heuristic methods, GA has been widely used in generating solutions to optimization and searching problems [189] [190]. Different from single-state methods (only one solution is evaluated at a time), e.g., simulated annealing, hill climbing, etc., GA is a “population” method that maintains a set of solutions evolving toward an optimal solution. The evolution usually starts from a population of randomly generated solutions. In each iteration (or so called “generation”), “survivor solutions” with larger values of optimization criteria, i.e., the average silhouette index, are selected to form a new generation of solutions. These survival solutions can be generated from the crossover, which produced new solutions by randomly combining two current solutions, mutations, which randomly changes new solutions

with a small probability, or from the initial population. The new generation of solutions is then used in the next iteration of the algorithm. Conventionally, the algorithm terminates when the best solution cannot be improved any further.

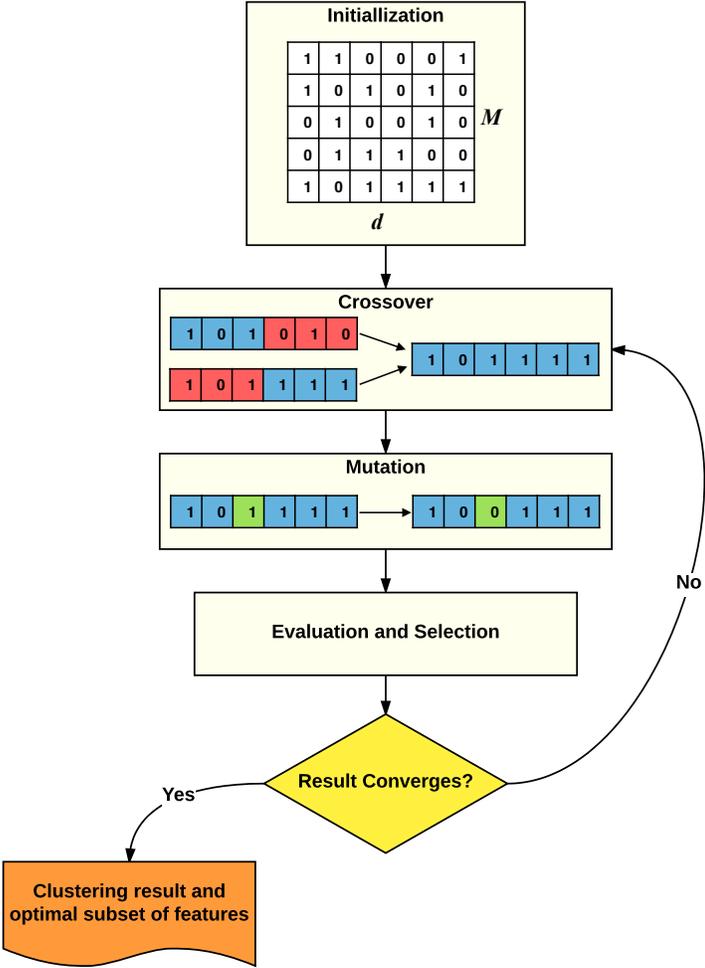


Figure 4.9 Flowchart of GA for feature selection. In the M -by- d matrix, each row represents a candidate solution, describing a subset of selected features. Each of the d bits in a row represents whether a feature is selected (1) or discarded (0).

In this study, an array of d bits was used to represent the selected subset of features and the population size is represented using M . Each bit in the array indicates the activation status of one

specific feature: 1 indicates selected and 0 indicated discarded. The complete procedure of GA is described below (Figure 4.9):

- 1) Initialization: 400 candidate solutions were generated by randomly setting 1 or 0 for each bit in vectors.
- 2) Crossover: two candidate solutions A and B were randomly selected from the current population. A value v between 1 and d was randomly selected. Then a new solution was formed by combining the feature bits 1 to v from A and feature bits $v + 1$ to d from B.
- 3) Mutation: for each new generated solution, a mutation was applied by reversing bits in the vector with a probability of 0.1.
- 4) Evaluation: the clustering method was applied on each candidate solution (i.e., a subset of selected features), and the average silhouette index was computed for each obtained partition.
- 5) Selection: 280 solutions resulting in high average silhouette index were selected along with 120 solutions randomly selected from the rest of the solution (to increase the diversity of the population).
- 6) If the result did converge, i.e., the average silhouette index of the best solution in the population keep increases, we iterated back to step 2. Otherwise, the clustering result with the largest average silhouette index and the corresponding selected subset of features were saved as the output.

4.2.6. Volume Level Clustering

In order to determine whether characterizing layer-specific microcircuits using ultra-high field fMRI provides any advantages over conventionally computed voxel-level connectivity, the same

clustering procedure enumerated above was also performed on the FC computed between hippocampal and DAN/DMN voxels during encoding/retrieval tasks. Specifically, let $\mathbf{Y} = \{\mathbf{Y}_1, \dots, \mathbf{Y}_i, \dots, \mathbf{Y}_N\}$ represent a set of N objects, i.e., number of hippocampal voxels. $\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \dots, Y_{id}) \in \mathbb{R}^d$, where d equals to the number of FC features computed between hippocampal voxels and voxels in the DAN/DMN, respectively. Subsequently, the same clustering and feature selection process was repeated on encoding and retrieval tasks, DAN and DMN ROIs, separately.

4.2.7 Comparison with Anatomical Anterior-Posterior Segmentation

In this study, clusters were identified based on the DAN/DMN-hippocampal FC during encoding/retrieval tasks, respectively. This provided a functional parcellation of the hippocampus, which was compared with anatomically delineated anterior and posterior hippocampal segments.

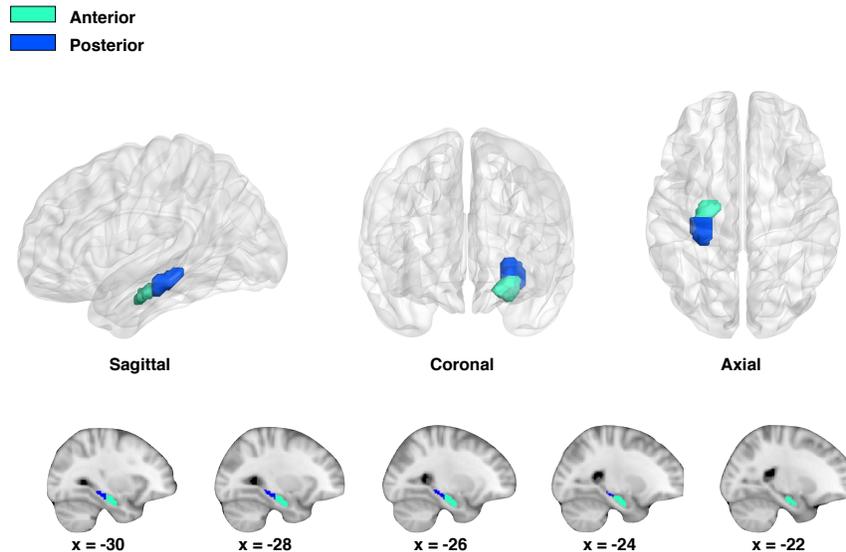


Figure 4.10 Anatomical anterior-posterior segmentation used in this study. (Coordinates are in MNI space)

The hippocampus can be anatomically separated into head, body, and tails, with the head and body being considered as anterior, and the tail being considered as the posterior part. Various methods, such as landmark-based segmentation [369], percentile-based axis segmentation [393] [394], a Talairach/MNI coordinate-based segmentation [369], have been used to define the anterior and posterior regions of the hippocampus. We employed the Talairach/MNI coordinate-based segmentation [369] which chose $y = -21$ in MNI space ($y = -20$ in Talairach space) as the border between anterior and posterior segmentations as shown in Figure 4.10. This coordinate corresponds to the uncus apex, which is considered as the end of the posterior portion of the hippocampus [395].

Let $A = \{A_1, A_2, \dots, A_M\}$ represent m anatomical parcels, and $F = \{F_1, F_2, \dots, F_M\}$ denote n functional parcels identified by applying clustering methods on FC features. The similarity between these two parcellations was then quantified using Torres' method [188]. The similarity matrix for A and F is an $m \times n$ matrix defined as:

$$S_{A,F} = \begin{bmatrix} S_{11} & \dots & S_{1j} & \dots & S_{1n} \\ & & \vdots & & \\ S_{i1} & \dots & S_{ij} & \dots & S_{in} \\ & & \vdots & & \\ S_{m1} & \dots & S_{mj} & \dots & S_{mn} \end{bmatrix} \quad (4.8)$$

where $S_{ij} = i/u$, which is Jaccard's Similarity Coefficient with i being the size of intersection and u being the size of the union of cluster sets A_i and F_j . The similarity of parcellations A and F is then defined as:

$$Sim(A, F) = \frac{\sum_{i \leq m, j \leq n} S_{ij}}{\max(m, n)} \quad (4.9)$$

From Equation (4.8) and (4.9), it can be seen that $0 \leq Sim(A, F) \leq 1$, and $Sim(A, F) = 1$ when two parcellations are identical.

The entire analysis pipeline proposed for investigating functional differentiation of the hippocampus and layer-specific microcircuitry between the hippocampus and the DAN and DMN during encoding/retrieval tasks are illustrated in Figure 4.2.

4.3 Results

The optimal values of each input parameter determined for the three clustering methods are presented Table 4.1 and Table 4.2 for layer-level and volume-level clustering, respectively. Using each clustering method, the hippocampal voxels were clustered into two different functional parcels based on their FC with 1) layer V of the DAN/DMN during the encoding task, and 2) layer II of the DAN/DMN during the retrieval task. This was true across methods. The obtained clusters were then mapped back to the image space and the resulting hippocampal parcels were overlaid on the anatomical image for the visualization.

Similar hippocampal parcellations were discovered using their FC with layer V of the DAN/DMN during the encoding task and layer II of the DAN/DMN during the retrieval task, and the average cluster similarity between DPC and hierarchical clustering, between DPC and OPTICS, and between hierarchical clustering and OPTICS, were 0.93, 0.95, and 0.92, respectively. For illustration, the clustering results obtained using the DPC method are shown in Figure 4.11. The clustering results obtained using hierarchical clustering and OPTICS are shown in Supplementary Information Figure B.1 and Figure B.2. From Figure 4.11, it can be seen that

the hippocampus showed an anterior-posterior gradient along the long-axis, which is consistent with the anatomical anterior-posterior segmentation (Figure 4.10).

Method Name	Parameter	Left Hippocampus			
		DAN		DMN	
		Encoding Layer V	Retrieval Layer II	Encoding Layer V	Retrieval Layer II
DPC	ρ	9.03	7.13	8.29	8.50
	δ	3.15	2.82	3.14	3.04
Hierarchical	h	1.16	1.15	1.15	1.15
OPTICS	s	0.55	0.50	0.53	0.25

Table 4.1 Estimated optimal values of each input parameter in layer-specific clustering. h : cutting height, and s : reachability threshold.

Method Name	Parameter	Left Hippocampus			
		DAN		DMN	
		Encoding	Retrieval	Encoding	Retrieval
DPC	ρ	7.59	6.38	5.61	7.93
	δ	0.09	0.13	0.18	0.12
Hierarchical	h	1.15	1.15	1.16	1.15
OPTICS	s	0.39	0.06	0.002	0.03

Table 4.2 Estimated optimal values of each input parameter in volume-level clustering. h : cutting height, and s : reachability threshold.

■ Anterior
■ Posterior

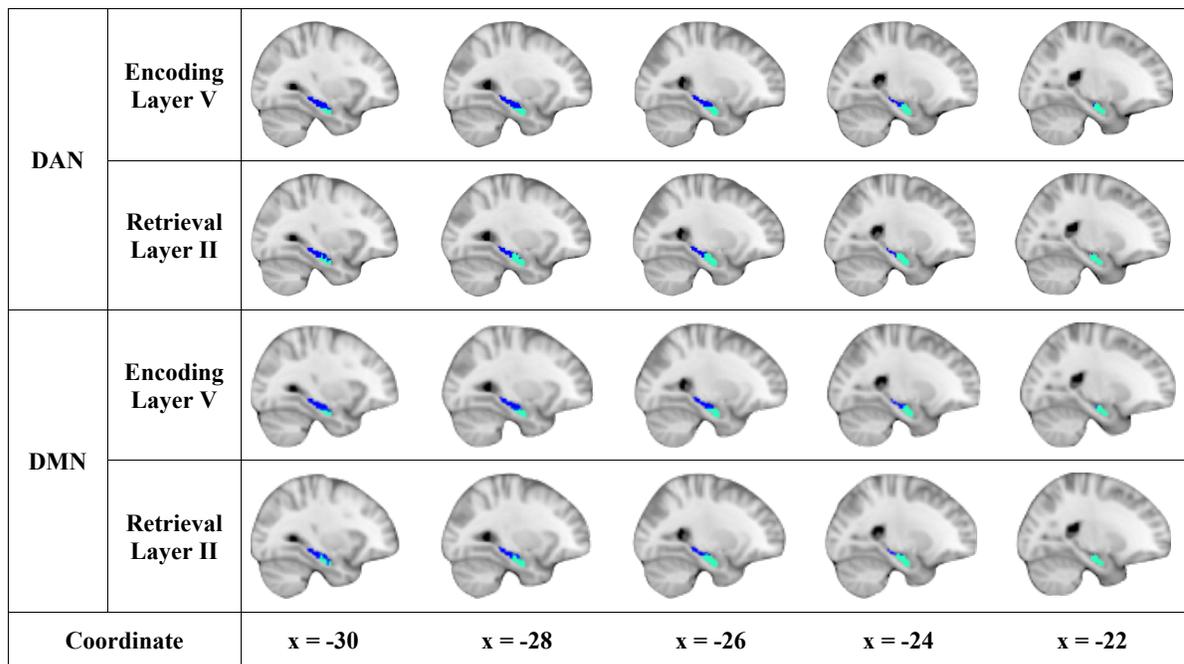


Figure 4.11 Clusters of hippocampal voxels determined (using the DPC method) based on their functional connectivity with 1) layer V of the DMN/DAN during the encoding task, and 2) layer II of the DMN/DAN during the retrieval task. (Coordinates are in MNI space)

Clustering Method	Feature Type	Task Name	DAN			DMN		
			Correlation		Sim.	Correlation		Sim.
			Anterior	Posterior		Anterior	Posterior	
DPC	Layer-specific	Encoding (Layer V)	-0.36	0.36	0.70	-0.13	0.16	0.70
		Retrieval (Layer II)	0.09	-0.06	0.73	-0.39	0.31	0.77
	Volume-level	Encoding	0.20	0.04	0.69	0.02	0.12	0.69
		Retrieval	0.30	0.33	0.74	-0.01	-0.02	0.74
Hier.	Layer-specific	Encoding (Layer V)	-0.36	0.35	0.70	-0.13	0.16	0.70

		Retrieval (Layer II)	-0.14	0.11	0.63	-0.41	0.34	0.77
	Volume-level	Encoding	0.48	-0.17	0.68	0.51	-0.18	0.69
		Retrieval	0.08	-0.05	0.72	0.01	0.02	0.74
OPTICS	Layer-specific	Encoding (Layer V)	-0.35	0.34	0.70	-0.09	0.10	0.70
		Retrieval (Layer II)	-0.14	0.07	0.73	-0.36	0.27	0.77
	Volume-level	Encoding	0.51	-0.18	0.69	0.02	0.03	0.71
		Retrieval	-0.33	0.41	0.57	0.21	-0.09	0.74

Table 4.3 Cluster similarity between functional and anatomical anterior-posterior parcellations using different clustering methods and mean correlation obtained within each cluster on the left side of the hippocampus.

To quantitatively characterize the identified clusters, the cluster similarity between our functionally obtained hippocampal parcels and anatomically defined anterior-posterior parcels (Figure 4.10) was computed. The mean correlation between hippocampal voxels and selected vertices (using GA-based feature selection method) in layer V of the DAN/DMN (during the encoding task) and the layer II of the DAN/DMN (during the retrieval task) was also computed within anterior and posterior hippocampal regions (Table 4.3).

For the left hippocampus, during the encoding task, the average cluster similarity, over different clustering methods, between functional and anatomical parcellations for layer V of the DAN and DMN was 0.70 and 0.70, respectively. The absolute correlation observed between layer V of the DAN and the hippocampus was significantly larger than that between layer V of the DMN and the hippocampus as the p-values shown in Table 4.4, which is in line with our second hypothesis. A one-tailed 2-sample t-test was conducted to test whether the correlation of

layer V of the DAN with the anterior hippocampal regions was significantly larger than that with the posterior hippocampal regions. The p-values obtained for different clustering methods were all close to 1, which did not provide evidence to support our third hypothesis for DAN part.

During the retrieval task, the average cluster similarity between functional and anatomical parcellations for layer II of the DAN and DMN was 0.70 and 0.77, respectively. The absolute correlation obtained between layer II of the DMN and the hippocampus was significantly larger than that between layer II of the DAN and the hippocampus as the p-values shown in Table 4.5, which was consistent with our second hypothesis. In addition, the correlation between layer II of the DMN and the posterior hippocampal regions was significantly larger than that with the anterior hippocampal regions (one-tailed 2-sample t-test; $p < 0.001$ for different clustering methods), considering the sign of the correlation. This result was in line with our third hypothesis for DMN part.

Feature Type	Clustering Method	Left Hippocampus					
		Mean Absolute Correlation				P-value	
		DAN		DMN			
		Anterior	Posterior	Anterior	Posterior	Anterior	Posterior
Layer-specific (Layer V)	DPC	0.39	0.39	0.34	0.35	< 0.001	0.005
	Hierarchical	0.38	0.39	0.34	0.35	0.004	0.009
	OPTICS	0.37	0.38	0.34	0.35	0.002	0.007
Volume-level	DPC	0.51	0.18	0.51	0.18	0.500	0.500
	Hierarchical	0.48	0.17	0.51	0.18	0.750	0.564
	OPTICS	0.51	0.18	0.45	0.19	0.170	0.507

Table 4.4 Comparison of the absolute correlations obtained between the DAN with the hippocampus and the DMN with the hippocampus during the encoding task by conducting one-

tailed two-sample t-test (H_0 : correlation of DAN and the hippocampus \leq correlation of the DMN and the hippocampus).

Feature Type	Clustering Method	Left Hippocampus					
		Mean Absolute Correlation				P-value	
		DAN		DMN			
		Anterior	Posterior	Anterior	Posterior	Anterior	Posterior
Layer-specific (Layer II)	DPC	0.37	0.37	0.51	0.41	< 0.001	0.002
	Hierarchical	0.34	0.38	0.51	0.41	< 0.001	0.023
	OPTICS	0.45	0.28	0.50	0.40	0.015	< 0.001
Volume-level	DPC	0.32	0.30	0.35	0.24	0.298	0.932
	Hierarchical	0.23	0.27	0.31	0.28	0.053	0.444
	OPTICS	0.33	0.41	0.32	0.26	0.540	0.973

Table 4.5 Comparison of the absolute correlations obtained between the DMN with the hippocampus and the DAN with the hippocampus during the retrieval task by conducting one-tailed two-sample t-test (H_0 : correlation of the DMN with the hippocampus \leq correlation of the DAN and the hippocampus).

For volume level analysis, i.e. clustering of hippocampal voxels based on their FC with voxels in DAN/DMN volume (as opposed to layer II and layer V of the DAN/DMN as before), similar parcellations were discovered using different clustering methods, and the average cluster similarity between DPC and hierarchical clustering, between DPC and OPTICS, and between hierarchical clustering and OPTICS, were 0.96, 0.88, and 0.87, respectively. For illustration, the clustering results using DPC are shown in Figure 4.12, and the clustering results using hierarchical clustering and OPTICS are shown in supplementary information B.3 and B.4.

■ Anterior
■ Posterior

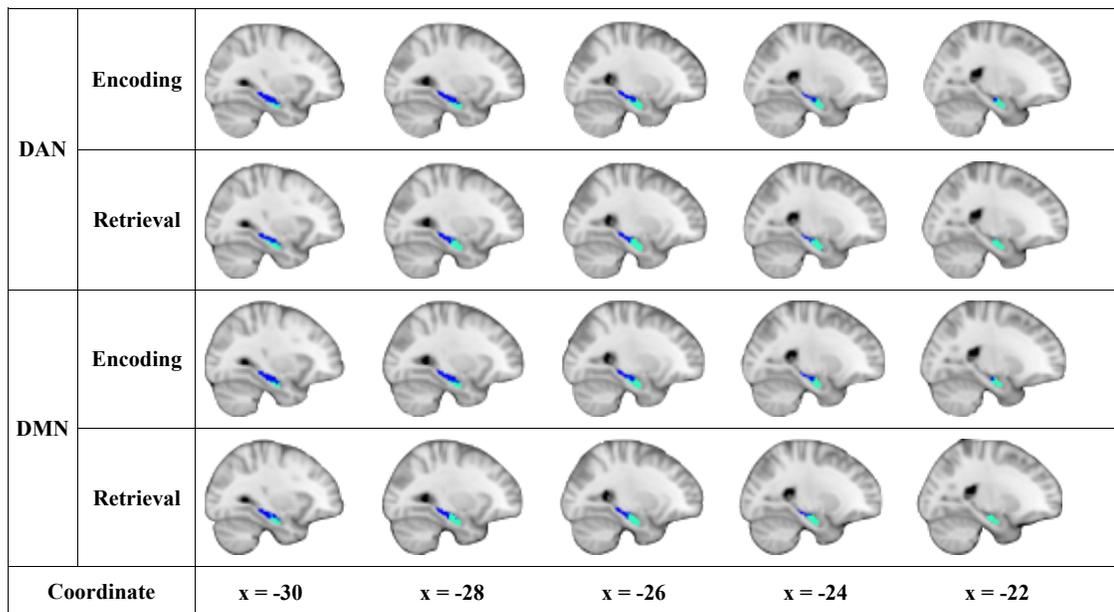


Figure 4.12 Clusters of hippocampal voxels determined (using the DPC method) based on their functional connectivity DMN/DAN volume during encoding and retrieval tasks. (Coordinates are in MNI space)

As shown in Figure 4.12, an anterior-posterior gradient was obtained as well. The cluster similarity between functional and anatomical parcellations, and the mean correlation within anterior and posterior regions were also computed for volume level clustering. As shown in Table 4.3, the average cluster similarity for the DAN (during the encoding task) and the DMN (during the retrieval task) were 0.69 and 0.74, respectively, which were qualitatively less than the value obtained using the layer-specific data. During the encoding task, the correlation between DAN and the hippocampus was not significantly greater than the correlation between DMN and the hippocampus ($p > 0.05$) for all three clustering methods (Table 4.4), whereas during the retrieval task, the correlation between DMN and the hippocampus was not significantly greater

than the correlation between DAN and the hippocampus for all three methods (Table 4.5). This result was different from the result obtained using layer-specific data. It was also contradictory to our second hypothesis. The correlations between anterior and posterior hippocampal regions with DAN/DMN for encoding/retrieval tasks were also compared using one-tailed 2-sample t-test. During the encoding task, the p-values obtained for DPC, hierarchical, and OTPICS were 0.019, < 0.001 , and 0.02, respectively, whereas during the retrieval task, the p-values obtained for these three methods were 1, 1, and 0.436, respectively. This result provided evidence for our third hypothesis for only the DAN part considering the sign of the correlation, not for the DMN part.

4.4 Discussion

In this work, we investigated functional differentiation of the hippocampus and the connectivity between the hippocampus and DAN/DMN regions during encoding/retrieval tasks. Given the predictions obtained from the HIPER/HERNET model and known anatomical pathways between these networks and the hippocampus, we tested four hypotheses as described in Introduction section. In order to do so, we investigated the connectivity between the hippocampus and layer V of the DAN during the encoding task, and between the hippocampus and layer II of the DMN during the retrieval task. The proposed first, second and fourth hypotheses were confirmed, whereas the third hypothesis was partially confirmed for DMN part (considering sign of the correlation), not for DAN part. The discussion of results is organized as follows. First, we discuss the results obtained by the validation of the anterior-posterior differentiation of the hippocampus using the FC between 1) hippocampal voxels and layer V of the DAN during the encoding task, and 2) hippocampal voxels and layer II of the DMN during the retrieval task. Second, we discuss the connectivity between the hippocampus and the DAN/DMN during encoding/retrieval tasks. Third, we discuss the layer-specific functional

microcircuits between the layer V of the DAN and the hippocampus during the encoding task, and between the layer II of the DMN and the hippocampus during the retrieval task. Forth, we discuss the importance of ultra-high field functional neuroimaging in developing accurate and robust models of functional connectivity. Finally, we discuss the functional differentiation of the hippocampus along the long-axis.

4.4.1 Anterior-Posterior Functional Differentiation of the Hippocampus

The idea that the anterior and posterior parts of the hippocampus may serve different functions emerged half a century ago [372] [373]. More recently, Robinson et al. [370] conducted meta-analyses and found an anterior-posterior long-axis segmentation on both the left and right hippocampi. Duarte et al. [396] studied functional specialization of the hippocampus using fMRI and a virtual reality 3D paradigm. They found a functional dichotomy whereby the anterior/posterior hippocampus shows antagonistic processing patterns for spatial encoding and retrieval of 3D spatial information. Prince et al. [397] also provided evidence for an anterior-posterior parcellation that corresponded to encoding/retrieval processes.

In this study, the functional specialization of the hippocampus has been investigated by using unsupervised clustering of functional connectivity between the hippocampus and layer II/V of the DAN/DMN during encoding/retrieval processes, respectively. Our results yielded a consistent anterior-posterior long-axis parcellation, which showed high cluster similarity across clustering methods based on completely different principles. This result indicates that during the encoding/retrieval processes, there is a robust anterior-posterior functional differentiation along the long-axis of the hippocampus.

4.4.2 Layer-specific Connectivity Between the DAN/DMN and Hippocampus

The anatomical pathways between the hippocampus and DAN/DMN have been studied using invasive animal models by many researchers [398] [389] [391]. During an encoding task, deeper infra-granular layers (specifically, layer V) of the DAN project to the hippocampus, whereas the hippocampal output reaches primarily more superficial supra-granular layers (specifically, layer II) of the DMN. This layer-specific organization has been studied by many anatomical studies. However, it has never been directly investigated using connectivity based non-invasive methods in humans.

This layer-specific pathway between the hippocampus and DAN/DMN is not exclusive since pathways which originate/terminate in other layers of the cortex may also contribute to the hippocampal input or output. This is not surprising given the highly complex underlying microcircuitry and given that signals between any two brain regions can relay via multiple structures including the thalamus. However, the pathways between the hippocampus and layers II and V of the DMN/DAN seem to be the dominant ones based on prior invasive animal literature [390] [391] and therefore we have chosen to test them in this study.

We investigated the connectivity between the hippocampus and the DAN/DMN using FC between layer V of the DAN/DMN and the hippocampus during the encoding task, and between layer II of the DAN/DMN and the hippocampus during the retrieval task. During the encoding task, the correlation observed between layer V of the DAN and the hippocampus was significantly larger than that between layer V of the DMN and the hippocampus. Contrarily, during the retrieval task, the correlation obtained between layer II of the DMN and the hippocampus was significantly larger than that between layer II of the DAN and the

hippocampus. Together, these findings provide evidence that the output from layer V of the DAN likely drives the hippocampus during encoding, whereas during a retrieval task, the output from the hippocampus is likely relayed as inputs to layer II of the DMN. Given that Pearson's correlation is a directionless quantity, our inference of directionality is indirect at best, based on known directionality of the underlying anatomical projections. It is also noteworthy that we have downplayed the sign of the correlation and have made inferences primarily based on its magnitude considering the fact that Hippocampus \leftrightarrow DAN/DMN microcircuits are complex and we do not have sufficient spatial resolution or sensitivity to take phase delays into account which can likely lead to strong negative correlations wherein signals are strongly related, and yet out of phase.

4.4.3 Layer-specific Functional Pathway Between the DAN/DMN and Hippocampus

Considering the directionality of signal projection during encoding/retrieval processes and the HERNET model, the information flow between different layers of the DAN/DMN and different regions of the hippocampus should follow the pattern: layer V of the DAN \rightarrow anterior hippocampus \rightarrow posterior hippocampus \rightarrow layer II of the DMN. Thus, we hypothesized that during an encoding task, layer V of the DAN must show stronger correlation with anterior hippocampal regions than with posterior hippocampal regions, whereas during retrieval task, layer II of the DMN must exhibit stronger correlation with posterior hippocampal regions than with anterior hippocampal regions. Our results provide partial support for this hypothesis, i.e., only the DMN part was confirmed during retrieval task considering the sign of the correlation (negative in anterior, positive in posterior), but not for DAN part during encoding task. However, if we only consider the magnitude of the correlation ignoring its sign, the third hypothesis was not true for both the DAN and DMN part. The linear association between signals is important

and we cannot over-emphasize the sign of the correlation between signals given that two signals may have a strong linear relationship and yet may exhibit negative correlation due to phase differences.

4.4.4 High-resolution Functional Imaging

Recent advances in ultra-high field fMRI have provided a non-invasive way of investigating cortical columns. This technique provides several advantages over conventional field strengths, e.g., improved spatial resolution, increased signal to noise ratio, etc. More importantly, this technique makes it feasible to examine layer-specific brain activation across different brain areas. Several recent studies have showed that investigating changes in fMRI activation as a function of laminar depth can lead to more precise results [342] [343]. In this study, we investigated functional differentiation of the hippocampus along the long-axis using unsupervised clustering of layer-specific functional connectivity between the hippocampus and the DAN/DMN regions. The same clustering process was also applied on the functional connectivity between the hippocampus and the DAN/DMN volume. As we expected, the layer-specific data led to more definitive results, i.e., the proposed first, second and forth hypotheses were confirmed, whereas the third hypothesis was partially confirmed for DMN part (considering sign of the correlation), not for DAN part. However, using volume-level data, only the first hypothesis was confirmed and the third hypothesis (only for DAN part with sign of the correlation being considered) was partially confirmed. Therefore, it is important to note the relevance of high-resolution functional neuroimaging as the field progresses toward developing more accurate and robust network models of brain function in general and hippocampal function in specific.

4.4.5 Long-axis Differentiation of the Hippocampus

Recent evidence has suggested that there is an anterior-posterior functional differentiation of the hippocampus along the long-axis. In Lepage's HIPER model [374] and more recently Kim's HERNET mode [378], the anterior and posterior hippocampus are posited to be more associated with encoding and retrieval process, respectively. This encoding/retrieval dichotomy faced conflicting evidence from some meta-analyses and studies. Schacter's [382] reviewed data from diverse fMRI studies and observed that both the anterior and posterior hippocampal regions were associated with the encoding activation. Kumaran [383], Peppen [384], and Zweynert [385] found that most encoding studies used novel stimuli, which were associated with the anterior hippocampus. On the other hand, Poppenk [386] [387] observed that familiar stimuli were associated with the posterior hippocampus and superior source memory. Together, these studies suggest that the encoding/retrieval differentiation cannot rely on findings that link the anterior/posterior hippocampus to novel/familiar stimuli.

Although the encoding/retrieval dichotomy is predominant in humans, many alternative specializations have also been proposed. A motivational processing model has been proposed with the anterior hippocampus is mainly engaged in "hot" processing (emotion/motivation), whereas the posterior hippocampus is mainly associated with "cold" processing (cognition) [388]. Robinson et al also found support for this hypothesis based on hippocampal parcellations obtained from meta-analyses, resting state fMRI connectivity and diffusion tensor imaging [370] [371]. However, Wolosin [399] found that the posterior hippocampus also contributed to negative emotional memory. Some other studies have proposed that the posterior hippocampus is especially important for spatial processing, whereas the anterior hippocampus may be important

for episodic memory or other functions [400] [401]. This model was undercut by evidence that the anterior hippocampus also plays a spatial role [402].

In this study, our results provided partial support for HIPER/HERNET model. The strength of evidence in favor of the HERNET model was superior with layer-specific data compared to conventional volume data. A consistent anterior to posterior long-axis segmentation was found during encoding/retrieval tasks. We also found that during an encoding task, the hippocampus was more correlated with the layer V of the DAN, whereas during a retrieval task, the hippocampus was more associated with layer II of the DMN. However, we did not find support for stronger correlation of layer V of the DAN and the anterior portions during the encoding task. Our results also did not support the prediction of stronger correlation of layer II of the DMN and the posterior hippocampal segments during the retrieval task (ignoring the sign of the correlation). Together, these results suggest that the underlying neurophysiology of the hippocampus and its interaction with the neocortex under various neurocognitive contexts is far more complex than relatively simplistic models currently available. Our study demonstrates that it will take better quality data in terms of spatial-temporal resolution in order to build better models of hippocampal function and specialization.

4.5 Limitation and Future Works

Present study has a few limitations that may be addressed in future research. First, in this study the cortical layers were delineated using equidistant model [366], i.e., the cortical layers were reconstructed at a relatively fixed distance to the cortical interface, assuming the thickness of cortical layers maintains the same between different layers. An alternative way is to solve the Laplace equation within the interfaces of the cortex and generate cortical profiles along the

gradient of the solution. This model was so called “Laplace model”, which was proposed by Jones et al. [367]. Recently, Waehnert et al. [368] proposed a novel equivolume model, which was claimed to have better performance than equidistant model and Laplace model. Future work should evaluate the performance of this equivolume model. If found to be superior, the results presented in this work need to be replicated with more accurate reconstructions of cortical layers.

Second, our results consistently showed an anterior-posterior gradient in the left hippocampus by using the functional connectivity between the hippocampus and layer II/V of DAN/DMN ROIs during encoding/retrieval processes, respectively. This pattern was consistent with encoding/retrieval dichotomy that was proposed by HIPER/HERNET model. However, this pattern faced conflicting evidence from other studies [382]. In addition, many alternative specializations have also been proposed, explaining the anterior-posterior segmentation in different perspectives, e.g., motivational processing model [388], etc. Given the complexity of underlying neurophysiology of the hippocampus and its interaction with neocortex under various neurocognitive contexts, it might be beneficial to find a way to reconcile different hippocampal models.

Third, in this study all hypotheses were validated only on the left hippocampus. Our task involved factual memories of objects, pictures of scenes, and words, which likely preferentially recruits the left hippocampus [346] [347]. However, it is necessary to replicate the same process on the right hippocampus to obtain more comprehensive and complete results. This should be done in future work.

4.6 Conclusion

As one of the most important components in the brain of different species, the hippocampus plays crucial role in episodic memory and spatial navigation. Recent neuroimaging evidences have suggested that there may be an anterior-posterior functional specialization of the hippocampus along the long-axis. Moreover, Lepage's HIPER model and Kim's HERNET model indicated that the anterior and posterior hippocampal regions were associated with memory encoding and retrieval, respectively. This model received supports from many meta-analyses studies, but it also faced conflicting evidence from the outset.

In this study, we investigated the functional differentiation of the hippocampus using their functional connectivity with layer II/V of the DAN/DMN ROIs during encoding and retrieval processes. Given HIPER/HERNET model and anatomical layer-specific connection between the hippocampus and the DAN/DMN, four hypotheses were also proposed. Our results revealed a consistent anterior-posterior pattern, which showed high cluster similarity with an anatomical anterior-posterior segmentation. Meanwhile, the results demonstrated that during an encoding task, the hippocampus was more correlated with the layer V of the DAN, whereas during a retrieval task, the hippocampus was more associated with layer II of the DMN. However, we did not find support for stronger correlation of layer V of the DAN and the anterior portions during the encoding task, neither with stronger correlation of layer II of the DMN and the posterior portions during the retrieval task (ignoring the sign of the correlation).

The same clustering process was also applied on volume-level data in comparison with layer-specific data. As we expected, the results provided by layer-specific data were more definitive,

which suggested that it is necessary to use a better quality data, in terms of spatial-temporal resolution, to build more precise and robust models of hippocampal function.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

Unsupervised clustering is one of the most popular techniques for fMRI analysis. The commonly used k-means clustering and semi-supervised clustering methods require the number of clusters to be predefined, which is difficult to be determined in the majority of real data. Thus, in this dissertation, three unsupervised clustering methods were specifically chosen, which did not require a priori specification of the number of clusters. We investigated the feasibility of these three methods in three different fMRI studies.

Study 1: many brain-based disorders are conventionally diagnosed based on clinical interviews and behavioral assessments, which are recognized largely imperfect. Therefore it is necessary to establish neuroimaging-based biomarkers to improve diagnostic precision and accuracy. Rs-fMRI have been used as a promising technique for characterization and classification of different disorders. However, most of existing methods are besieged with several methodological issues such as a priori choice of clusters needed in k-means, a stop criterion needed in hierarchical clustering, etc. To address these issues, a general pipeline was derived to identify different brain based disorders using selected unsupervised clustering methods along with two supplementary analyses, i.e., site-specific analysis and outlier subject elimination, to improve the clustering accuracy. The effectiveness of proposed pipeline was verified on five different disorders: ADHD, AD, ASD, PTSD, and PCS. For ADHD and AD, highest similarity was achieved between connectivity and phenotypic clusters, whereas for ASD and PTSD/PCS, highest

similarity was achieved between connectivity and clinical diagnostic clusters. The result suggests that neurobiological and phenotypic biomarkers could potentially be used as an aid by the clinician, in addition to currently available subjective markers, to improve diagnostic precision.

Study 2: the connection between the hippocampus and EC has been investigated by previous studies. There is anatomical and electrophysiological evidence that during an encoding task, superficial layers of EC appear to be connected to different subfields of the hippocampus through anatomically characterized circuits, which is referred to as perforant pathway. In this study, we investigated the functional perforant pathway using the connectivity between layer II of EC with different subfields of the hippocampus during an encoding task. The hippocampus was parcellated into proximal and distal regions along perforant pathway. This parcellation was based on our observation of stronger connectivity between layer II of EC with hippocampal subfields such as DG/CA4/CA3/CA2 which are proximal to the EC along the perforant pathway, compared to subfields such as CA1/Subiculum which are distal. In addition we found evidence for hemisphere specialization with the left hippocampus is associated with more verbal memory task. The result also suggests that examining functional connectivity as a function of laminar depth can lead to more precise and robust result.

Study 3: recent neuroimaging evidences have suggested that there may be an anterior-posterior functional specialization of the hippocampus along the long-axis. Moreover, Kim's HERNET model indicates the anterior hippocampus is more connected to the DAN during memory encoding, and the posterior portions are more connected to the DMN) during retrieval. In this study, the functional specialization of the hippocampus was investigated using their connectivity with layer II/V of the DAN/DMN ROIs during encoding/retrieval processes. Our results support

some predictions of the HERNET model including anterior-posterior gradient along the long-axis of the hippocampus. While preferential relationships between the entire hippocampus and DAN/DMN during encoding/retrieval, respectively, were observed as predicted, anterior-posterior specificity in these network relationships could not be confirmed. The strength and clarity of evidence for/against the HERNET model were superior with layer-specific data compared to conventional volume data.

5.2 Future Work

The works in this dissertation have a few limitations. This chapter discusses possible solutions to address them as an extension of this research in the future.

In Chapter 3 and Chapter 4, different cortical layers were delineated using equidistant model [366], i.e., the cortical layers were reconstructed at a relatively fixed distance to the cortical interface, assuming that the thickness maintains the same between different cortical layers. An alternative way is to solve Laplace equation within the interfaces of the cortex and generate cortical profiles along the gradient of the solution. This model is referred to as “Laplace model” [367]. Previous study indicates that both these two models may fail to precisely follow anatomical layers all along the cortex [368]. To solve this issue, a novel equivolume model has been proposed [368], which showed potentials of better performance than previous equidistant model and Laplace model. Future work should evaluate the accuracy of this novel model. If it is found to be superior, the results presented in Chapter 3 and Chapter 4 need to be replicated with more accurate layer reconstruction model.

In chapter 3, the major pathway, i.e. perforant pathway, between the hippocampus and EC during the encoding process was investigated. Previous studies have shown that there are some

other pathways, e.g., indirect pathway, between different layers of EC and subfields of the hippocampus during encoding and retrieval processes [330] [331]. This study demonstrates the possibility to investigate layer-specific microcircuits in both healthy and clinical human populations, using non-invasive ultra-high fields fMRI. Thus, further fine-grained investigation can be applied in the future research.

In chapter 4, we produced partial results to support HERNET model. A consistent anterior-posterior gradient has been found in the left hippocampus by applying clustering FC between hippocampus and layer II/V of DAN/DMN during encoding/retrieval processes. In addition, it has been demonstrated that the hippocampus is primarily associated with layer V of the DAN during the encoding, whereas the hippocampus is mainly related to layer II of the DMN during the retrieval. However, we did not find evidence to support the proposed third hypothesis. Together, these results suggest that the underlying neurophysiology of the hippocampus and its interaction with DAN/DMN is far more complex than existing simplistic models. Therefore, it might be beneficial to find a way to reconcile different hippocampal specializations. Moreover, all proposed hypotheses were validated only on the left hippocampus. Our encoding task involved factual memories of objects, pictures of scenes, and works, which likely preferentially recruits the left hippocampus [346] [347]. However, it is necessary to replicate the same procedure on the right hippocampus to generate more comprehensive and complete results. These can be part of the future work.

Bibliography

- [1] S.-G. Kim and K. Ugurbil, “Functional magnetic resonance imaging of the human brain,” *J. Neurosci. Methods*, vol. 74, no. 2, pp. 229–243, 1997.
- [2] Scott A. Huettel; Allen W. Song; Gregory McCarthy, A. Huettel, Scott; Song, C. F. Beckmann, S. M. Smith, S. a. Huettel, A. W. Song, and G. McCarthy, *Functional Magnetic Resonance Imaging*, vol. 23, no. 2. 2004.
- [3] M. Smits, E. Visch-Brink, C. K. Schraa-Tam, P. J. Koudstaal, and A. van der Lugt, “Functional MR imaging of language processing: an overview of easy-to-implement paradigms for patient care and clinical research,” *Radiographics*, vol. 26 Suppl 1, pp. S145-58, 2006.
- [4] R. Sitaram, A. Caria, R. Veit, T. Gaber, G. Rota, A. Kuebler, and N. Birbaumer, “fMRI brain-computer interface: A tool for neuroscientific research and treatment,” *Comput. Intell. Neurosci.*, vol. 2007, 2007.
- [5] N. Weiskopf, “Real-time fMRI and its application to neurofeedback,” *NeuroImage*, vol. 62, no. 2. pp. 682–692, 2012.
- [6] M. H. Lee, C. D. Smyser, and J. S. Shimony, “Resting-state fMRI: A review of methods and clinical applications,” *Am. J. Neuroradiol.*, vol. 34, no. 10, pp. 1866–1872, 2013.
- [7] M. D. Fox and M. D. Greicius, “Clinical applications of resting state functional connectivity,” *Front. Syst. Neurosci.*, vol. 4, p. 19, 2010.
- [8] S. Ogawa, T. M. Lee, A. R. Kay, and D. W. Tank, “Brain magnetic resonance imaging with contrast dependent on blood oxygenation.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 87,

- no. 24, pp. 9868–72, 1990.
- [9] Y. Duchin, A. Abosch, E. Yacoub, G. Sapiro, and N. Harel, “Feasibility of using ultra-high field (7 T) MRI for clinical surgical targeting,” *PLoS One*, vol. 7, no. 5, 2012.
- [10] S. Yun, S. S. Oh, Y. Han, and H. Par, “High-resolution fMRI with higher-Order generalized series imaging and parallel imaging techniques (HGS-Parallel),” *J. Magn. Reson. Imaging*, vol. 29, no. 4, pp. 924–936, 2009.
- [11] D. A. Feinberg and K. Setsompop, “Ultra-fast MRI of the human brain with simultaneous multi-slice imaging,” *J. Magn. Reson.*, vol. 229, pp. 90–100, 2013.
- [12] R. Sladky, P. Baldinger, G. S. Kranz, J. Tr??stl, A. H??flich, R. Lanzenberger, E. Moser, and C. Windischberger, “High-resolution functional MRI of the human amygdala at 7 T,” *Eur. J. Radiol.*, vol. 82, no. 5, pp. 728–733, 2013.
- [13] K. Friston and J. Ashburner, “Statistical parametric mapping,” *Funct. neuroimaging Tech. Found.*, pp. 1–74, 1994.
- [14] Fsl, “FMRIB Software Library,” *University of Oxford*, 2006. [Online]. Available: <http://www.fmrib.ox.ac.uk/fsl>.
- [15] R. W. Cox, “AFNI: software for analysis and visualization of functional magnetic resonance neuroimages,” *Comput. Biomed. Res.*, vol. 29, no. 3, pp. 162–173, 1996.
- [16] P. Jezzard and R. S. Balaban, “Correction for Geometric Distortion in Echo-Planar Images from B-0 Field Variations,” *Magn. Reson. Med.*, vol. 34, no. 1, pp. 65–73, 1995.
- [17] H. L. Jin, S. O. Dumoulin, E. U. Saritas, G. H. Glover, B. A. Wandell, D. G. Nishimura, and J. M. Pauly, “Full-brain coverage and high-resolution imaging capabilities of passband b-SSFP fMRI at 3T,” *Magn. Reson. Med.*, vol. 59, no. 5, pp. 1099–1110, 2008.
- [18] R. Sladky, K. J. Friston, J. Tr??stl, R. Cunnington, E. Moser, and C. Windischberger,

- “Slice-timing effects and their correction in functional MRI,” *Neuroimage*, vol. 58, no. 2, pp. 588–594, 2011.
- [19] R. P. Woods, S. R. Cherry, and J. C. Mazziotta, “Rapid automated algorithm for aligning and reslicing PET images,” *J. Comput. Assist. Tomogr.*, vol. 16, no. 4, pp. 620–633, 1992.
- [20] R. P. Woods, J. C. Mazziotta, and S. R. Cherry, “MRI-PET registration with automated algorithm,” *J. Comput. Assist. Tomogr.*, vol. 17, no. 4, pp. 536–546, 1993.
- [21] A. Jiang, D. N. Kennedy, J. R. Baker, R. M. Weisskoff, R. B. H. Tootell, R. P. Woods, R. Benson, K. K. Kwong, T. J. Brady, B. R. Rosen, and J. W. Belliveau, “Motion detection and correction in functional MR imaging,” *Hum. Brain Mapp.*, vol. 3, no. 3, pp. 224–235, 1995.
- [22] J. L. Ostuni, A. K. Santha, V. S. Mattay, D. R. Weinberger, R. L. Levin, and J. A. Frank, “Analysis of interpolation effects in the reslicing of functional MR images,” *J Comput Assist Tomogr*, vol. 21, no. 5, pp. 803–810, 1997.
- [23] J. L. Lancaster, D. Tordesillas-Gutiérrez, M. Martinez, F. Salinas, A. Evans, K. Zilles, J. C. Mazziotta, and P. T. Fox, “Bias between MNI and talairach coordinates analyzed using the ICBM-152 brain template,” *Hum. Brain Mapp.*, vol. 28, no. 11, pp. 1194–1205, 2007.
- [24] G. Chen, B. D. Ward, C. Xie, W. Li, G. Chen, J. S. Goveas, P. G. Antuono, and S. J. Li, “A clustering-based method to detect functional connectivity differences,” *Neuroimage*, vol. 61, no. 1, pp. 56–61, 2012.
- [25] R. Xu and D. Wunsch, “Survey of clustering algorithms,” *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 645–678, 2005.
- [26] B. S. Everitt, S. Landau, M. Leese, and D. Stahl, *Cluster Analysis*, vol. 14, 2011.
- [27] A. Baraldi and E. Alpaydin, “Constructive feedforward ART clustering networks - Part

- II,” *IEEE Trans. Neural Networks*, vol. 13, no. 3, pp. 662–677, 2002.
- [28] A. K. Jain and R. C. Dubes, “Algorithms for Clustering Data,” *Prentice Hall*, vol. 355. p. 320, 1988.
- [29] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: a review,” *ACM Comput. Surv.*, vol. 31, no. 3, pp. 264–323, 1999.
- [30] C. M. Bishop, “Neural networks for pattern recognition,” *J. Am. Stat. Assoc.*, vol. 92, p. 482, 1995.
- [31] G. Varadhan and D. Manocha, “Accurate Minkowski sum approximation of polyhedral models,” *Graph. Models*, vol. 68, no. 4, pp. 343–355, 2006.
- [32] R. J. Hathaway, J. C. Bezdek, and Y. Hu, “Generalized fuzzy c-means clustering strategies using Lp norm distances,” *IEEE Trans. Fuzzy Syst.*, vol. 8, no. 5, pp. 576–582, 2000.
- [33] J. Macqueen, “Some methods for classification and analysis of multivariate observations,” *Proc. Fifth Berkeley Symp. Math. Stat. Probab.*, vol. 1, pp. 281–297, 1967.
- [34] E. W. Forgy, “Cluster analysis of multivariate data: efficiency versus interpretability of classifications,” *Biometrics*, vol. 21, no. 3, pp. 768–769, 1965.
- [35] R. De Maesschalck, D. Jouan-Rimbaud, and D. L. Massart, “The Mahalanobis distance,” *Chemometrics and Intelligent Laboratory Systems*, vol. 50, no. 1. pp. 1–18, 2000.
- [36] G. C. Anagnostopoulos and M. Georgiopoulos, “Hypersphere ART and ARTMAP for unsupervised and supervised, incremental learning,” *Neural Networks, 2000. IJCNN 2000, Proc. IEEE-INNS-ENNS Int. Jt. Conf.*, pp. 59–64, 2000.
- [37] J. Mao and A. K. Jain, “A self-organizing network for hyperellipsoidal clustering (HEC),” *IEEE Trans. Neural Networks*, vol. 7, no. 1, pp. 16–29, 1996.

- [38] J. Ye, “Cosine similarity measures for intuitionistic fuzzy sets and their applications,” *Math. Comput. Model.*, vol. 53, no. 1–2, pp. 91–97, 2011.
- [39] M. Steinbach, G. Karypis, and V. Kumar, “A Comparison of Document Clustering Techniques,” in *KDD workshop on text mining*, 2000, vol. 400, no. X, pp. 1–2.
- [40] M. Norouzi, D. J. D. D. J. Fleet, R. Salakhutdinov, and D. M. Blei, “Hamming distance metric learning,” *Adv. Neural Inf. Process. Syst.*, pp. 1–9, 2012.
- [41] L. Hamers, Y. Hemeryck, G. Herweyers, M. Janssen, H. Keters, R. Rousseau, and A. Vanhoutte, “Similarity measures in scientometric research: The Jaccard index versus Salton’s cosine formula,” *Inf. Process. Manag.*, vol. 25, no. 3, pp. 315–318, 1989.
- [42] V. Estivill-Castro, “Why so many clustering algorithms: a position paper,” *ACM SIGKDD Explor. Newsl.*, vol. 4, no. 1, pp. 65–75, 2002.
- [43] S. C. Johnson, “Hierarchical clustering schemes,” *Psychometrika*, vol. 32, no. 3, pp. 241–254, 1967.
- [44] F. Murtagh and P. Contreras, “Methods of Hierarchical Clustering,” *Computer (Long. Beach. Calif.)*, vol. 38, no. 2, pp. 1–21, 2011.
- [45] P. Cimiano, A. Hotho, and S. Staab, “Comparing Conceptual , Divisive and Agglomerative Clustering for Learning Taxonomies from Text,” *Text*, vol. 16, pp. 435–439, 2004.
- [46] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis (Wiley Series in Probability and Statistics)*. 1990.
- [47] P. P. Rodrigues, J. Gama, and J. P. Pedroso, “Hierarchical clustering of time series data streams,” *IEEE Trans Knowl Data Engin*, vol. 5, p. XX, 2008.
- [48] J. A. Hartigan and M. A. Wong, “A K-Means Clustering Algorithm,” *Appl. Stat.*, vol. 28,

- no. 1, pp. 100–108, 1979.
- [49] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, “An efficient k-means clustering algorithm: analysis and implementation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 881–892, 2002.
- [50] S. P. Preheim, A. R. Perrott, A. M. Martin-Platero, A. Gupta, and E. J. Alm, “Distribution-based clustering: Using ecology to refine the operational taxonomic unit,” *Appl. Environ. Microbiol.*, vol. 79, no. 21, pp. 6593–6603, 2013.
- [51] X. Xu, M. Ester, H. Kriegel, and J. Sander, “A Distribution-Based Clustering Algorithm for Mining in Large Spatial Databases,” *14th Int. Conf. Data Eng. (ICDE '98)*, 1998.
- [52] J. Graca, K. Ganchev, and B. Taskar, “Expectation Maximization and Posterior Constraints,” *Adv. Neural Inf. Process. Syst.* 20, pp. 1–8, 2008.
- [53] B. Safarinejadian, M. Menhaj, and M. Karrari, “Distributed Data Clustering Using Expectation Maximization Algorithm.pdf,” *J. Appl. Sci.*, vol. 9, no. 5, pp. 854–864, 2009.
- [54] M. Daszykowski and B. Walczak, “Density-Based Clustering Methods,” in *Comprehensive Chemometrics*, vol. 2, 2010, pp. 635–654.
- [55] H. P. Kriegel, P. Kröger, J. Sander, and A. Zimek, “Density-based clustering,” *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.*, vol. 1, no. 3, pp. 231–240, 2011.
- [56] M. Ester, H. P. Kriegel, J. Sander, and X. Xu, “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise,” in *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining*, 1996, pp. 226–231.
- [57] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander, “Optics: Ordering points to identify the clustering structure,” *ACM Sigmod Rec.*, pp. 49–60, 1999.
- [58] R. Thorndike, “Who belongs in the family?,” *Psychometrika*, vol. 18, pp. 267–276, 1953.

- [59] D. J. Ketchen Jr. and C. L. Shook, "The application of cluster analysis in strategic management research: An analysis and critique," *Strateg. Manag. J.*, vol. 17, no. 6, pp. 441–458, 1996.
- [60] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, no. C, pp. 53–65, 1987.
- [61] T. Calinski and J. Harabasz, "A Dendrite Method for Cluster Analysis," *Commun. Stat. - Simul. Comput.*, vol. 3, no. 1, pp. 1–27, 1974.
- [62] J. F. Hair, W. C. Black, B. J. Babin, and R. E. Anderson, "Multivariate Data Analysis," *Vectors*. p. 816, 2010.
- [63] D. J. Ketchen, J. B. Thomas, and C. C. Snow, "Organizational configurations and performance: a comparison of theoretical approaches," *Acad. Manag. J.*, vol. 36, no. 6, pp. 1278–1313, 1993.
- [64] J. Han and M. Kamber, *Data Mining: Concepts and Techniques*, vol. 54, no. Second Edition. 2006.
- [65] R. Baumgartner, G. Scarth, C. Teichtmeister, R. Somorjai, and E. Moser, "Fuzzy clustering of gradient-echo functional MRI in the human visual cortex. Part I: reproducibility," *J. Magn. Reson. Imaging*, vol. 7, no. 6, pp. 1094–101, 1997.
- [66] C. Goutte, P. Toft, E. Rostrup, F. Å. Nielsen, and L. K. Hansen, "On Clustering fMRI Time Series," *Neuroimage*, vol. 9, no. 3, pp. 298–310, 1999.
- [67] A. Mezer, Y. Yovel, O. Pasternak, T. Gorfine, and Y. Assaf, "Cluster analysis of resting-state fMRI time series," *Neuroimage*, vol. 45, no. 4, pp. 1117–1125, 2009.
- [68] G. Flandin, F. Kherif, X. Pennec, D. Rivière, N. Ayacke, and J. B. Potine, "Parcellation of brain images with anatomical and functional constraints for fMRI data analysis," in

- Proceedings - International Symposium on Biomedical Imaging*, 2002, vol. 2002–Janua, pp. 907–910.
- [69] B. T. 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, B. Fischl, H. Liu, and R. L. Buckner, “The organization of the human cerebral cortex estimated by intrinsic functional connectivity.,” *J. Neurophysiol.*, vol. 106, no. 3, pp. 1125–65, 2011.
- [70] Y. Zhang, L. Fan, Y. Zhang, J. Wang, M. Zhu, Y. Zhang, C. Yu, and T. Jiang, “Connectivity-based parcellation of the human posteromedial cortex,” *Cereb. Cortex*, vol. 24, no. 3, pp. 719–727, 2014.
- [71] S. B. Eickhoff, D. Bzdok, A. R. Laird, C. Roski, S. Caspers, K. Zilles, and P. T. Fox, “Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation,” *Neuroimage*, vol. 57, no. 3, pp. 938–949, 2011.
- [72] V. Michel, A. Gramfort, G. Varoquaux, E. Eger, C. Keribin, and B. Thirion, “A supervised clustering approach for fMRI-based inference of brain states,” *Pattern Recognit.*, vol. 45, no. 6, pp. 2041–2049, 2012.
- [73] J. H. Ward, “Hierarchical Grouping to Oprimize an Objective Function,” *J. Am. Stat. Assoc.*, vol. 58, no. 301, pp. 236–244, 1963.
- [74] B. Thirion, G. Flandin, P. Pinel, A. Roche, P. Ciuciu, and J. B. Poline, “Dealing with the shortcomings of spatial normalization: Multi-subject parcellation of fMRI datasets,” *Hum. Brain Mapp.*, vol. 27, no. 8, pp. 678–693, 2006.
- [75] H. Chen, K. Li, D. Zhu, X. Jiang, Y. Yuan, P. Lv, T. Zhang, L. Guo, D. Shen, and T. Liu, “Inferring group-wise consistent multimodal brain networks via multi-view spectral clustering,” *IEEE Trans. Med. Imaging*, vol. 32, no. 9, pp. 1576–1586, 2013.

- [76] S. J. Hanson, R. Rebecchi, C. Hanson, and Y. O. Halchenko, “Dense mode clustering in brain maps,” *Magn. Reson. Imaging*, vol. 25, no. 9, pp. 1249–1262, 2007.
- [77] 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, 2008.
- [78] L. J. Larson-Prior, J. M. Zempel, T. S. Nolan, F. W. Prior, A. Z. Snyder, and M. E. Raichle, “Cortical network functional connectivity in the descent to sleep,” *Proc. Natl. Acad. Sci.*, vol. 106, no. 11, pp. 4489–4494, 2009.
- [79] K. J. Friston, “The disconnection hypothesis,” in *Schizophrenia Research*, 1998, vol. 30, no. 2, pp. 115–125.
- [80] C. F. Beckmann, M. DeLuca, J. T. Devlin, and S. M. Smith, “Investigations into resting-state connectivity using independent component analysis.,” *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 360, no. 1457, pp. 1001–13, 2005.
- [81] D. Cordes, V. Haughton, J. D. Carew, K. Arfanakis, and K. Maravilla, “Hierarchical clustering to measure connectivity in fMRI resting-state data,” *Magn. Reson. Imaging*, vol. 20, no. 4, pp. 305–317, 2002.
- [82] R. Salvador, J. Suckling, M. R. Coleman, J. D. Pickard, D. Menon, and E. Bullmore, “Neurophysiological architecture of functional magnetic resonance images of human brain.,” *Cereb. Cortex*, vol. 15, no. 9, pp. 1332–42, 2005.
- [83] R. Filipovych, S. M. Resnick, and C. Davatzikos, “JointMMCC: Joint maximum-margin classification and clustering of imaging data,” *IEEE Trans. Med. Imaging*, vol. 31, no. 5, pp. 1124–1140, 2012.
- [84] M. Plitt, K. A. Barnes, and A. Martin, “Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards,”

- NeuroImage Clin.*, vol. 7, pp. 359–366, 2014.
- [85] A. Khazaei, A. Ebrahimzadeh, and A. Babajani-Feremi, “Identifying patients with Alzheimer’s disease using resting-state fMRI and graph theory,” *Clin. Neurophysiol.*, vol. 126, no. 11, pp. 2132–2141, 2015.
- [86] G. Deshpande, K. Sathian, and X. Hu, “Effect of hemodynamic variability on Granger causality analysis of fMRI,” *Neuroimage*, vol. 52, no. 3, pp. 884–896, 2010.
- [87] G. Deshpande, L. E. Libero, K. R. Sreenivasan, H. D. Deshpande, and R. K. Kana, “Identification of neural connectivity signatures of autism using machine learning,” *Front Hum Neurosci*, vol. 7, no. October, p. 670, 2013.
- [88] G. Deshpande, P. Wang, D. Rangaprakash, and B. Wilamowski, “Fully connected cascade artificial neural network architecture for attention deficit hyperactivity disorder classification from functional magnetic resonance imaging data,” *IEEE Trans. Cybern.*, vol. 45, no. 12, pp. 2668–2679, 2015.
- [89] L. E. Libero, T. P. DeRamus, A. C. Lahti, G. Deshpande, and R. K. Kana, “Multimodal neuroimaging based classification of autism spectrum disorder using anatomical, neurochemical, and white matter correlates,” *Cortex*, vol. 66, pp. 46–59, 2015.
- [90] S. V. Guttula, A. Allam, and R. S. Gumpeny, “Analyzing microarray data of Alzheimer’s using cluster analysis to identify the biomarker genes,” *International Journal of Alzheimer’s Disease*. 2012.
- [91] X. J. Wang and J. H. Krystal, “Computational psychiatry,” *Neuron*, vol. 84, no. 3. pp. 638–654, 2014.
- [92] A. Abraham, M. Milham, and A. D. Martino, “Deriving robust biomarkers from multi-site resting-state data: An Autism-based example,” *bioRxiv*, 2016.

- [93] Z. Yao, “Resting-State Time-Varying Analysis Reveals Aberrant Variations of Functional Connectivity in Autism,” *Front. Hum. Neurosci.*, vol. 10, 2016.
- [94] J. Ellegood, E. Anagnostou, B. A. Babineau, J. N. Crawley, L. Lin, M. Genestine, E. DiCicco-Bloom, J. K. Y. Lai, J. A. Foster, O. Peñagarikano, D. H. Geschwind, L. K. Pacey, D. R. Hampson, C. L. Laliberté, A. A. Mills, E. Tam, L. R. Osborne, M. Kouser, F. Espinosa-Becerra, Z. Xuan, C. M. Powell, A. Raznahan, D. M. Robins, N. Nakai, J. Nakatani, T. Takumi, M. C. van Eede, T. M. Kerr, C. Muller, R. D. Blakely, J. Veenstra-VanderWeele, R. M. Henkelman, and J. P. Lerch, “Clustering autism: using neuroanatomical differences in 26 mouse models to gain insight into the heterogeneity,” *Mol. Psychiatry*, vol. 20, no. 1, pp. 118–125, 2014.
- [95] M. Hrdlicka, I. Dudova, I. Beranova, J. Lisy, T. Belsan, J. Neuwirth, V. Komarek, L. Faladova, M. Havlovicova, Z. Sedlacek, M. Blatny, and T. Urbanek, “Subtypes of autism by cluster analysis based on structural MRI data.,” *Eur. Child Adolesc. Psychiatry*, vol. 14, no. 3, pp. 138–144, 2005.
- [96] C. D. Sloan, L. Shen, J. D. West, H. A. Wishart, L. A. Flashman, L. A. Rabin, R. B. Santulli, S. J. Guerin, C. H. Rhodes, G. J. Tsongalis, T. W. McAllister, T. A. Ahles, S. L. Lee, J. H. Moore, and A. J. Saykin, “Genetic pathway-based hierarchical clustering analysis of older adults with cognitive complaints and amnesic mild cognitive impairment using clinical and neuroimaging phenotypes,” *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.*, vol. 153, no. 5, pp. 1060–1069, 2010.
- [97] G. V. Polanczyk, E. G. Willcutt, G. a. Salum, C. Kieling, and L. a. Rohde, “ADHD prevalence estimates across three decades: An updated systematic review and meta-regression analysis,” *Int. J. Epidemiol.*, vol. 43, no. 2, pp. 434–442, 2014.

- [98] G. B. LeFever, A. P. Arcona, and D. O. Antonuccio, “ADHD among American Schoolchildren: Evidence of Overdiagnosis and Overuse of Medication,” *Sci. Rev. Ment. Heal. Pract. Object. Investig. Controv. Unorthodox Claims Clin. Psychol. Psychiatry, Soc. Work*, vol. 2, no. 1, pp. 49–60, 2003.
- [99] C. Reitz, C. Brayne, and R. Mayeux, “Epidemiology of Alzheimer disease,” *Nat. Rev. Neurol.*, vol. 7, no. 3, pp. 137–52, 2011.
- [100] R. C. Petersen, G. E. Smith, S. C. Waring, R. J. Ivnik, E. G. Tangalos, and E. Kokmen, “Mild cognitive impairment: clinical characterization and outcome,” *Arch. Neurol.*, vol. 56, no. 3, pp. 303–308, 1999.
- [101] L. J. Launer, G. W. Ross, H. Petrovitch, K. Masaki, D. Foley, L. R. White, and R. J. Havlik, “Midlife blood pressure and dementia: the Honolulu–Asia aging study,” *Neurobiol. Aging*, vol. 21, no. 1, pp. 49–55, 2000.
- [102] S. D. Mayes, S. L. Calhoun, and D. L. Crites, “Does DSM-IV Asperger’s disorder exist?” *J. Abnorm. Child Psychol.*, vol. 29, no. 3, pp. 263–271, 2001.
- [103] J. N. Miller and S. Ozonoff, “The external validity of Asperger disorder: lack of evidence from the domain of neuropsychology,” *J. Abnorm. Psychol.*, vol. 109, no. 2, pp. 227–238, 2000.
- [104] J. C. McPartland, B. Reichow, and F. R. Volkmar, “Sensitivity and specificity of proposed DSM-5 diagnostic criteria for autism spectrum disorder,” *J. Am. Acad. Child Adolesc. Psychiatry*, vol. 51, no. 4, pp. 368–383, 2012.
- [105] L. Chossegros, M. Hours, P. Charnay, M. Bernard, E. Fort, D. Boisson, P.-O. Sancho, S. N. Yao, and B. Laumon, “Predictive factors of chronic post-traumatic stress disorder 6 months after a road traffic accident,” *Accid. Anal. Prev.*, vol. 43, no. 1, pp. 471–7, 2011.

- [106] American Psychiatric Association, *American Psychiatric Association, 2013. Diagnostic and statistical manual of mental disorders (5th ed.)*. 2013.
- [107] M. N. Dretsch, K. Williams, T. Emmerich, G. Crynen, G. Ait-Ghezala, H. Chaytow, V. Mathura, F. C. Crawford, and G. L. Iverson, “Brain-derived neurotropic factor polymorphisms, traumatic stress, mild traumatic brain injury, and combat exposure contribute to postdeployment traumatic stress,” *Brain Behav.*, vol. 6, no. 1, pp. 1–12, 2016.
- [108] J. A. Greco and I. Liberzon, “Neuroimaging of Fear-Associated Learning.” *Neuropsychopharmacology*, vol. 41, no. 1, pp. 320–34, 2016.
- [109] M. R. Milad, C. I. Wright, S. P. Orr, R. K. Pitman, G. J. Quirk, and S. L. Rauch, “Recall of fear extinction in humans activates the ventromedial prefrontal cortex and hippocampus in concert,” *Biol Psychiatry*, vol. 62, no. 5, pp. 446–454, 2007.
- [110] M. N. Dretsch, K. J. Thiel, J. R. Athy, C. R. Irvin, B. Sirmon-Fjordbak, and A. Salvatore, “Mood symptoms contribute to working memory decrement in active-duty soldiers being treated for posttraumatic stress disorder,” *Brain Behav.*, vol. 2, no. 4, pp. 357–364, 2012.
- [111] M. N. Dretsch, N. D. Silverberg, and G. L. Iverson, “Multiple past concussions are associated with ongoing post-concussive symptoms but not cognitive impairment in active-duty Army soldiers.” *J. Neurotrauma*, vol. 6, no. 17, pp. 1–6, 2015.
- [112] M. Dretsch, J. Blieberg, K. Amador, J. Caban, J. Kelly, G. Grammer, and T. DeGraba, “Three scoring approaches to the Neurobehavioral Symptom Inventory for measuring clinical change in service members receiving intensive treatment for combat-related mTBI,” *J. Head Trauma Rehabil.*, vol. 31, pp. 23–29, 2015.
- [113] M. M. Pape, K. Williams, P. N. Kodosky, and M. Dretsch, “The Community Balance and

- Mobility Scale,” *J. Head Trauma Rehabil.*, p. 1, 2015.
- [114] M. N. Dretsch, K. H. Wood, T. A. Daniel, J. Katz, A. Deshpande, G., Goodman, M. D. Wheelock, K. B. Wood, T. Denney, S. Traynham, and D. C. Knight, “Exploring the neurocircuitry underpinning predictability to threat in soldiers with PTSD compared to deployment exposed controls,” *Open Neuroimag. J.*, vol. 10, pp. 111–124, 2016.
- [115] G. Chen, B. D. Ward, C. Xie, W. Li, Z. Wu, J. L. Jones, M. Franczak, P. Antuono, and S.-J. Li, “Classification of Alzheimer disease, mild cognitive impairment, and normal cognitive status with large-scale network analysis based on resting-state functional MR imaging,” *Radiology*, vol. 259, no. 1, pp. 213–21, 2011.
- [116] 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,” *J. Neurosci.*, vol. 32, no. 26, pp. 8890–8899, 2012.
- [117] F. Liu, W. Guo, J.-P. Fouche, Y. Wang, W. Wang, J. Ding, L. Zeng, C. Qiu, Q. Gong, W. Zhang, and H. Chen, “Multivariate classification of social anxiety disorder using whole brain functional connectivity,” *Brain Struct. Funct.*, vol. 220, no. 1, pp. 101–115, 2015.
- [118] Y. Tang, W. Jiang, J. Liao, W. Wang, and A. Luo, “Identifying Individuals with Antisocial Personality Disorder Using Resting-State fMRI,” *PLoS One*, vol. 8, no. 4, p. e60652, 2013.
- [119] L.-L. Zeng, H. Shen, L. Liu, and D. Hu, “Unsupervised classification of major depression using functional connectivity MRI,” *Hum. Brain Mapp.*, vol. 35, no. 4, pp. 1630–41, 2014.
- [120] V. Ashikh, G. Deshpande, D. Rangaprakash, and D. N. Dutt, “Clustering of Dynamic

- Functional Connectivity Features Obtained from Functional Magnetic Resonance Imaging Data,” 2015, pp. 308–312.
- [121] B. B. Biswal, “Resting state fMRI: A personal history,” *Neuroimage*, vol. 62, no. 2, pp. 938–944, 2012.
- [122] R. L. Buckner, F. M. Krienen, and B. T. Yeo, “Opportunities and limitations of intrinsic functional connectivity MRI,” *Nat. Neurosci.*, vol. 16, no. 7, pp. 832–837, 2013.
- [123] S. Dasgupta and P. M. Long, “Performance guarantees for hierarchical clustering,” *J. Comput. Syst. Sci.*, vol. 70, pp. 555–569, 2005.
- [124] A. Rodriguez and A. Laio, “Machine learning. Clustering by fast search and find of density peaks,” *Science*, vol. 344, no. 6191, pp. 1492–6, 2014.
- [125] E. A. Allen, E. Damaraju, S. M. Plis, E. B. Erhardt, T. Eichele, and V. D. Calhoun, “Tracking whole-brain connectivity dynamics in the resting state,” *Cereb. Cortex*, vol. 24, no. 3, pp. 663–676, 2014.
- [126] A. Venkataraman, K. R. A. Van Dijk, R. L. Buckner, and P. Golland, “Exploring functional connectivity in fMRI via clustering,” in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 441–444.
- [127] J. G. Dy and C. E. Brodley, “Feature Selection for Unsupervised Learning,” *J. Mach. Learn. Res.*, vol. 5, pp. 845–889, 2004.
- [128] P. S. Bradley and O. L. Mangasarian, “Feature selection via mathematical programming,” *INFORMS J. Comput.*, vol. 10, no. 2, pp. 209–217, 1998.
- [129] C. Sharp, D. Skinner, M. Serekoane, and M. W. Ross, “A qualitative study of the cultural appropriateness of the Diagnostic Interview Schedule for Children (DISC-IV) in South Africa,” *Soc. Psychiatry Psychiatr. Epidemiol.*, vol. 46, pp. 743–751, 2011.

- [130] J. Kaufman, B. Birmaher, D. Brent, U. Rao, C. Flynn, P. Moreci, D. Williamson, and N. Ryan, "Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime Version (K-SADS-PL): initial reliability and validity data.," *J. Am. Acad. Child Adolesc. Psychiatry*, vol. 36, no. 7, pp. 980–988, 1997.
- [131] D. Pappas, "ADHD Rating Scale-IV: Checklists, Norms, and Clinical Interpretation," *J. Psychoeduc. Assess.*, vol. 24, no. 2, pp. 172–178, 2006.
- [132] K. E. McGoey, G. J. DuPaul, E. Haley, and T. L. Shelton, "Parent and teacher ratings of attention-deficit/hyperactivity disorder in preschool: The ADHD rating scale-IV preschool version," *J. Psychopathol. Behav. Assess.*, vol. 29, no. 4, pp. 269–276, 2007.
- [133] D. Wechsler, *Wechsler Intelligence Scale for Children-Revised (WISC-R)*,. 1974.
- [134] W. Reich, "Diagnostic interview for children and adolescents (DICA)," *J. Am. Acad. Child Adolesc. Psychiatry*, vol. 39, no. 1, pp. 59–66, 2000.
- [135] C. K. Conners, G. Sitarenios, J. D. A. Parker, and J. N. Epstein, "The Revised Conners' Parent Rating Scale (CPRS-R): Factor structure, reliability, and criterion validity," *J. Abnorm. Child Psychol.*, vol. 26, no. 4, pp. 257–268, 1998.
- [136] D. Wechsler, *The Wechsler intelligence scale for children—fourth edition*. 2004.
- [137] D. R. Smith, "Wechsler Individual Achievement Test," in *Handbook of psychoeducational assessment: Ability, achievement, and behavior in children. A volume in the educational psychology series*, 2001, pp. 169–193.
- [138] G. L. Canivez, T. R. Konold, J. M. Collins, and G. Wilson, "Construct validity of the Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test: Convergent and structural validity.," *Sch. Psychol. Quarterly*, vol. 24, no. 4, pp. 252–265, 2009.

- [139] G. L. Malfa, S. Lassi, M. Bertelli, S. Pallanti, and G. Albertini, "Detecting attention-deficit/hyperactivity disorder (ADHD) in adults with intellectual disability. The use of Conners' Adult ADHD Rating Scales (CAARS)," *Res. Dev. Disabil.*, vol. 29, no. 2, pp. 158–164, 2008.
- [140] D. J. Bridgett and M. E. Walker, "Intellectual functioning in adults with ADHD: a meta-analytic examination of full scale IQ differences between adults with and without ADHD.," *Psychol. Assess.*, vol. 18, no. 1, pp. 1–14, 2006.
- [141] D. Galasko, M. R. Klauber, C. R. Hofstetter, D. P. Salmon, B. Lasker, and L. J. Thal, "The Mini-Mental State Examination in the early diagnosis of Alzheimer's disease.," *Arch. Neurol.*, vol. 47, no. 1, pp. 49–52, 1990.
- [142] W. J. Burke, M. J. Houston, S. J. Boust, and W. H. Roccaforte, "Use of the Geriatric Depression Scale in dementia of the Alzheimer type," *J Am Geriatr Soc*, vol. 37, no. 9, pp. 856–860, 1989.
- [143] G. McKhann, D. Drachman, M. Folstein, R. Katzman, D. Price, and E. M. Stadlan, "Clinical diagnosis of Alzheimer's disease: report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease.," *Neurology*, vol. 34, no. 7, pp. 939–944, 1984.
- [144] J. L. Cummings, "The Neuropsychiatric Inventory: assessing psychopathology in dementia patients.," *Neurology*, vol. 48, no. 5 Suppl 6, pp. S10–S16, 1997.
- [145] M. N. Sabbagh, M. Malek-Ahmadi, R. Kataria, C. M. Belden, D. J. Connor, C. Pearson, S. Jacobson, K. Davis, R. Yaari, and U. Singh, "The Alzheimer's questionnaire: A proof of concept study for a new informant-based dementia assessment," *J. Alzheimer's Dis.*, vol. 22, no. 3, pp. 1015–1021, 2010.

- [146] J. Kim, J. M. Basak, and D. M. Holtzman, “The role of apolipoprotein E in Alzheimer’s disease.,” *Neuron*, vol. 63, no. 3, pp. 287–303, 2009.
- [147] American Psychiatric Association, *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision*, vol. Washington. 2000.
- [148] C. Lord, S. Risi, L. Lambrecht, E. H. Cook, B. L. Leventhal, P. C. DiLavore, A. Pickles, and M. Rutter, *Autism Diagnostic Observation Schedule (ADOS)*, vol. 30, no. 3. 2000.
- [149] C. Lord, M. Rutter, and A. Couteur, “Autism Diagnostic Interview-Revised;,” *Journal Autism Developmental Disorders.*, vol. 24. pp. 659–85, 1994.
- [150] B. D. Dickstein, F. W. Weathers, A. C. Angkaw, C. M. Nievergelt, K. Yurgil, W. P. Nash, D. G. Baker, and B. T. Litz, “Diagnostic Utility of the Posttraumatic Stress Disorder (PTSD) Checklist for Identifying Full and Partial PTSD in Active-Duty Military.,” *Assessment*, vol. 22, no. 3, pp. 289–97, 2015.
- [151] K. D. Cicerone and K. Kalmar, “Persistent postconcussion syndrome: The structure of subjective complaints after mild traumatic brain injury,” *The Journal of Head Trauma Rehabilitation*, vol. 10. pp. 1–17, 1995.
- [152] C. T. Gualtieri and L. G. Johnson, “Reliability and validity of a computerized neurocognitive test battery, CNS Vital Signs,” *Arch. Clin. Neuropsychol.*, vol. 21, no. 7, pp. 623–643, 2006.
- [153] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. J. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, R. K. Niazy, J. Saunders, J. Vickers, Y. Zhang, N. De Stefano, J. M. Brady, and P. M. Matthews, “Advances in functional and structural MR image analysis and implementation as FSL,” *Neuroimage*, vol. 23, no. SUPPL. 1, pp. 208–219, 2004.

- [154] Y. Chao-Gan and Z. Yu-Feng, “DPARF: A MATLAB Toolbox for ‘Pipeline’ Data Analysis of Resting-State fMRI,” *Front. Syst. Neurosci.*, vol. 4, no. May, p. 13, 2010.
- [155] R. C. Craddock, G. A. James, P. E. Holtzheimer, X. P. Hu, and H. S. Mayberg, “A whole brain fMRI atlas generated via spatially constrained spectral clustering,” *Hum. Brain Mapp.*, vol. 33, no. 8, pp. 1914–28, 2012.
- [156] G.-R. Wu, W. Liao, S. Stramaglia, J.-R. Ding, H. Chen, and D. Marinazzo, “A blind deconvolution approach to recover effective connectivity brain networks from resting state fMRI data,” *Med. Image Anal.*, vol. 17, no. 3, pp. 365–74, 2013.
- [157] G. Deshpande and X. Hu, “Investigating effective brain connectivity from fMRI data: past findings and current issues with reference to Granger causality analysis,” *Brain Connect.*, vol. 2, no. 5 PG-235-245, pp. 235–245, 2012.
- [158] G. Deshpande, K. Sathian, X. Hu, and J. a Buckhalt, “A rigorous approach for testing the constructionist hypotheses of brain function,” *Behav. Brain Sci.*, vol. 35, no. 3, pp. 148–9, 2012.
- [159] K. Sathian, G. Deshpande, and R. Stilla, “Neural Changes with Tactile Learning Reflect Decision-Level Reweighting of Perceptual Readout,” *J. Neurosci.*, vol. 33, no. 12, pp. 5387–5398, 2013.
- [160] N. L. Hutcherson, K. R. Sreenivasan, G. Deshpande, M. A. Reid, J. Hadley, D. M. White, L. Ver Hoef, and A. C. Lahti, “Effective connectivity during episodic memory retrieval in schizophrenia participants before and after antipsychotic medication,” *Hum. Brain Mapp.*, vol. 36, no. 4, pp. 1442–1457, 2015.
- [161] D. a. Handwerker, J. M. Ollinger, and M. D’Esposito, “Variation of BOLD hemodynamic responses across subjects and brain regions and their effects on statistical analyses,”

- Neuroimage*, vol. 21, no. 4, pp. 1639–1651, 2004.
- [162] G. Deshpande, D. Rangaprakash, W. Yan, J. S. Katz, J. T. S. Denney, and M. N. Dretsch, “Hemodynamic Alterations in Post-traumatic Stress Disorder and Mild Traumatic Brain Injury,” *Proc. Annu. Meet. Int. Soc. Magn. Reson. Med. Singapore*, vol. 24, p. 222, 2016.
- [163] H. Jia, X. Hu, and G. Deshpande, “Behavioral relevance of the dynamics of the functional brain connectome.,” *Brain Connect.*, vol. 4, no. 9, pp. 741–59, 2014.
- [164] G. Deshpande, P. Santhanam, and X. Hu, “Instantaneous and causal connectivity in resting state brain networks derived from functional MRI data,” *Neuroimage*, vol. 54, no. 2, pp. 1043–1052, 2011.
- [165] S. Lacey, H. Hagtvedt, V. M. Patrick, A. Anderson, R. Stilla, G. Deshpande, X. Hu, J. R. Sato, S. Reddy, and K. Sathian, “Art for reward’s sake: Visual art recruits the ventral striatum,” *Neuroimage*, vol. 55, no. 1, pp. 420–433, 2011.
- [166] F. Krueger, S. Landgraf, E. Van Der Meer, G. Deshpande, and X. Hu, “Effective connectivity of the multiplication network: A functional MRI and multivariate granger causality mapping study,” *Hum. Brain Mapp.*, vol. 32, no. 9, pp. 1419–1431, 2011.
- [167] F. Preusse, van der Meer Elke, G. Deshpande, F. Krueger, and I. Wartenburger, “Fluid intelligence allows flexible recruitment of the parieto-frontal network in analogical reasoning.,” *Front. Hum. Neurosci.*, vol. 5, no. March, p. 22, 2011.
- [168] M. M. Grant, K. Wood, K. Sreenivasan, M. Wheelock, D. White, J. Thomas, D. C. Knight, and G. Deshpande, “Influence of Early Life Stress on Intra- and Extra-Amygdaloid Causal Connectivity.,” *Neuropsychopharmacology*, vol. 40, no. 7, pp. 1–12, 2015.
- [169] B. M. Hampstead, M. Khoshnoodi, W. Yan, G. Deshpande, and K. Sathian, “Patterns of

- effective connectivity during memory encoding and retrieval differ between patients with mild cognitive impairment and healthy older adults,” *Neuroimage*, vol. 124, pp. 997–1008, 2016.
- [170] C. Feng, G. Deshpande, C. Liu, R. Gu, Y. J. Luo, and F. Krueger, “Diffusion of responsibility attenuates altruistic punishment: A functional magnetic resonance imaging effective connectivity study,” *Hum. Brain Mapp.*, vol. 37, no. 2, pp. 663–677, 2016.
- [171] Ü. Sakoğlu, G. D. Pearlson, K. A. Kiehl, Y. M. Wang, A. M. Michael, and V. D. Calhoun, “A method for evaluating dynamic functional network connectivity and task-modulation: Application to schizophrenia,” *Magn. Reson. Mater. Physics, Biol. Med.*, vol. 23, no. 5–6, pp. 351–366, 2010.
- [172] H. Jia, X. Hu, and G. Deshpande, “Behavioral relevance of the dynamics of functional brain connectome,” *Brain Connect.*, vol. 4, no. 9, pp. 741–759, 2014.
- [173] R. M. Hutchison, T. Womelsdorf, E. a. Allen, P. a. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, D. a. Handwerker, S. Keilholz, V. Kiviniemi, D. a. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang, “Dynamic functional connectivity: Promise, issues, and interpretations,” *Neuroimage*, vol. 80, pp. 360–378, 2013.
- [174] G. Deshpande, S. LaConte, G. A. James, S. Peltier, and X. Hu, “Multivariate granger causality analysis of fMRI data,” *Hum. Brain Mapp.*, vol. 30, no. 4, pp. 1361–1373, 2009.
- [175] B. M. Hampstead, A. Y. Stringer, R. F. Stilla, G. Deshpande, X. Hu, A. B. Moore, and K. Sathian, “Activation and effective connectivity changes following explicit-memory training for face-name pairs in patients with mild cognitive impairment: a pilot study,” *Neurorehabil. Neural Repair*, vol. 25, no. 3, pp. 210–22, 2011.

- [176] K. Sathian, S. Lacey, R. Stilla, G. O. Gibson, G. Deshpande, X. Hu, S. LaConte, and C. Glielmi, “Dual pathways for haptic and visual perception of spatial and texture information,” *Neuroimage*, vol. 57, no. 2, pp. 462–475, 2011.
- [177] D. Kapogiannis, G. Deshpande, F. Krueger, M. P. Thornburg, and J. H. Grafman, “Brain networks shaping religious belief,” *Brain Connect.*, vol. 4, no. 1, pp. 70–9, 2014.
- [178] K. Goodyear, R. Parasuraman, S. Chernyak, E. Visser, P. Madhavan, G. Deshpande, and F. Krueger, “An fMRI and effective connectivity study investigating miss errors during advice utilization from human and machine agents,” *Soc. Neurosci.*, 2016.
- [179] P. Liang, G. Deshpande, S. Zhao, J. Liu, X. Hu, and K. Li, “Altered directional connectivity between emotion network and motor network in Parkinson’s disease with depression,” *Medicine (Baltimore)*, 2016.
- [180] M. M. Grant, D. White, J. Hadley, N. Hutcheson, R. Shelton, K. Sreenivasan, and G. Deshpande, “Early life trauma and directional brain connectivity within major depression,” *Hum. Brain Mapp.*, vol. 35, no. 9, pp. 4815–4826, 2014.
- [181] S. Lacey, R. Stilla, K. Sreenivasan, G. Deshpande, and K. Sathian, “Spatial imagery in haptic shape perception,” *Neuropsychologia*, vol. 60, no. 1, pp. 144–158, 2014.
- [182] G. Bellucci, S. Chernyak, H. Hoffman, G. Deshpande, O. D. Monte, K. M. Knutson, J. Grafman, and F. Krueger, “Effective connectivity of brain regions underlying third party punishment: functional MRI and Granger causality evidence,” *Soc. Neurosci.*, 2016.
- [183] M. D. Wheelock, K. R. Sreenivasan, K. H. Wood, L. W. Ver Hoef, G. Deshpande, and D. C. Knight, “Threat-related learning relies on distinct dorsal prefrontal cortex network connectivity,” *Neuroimage*, vol. 102, no. P2, pp. 904–912, 2014.
- [184] W. Liao, H. Chen, Q. Yang, and X. Lei, “Analysis of fMRI data using improved self-

- organizing mapping and spatio-temporal metric hierarchical clustering,” *IEEE Trans. Med. Imaging*, vol. 27, no. October, pp. 1472–1483, 2008.
- [185] D. Cheng, R. Kannan, S. Vempala, and G. Wang, “A divide-and-merge methodology for clustering,” *ACM Trans. Database Syst.*, vol. 31, no. 4, pp. 1499–1525, 2006.
- [186] B. Andreopoulos, A. An, X. Wang, and M. Schroeder, “A roadmap of clustering algorithms: finding a match for a biomedical application,” *Brief. Bioinform.*, vol. 10, no. 3, pp. 297–314, 2008.
- [187] J. Bergstra and Y. Bengio, “Random Search for Hyper-Parameter Optimization,” *J. Mach. Learn. Res.*, vol. 13, pp. 281–305, 2012.
- [188] G. Torres, R. Basnet, and A. Sung, “A similarity measure for clustering and its applications,” *Proc. World Acad. Sci. Eng. Technol.*, vol. 31, no. 1307–6884, pp. 490–496, 2008.
- [189] J. Yang and V. Honavar, “Feature Subset Selection Using A Genetic Algorithm,” *Pattern Recognit.*, vol. 13, no. 2, p. 380, 1997.
- [190] H. Shahamat and A. A. Pouyan, “Feature selection using genetic algorithm for classification of schizophrenia using fMRI data,” *J. Artif. Intell. Data Min.*, vol. 3, no. 1, pp. 30–37, 2015.
- [191] A. L. N. Fred and A. K. Jain, “Combining multiple clusterings using evidence accumulation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 6, pp. 835–850, 2005.
- [192] A. L. N. Fred and A. K. Jain, “Data clustering using evidence accumulation,” *Object Recognit. Support. by user Interact. Serv. Robot.*, vol. 4, pp. 276–280, 2002.
- [193] R. E. Tarjan, “Efficiency of a Good But Not Linear Set Union Algorithm,” *J. ACM*, vol.

- 22, no. 2, pp. 215–225, 1975.
- [194] A. Maron-Katz, D. Amar, E. Ben Simon, T. Hendler, and R. Shamir, “RichMind: A tool for improved inference from large-scale neuroimaging results,” *PLoS One*, vol. 11, no. 7, 2016.
- [195] M. Xia, J. Wang, and Y. He, “BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics,” *PLoS One*, vol. 8, no. 7, 2013.
- [196] S. Whitfield-Gabrieli and J. M. Ford, “Default Mode Network Activity and Connectivity in Psychopathology,” *Annu. Rev. Clin. Psychol.*, vol. 8, no. 1, pp. 49–76, 2012.
- [197] V. Menon, “Large-scale brain networks and psychopathology: A unifying triple network model,” *Trends Cogn. Sci.*, vol. 15, no. 10, pp. 483–506, 2011.
- [198] R. L. Buckner, J. R. Andrews-Hanna, and D. L. Schacter, “The Brain’s Default Network,” *Ann. N. Y. Acad. Sci.*, vol. 1124, no. 1, pp. 1–38, 2008.
- [199] C. Sestieri, M. Corbetta, G. L. Romani, and G. L. Shulman, “Episodic memory retrieval, parietal cortex, and the default mode network: functional and topographic analyses.,” *J. Neurosci.*, vol. 31, no. 12, pp. 4407–20, 2011.
- [200] P. Vannini, J. O’Brien, K. O’Keefe, M. Pihlajamäki, P. Laviolette, and R. A. Sperling, “What goes down must come up: role of the posteromedial cortices in encoding and retrieval.,” *Cereb. cortex*, vol. 21, no. 1, pp. 22–34, 2011.
- [201] R. N. Spreng, R. a Mar, and A. S. N. Kim, “The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis.,” *J. Cogn. Neurosci.*, vol. 21, no. 3, pp. 489–510, 2009.
- [202] J. R. Binder, R. H. Desai, W. W. Graves, and L. L. Conant, “Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies,”

- Cereb. Cortex*, vol. 19, no. 12, pp. 2767–2796, 2009.
- [203] D. M. Amodio and C. D. Frith, “Meeting of minds: the medial frontal cortex and social cognition.,” *Nat. Rev. Neurosci.*, vol. 7, no. 4, pp. 268–277, 2006.
- [204] A. Rangel, C. Camerer, and P. R. Montague, “A framework for studying the neurobiology of value-based decision making.,” *Nat. Rev. Neurosci.*, vol. 9, no. 7, pp. 545–556, 2008.
- [205] A. Etkin, T. Egner, and R. Kalisch, “Emotional processing in anterior cingulate and medial prefrontal cortex.,” *Trends Cogn. Sci.*, vol. 15, no. 2, pp. 85–93, 2010.
- [206] A. Lagioia, D. Van De Ville, M. Debbané, F. Lazeyras, and S. Eliez, “Adolescent resting state networks and their associations with schizotypal trait expression.,” *Front. Syst. Neurosci.*, vol. 4, no. August, pp. 1–12, 2010.
- [207] D. Mantini, M. G. Perrucci, C. Del Gratta, G. L. Romani, and M. Corbetta, “Electrophysiological signatures of resting state networks in the human brain.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 32, pp. 13170–5, 2007.
- [208] G. Ganis, W. L. Thompson, and S. M. Kosslyn, “Brain areas underlying visual mental imagery and visual perception: An fMRI study,” *Cogn. Brain Res.*, vol. 20, no. 2, pp. 226–241, 2004.
- [209] C. Schiltz, B. Sorger, R. Caldara, F. Ahmed, E. Mayer, R. Goebel, and B. Rossion, “Impaired face discrimination in acquired prosopagnosia is associated with abnormal response to individual faces in the right middle fusiform gyrus,” *Cereb. Cortex*, vol. 16, no. 4, pp. 574–586, 2006.
- [210] L. A. Renier, I. Anurova, A. G. De Volder, S. Carlson, J. VanMeter, and J. P. Rauschecker, “Preserved functional specialization for spatial processing in the middle occipital gyrus of the early blind,” *Neuron*, vol. 68, no. 1, pp. 138–148, 2010.

- [211] S. Vanni, T. Tanskanen, M. Seppä, K. Uutela, and R. Hari, “Coinciding early activation of the human primary visual cortex and anteromedial cuneus.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 98, no. 5, pp. 2776–2780, 2001.
- [212] E. R. Kandel, J. H. Schwartz, and T. M. Jessell, *Principles of Neural Science*, vol. 3. 2000.
- [213] H. H. Kornhuber, “Cortex, basal ganglia and cerebellum in motor control.,” *Electroencephalogr. Clin. Neurophysiol. Suppl.*, no. 34, pp. 449–455, 1978.
- [214] O. Hikosaka, Y. Takikawa, and R. Kawagoe, “Role of the basal ganglia in the control of purposive saccadic eye movements.,” *Physiol. Rev.*, vol. 80, no. 3, pp. 953–78, 2000.
- [215] A. Stocco, C. Lebiere, and J. R. Anderson, “Conditional routing of information to the cortex: a model of the basal ganglia’s role in cognitive coordination.,” *Psychol. Rev.*, vol. 117, no. 2, pp. 541–74, 2010.
- [216] M. Bennett, D. Dennett, P. Hacker, J. Searle, J. R. Hurford, M. Bennett, D. Dennett, P. Hacker, and J. Searle, “Neuroscience and Philosophy: Brain, Mind, and Language,” *Q. Rev. Biol.*, vol. 82, no. 4, pp. 439–440, 2007.
- [217] J. M. Soares, A. Sampaio, L. M. Ferreira, N. C. Santos, P. Marques, F. Marques, J. A. Palha, J. J. Cerqueira, and N. Sousa, “Stress Impact on Resting State Brain Networks,” *PLoS One*, vol. 8, no. 6, 2013.
- [218] C. Habas, N. Kamdar, D. Nguyen, K. Prater, C. F. Beckmann, V. Menon, and M. D. Greicius, “Distinct Cerebellar Contributions to Intrinsic Connectivity Networks,” *J. Neurosci.*, vol. 29, no. 26, pp. 8586–8594, 2009.
- [219] 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.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 100, no. 1, pp. 253–8, 2003.

- [220] S. Chenji, S. Jha, D. Lee, M. Brown, P. Seres, D. Mah, and S. Kalra, “Investigating default mode and sensorimotor network connectivity in amyotrophic lateral sclerosis,” *PLoS One*, vol. 11, no. 6, 2016.
- [221] A. D. Friederici, “The brain basis of language processing: From structure to function,” *Physiol. Rev.*, vol. 91, no. 4, pp. 1357–92, 2011.
- [222] A. D. Friederici and S. M. E. Gierhan, “The language network,” *Current Opinion in Neurobiology*, vol. 23, no. 2, pp. 250–254, 2013.
- [223] K. Emmorey, “The Role of Brocas’s Area in Sign Language,” in *Broca’s Region*, no. May 2012, 2006.
- [224] R. A. Mason, C. S. Prat, and M. A. Just, “Neurocognitive brain response to transient impairment of wernicke’s area,” *Cereb. Cortex*, vol. 24, no. 6, pp. 1474–1484, 2014.
- [225] D. J. Acheson and P. Hagoort, “Stimulating the brain’s language network: syntactic ambiguity resolution after TMS to the inferior frontal gyrus and middle temporal gyrus,” *J Cogn Neurosci*, vol. 25, no. 10, pp. 1664–1677, 2013.
- [226] M. D. Fox, M. Corbetta, A. Z. Snyder, J. L. Vincent, and M. E. Raichle, “Spontaneous neuronal activity distinguishes human dorsal and ventral attention systems,” *Proc. Natl. Acad. Sci.*, vol. 103, no. 26, pp. 10046–10051, 2006.
- [227] S. Yantis, “Goal-directed and stimulus-driven determinants of attentional control,” *Control Cogn. Process. Atten. Perform. Xviii*, pp. 73–103, 2000.
- [228] X. Duan, W. Liao, D. Liang, L. Qiu, Q. Gao, C. Liu, Q. Gong, and H. Chen, “Large-scale brain networks in board game experts: Insights from a domain-related task and task-free resting state,” *PLoS One*, vol. 7, no. 3, 2012.
- [229] S. Majerus, L. Attout, A. D’Argembeau, C. Degueldre, W. Fias, P. Maquet, T. Martinez

- Perez, D. Stawarczyk, E. Salmon, M. Van Der Linden, C. Phillips, and E. Balteau, “Attention supports verbal short-term memory via competition between dorsal and ventral attention networks,” *Cereb. Cortex*, vol. 22, no. 5, pp. 1086–1097, 2012.
- [230] M. D. Fox and M. E. Raichle, “Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging,” *Nat Rev Neurosci*, vol. 8, no. 9, pp. 700–711, 2007.
- [231] P. Shaw, J. Lerch, D. Greenstein, W. Sharp, L. Clasen, A. Evans, J. Giedd, F. X. Castellanos, and J. Rapoport, “Longitudinal mapping of cortical thickness and clinical outcome in children and adolescents with attention-deficit/hyperactivity disorder.,” *Arch. Gen. Psychiatry*, vol. 63, no. 5, pp. 540–549, 2006.
- [232] A. Cubillo and K. Rubia, “Structural and functional brain imaging in adult attention-deficit/hyperactivity disorder.,” *Expert Rev. Neurother.*, vol. 10, no. 4, pp. 603–620, 2010.
- [233] K. Rubia, R. Halari, A. Cubillo, A. M. Mohammad, S. Scott, and M. Brammer, “Disorder-specific inferior prefrontal hypofunction in boys with pure attention-deficit/hyperactivity disorder compared to boys with pure conduct disorder during cognitive flexibility,” *Hum. Brain Mapp.*, vol. 31, no. 12, pp. 1823–1833, 2010.
- [234] E. B. Liddle, C. Hollis, M. J. Batty, M. J. Groom, J. J. Totman, M. Liotti, G. Scerif, and P. F. Liddle, “Task-related default mode network modulation and inhibitory control in ADHD: effects of motivation and methylphenidate.,” *J. Child Psychol. Psychiatry.*, vol. 52, no. 7, pp. 761–771, 2011.
- [235] K. Konrad and S. B. Eickhoff, “Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder,” *Hum. Brain Mapp.*, vol. 31, no. 6, pp. 904–916, 2010.

- [236] L. Sun, Q. Cao, X. Long, M. Sui, X. Cao, C. Zhu, X. Zuo, L. An, Y. Song, Y. Zang, and Y. Wang, “Abnormal functional connectivity between the anterior cingulate and the default mode network in drug-naïve boys with attention deficit hyperactivity disorder.” *Psychiatry Res.*, vol. 201, no. 2, pp. 120–7, 2012.
- [237] X. Cao, Q. Cao, X. Long, L. Sun, M. Sui, C. Zhu, X. Zuo, Y. Zang, and Y. Wang, “Abnormal resting-state functional connectivity patterns of the putamen in medication-naive children with attention deficit hyperactivity disorder,” *Brain Res*, vol. 1303, pp. 195–206, 2009.
- [238] G. Bush, “Attention-deficit/hyperactivity disorder and attention networks.” *Neuropsychopharmacology*, vol. 35, no. 1, pp. 278–300, 2010.
- [239] F. X. Castellanos, D. S. Margulies, C. Kelly, L. Q. Uddin, M. Ghaffari, A. Kirsch, D. Shaw, Z. Shehzad, A. Di Martino, B. Biswal, E. J. S. Sonuga-Barke, J. Rotrosen, L. A. Adler, and M. P. Milham, “Cingulate-Precuneus Interactions: A New Locus of Dysfunction in Adult Attention-Deficit/Hyperactivity Disorder,” *Biol. Psychiatry*, vol. 63, no. 3, pp. 332–337, 2008.
- [240] K. Rubia, “‘Cool’ inferior frontostriatal dysfunction in attention-deficit/hyperactivity disorder versus ‘hot’ ventromedial orbitofrontal-limbic dysfunction in conduct disorder: A review,” *Biological Psychiatry*, vol. 69, no. 12, 2011.
- [241] G. Bush, “Cingulate, frontal, and parietal cortical dysfunction in attention-deficit/hyperactivity disorder.” *Biol. Psychiatry*, vol. 69, no. 12, pp. 1160–7, 2011.
- [242] F. X. Castellanos and E. Proal, “Large-scale brain systems in ADHD: beyond the prefrontal-striatal model.” *Trends Cogn. Sci.*, vol. 16, no. 1, pp. 17–26, 2012.
- [243] A. Cubillo, R. Halari, C. Ecker, V. Giampietro, E. Taylor, and K. Rubia, “Reduced

- activation and inter-regional functional connectivity of fronto-striatal networks in adults with childhood Attention-Deficit Hyperactivity Disorder (ADHD) and persisting symptoms during tasks of motor inhibition and cognitive switching,” *J Psychiatr Res*, vol. 44, no. 10, pp. 629–639, 2010.
- [244] A. Köchel, F. Schöngäßner, S. Feierl-Gsodam, and A. Schienle, “Processing of affective prosody in boys suffering from attention deficit hyperactivity disorder: A near-infrared spectroscopy study,” *Soc. Neurosci.*, vol. 919, no. March, pp. 1–9, 2015.
- [245] D. A. Kessler, M. A. Angstadt, R. C. A. C. B. C. Welsh, C. Sripada, and C. . Sr.ipada, “Modality-Spanning Deficits in Attention-Deficit/Hyperactivity Disorder in Functional Networks, Gray Matter, and White Matter,” *J. Neurosci.*, vol. 34, no. 50, pp. 16555–16566, 2014.
- [246] S. G. Dickstein, K. Bannon, F. Xavier Castellanos, and M. P. Milham, “The neural correlates of attention deficit hyperactivity disorder: An ALE meta-analysis,” *J. Child Psychol. Psychiatry Allied Discip.*, vol. 47, no. 10, pp. 1051–1062, 2006.
- [247] L. Tian, T. Jiang, Y. Wang, Y. Zang, Y. He, M. Liang, M. Sui, Q. Cao, S. Hu, M. Peng, and Y. Zhuo, “Altered resting-state functional connectivity patterns of anterior cingulate cortex in adolescents with attention deficit hyperactivity disorder,” *Neurosci. Lett.*, vol. 400, no. 1–2, pp. 39–43, 2006.
- [248] L. J. Seidman, E. M. Valera, and N. Makris, “Structural brain imaging of attention-deficit/hyperactivity disorder,” *Biol. Psychiatry*, vol. 57, no. 11, pp. 1263–1272, 2005.
- [249] X. Li, A. Sroubek, M. S. Kelly, I. Lesser, E. Sussman, Y. He, C. Branch, and J. J. Foxe, “Atypical pulvinar-cortical pathways during sustained attention performance in children with attention-deficit/hyperactivity disorder.” *J. Am. Acad. Child Adolesc. Psychiatry*,

- vol. 51, no. 11, p. 1197–1207.e4, 2012.
- [250] T. S. Hale, A. M. Kane, O. Kaminsky, K. L. Tung, J. F. Wiley, J. J. McGough, S. K. Loo, and J. T. Kaplan, “Visual network asymmetry and default mode network function in ADHD: An fMRI study,” *Front. Psychiatry*, vol. 5, no. JUL, 2014.
- [251] M. Ewers, R. a. Sperling, W. E. Klunk, M. W. Weiner, and H. Hampel, “Neuroimaging markers for the prediction and early diagnosis of Alzheimer’s disease dementia,” *Trends Neurosci.*, vol. 34, no. 8, pp. 430–442, 2011.
- [252] M. D. Greicius, G. Srivastava, A. L. Reiss, and V. Menon, “Default-mode network activity distinguishes Alzheimer’s disease from healthy aging: evidence from functional MRI,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 101, no. 13, pp. 4637–42, 2004.
- [253] K. Supekar, V. Menon, D. Rubin, M. Musen, and M. D. Greicius, “Network analysis of intrinsic functional brain connectivity in Alzheimer’s disease,” *PLoS Comput. Biol.*, vol. 4, no. 6, p. e1000100, 2008.
- [254] 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.
- [255] J. Zhou, M. D. Greicius, E. D. Gennatas, M. E. Growdon, J. Y. Jang, G. D. Rabinovici, J. H. Kramer, M. Weiner, B. L. Miller, and W. W. Seeley, “Divergent network connectivity changes in behavioural variant frontotemporal dementia and Alzheimer’s disease,” *Brain*, vol. 133, no. Pt 5, pp. 1352–67, 2010.
- [256] J. Gomez-Ramirez and J. Wu, “Network-based biomarkers in Alzheimer’s disease: Review and future directions,” *Front. Aging Neurosci.*, vol. 6, no. FEB, pp. 1–9, 2014.
- [257] 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.,” *J. Neurosci.*, vol. 29, no. 6, pp. 1860–73, 2009.
- [258] R. A. Sperling, B. C. Dickerson, M. Pihlajamaki, P. Vannini, P. S. LaViolette, O. V Vitolo, T. Hedden, J. A. Becker, D. M. Rentz, D. J. Selkoe, and K. A. Johnson, “Functional alterations in memory networks in early Alzheimer’s disease.,” *Neuromolecular Med.*, vol. 12, no. 1, pp. 27–43, 2010.
- [259] M. M. Machulda, H. A. Ward, B. Borowski, J. L. Gunter, R. H. Cha, P. C. O’Brien, R. C. Petersen, B. F. Boeve, D. Knopman, D. F. Tang-Wai, R. J. Ivnik, G. E. Smith, E. G. Tangalos, and C. R. Jack, “Comparison of memory fMRI response among normal, MCI, and Alzheimer’s patients.,” *Neurology*, vol. 61, no. 4, pp. 500–6, 2003.
- [260] S. C. Johnson, T. W. Schmitz, M. a Trivedi, M. L. Ries, B. M. Torgerson, C. M. Carlsson, S. Asthana, B. P. Hermann, and M. a Sager, “The influence of Alzheimer disease family history and apolipoprotein E epsilon4 on mesial temporal lobe activation.,” *J. Neurosci.*, vol. 26, no. 22, pp. 6069–6076, 2006.
- [261] F. Agosta, M. Pievani, C. Geroldi, M. Copetti, G. B. Frisoni, and M. Filippi, “Resting state fMRI in Alzheimer’s disease: Beyond the default mode network,” *Neurobiol. Aging*, vol. 33, no. 8, pp. 1564–1578, 2012.
- [262] A. L. Foundas, C. M. Leonard, S. M. Mahoney, O. F. Agee, and K. M. Heilman, “Atrophy of the hippocampus, parietal cortex, and insula in Alzheimer’s disease: a volumetric magnetic resonance imaging study,” *Neuropsychiatry, neuropsychology, and behavioral neurology*, vol. 10, no. 2, pp. 81–89, 1997.
- [263] S. A. R. B. Rombouts, F. Barkhof, M. P. Witter, and P. Scheltens, “Unbiased whole-brain

- analysis of gray matter loss in Alzheimer's disease," *Neurosci. Lett.*, vol. 285, no. 3, pp. 231–233, 2000.
- [264] G. B. Karas, P. Scheltens, S. A. R. B. Rombouts, P. J. Visser, R. A. Van Schijndel, N. C. Fox, and F. Barkhof, "Global and local gray matter loss in mild cognitive impairment and Alzheimer's disease," *Neuroimage*, vol. 23, no. 2, pp. 708–716, 2004.
- [265] R. Li, X. Wu, K. Chen, A. S. Fleisher, E. M. Reiman, and L. Yao, "Alterations of directional connectivity among resting-state networks in Alzheimer disease," *Am. J. Neuroradiol.*, vol. 34, no. 2, pp. 340–345, 2013.
- [266] C. S. Monk, S. J. Peltier, J. L. Wiggins, S.-J. Weng, M. Carrasco, S. Risi, and C. Lord, "Abnormalities of intrinsic functional connectivity in autism spectrum disorders.," *Neuroimage*, vol. 47, no. 2, pp. 764–72, 2009.
- [267] M. Assaf, K. Jagannathan, V. D. Calhoun, L. Miller, M. C. Stevens, R. Sahl, J. G. O'Boyle, R. T. Schultz, and G. D. Pearlson, "Abnormal functional connectivity of default mode sub-networks in autism spectrum disorder patients," *Neuroimage*, vol. 53, no. 1, pp. 247–256, 2010.
- [268] J. Radua, E. Via, M. Catani, and D. Mataix-Cols, "Voxel-based meta-analysis of regional white-matter volume differences in autism spectrum disorder versus healthy controls.," *Psychol. Med.*, vol. 41, no. 7, pp. 1539–1550, 2011.
- [269] A. Martino, K. Ross, L. Q. Uddin, A. B. Sklar, F. X. Castellanos, and M. P. Milham, "Functional Brain Correlates of Social and Nonsocial Processes in Autism Spectrum Disorders: An Activation Likelihood Estimation Meta-Analysis," *Biol. Psychiatry*, vol. 65, no. 1, pp. 63–74, 2009.
- [270] S.-J. Weng, J. L. Wiggins, S. J. Peltier, M. Carrasco, S. Risi, C. Lord, and C. S. Monk,

- “Alterations of resting state functional connectivity in the default network in adolescents with autism spectrum disorders,” *Brain Res.*, vol. 1313, no. May, pp. 202–214, 2010.
- [271] M. Mody, D. S. Manoach, F. H. Guenther, T. Kenet, K. A. Bruno, C. J. McDougle, and K. A. Stigler, “Speech and language in autism spectrum disorder: a view through the lens of behavior and brain imaging,” *Neuropsychiatry (London)*, vol. 3, pp. 223–232, 2013.
- [272] M. Verly, J. Verhoeven, I. Zink, D. Mantini, R. Peeters, S. Deprez, L. Emsell, B. Boets, I. Noens, J. Steyaert, L. Lagae, P. De Cock, N. Rommel, and S. Sunaert, “Altered functional connectivity of the language network in ASD: Role of classical language areas and cerebellum,” *NeuroImage Clin.*, vol. 4, pp. 374–382, 2014.
- [273] M. G. Peeva, J. A. Tourville, Y. Agam, B. Holland, D. S. Manoach, and F. H. Guenther, “White matter impairment in the speech network of individuals with autism spectrum disorder,” *NeuroImage Clin.*, vol. 3, pp. 234–241, 2013.
- [274] C. S. Prat, A. Stocco, E. Neuhaus, and N. M. Kleinmans, “Basal Ganglia Impairments in Autism Spectrum Disorder Are Related to Abnormal Signal Gating to Prefrontal Cortex,” *Neuropsychologia*, 2016.
- [275] L. Q. Uddin and V. Menon, “The anterior insula in autism: under-connected and under-examined,” *Neurosci. Biobehav. Rev.*, vol. 33, no. 8, pp. 1198–203, 2009.
- [276] 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. 869–79, 2013.
- [277] R. M. Birn, R. Patriat, M. L. Phillips, A. Germain, and R. J. Herringa, “Childhood maltreatment and combat posttraumatic stress differentially predict fear-related fronto-subcortical connectivity,” *Depress Anxiety*, vol. 31, no. 10, pp. 880–892, 2014.

- [278] A. C. Chen and A. Etkin, “Hippocampal network connectivity and activation differentiates post-traumatic stress disorder from generalized anxiety disorder.,” *Neuropsychopharmacology*, vol. 38, no. 10, pp. 1889–98, 2013.
- [279] J. M. Cisler, J. S. Steele, J. K. Lenow, S. Smitherman, B. Everett, E. Messias, and C. D. Kilts, “Functional reorganization of neural networks during repeated exposure to the traumatic memory in posttraumatic stress disorder: An exploratory fMRI study,” *J. Psychiatr. Res.*, vol. 48, no. 1, pp. 47–55, 2014.
- [280] K. Braun, “The Prefrontal-Limbic System: Development, Neuroanatomy, Function, and Implications for Socioemotional Development,” *Clin. Perinatol.*, vol. 38, no. 4, pp. 685–702, 2011.
- [281] R. L. Bluhm, P. C. Williamson, E. a. Osuch, P. a. Frewen, T. K. Stevens, K. Boksman, R. W. J. Neufeld, J. Théberge, and R. a. Lanius, “Alterations in default network connectivity in posttraumatic stress disorder related to early-life trauma,” *J. Psychiatry Neurosci.*, vol. 34, no. 3, pp. 187–194, 2009.
- [282] Z. Long, X. Duan, B. Xie, H. Du, R. Li, Q. Xu, L. Wei, S. X. Zhang, Y. Wu, Q. Gao, and H. Chen, “Altered brain structural connectivity in post-traumatic stress disorder: A diffusion tensor imaging tractography study,” *J. Affect. Disord.*, vol. 150, no. 3, pp. 798–806, 2013.
- [283] S. F. White, M. E. Costanzo, J. R. Blair, and M. J. Roy, “PTSD symptom severity is associated with increased recruitment of top-down attentional control in a trauma-exposed sample.,” *NeuroImage. Clin.*, vol. 7, pp. 19–27, 2015.
- [284] C. A. Rabinak, M. Angstadt, R. C. Welsh, A. E. Kenndy, M. Lyubkin, B. Martis, and K. Luan Phan, “Altered amygdala resting-state functional connectivity in post-traumatic

- stress disorder,” *Front. Psychiatry*, vol. 2, no. NOV, pp. 1–8, 2011.
- [285] S. Ebdlahad, E. A. Nofzinger, J. A. James, D. J. Buisse, J. C. Price, and A. Germain, “Comparing neural correlates of REM sleep in posttraumatic stress disorder and depression: a neuroimaging study.,” *Psychiatry Res.*, vol. 214, no. 3, pp. 422–8, 2013.
- [286] A. Germain, J. James, S. Insana, R. J. Herringa, O. Mammen, J. Price, and E. Nofzinger, “A window into the invisible wound of war: Functional neuroimaging of REM sleep in returning combat veterans with PTSD,” *Psychiatry Res. - Neuroimaging*, vol. 211, no. 2, pp. 176–179, 2013.
- [287] R. A. Lanius, R. Bluhm, U. Lanius, and C. Pain, “A review of neuroimaging studies in PTSD: Heterogeneity of response to symptom provocation,” *J. Psychiatr. Res.*, vol. 40, no. 8, pp. 709–729, 2006.
- [288] E. A. Stark, C. E. Parsons, T. J. Van Hartevelt, M. Charquero-Ballester, H. McManners, A. Ehlers, A. Stein, and M. L. Kringelbach, “Post-traumatic stress influences the brain even in the absence of symptoms: A systematic, quantitative meta-analysis of neuroimaging studies,” *Neuroscience and Biobehavioral Reviews*, vol. 56. pp. 207–221, 2015.
- [289] R. Herringa, M. Phillips, J. Almeida, S. Insana, and A. Germain, “Post-traumatic stress symptoms correlate with smaller subgenual cingulate, caudate, and insula volumes in unmedicated combat veterans,” *Psychiatry Res. - Neuroimaging*, vol. 203, no. 2–3, pp. 139–145, 2012.
- [290] Y. Yin, C. Jin, L. T. Eyler, H. Jin, X. Hu, L. Duan, H. Zheng, B. Feng, X. Huang, B. Shan, Q. Gong, and L. Li, “Altered regional homogeneity in post-traumatic stress disorder: A restingstate functional magnetic resonance imaging study,” *Neurosci. Bull.*, vol. 28, no. 5,

pp. 541–549, 2012.

- [291] D. Rangaprakash, G. Deshpande, T. Daniel, A. Goodman, J. Katz, N. Salibi, T. D. Jr., and M. Dretsch, “Static and Dynamic Functional Connectivity Impairments in Concussed Soldiers with and without PTSD,” *Proc. 23rd Annu. Meet. Int. Soc. Magn. Reson. Med.*, vol. 23, p. 4402, 2015.
- [292] D. E. Faries, I. Yalcin, D. Harder, and J. H. Heiligenstein, “Validation of the ADHD Rating Scale as a clinician administered and scored instrument,” *J. Atten. Disord.*, vol. 5, no. 2, pp. 107–115, 2001.
- [293] S. Zhang, D. E. Faries, M. Vowles, and D. Michelson, “ADHD Rating Scale IV: Psychometric properties from a multinational study as a clinician-administered instrument,” *Int. J. Methods Psychiatr. Res.*, vol. 14, no. 4, pp. 186–201, 2005.
- [294] N. S. Thaler, D. T. Bello, and L. M. Etcoff, “WISC-IV profiles are associated with differences in symptomatology and outcome in children with ADHD,” *J. Atten. Disord.*, vol. 17, no. 4, pp. 291–301, 2013.
- [295] C.-C. Liu, C.-C. Liu, T. Kanekiyo, H. Xu, and G. Bu, “Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy,” *Nat. Rev. Neurol.*, vol. 9, no. 2, pp. 106–18, 2013.
- [296] J. C. Morris, “Clinical dementia rating: a reliable and valid diagnostic and staging measure for dementia of the Alzheimer type,” *Int. Psychogeriatr.*, vol. 9 Suppl 1, pp. 173-176-178, 1997.
- [297] R. Duara, D. a Loewenstein, M. T. Greig-Custo, A. Raj, W. Barker, E. Potter, E. Schofield, B. Small, J. Schinka, Y. Wu, and H. Potter, “Diagnosis and staging of mild cognitive impairment, using a modification of the clinical dementia rating scale: the

- mCDR.," *Int. J. Geriatr. Psychiatry*, vol. 25, no. 3, pp. 282–9, 2010.
- [298] M. M. Williams, M. Storandt, C. M. Roe, and J. C. Morris, "Progression of Alzheimer's disease as measured by Clinical Dementia Rating Sum of Boxes scores.," *Alzheimers. Dement.*, vol. 9, no. 1 Suppl, pp. S39-44, 2013.
- [299] E. Teng, B. W. Becker, E. Woo, D. S. Knopman, J. L. Cummings, and P. H. Lu, "Utility of the Functional Activities Questionnaire for Distinguishing Mild Cognitive Impairment From Very Mild Alzheimer Disease," *Alzheimer Dis. Assoc. Disord.*, vol. 24, no. 4, pp. 348–353, 2010.
- [300] J. a. Bastiaansen, H. Meffert, S. Hein, P. Huizinga, C. Ketelaars, M. Pijnenborg, A. Bartels, R. Minderaa, C. Keyzers, and A. De Bildt, "Diagnosing autism spectrum disorders in adults: The use of Autism Diagnostic Observation Schedule (ADOS) module 4," *J. Autism Dev. Disord.*, vol. 41, no. 9, pp. 1256–1266, 2011.
- [301] K. Gotham, S. Risi, A. Pickles, and C. Lord, "The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity," *J. Autism Dev. Disord.*, vol. 37, no. 4, pp. 613–627, 2007.
- [302] G. Goldstein, N. J. Minshew, D. N. Allen, and B. E. Seaton, "High-functioning autism and schizophrenia: a comparison of an early and late onset neurodevelopmental disorder.," *Arch. Clin. Neuropsychol.*, vol. 17, no. 5, pp. 461–75, 2002.
- [303] D. J. Siegel, N. J. Minshew, and G. Goldstein, "Wechsler IQ profiles in diagnosis of high-functioning autism," *J. Autism Dev. Disord.*, vol. 26, no. 4, pp. 389–406, 1996.
- [304] C. Lord, M. Rutter, S. Goode, J. Heemsbergen, H. Jordan, L. Mawhood, and E. Schopler, "Autism diagnostic observation schedule: A standardized observation of communicative and social behavior," *J. Autism Dev. Disord.*, vol. 19, no. 2, pp. 185–212, 1989.

- [305] R. A. Bryant, “Early predictors of posttraumatic stress disorder,” *Biol. Psychiatry*, vol. 53, no. 9, pp. 789–795, 2003.
- [306] L. Burriss, E. Ayers, J. Ginsberg, and D. a. Powell, “Learning and memory impairment in PTSD: Relationship to depression,” *Depress. Anxiety*, vol. 25, pp. 149–157, 2008.
- [307] T. M. Keane, J. A. Fairbank, J. M. Caddell, R. T. Zimering, K. L. Taylor, and C. A. Mora, “Clinical evaluation of a measure to assess combat exposure.,” *Psychol. Assess.*, vol. 1, no. 1, pp. 53–55, 1989.
- [308] K. C. Koenen, M. J. Lyons, J. Goldberg, J. Simpson, W. M. Williams, R. Toomey, S. a Eisen, W. True, and M. T. Tsuang, “Co-twin control study of relationships among combat exposure, combat-related PTSD, and other mental disorders.,” *J. Trauma. Stress*, vol. 16, no. 5, pp. 433–8, 2003.
- [309] M. J. Gray, B. T. Litz, J. L. Hsu, and T. W. Lombardo, “Psychometric properties of the life events checklist.,” *Assessment*, vol. 11, no. 4, pp. 330–341, 2004.
- [310] S. R. Das, D. Mechanic-Hamilton, J. Pluta, M. Korczykowski, J. A. Detre, and P. A. Yushkevich, “Heterogeneity of functional activation during memory encoding across hippocampal subfields in temporal lobe epilepsy,” *Neuroimage*, vol. 58, no. 4, pp. 1121–1130, 2011.
- [311] P. A. Yushkevich, H. Wang, J. Pluta, S. R. Das, C. Craige, B. B. Avants, M. W. Weiner, and S. Mueller, “Nearly automatic segmentation of hippocampal subfields in in vivo focal T2-weighted MRI,” *Neuroimage*, vol. 53, no. 4, pp. 1208–1224, 2010.
- [312] R. A. Heckemann, S. Keihaninejad, P. Aljabar, K. R. Gray, C. Nielsen, D. Rueckert, J. V. Hajnal, and A. Hammers, “Automatic morphometry in Alzheimer’s disease and mild cognitive impairment,” *Neuroimage*, vol. 56, no. 4, pp. 2024–2037, 2011.

- [313] W. J. Lukiw, "Micro-RNA speciation in fetal, adult and Alzheimer's disease hippocampus.," *Neuroreport*, vol. 18, no. 3, pp. 297–300, 2007.
- [314] Y. Zhou, J. H. Dougherty, K. F. Hubner, B. Bai, R. L. Cannon, and R. K. Hutson, "Abnormal connectivity in the posterior cingulate and hippocampus in early Alzheimer's disease and mild cognitive impairment," *Alzheimers. Dement.*, vol. 4, no. 4, pp. 265–270, 2008.
- [315] S. W. Scheff and D. a Price, "Alzheimer's disease-related alterations in synaptic density: neocortex and hippocampus.," *J. Alzheimers. Dis.*, vol. 9, pp. 101–115, 2006.
- [316] S. Campbell and G. MacQueen, "The role of the hippocampus in the pathophysiology of major depression," *Journal of Psychiatry and Neuroscience*, vol. 29, no. 6. pp. 417–426, 2004.
- [317] C. A. Stockmeier, G. J. Mahajan, L. C. Konick, J. C. Overholser, G. J. Jurjus, H. Y. Meltzer, H. B. M. Uylings, L. Friedman, and G. Rajkowska, "Cellular changes in the postmortem hippocampus in major depression," *Biol. Psychiatry*, vol. 56, no. 9, pp. 640–650, 2004.
- [318] I. M. Rosso, C. M. Cintron, R. J. Steingard, P. F. Renshaw, A. D. Young, and D. A. Yurgelun-Todd, "Amygdala and hippocampus volumes in pediatric major depression," *Biol. Psychiatry*, vol. 57, no. 1, pp. 21–26, 2005.
- [319] G. MacQueen and T. Frodl, "The hippocampus in major depression: evidence for the convergence of the bench and bedside in psychiatric research?," *Mol. Psychiatry*, vol. 16, no. 3, pp. 252–264, 2011.
- [320] R. S. Astur, S. a St Germain, D. Tolin, J. Ford, D. Russell, and M. Stevens, "Hippocampus function predicts severity of post-traumatic stress disorder.," *Cyberpsychol. Behav.*, vol. 9,

- no. 2, pp. 234–240, 2006.
- [321] H. Javidi and M. Yadollahie, “Post-traumatic Stress Disorder,” *Int. J. Occup. Environ. Med.*, vol. 3, no. 1, pp. 2–9, 2012.
- [322] P. J. Harrison, “The hippocampus in schizophrenia: a review of the neuropathological evidence and its pathophysiological implications,” *Psychopharmacology (Berl.)*, vol. 174, no. 1, pp. 151–62, 2004.
- [323] S. Heckers, “Neuroimaging studies of the hippocampus in schizophrenia,” *Hippocampus*, vol. 11, no. 5, pp. 520–528, 2001.
- [324] A. A. Grace, “Dopamine system dysregulation by the hippocampus: Implications for the pathophysiology and treatment of schizophrenia,” *Neuropharmacology*, vol. 62, no. 3, pp. 1342–1348, 2012.
- [325] H. M. Duvernoy, F. Cattin, P. Y. Risold, J. L. Vannson, and M. Gaudron, *The human hippocampus: Functional anatomy, vascularization and serial sections with MRI, fourth edition*. 2013.
- [326] G. M. Fatterpekar, T. P. Naidich, B. N. Delman, J. G. Aguinaldo, S. H. Gultekin, C. C. Sherwood, P. R. Hof, B. P. Drayer, and Z. A. Fayad, “Cytoarchitecture of the human cerebral cortex: MR microscopy of excised specimens at 9.4 Tesla,” *Am. J. Neuroradiol.*, vol. 23, no. 8, pp. 1313–1321, 2002.
- [327] S. Card, *The Hippocampus book*. 2007.
- [328] C. B. Canto, F. G. Wouterlood, and M. P. Witter, “What does the anatomical organization of the entorhinal cortex tell us?,” *Neural Plasticity*, vol. 2008. 2008.
- [329] P. S. F. Bellgowan, E. a Buffalo, J. Bodurka, and A. Martin, “Lateralized spatial and object memory encoding in entorhinal and perirhinal cortices.,” *Learn. Mem.*, vol. 16, no.

- 7, pp. 433–438, 2009.
- [330] N. M. V. Strien, N. L. M. Cappaert, and M. P. Witter, “The anatomy of memory: an interactive overview of the parahippocampal-hippocampal network,” *Nat. Rev. Neurosci.*, vol. 10, no. 4, pp. 272–82, 2009.
- [331] A. Maass, H. Schütze, O. Speck, A. Yonelinas, C. Tempelmann, H.-J. Heinze, D. Berron, A. Cardenas-Blanco, K. H. Brodersen, K. Enno Stephan, and E. Düzel, “Laminar activity in the hippocampus and entorhinal cortex related to novelty and episodic encoding,” *Nat. Commun.*, vol. 5, no. March 2016, p. 5547, 2014.
- [332] K. M. Igarashi, H. T. Ito, E. I. Moser, and M. B. Moser, “Functional diversity along the transverse axis of hippocampal area CA1,” *FEBS Letters*, vol. 588, no. 15, pp. 2470–2476, 2014.
- [333] J. Suh, A. J. Rivest, T. Nakashiba, T. Tominaga, and S. Tonegawa, “Entorhinal cortex layer III input to the hippocampus is crucial for temporal association memory,” *Science (80-.)*, vol. 334, pp. 1415–1420, 2011.
- [334] M. P. Witter, P. A. Naber, T. Van Haeften, W. C. M. Machielsen, S. A. R. B. Rombouts, F. Barkhof, P. Scheltens, and F. H. Lopes Da Silva, “Cortico-hippocampal communication by way of parallel parahippocampal-subicular pathways,” *Hippocampus*, vol. 10, no. 4, pp. 398–410, 2000.
- [335] J. T. Vaughan, M. Garwood, C. M. Collins, W. Liu, L. Delabarre, G. Adriany, P. Andersen, H. Merkle, R. Goebel, M. B. Smith, and K. Ugurbil, “7T vs. 4T: RF power, homogeneity, and signal-to-noise comparison in head images,” *Magn. Reson. Med.*, vol. 46, no. 1, pp. 24–30, 2001.
- [336] X. Zhang, K. Ugurbil, and W. Chen, “Microstrip RF surface coil design for extremely

- high-field MRI and spectroscopy,” *Magn. Reson. Med.*, vol. 46, no. 3, pp. 443–450, 2001.
- [337] E. Yacoub, A. Shmuel, J. Pfeuffer, P. F. Van De Moortele, G. Adriany, P. Andersen, J. T. Vaughan, H. Merkle, K. Ugurbil, and X. Hu, “Imaging brain function in humans at 7 Tesla,” *Magn. Reson. Med.*, vol. 45, no. 4, pp. 588–94, 2001.
- [338] K. Ugurbil, G. Adriany, P. Andersen, W. Chen, M. Garwood, R. Gruetter, P. G. Henry, S. G. Kim, H. Lieu, I. Tkac, T. Vaughan, P. F. Van De Moortele, E. Yacoub, and X. H. Zhu, “Ultrahigh field magnetic resonance imaging and spectroscopy,” *Magnetic Resonance Imaging*, vol. 21, no. 10, pp. 1263–1281, 2003.
- [339] C.-H. Moon, M. Fukuda, S.-H. Park, and S.-G. Kim, “Neural Interpretation of Blood Oxygenation Level-Dependent fMRI Maps at Submillimeter Columnar Resolution,” *J. Neurosci.*, vol. 27, no. 26, pp. 6892–6902, 2007.
- [340] S. M. Smirnakis, M. C. Schmid, B. Weber, A. S. Tolias, M. Augath, and N. K. Logothetis, “Spatial specificity of BOLD versus cerebral blood volume fMRI for mapping cortical organization,” *J. Cereb. Blood Flow Metab.*, vol. 27, no. 6, pp. 1248–1261, 2007.
- [341] T. Jin and S.-G. Kim, “Cortical layer-dependent dynamic blood oxygenation, cerebral blood flow and cerebral blood volume responses during visual stimulation,” *Neuroimage*, vol. 43, no. 1, pp. 1–9, 2008.
- [342] C. A. Olman, N. Harel, D. A. Feinberg, S. He, P. Zhang, K. Ugurbil, and E. Yacoub, “Layer-specific fmri reflects different neuronal computations at different depths in human V1,” *PLoS One*, vol. 7, no. 3, 2012.
- [343] P. Kok, L. J. Bains, T. Van Mourik, D. G. Norris, and F. P. De Lange, “Selective activation of the deep layers of the human primary visual cortex by top-down feedback,” *Curr. Biol.*, vol. 26, no. 3, pp. 371–376, 2016.

- [344] P. J. Koopmans, S. Orzada, M. Barth, and D. G. Norris, “Distinguishing pial and laminar gradient-echo BOLD signals at 7 Tesla,” *Proc. Int. Soc. Magn. Reson. Med.*, vol. 2009, p. 4064, 2009.
- [345] E. Yacoub, K. Ugurbil, and C. Olman, “Feasibility of detecting differential layer specific activations in humans using SE BOLD fMRI at 7 T,” *Proc. 17th Sci. Meet. Int. Soc. Magn. Reson. Med.*, vol. Honolulu, p. 1607, 2009.
- [346] N. Burgess, E. A. Maguire, and J. O’Keefe, “The human hippocampus and spatial and episodic memory,” *Neuron*, vol. 35, no. 4, pp. 625–641, 2002.
- [347] G. Glosser, A. J. Saykin, G. K. Deutsch, and M. J. O’Connor, “Neural organization of material-specific memory functions in temporal lobe epilepsy patients as assessed by the intracarotid amobarbital test.,” *Neuropsychology*, vol. 9, no. 4, pp. 449–456, 1995.
- [348] D. M. Branco, R. O. Suarez, S. Whalen, J. P. O’Shea, A. P. Nelson, J. C. da Costa, and A. J. Golby, “Functional MRI of memory in the hippocampus: Laterality indices may be more meaningful if calculated from whole voxel distributions,” *Neuroimage*, vol. 32, no. 2, pp. 592–602, 2006.
- [349] J. Ashburner, “SPM: A history,” *NeuroImage*, vol. 62, no. 2, pp. 791–800, 2012.
- [350] B. Fischl, “FreeSurfer,” *NeuroImage*, vol. 62, no. 2, pp. 774–781, 2012.
- [351] A. M. Dale, B. Fischl, and M. I. Sereno, “Cortical Surface-Based Analysis: I. Segmentation and Surface Reconstruction,” *Neuroimage*, vol. 9, no. 2, pp. 179–194, 1999.
- [352] B. Fischl, M. I. Sereno, and A. M. Dale, “Cortical Surface-Based Analysis: II: Inflation, Flattening, and a Surface-Based Coordinate System,” *Neuroimage*, vol. 9, no. 2, pp. 195–207, 1999.
- [353] B. Fischl and A. M. Dale, “Measuring the thickness of the human cerebral cortex from

- magnetic resonance images.,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 97, no. 20, pp. 11050–5, 2000.
- [354] X. Han, J. Jovicich, D. Salat, A. van der Kouwe, B. Quinn, S. Czanner, E. Busa, J. Pacheco, M. Albert, R. Killiany, P. Maguire, D. Rosas, N. Makris, A. Dale, B. Dickerson, and B. Fischl, “Reliability of MRI-derived measurements of human cerebral cortical thickness: the effects of field strength, scanner upgrade and manufacturer,” *Neuroimage*, vol. 32, no. 1, pp. 180–194, 2006.
- [355] D. N. Greve and B. Fischl, “Accurate and robust brain image alignment using boundary-based registration,” *Neuroimage*, vol. 48, no. 1, pp. 63–72, 2009.
- [356] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, M. S. Albert, and R. J. Killiany, “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest,” *Neuroimage*, vol. 31, no. 3, pp. 968–980, 2006.
- [357] R. La Joie, M. Fouquet, F. Mézenge, B. Landeau, N. Villain, K. Mevel, A. Pélerin, F. Eustache, B. Desgranges, and G. Chételat, “Differential effect of age on hippocampal subfields assessed using a new high-resolution 3T MR sequence,” *Neuroimage*, vol. 53, no. 2, pp. 506–514, 2010.
- [358] P. A. Yushkevich, B. B. Avants, J. Pluta, S. Das, D. Minkoff, D. Mechanic-Hamilton, S. Glynn, S. Pickup, W. Liu, J. C. Gee, M. Grossman, and J. A. Detre, “A high-resolution computational atlas of the human hippocampus from postmortem magnetic resonance imaging at 9.4T,” *Neuroimage*, vol. 44, no. 2, pp. 385–398, 2009.
- [359] J. L. Winterburn, J. C. Pruessner, S. Chavez, M. Schira, N. J. Lobaugh, A. N. Voineskos, and M. M. Chakravarty, “A novel in vivo atlas of human hippocampal subfields using

- high-resolution 3T magnetic resonance imaging,” *Neuroimage*, vol. 74, pp. 254–265, 2013.
- [360] L. W. Swanson and W. M. Cowan, “An autoradiographic study of the organization of the efferent connections of the hippocampal formation in the rat,” *J. Comp. Neurol.*, vol. 172, no. 1, pp. 49–84, 1977.
- [361] C. Köhler, “Intrinsic connections of the retrohippocampal region in the rat brain. II. The medial entorhinal area,” *J. Comp. Neurol.*, vol. 246, no. 2, pp. 149–169, 1986.
- [362] S. C. Furtak, S. M. Wei, K. L. Agster, and R. D. Burwell, “Functional neuroanatomy of the parahippocampal region in the rat: The perirhinal and postrhinal cortices,” *Hippocampus*, vol. 17, no. 9, pp. 709–722, 2007.
- [363] M. R. Hunsaker, J. S. Rosenberg, and R. P. Kesner, “The role of the dentate gyrus, CA3a,b, and CA3c for detecting spatial and environmental novelty,” *Hippocampus*, vol. 18, no. 10, pp. 1064–1073, 2008.
- [364] O. A. Shipton, M. El-Gaby, J. Apergis-Schoute, K. Deisseroth, D. M. Bannerman, O. Paulsen, and M. M. Kohl, “Left-right dissociation of hippocampal memory processes in mice,” *Proc. Natl. Acad. Sci.*, vol. 111, no. 42, pp. 15238–43, 2014.
- [365] A. C. Papanicolaou, P. G. Simos, E. M. Castillo, J. I. Breier, J. S. Katz, and A. a. Wright, “The Hippocampus and Memory of Verbal and Pictorial Material,” *Learn. Mem.*, vol. 9, pp. 99–104, 2002.
- [366] R. Khan, Q. Zhang, S. Darayan, S. Dhandapani, S. Katyal, C. Greene, C. Bajaj, and D. Ress, “Surface-based analysis methods for high-resolution functional magnetic resonance imaging,” in *Graphical Models*, 2011, vol. 73, no. 6, pp. 313–322.
- [367] S. E. Jones, B. R. Buchbinder, and I. Aharon, “Three-dimensional mapping of cortical

- thickness using Laplace's equation," *Hum. Brain Mapp.*, vol. 11, no. 1, pp. 12–32, 2000.
- [368] M. D. Waehnert, J. Dinse, M. Weiss, M. N. Streicher, P. Waehnert, S. Geyer, R. Turner, and P. L. Bazin, "Anatomically motivated modeling of cortical laminae," *Neuroimage*, vol. 93, pp. 210–220, 2014.
- [369] J. Poppenk, H. R. Evensmoen, M. Moscovitch, and L. Nadel, "Long-axis specialization of the human hippocampus," *Trends in Cognitive Sciences*, vol. 17, no. 5, pp. 230–240, 2013.
- [370] J. L. Robinson, D. S. Barron, L. A. J. Kirby, K. L. Bottenhorn, A. C. Hill, J. E. Murphy, J. S. Katz, N. Salibi, S. B. Eickhoff, and P. T. Fox, "Neurofunctional topography of the human hippocampus," *Hum. Brain Mapp.*, vol. 36, no. 12, pp. 5018–5037, 2015.
- [371] J. L. Robinson, N. Salibi, and G. Deshpande, "Functional connectivity of the left and right hippocampi: Evidence for functional lateralization along the long-axis using meta-analytic approaches and ultra-high field functional neuroimaging," *Neuroimage*, vol. 135, pp. 64–78, 2016.
- [372] W. B. Scoville and B. Milner, "Loss of recent memory after bilateral hippocampal lesions.," *J. Neuropsychiatry Clin. Neurosci.*, vol. 20, no. 11, pp. 11–21, 1957.
- [373] L. Nadel, "Dorsal and Ventral Hippocampal Lesions and Behavior'," *Physiol. Behav.*, vol. 3, pp. 891–900, 1968.
- [374] M. Lepage, R. Habib, and E. Tulving, "Hippocampal PET activations of memory encoding and retrieval: The HIPER model," *Hippocampus*, vol. 8, no. 4, pp. 313–322, 1998.
- [375] J. Spaniol, P. S. R. Davidson, A. S. N. Kim, H. Han, M. Moscovitch, and C. L. Grady, "Event-related fMRI studies of episodic encoding and retrieval: Meta-analyses using

- activation likelihood estimation,” *Neuropsychologia*, vol. 47, no. 8–9, pp. 1765–1779, 2009.
- [376] L. Nadel, S. Hoscheidt, and L. R. Ryan, “Spatial cognition and the hippocampus: the anterior-posterior axis,” *J. Cogn. Neurosci.*, vol. 25, no. 1, pp. 22–28, 2013.
- [377] O. Baumann and J. B. Mattingley, “Dissociable representations of environmental size and complexity in the human hippocampus,” *J. Neurosci.*, vol. 33, no. 25, pp. 10526–10533, 2013.
- [378] H. Kim, “Encoding and retrieval along the long axis of the hippocampus and their relationships with dorsal attention and default mode networks: The HERNET model,” *Hippocampus*, vol. 25, no. 4, pp. 500–510, 2015.
- [379] H. Kim, “Dissociating the roles of the default-mode, dorsal, and ventral networks in episodic memory retrieval,” *Neuroimage*, vol. 50, no. 4, pp. 1648–1657, 2010.
- [380] M. M. Chun, J. D. Golomb, and N. B. Turk-Browne, “A taxonomy of external and internal attention,” *Annu. Rev. Psychol.*, vol. 62, pp. 73–101, 2011.
- [381] M. Corbetta and G. L. Shulman, “Control of goal-directed and stimulus-driven attention in the brain,” *Nat. Rev. Neurosci.*, vol. 3, no. 3, pp. 201–15, 2002.
- [382] D. L. Schacter and A. D. Wagner, “Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval,” *Hippocampus*, vol. 9, no. 1, pp. 7–24, 1999.
- [383] D. Kumaran and E. A. Maguire, “An unexpected sequence of events: Mismatch detection in the human hippocampus,” *PLoS Biol.*, vol. 4, no. 12, pp. 2372–2382, 2006.
- [384] J. Poppenk, G. Walia, A. R. McIntosh, M. F. Joannisse, D. Klein, and S. Köhler, “Why is the meaning of a sentence better remembered than its form? An fMRI study on the role of novelty-encoding processes,” *Hippocampus*, vol. 18, no. 9, pp. 909–918, 2008.

- [385] S. Zweynert, J. P. Pade, T. Wüstenberg, P. Sterzer, H. Walter, C. I. Seidenbecher, A. Richardson-Klavehn, E. Düzel, and B. H. Schott, “Motivational salience modulates hippocampal repetition suppression and functional connectivity in humans.,” *Front. Hum. Neurosci.*, vol. 5, no. November, p. 144, 2011.
- [386] J. Poppenk, S. Köhler, and M. Moscovitch, “Revisiting the novelty effect: when familiarity, not novelty, enhances memory.,” *J. Exp. Psychol. Mem. Cogn.*, vol. 36, no. 5, pp. 1321–1330, 2010.
- [387] J. Poppenk, A. R. McIntosh, F. I. M. Craik, and M. Moscovitch, “Past Experience Modulates the Neural Mechanisms of Episodic Memory Formation,” *J. Neurosci.*, vol. 30, no. 13, pp. 4707–4716, 2010.
- [388] V. P. Murty, M. Ritchey, R. A. Adcock, and K. S. LaBar, “Reprint of: fMRI studies of successful emotional memory encoding: A quantitative meta-analysis,” *Neuropsychologia*, vol. 49, no. 4, pp. 695–705, 2011.
- [389] A. M. Thomson and a P. Bannister, “Interlaminar connections in the neocortex.,” *Cereb. Cortex*, vol. 13, no. 1, pp. 5–14, 2003.
- [390] G. M. Shepard and S. Grillner, “Handbook of Brain Microcircuits,” *Handbook of Brain Microcircuits*. p. 46, 2010.
- [391] A. M. Thomson and C. Lamy, “Functional maps of neocortical local circuitry.,” *Front. Neurosci.*, vol. 1, no. 1, pp. 19–42, 2007.
- [392] Y. Behzadi, K. Restom, J. Liau, and T. T. Liu, “A component based noise correction method (CompCor) for BOLD and perfusion based fMRI,” *Neuroimage*, vol. 37, no. 1, pp. 90–101, 2007.
- [393] V. H. Hackert, T. den Heijer, M. Oudkerk, P. J. Koudstaal, A. Hofman, and M. M.

- Breteler, “Hippocampal head size associated with verbal memory performance in nondemented elderly,” *Neuroimage*, vol. 17, no. 3, pp. 1365–1372, 2002.
- [394] M. D. Greicius, B. Krasnow, J. M. Boyett-Anderson, S. Eliez, A. F. Schatzberg, A. L. Reiss, and V. Menon, “Regional analysis of hippocampal activation during memory encoding and retrieval: fMRI study,” *Hippocampus*, vol. 13, no. 1, pp. 164–174, 2003.
- [395] C. Destrieux, D. Bourry, and S. Velut, “Surgical anatomy of the hippocampus,” *Neurochirurgie*, vol. 59, no. 4–5, pp. 149–158, 2013.
- [396] I. C. Duarte, C. Ferreira, J. Marques, and M. Castelo-Branco, “Anterior/posterior competitive deactivation/activation dichotomy in the human hippocampus as revealed by a 3D navigation task,” *PLoS One*, vol. 9, no. 1, 2014.
- [397] S. E. Prince, S. M. Daselaar, and R. Cabeza, “Neural correlates of relational memory: successful encoding and retrieval of semantic and perceptual associations,” *J Neurosci*, vol. 25, no. 5, pp. 1203–1210, 2005.
- [398] J. Sugar, M. P. Witter, N. M. van Strien, and N. L. M. Cappaert, “The retrosplenial cortex: intrinsic connectivity and connections with the (para)hippocampal region in the rat. An interactive connectome,” *Front. Neuroinform.*, vol. 5, no. July, p. 7, 2011.
- [399] S. M. Wolosin, D. Zeithamova, and A. R. Preston, “Reward Modulation of Hippocampal Subfield Activation during Successful Associative Encoding and Retrieval,” *J. Cogn. Neurosci.*, vol. 24, no. 7, pp. 1532–1547, 2012.
- [400] L. Ryan, C.-Y. Lin, K. Ketcham, and L. Nadel, “The role of medial temporal lobe in retrieving spatial and nonspatial relations from episodic and semantic memory,” *Hippocampus*, vol. 20, no. 1, pp. 11–18, 2010.
- [401] M. Hirshhorn, C. Grady, R. S. Rosenbaum, G. Winocur, and M. Moscovitch, “The

hippocampus is involved in mental navigation for a recently learned, but not a highly familiar environment: A longitudinal fMRI study,” *Hippocampus*, vol. 22, no. 4, pp. 842–852, 2012.

[402] K. Woollett and E. A. Maguire, “Exploring anterograde associative memory in London taxi drivers,” *Neuroreport*, vol. 23, no. 15, pp. 885–888, 2012.

Appendix A

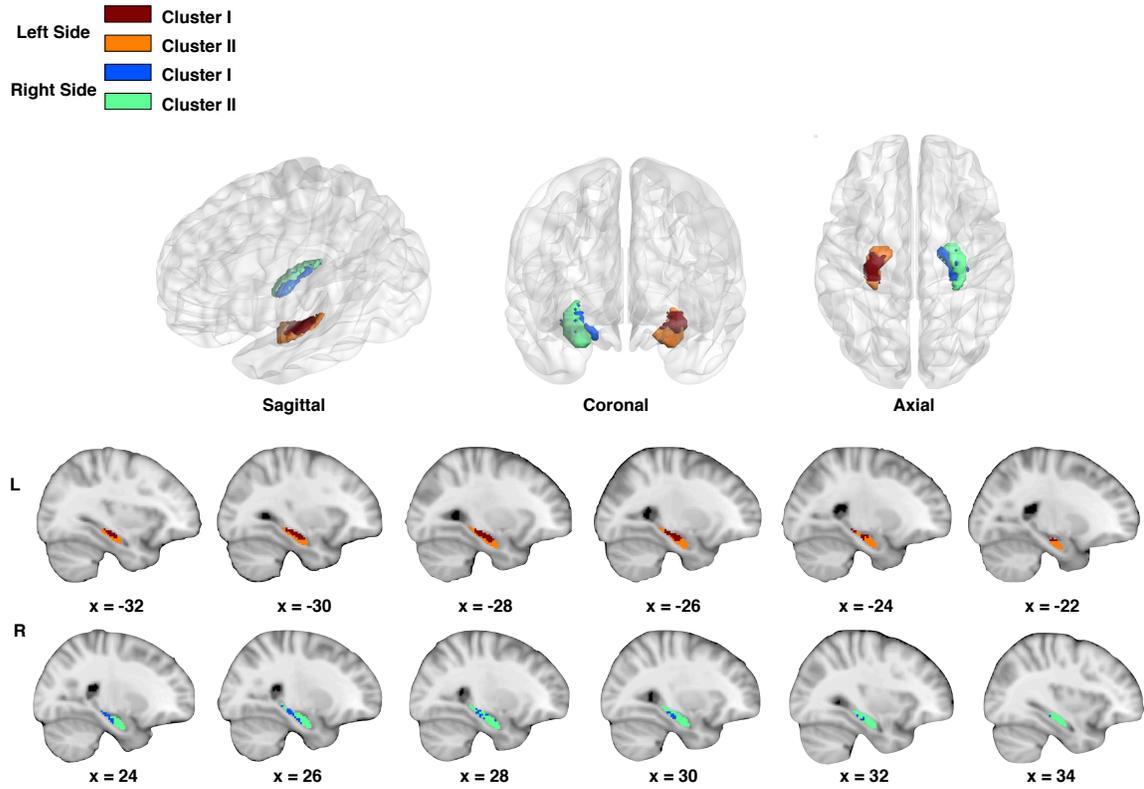


Figure A.1 Clusters of hippocampal voxels determined (using the hierarchical clustering method) based on their functional connectivity with layer II of the entorhinal cortex during the encoding task. (Coordinates are in MNI space)

Left Side Cluster I (dark red)
 Cluster II (orange)
 Right Side Cluster I (blue)
 Cluster II (green)

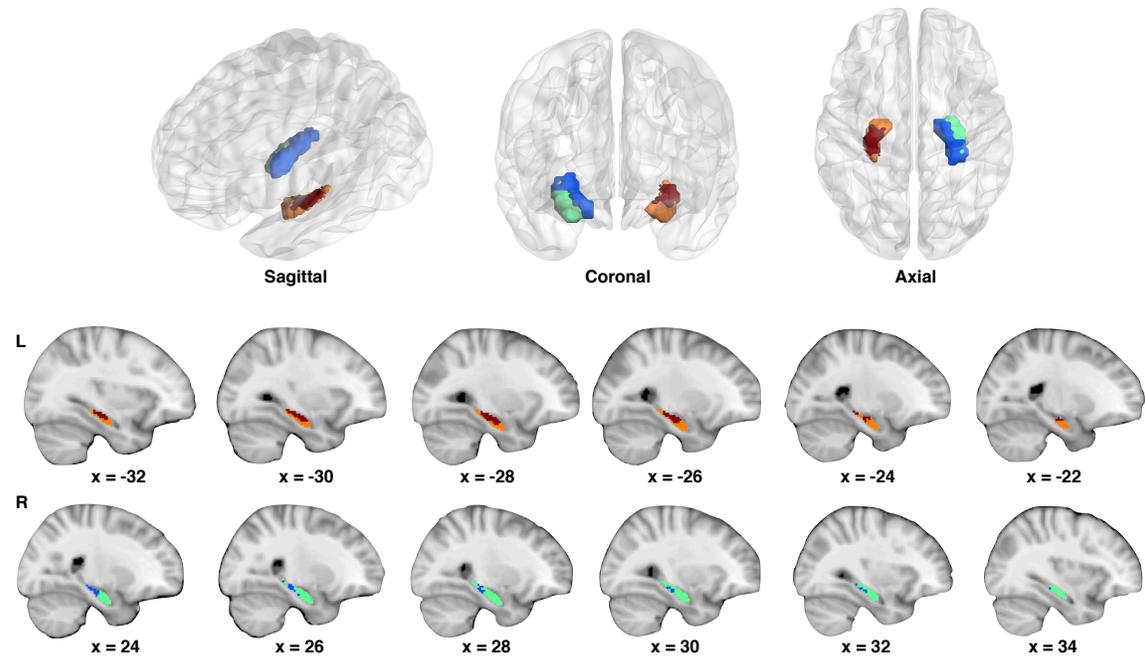


Figure A.2 Clusters of hippocampal voxels determined (using the OPTICS method) based on their functional connectivity with layer II of the entorhinal cortex during the encoding task. (Coordinates are in MNI space)

Left Side ■ Cluster I
 ■ Cluster II
 Right Side ■ Cluster I
 ■ Cluster II

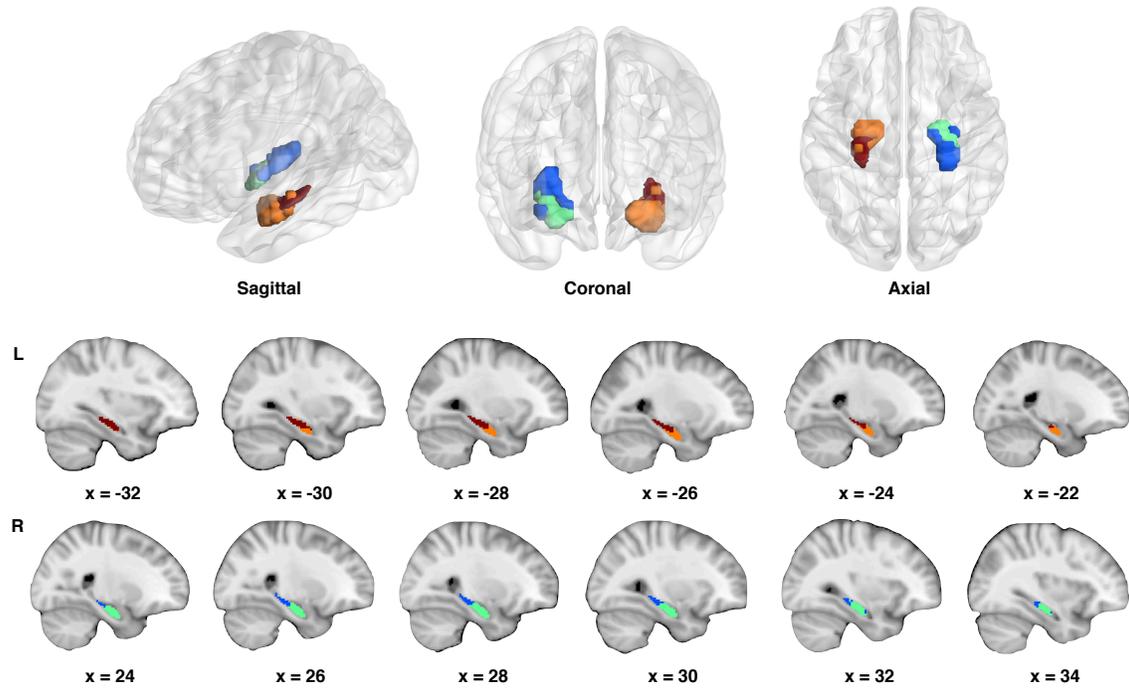


Figure A.3 Clusters of hippocampal voxels determined (using the hierarchical clustering method) based on their functional connectivity with voxels in the entorhinal cortex volume during the encoding task. (Coordinates are in MNI space)

Left Side Cluster I (dark red)
 Cluster II (orange)
 Right Side Cluster I (blue)
 Cluster II (green)

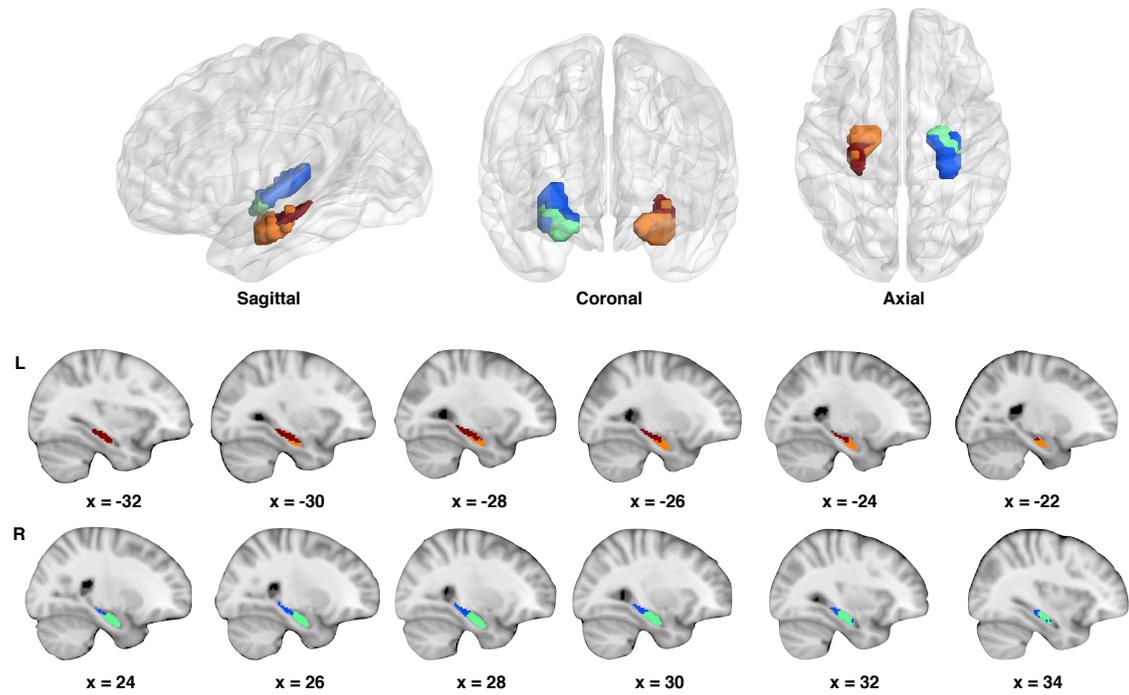


Figure A.4 Clusters of hippocampal voxels determined (using the hierarchical clustering method) based on their functional connectivity with voxels in the entorhinal cortex volume during the encoding task. (Coordinates are in MNI space).

Appendix B

■ Anterior
■ Posterior

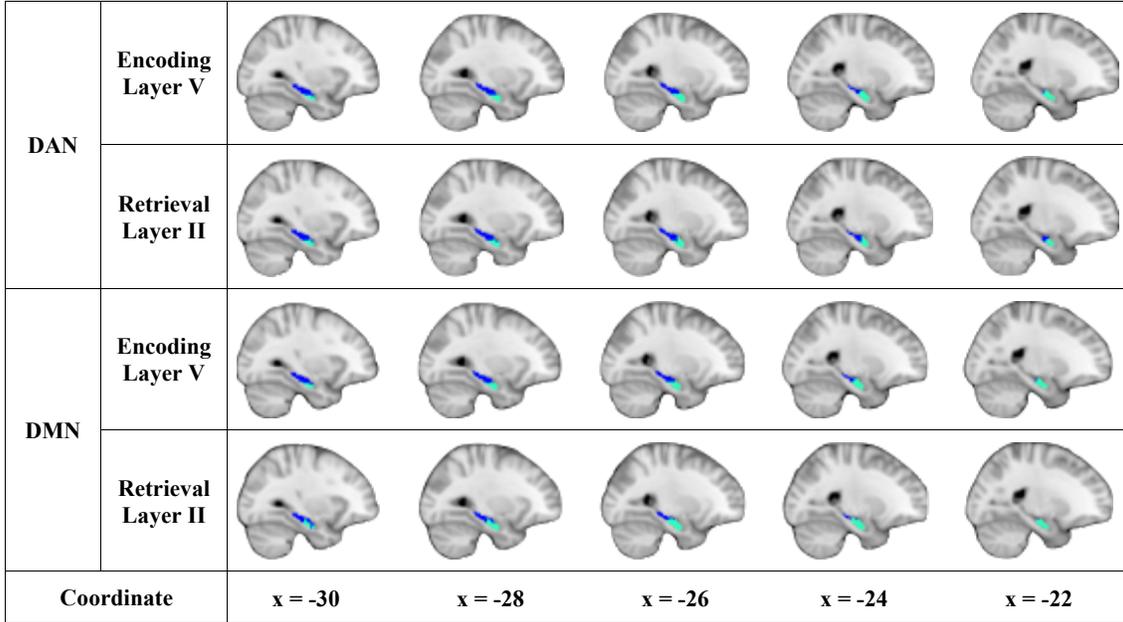


Figure B.1 Clusters of hippocampal voxels determined (using the hierarchical method) based on their functional connectivity with 1) layer V of the DMN/DAN during the encoding task, and 2) layer II of the DMN/DAN during the retrieval task. (Coordinates are in MNI space)

■ Anterior
■ Posterior

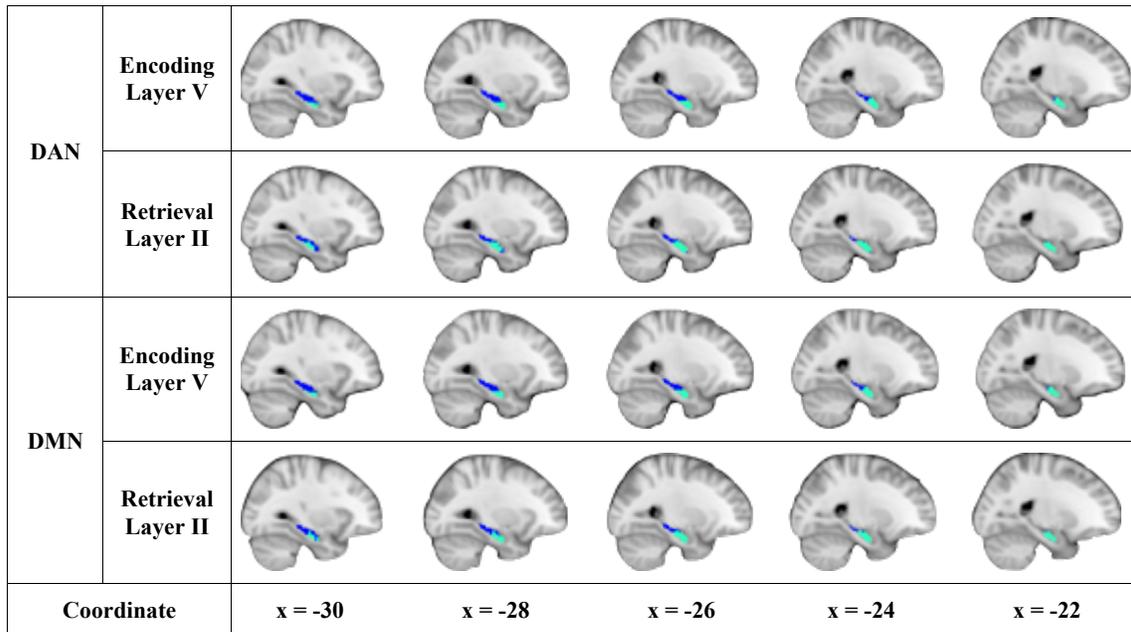


Figure B.2 Clusters of hippocampal voxels determined (using the OPTICS method) based on their functional connectivity with 1) layer V of the DMN/DAN during the encoding task, and 2) layer II of the DMN/DAN during the retrieval task. (Coordinates are in MNI space)

■ Anterior
■ Posterior

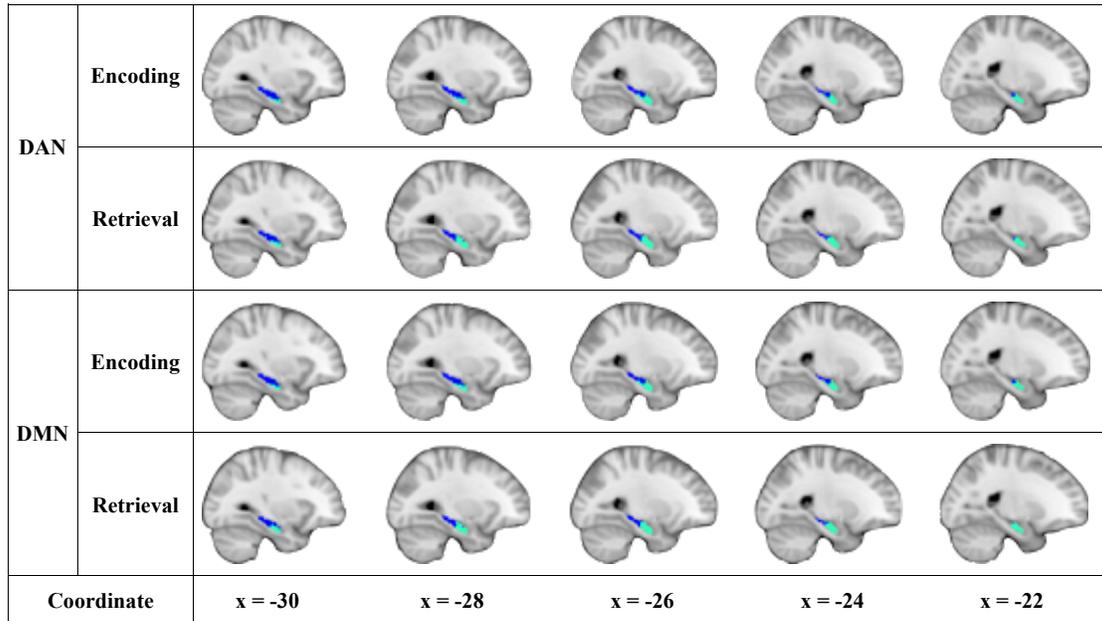


Figure B.3 Clusters of hippocampal voxels determined (using the DPC method) based on their functional connectivity DMN/DAN volume during encoding and retrieval tasks. (Coordinates are in MNI space)

■ Anterior
■ Posterior

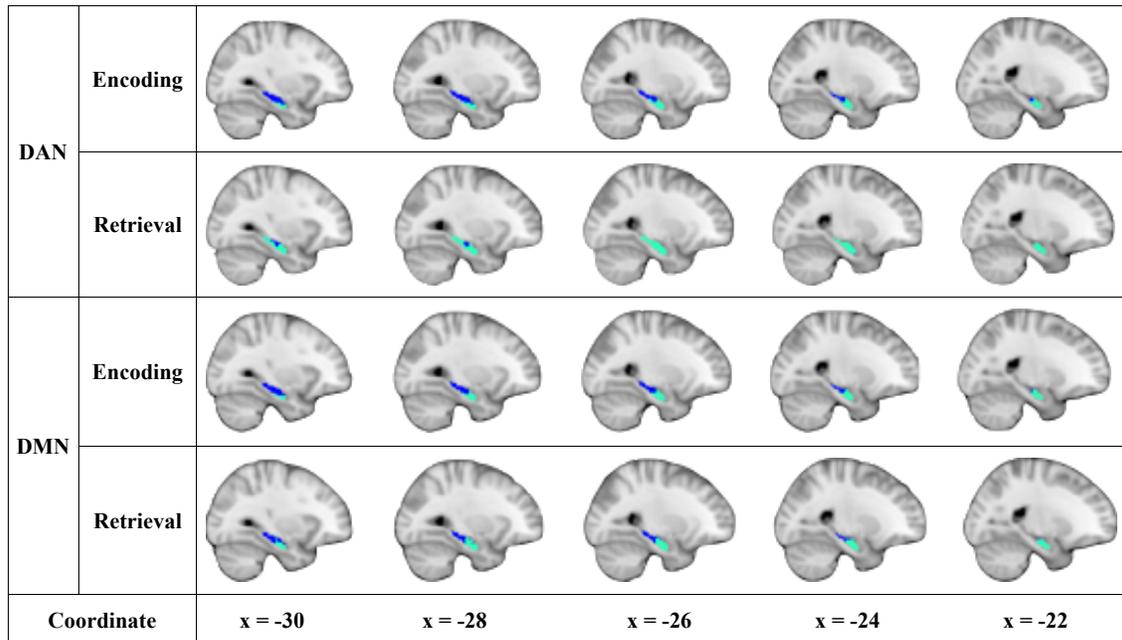


Figure B.4 Clusters of hippocampal voxels determined (using the OPTICS method) based on their functional connectivity DMN/DAN volume during encoding and retrieval tasks. (Coordinates are in MNI space)