UNIVERSITY OF CALIFORNIA
RIVERSIDE

Neural Network Based Analysis of Resting-State Functional Magnetic Resonance Imaging Data

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical Engineering

by

Lebo Wang

September 2020

Dissertation Committee:

Dr. Xiaoping P. Hu, Chairperson
Dr. Amit K. Roy-Chowdhury
Dr. Megan Peters

The Dissertation of Lebo Wang is approved:

Committee Chairperson

University of California, Riverside

Acknowledgments

First of all, I would like to thank my advisor, Dr. Xiaoping P. Hu, for the great opportunity to pursue a Ph.D. under his supervision. I am deeply grateful for his continuous support and encouragement. His intelligence and diligence also inspire me a lot.

I would like to express my thanks to my thesis committee members, Prof. Amit K. Roy-Chowdhury and Prof. Megan Peters for their direction and suggestions.

I am very grateful for helpful guidance from Dr. Kaiming Li during these years. I have learned a lot from his valuable suggestions and enjoyed insightful discussions with him. Many thanks also go to Dr. Xu Chen for his great help in every aspect, especially the valuable guidance from the very beginning. I would like to thank my labmates for the great help and fantastic memories, especially Dr. Jason Langley, Abby Barlow, Alexandra Reardon, Kaiqing Chen, Zhenhai Zhang, Tonglu Dou and Chengchun Gao. Special thanks to Sana Hussain for her professional proofreading and valuable suggestions.

Thanks to my near and far friends, for their support and encouragement during this wonderful journey: Jilong, Ruofan, Jie, Zhenyang, Sihan and others.

I want to acknowledge the journal that makes publication of my work possible. Chapter 3 is a version of: Lebo Wang, Kaiming Li, Xu Chen, and Xiaoping P. Hu. Application of Convolutional Recurrent Neural Network for Individual Recognition Based on Resting State fMRI Data. 13, 2019.

Finally, I would like to thank my family, especially my parents and my fiancée, Shiqi, for their love, support and encouragement.

To my family and Shiqi for their love and support.

ABSTRACT OF THE DISSERTATION

Neural Network Based Analysis of Resting-State Functional Magnetic Resonance Imaging Data

by

Lebo Wang

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, September 2020
Dr. Xiaoping P. Hu, Chairperson

Functional magnetic resonance imaging (fMRI) measures brain activity through the blood-oxygen-level-dependent (BOLD) signal, which has been widely used to study the human brain in neuroimaging studies. Resting-state fMRI has been acquired to study functional connectivity between brain regions without any specific tasks being involved during the scan. Raw fMRI data, a time series data with three-dimensional images in stereotaxic space, provide rich information related to neural activities. Based on recurrent neural networks (RNNs), the temporal dynamics behind the brain network has been characterized. To address the functional connectivity between spatially distant regions within/across functional networks relying on regions-of-interest (ROIs) based fMRI data, we propose two deep learning architectures for fMRI analysis. First, Convolutional RNN (ConvRNN) has been introduced with convolutions for spatial feature extraction on ROI-based fMRI data. Local interactions among ROIs from the same functional network have been extracted, and temporal features have been processed by the RNN-based architecture. Better classification performance has been achieved over previous studies. The in-place visualization based on

ConvRNN reveals the informative regions related to individual identification, leading to the same conclusions from previous studies. Second, we introduce a connectivity-based graph convolution network (cGCN) architecture for fMRI analysis. fMRI data are represented as the k-nearest neighbors graph based on the group functional connectivity, and spatial features are extracted from connectomic neighborhoods through Graph Convolution Networks. We have demonstrated our cGCN architecture on two scenarios with improved classification accuracy. cGCN on the graph-represented data can be extended to fMRI data in other data representations, which provides a promising deep learning architecture for fMRI analysis.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Functional MRI (fMRI) has been playing an important role in the neuroimaging study. It measures brain activity by detecting changes associated with the blood-oxygen-level-dependent (BOLD) signal. Resting-state fMRI (rs-fMRI) has been widely used to investigate the intrinsic functional architecture of the brain in the absence of explicit tasks. The temporal correlation of low frequency fluctuations within different brain regions has manifested the functional connectivity of the brain [15]. It has provided insight into the functional architecture of the human brain and variability in behavior [48]. Researchers have been developing analytical approaches with tremendous efforts for fMRI analysis, which has facilitated our understanding of brain activity [57, 77, 102, 166]. It is also promising in clinical applications to provide diagnostic and prognostic information based on fMRI scans for neurologic and psychiatric diseases [98].

Deep learning has achieved astonishing success in the past decade in general data analysis. It has emerged as the leading machine learning tool with top performance in

almost every field. Deep learning architecture can solve complex problems based on raw data with automatic feature extraction. And the end-to-end learning paradigm is deployed by blurring the boundaries of complex processing pipelines to avoid feature engineering. To benefit from powerful deep learning techniques, neuroimaging researchers across the world have been entering the field and applying deep learning methodologies to a wide variety of neuroimaging studies [57, 106].

The work in this thesis is directed at classification tasks with neural network models on rs-fMRI data. Rather than reapplying existing deep learning architectures to fMRI data, we propose applying new neural network based architectures to study spatiotemporal features behind fMRI data. We have achieved not only improved accuracy on neuroimaging classification tasks, but also interpretable visualization on large-scale fMRI datasets. Therefore, these deep learning architectures can provide us powerful models to study the human brain with fMRI data from a new perspective.

The rest of this introductory chapter is organized as follows. The background and motivation of this work is presented in Section 1.1. We summarize contributions of the thesis in Section 1.2. Some research questions are proposed in Section 1.3. The organization of the rest of this thesis is given in Section 1.4.

1.1 Background and Motivation

Many analytical approaches have been proposed for fMRI analysis. The initial assumption is that the brain is stationary during the whole scan and each time acquisition can be considered as an independent sample. Traditional approaches have been introduced

to analyze fMRI data, including correlation-based methods [139, 159] and graph-based methods [85, 99, 166]. Matrix factorization algorithms have also been adopted to decompose the complex brain signals into orthogonal and independent components such as principal component analysis (PCA) [5] and spatial independent component analysis (ICA) [12, 110, 157]. Even though static descriptions of functional connectivity (FC) have provided valuable information related to human brains, dynamic FC investigations explore temporal changes in neural activity underlying critical aspects of cognition and behavior [77]. Dynamic FC patterns have been evaluated based on new approaches such as sliding-window analysis [141], time-frequency analysis [20] and temporal ICA [151]. The spatiotemporal dynamic pattern of brain activity behind fMRI data has been utilized to understand the spontaneous BOLD fluctuations [107, 156].

Meanwhile, machine learning methods have been successfully applied for neuroimaging tasks. Decision trees have been applied to decode brain states with interpretable results [137]. Support vector machines have been applied for the estimation of informative spatial pattern [35], activation map extraction [97], and disease classification [86, 121]. The spatiotemporal patterns and transition characteristics of fMRI data have been modeled by the Gaussian hidden Markov model across human brains [25]. In recent years, deep learning has become more popular in fMRI analysis. Along with its applications to different fMRI tasks, it is also promising in obtaining satisfactory interpretations of fMRI data to realize computer-aided diagnosis and accurate treatment [148].

Several data representations have also been adopted for fMRI data. Raw fMRI data are originally in the volumetric representation. But when focusing on cortical surface

and subcortical regions, fMRI data are given in the surface representation. To further reduce the number of brain regions for analysis, regions of interest (ROIs) based analysis is often performed. By selecting hundreds of ROIs and ignoring structured grids in the original data representation, ROI-based analysis can provide substantial insight into the brain networks with efficient computation compared with the whole-brain analysis [129]. For example, FC analysis successfully measures the pairwise correlation between brain regions and reflect the functional connectivity in the FC matrix. Therefore, nontraditional data representations should account for brain networks beyond structured image grids.

This work aims to develop neural network based architectures for fMRI analysis. Learning from traditional analytical methods and data representations, we propose to extract spatiotemporal features by reflecting the functional connectivity behind fMRI data with the help of neural network models. For classification tasks, we obtain not only improved performance, but also reasonable interpretation to gain deeper understanding from fMRI data. We believe that our neural network models can provide large-scale analysis on fMRI data, leading to accurate diagnosis and data-driven interpretation.

1.2 Contributions of This Thesis

We introduce deep learning techniques for large-scale fMRI analysis. Two neural network models have been proposed to extract spatiotemporal features from fMRI data based on the Recurrent Neural Network (RNN) architecture and convolutional layers. These are the convolutional RNN (ConvRNN) and the connectivity-based graph convolution network (cGCN). Superior performance has been achieved on classification tasks over previous

studies, along with interpretable visualization on classification tasks. Due to the satisfactory compatibility and scalability, it is promising to deploy our neural network models with different data representations for fMRI analysis.

We introduce ConvRNN for ROI-based fMRI data. Rather than deriving spatial features with fully-connected layers, convolutional layers are adopted to extract spatial features related to the functional connectivity between connectomic brain regions. Shared convolutional kernels achieve great generalization by largely reducing the number of model parameters to avoid overfitting. Sufficient performance improvement by ConvRNN demonstrates that convolutions are effective for spatial feature extraction. Also, the visualization based on convolutional models provides an in-place interpretation of informative brain regions regarding classification tasks.

In addition, the graph-represented fMRI data have been proposed to reflect the functional networks of the human brain, which defines the convolutional neighborhood with top functional neighborhood for the spatial feature extraction. Our cGCN builds graphs with BOLD signals and shared graph topology based on the group FC matrix. The k-nearest neighbors (k-NN) graph is applied to reduce the total number of edges in the graph with feasible scalability. Spatial Graph Convolutional Networks (GCNs) are applied to extend traditional Convolutional Neural Networks (CNNs) on graph-represented fMRI data. Compared with traditional graph-theory based analyses and machine learning approaches based on feature engineering, our cGCN architecture effectively combines GCNs with graph-represented fMRI data, leading to a promising deep learning architecture for large-scale fMRI analysis.

1.3 Research Questions

Focusing on ROI-based fMRI data, we have introduced two neural network models for classification tasks. Inspired by previous studies, this thesis is guided by the following research questions:

Research Question 1: *Which type of neural network based architectures is applicable on fMRI time series?*

Deep learning architectures for MRI analysis are reviewed in Section 2.2. We use BOLD signals in time series rather than connectivity matrices as input data. Considering temporal features behind fMRI time series, the RNN-based architecture is adopted for the spatiotemporal feature extraction, including spatial features from fMRI frames and temporal features between frames. Both ConvRNN (described in Chapter 3) and cGCN (described in Chapter 4) are RNN-based architectures.

Research Question 2: *Can we apply convolutions on fMRI data to extract spatial features that reflect brain networks?*

Our previous studies adopted fully-connected layers on ROI-based fMRI data to extract spatial features [24]. To address the functional configuration behind the brain networks and the local interaction between brain regions, we propose to adopt convolutions for the spatial feature extraction. Both ConvRNN in Chapter 3 and cGCN in Chapter 4 are built with convolutional layers, which also aligns with recent deep learning architectures based on CNNs. Thus, deep learning techniques (such as Batch normalization [79]) and advanced visualization techniques [181] can be easily applied for fMRI analysis.

Research Question 3: *How can fMRI data be represented to reflect brain networks for deep learning models?*

With fMRI data in different representations, we have reviewed several deep learning architectures in Section 2.2. For ConvRNN in Chapter 3, ROI-based fMRI data are time series data with each frame in a 1D array. The order of 1D arrays is defined by the functional atlas, which reflects the functional configuration of the human brain network. Considering the implicit graph behind brain networks, we adopt the graph representation in Chapter 4. Our graph-represented fMRI data highlights top connectomic neighborhood between brain regions, which is applied for the spatial feature extraction reflecting functional networks. The compatibility and scalability of the graph-represented fMRI data are of great importance for large-scale fMRI analysis with deep learning architectures.

1.4 Thesis Organization

The rest of this thesis is organized as follows.

In chapter 2, we mainly review deep learning learning techniques for medical imaging applications, focusing on neuroimaging studies with brain MRI data. First, some popular large-scale MRI datasets are introduced with their features and differences. Then, we mainly discuss deep learning applications on MRI data. According to MRI data representations, deep learning architectures with volumetric representation, surface representation and other non-Euclidean representations have been discussed with some representative examples. Differences in data representations and deep learning architectures reflects the progress in studying functional brain networks. Deep learning techniques could not only

improve MRI applications, but also provide a comprehensive understanding of MRI data from a new angle.

In chapter 3, ConvRNN architecture is introduced for individual identification using rs-fMRI data. Rather than deploying fully-connected layers for spatial features, our ConvRNN deploys convolutional layers to extract spatial features between ROIs from the same functional network. Compared with previous traditional RNN models and other traditional analytical approaches, our ConvRNN architecture could achieve better identification performance. We applied in-place visualization to determine informative regions related to the individual identification task, including default mode network and frontoparietal network. The success of ConvRNN architecture for individual identification indicates that the RNN-based convolutional architecture is promising to analyze fMRI data.

In chapter 4, we introduce a cGCN architecture for fMRI analysis. fMRI data are represented in graphs based on functional connectivity, and spatial features could be extracted from connectomic neighborhoods rather than from Euclidean ones. We have tested our cGCN architecture on two classification tasks with resting-state fMRI data. The improved classification accuracy demonstrates that deep learning architecture based on graph-represented fMRI data is effective in capturing functional connectivity features. Our results indicate that cGCN is effective and promising in fMRI analysis. ConvRNN stored ROI data in 1D array with pre-defined order from the functional atlas, while cGCN treated ROI data in graphs. Thus, connectomic information could be better represented in graphs and extracted by cGCN model.

In chapter 5, we summarize the similarities and differences between ConvRNN and cGCN architectures in fMRI analysis. Both of our neural network models take ROI-based fMRI data as inputs and adopt convolutions for spatial feature extraction, but ROI data are presented in different data representations. The contribution and limitations of our proposed architectures are discussed for future work.

Chapter 2

Deep Learning for Magnetic Resonance Imaging

Deep learning has achieved great success in the area of computer vision. Due to the similarity between images and MRI data, the ultimate goals with MRI images highly coincide with many image processing tasks, such as image classification, semantic segmentation and image reconstruction. Promising results have been demonstrated with deep learning techniques on medical images [57, 106]. Considering various types of MRI data for diverse applications, we will focus on the introduction of popular large-scale MRI datasets and deep learning architectures for brain MRI analysis.

The rest of this chapter is organized as follows. Some popular large-scale MRI datasets are briefly introduced in Section 2.1. MRI applications with deep learning techniques are shown in Section 2.2. In terms of data representations, we present deep learning applications on volume data in Section 2.2.1, surface data in Section 2.2.2 and other repre-

sentations in Section 2.2.3. Deep learning architectures on fMRI timeseries are separately discussed in Section 2.2.4.

2.1 Large-scale MRI Datasets

MRI techniques have been introduced and developed in the past 50 years. Despite the popularity of MRI techniques, it is still relatively costly to carry out MRI experiments, particularly with high-quality large-scale MRI datasets, as a lot of collaborative effort is always required from data acquisition to data preprocessing. Therefore, it is quite valuable to have some public datasets available for easy access, and analytical approaches can be rapidly prototyped and easily developed with a fair comparison. The following describes some popular MRI datasets for neuroimaging studies.

The Human Connectome Project (HCP) is one of the most popular MRI datasets [162]. There are over 1100 healthy young adults (22 to 35 years old). Each subject has structural data, rs-fMRI and task fMRI scans, which provide valuable resources to study human brains. There are two attractive assets for our study within the HCP dataset. First, the homogeneity of the whole dataset could reduce experimental challenges to design new models. Second, each rs-fMRI scan is relatively long (1200 frames in 15 minutes), which is helpful in studying the temporal dynamics behind fMRI sequence.

Autism Brain Imaging Data Exchange (ABIDE) is a large-scale consortium of MRI data, consisting of Autism spectrum disorders (ASD) patients and normal controls. There are more than 2000 subjects involved from multiple imaging sites (17 imaging sites in ABIDE I [38] and 18 sites in ABIDE II [37]. With different protocols, the huge heterogeneity

across imaging sites exists in many aspects, including imaging parameters, the length of fMRI scans and psychiatric variables. It is anticipated that analytical approaches can be developed with superior performance across imaging sites to achieve a better understanding of ASD on this heterogeneous dataset.

In addition, UK Biobank is a collection of brain MRI data from over 40000 subjects [111]. Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu) was launched with MRI data to investigate mild cognitive impairment (MCI) and Alzheimer's disease (AD). The Center for Biomedical Research Excellence (COBRE) provides a MRI dataset with 147 subjects to study schizophrenia [19, 108]. Recently, Facebook and NYU Langone Health released a large MRI database for MR reconstruction (the current dataset only consists of knee scans) [180]. OpenNEURO (https://openneuro.org/) is a free and open platform for sharing neuroimaging data. With the increasing availability of large-scale MRI datasets, deep learning based approaches are increasingly applied and rapidly developing for MRI analysis.

2.2 Deep Learning on MRI Data

There have been numerous types of deep learning applications with MRI data. Classification and prediction tasks have been realized, such as disease classification [69, 74, 92], individual identification [24, 167] and age prediction [76]. Segmentation tasks have been carried out, such as lesion segmentation [10, 81], volumetric brain segmentation [22] and sub-cortical structure segmentation [39]. The preprocessing pipeline of MRI data has also been improved with deep learning techniques in terms of speed, accuracy and reliability. [131]

performed skull stripping and brain extraction with improved speed and reliability. [176] developed robust motion-related regressors to reduce motion-related artifacts. [3] proposed a promising registration strategy based on T1-fMRI translation with synthesized fMRI volumes. [26] applied an unsupervised learning approach to realize cortical parcellation and functional area alignment. [54] demonstrated that cortical parcellation based on spectral embeddings could gain drastic speed improvement over conventional methods. Graph convolutional adversarial networks were also applied to predict missing diffusion MRI data in longitudinal studies [73]. Furthermore, improved image reconstruction and faster MR imaging techniques have been developed through deep learning architectures [134, 150, 174]. We will introduce different deep learning architectures according to the data representation of MRI data.

2.2.1 Deep Learning on Volumetric Representation

When considering volumetric MRI data as 3D grid images, it is straightforward to apply 3D CNN architectures on MRI volumes. [39] applied a 3D CNN model to realize subcortical segmentation without traditional registration/normalization steps, as shown in Figure 2.1, and the enhanced speed compared to atlas-based methods and robustness on heterogeneous datasets has been demonstrated. [92] adopted 3D CNNs to classify AD from 3D structural MRI data. With a 3D fMRI sequence, [101, 123] adopted 3D CNN architectures to realize disease classification with different preprocessing strategies on the temporal axis of fMRI data. [184] used 3D CNN on whole-brain fMRI data for automatic functional brain networks reconstruction. [88, 185] applied 3D CNN to extract connectome-based fea-

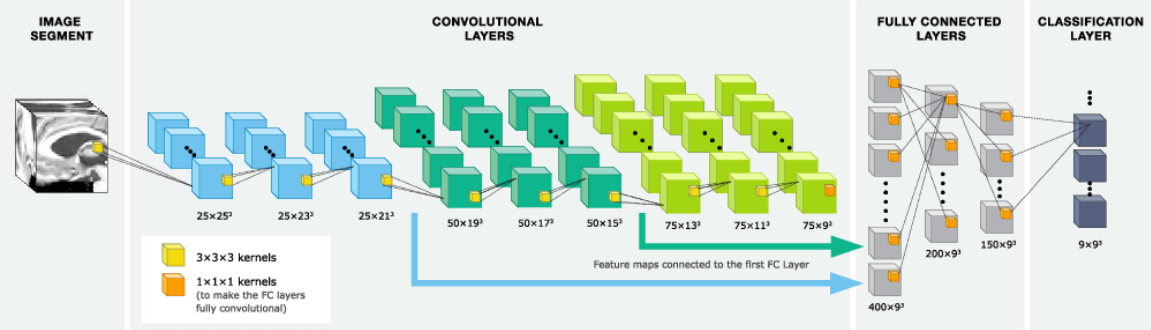


Figure 2.1: The 3D CNN architecture on volumetric MRI data for brain structure segmentation, reprinted from [39]. Nine 3D convolutional layers were used to extract spatial features.

tures for ASD classification, with different mapping strategies between the functional brain networks and fMRI voxels.

It is obvious that traditional spatial feature (e.g., edges, corners and texture) can be extracted on volumetric MRI data through 3D CNN models. But for the structural and functional networks behind MRI data, original volumetric representations of MRI data should be thoughtfully processed. Moreover, the training and inference of 3D CNN models are computationally intensive. It could become more intractable to unveil the temporal dynamics behind 4D MRI data.

2.2.2 Geometric Deep Learning on Surface Representations

In contrast to the volumetric representation of brain MRI data, the brain functionality in cortical surfaces and subcortical regions can be represented in cortical surface geometry with great advantages over volumetric representations [161, 163]. The surface-based representation has the topology of a 2D sheet with a highly folded geometry to reflect the geometric structure of the cerebral cortex [34]. Considering the characterizations of

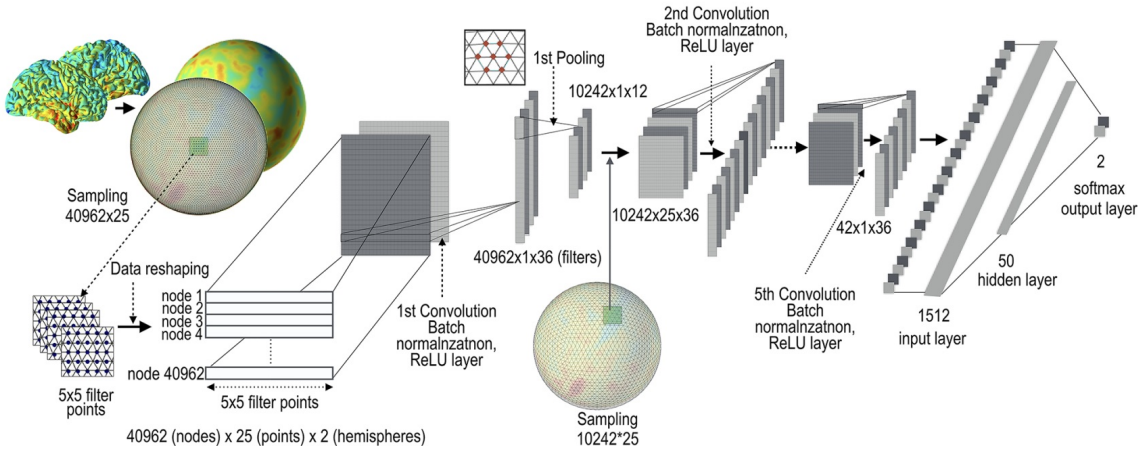


Figure 2.2: Architecture of geometric convolutional neural network (gCNN) on cortical thickness information for sex classification, reprinted from [146]. The cortical thickness maps in surface-based representation are transformed into spherical space. Specific data sampling and reshaping methods are defined for convolutions on a mesh surface to achieve spatial features with position invariance.

surface-based MRI data, geometric deep learning has been widely applied. A geometric CNN on the projected spherical surface, as shown in Figure 2.2, was proposed for sex classification using the cortical thickness information [146]. Similarly, Spherical CNN was employed to parcellate cortical surfaces on spherical surface meshes [183].

Rather than transforming surface points into volumetric representation, surface-based MRI data are represented in cortical surface geometry. By projecting surface data onto a spherical surface, geometric deep learning has proposed modified operations (e.g., convolution, pooling) on non-Euclidean manifolds of the brain surface. In fact, geometric deep learning models spatial features of the geodesic shape with the non-Euclidean representations of MRI data. Improved performance with efficient computation has been achieved over traditional deep learning architectures on volumetric representations.

2.2.3 Deep Learning on Other Representations

With the exception of the volumetric representation and surface representations, original MRI data have been processed into various representations to study the brain network, such as connectivity matrices and individual components by ICA. These data representations commonly ignore original image grids but highlight the connectomic information. Previous studies have adopted graph theory to study the topological property behind these data representations. For example, the structural and functional connectivity matrices are equivalent to the adjacency matrix. Graph related metrics have been utilized to understand brain mechanism, such as network efficiency, modularity, hierarchy [166] and small-world network [169]. Despite the great success of graph-based metrics, it is of great necessity to build deep learning architectures without feature engineering. The features that are automatically generated from MRI data might facilitate MRI applications and enable a novel understanding of MRI data. Next, deep learning architectures for MRI data based on other representations will be reviewed.

Deep Learning on Connectivity Matrices

Connectivity matrices are widely applied to represent structural and functional connectivity, where pairwise connectivity is evaluated between two brain regions. To study the brain connectome, deep learning models have been applied to extract spatial patterns from the connectivity matrix. [116] applied the autoencoder model with fully-connected layers to extract features from structural connectome data to predict treatment outcomes in epilepsy. Similar autoencoder architectures were also applied to functional connectome

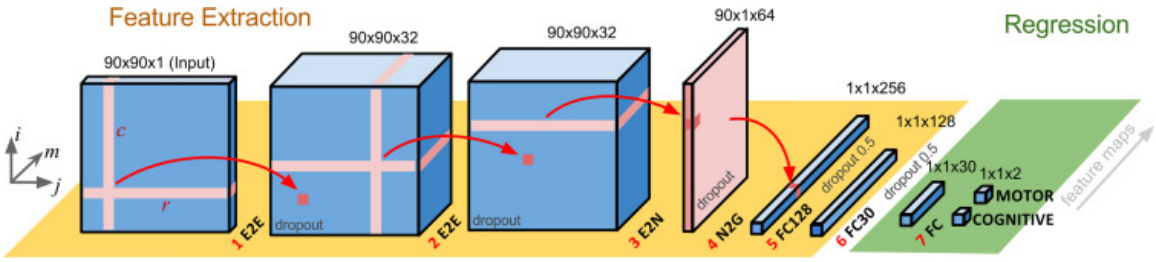


Figure 2.3: The BrainNetCNN architecture on structural brain connectivity networks for neurodevelopmental outcome prediction, reprinted from [83]. Rather than considering the adjacency matrix as an image, specific convolutional operations on the structural brain network have been proposed to consider spatial locality between elements from the adjacency matrix, including edge-to-edge (E2E), edge-to-node (E2N) and node-to-graph (N2G) layers.

data to classify ASD patients [28, 62, 69]. Considering the grid structure of the structural connectivity matrix, the BrainNetCNN architecture was proposed with special convolutional layers to leverage the topological locality of the connectivity matrix. As shown in Figure 2.3, BrainNetCNN can be seen as a modified CNN model in the structural connectivity matrix, which achieved satisfactory performance in predicting neural development for infant brains [83].

Traditional analytical approaches derive topological characterizations to reveal the brain connectome based on connectivity matrices. With deep learning models, spatial features based on connectivity matrices can be further generated beyond original image grids. However, there are some problems behind these deep learning methods. First, the theoretical explanation of the spatial features that are extracted by CNN models on the connectivity matrix (not raw MRI data) has not been well established. Second, the FC matrix is regarded as a temporal summary of fMRI sequence over a period of time, which makes it difficult to determine the temporal feature behind fMRI time sequence with the collapsed temporal axis.

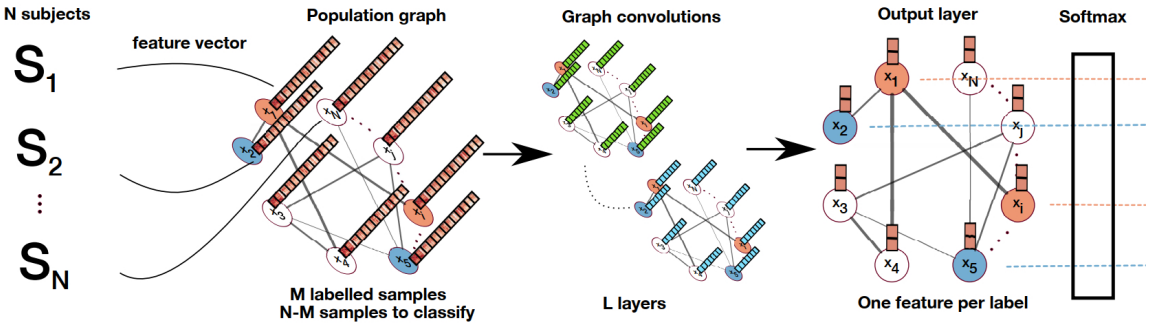


Figure 2.4: Graph Convolutional Network on the population graph for the disease classification task, reprinted from [122]. In the population graph, nodes are subjects with image-based feature vectors, and edges reflect the similarity between subjects evaluated by phenotypic information. The semi-supervised node-level classification on the population graph is deployed to classify patients from normal controls.

Deep Learning on Graph-Represented Data

Considering the complex interaction between brain regions, brain networks among distant brain regions should be presented in other data representations. Especially for ROI-based and ICA-based MRI data, the graph-based representation can be one of many promising choices. Meanwhile, traditional deep learning techniques have been successfully extended to graphs in recent years. Novel deep learning pipelines should be proposed for MRI analysis, including data representations and deep learning architectures.

Due to the implicit graph behind the brain connectome, MRI data have been constructed into various types of graphs. [122] proposed to achieve disease prediction on a population graph, as shown in Figure 2.4. In the population graph, nodes represent different subjects with image-related features, and weighted edges quantify the similarity between subjects according to phenotypic information. Semi-supervised learning was employed to achieve node-level classification for psychiatric disorders (ASD and AD). [6] improved the

construction of the population graph and deployed graph convolutional neural networks to achieve better classification performance. [96] constructed two types of graphs: a structural graph based on anatomical data and a functional graph from the FC matrix (in Figure 2.5), both of which achieved significant performance improvement. [178] also took functional connectome data over short time-intervals as graphs and generated low-dimensional graph embeddings to identify brain functional states. As shown in Figure 2.6, time-variant nodes from independent components by ICA and binary edges derived from the Ledoit Wolf covariance with threshold values are built into undirected graphs [7]. The node embeddings were generated by temporal convolutions on each node sequence. GCN and pooling operations were applied on nodes to extract graph-level features for classification. [80] defined graph topology based on anatomical data with time-variant node values from rs-fMRI data, in which weighted edges were quantified by the topological distance. For the cortical parcellation task with surface data, spectral embeddings based on GCNs have been extracted to achieve significant improvement in accuracy and in speed over Euclidean approaches and geometric deep learning approaches [54].

For GCNs applied to graph-represented data, two categories of GCNs have been proposed: spectral GCNs and spatial GCNs. For spectral GCNs, graphs are decomposed into spectral bases associated with graph-level information according to the graph spectral theory [17]. Because of its solid theoretical interpretation, spectral GCNs have been widely applied for neuroimaging studies [6, 7, 55, 73, 96, 122]. In contrast, spatial GCNs imitate the Euclidean convolution on grid data to aggregate spatial patterns centering on each node and its connected neighbors. With shared convolutional parameters, node-centered convo-

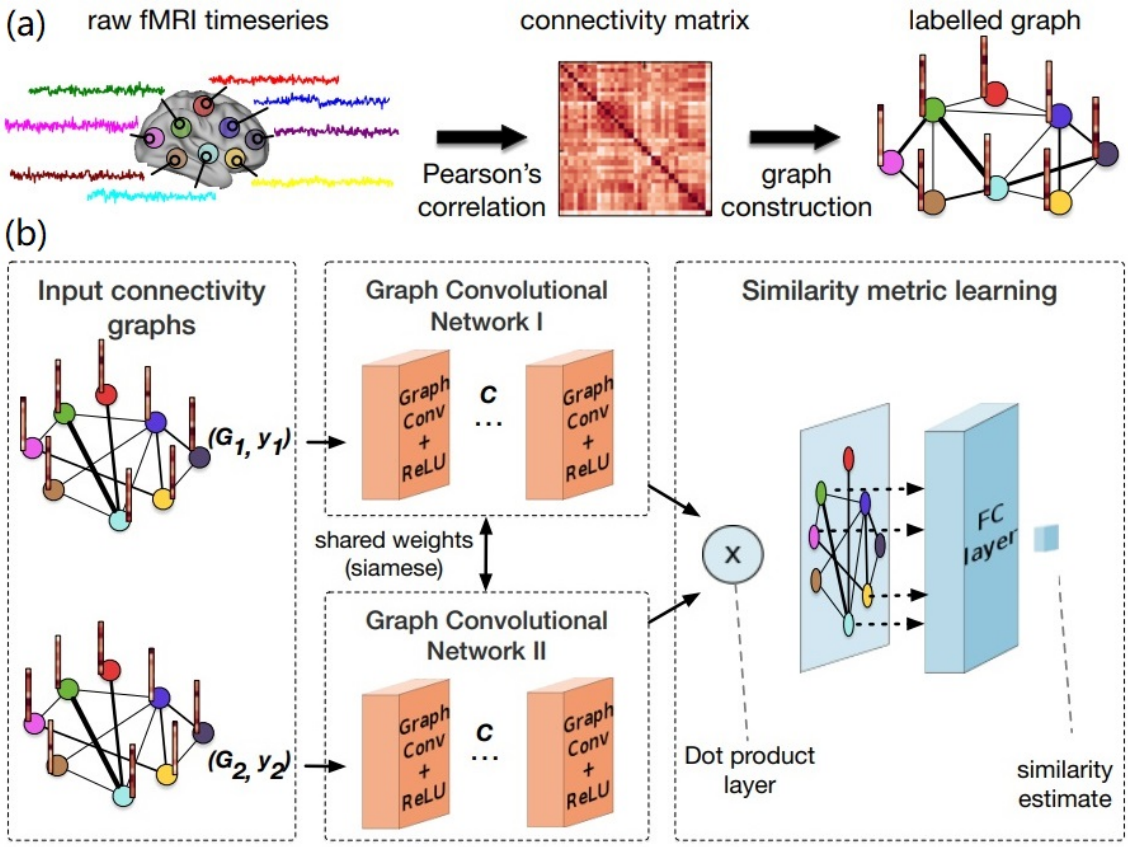


Figure 2.5: Overview of the pipeline of the graph-represented fMRI data for disease classification, reprinted from [96]. (a) The graph representation of fMRI data. The graph topology can be derived from the anatomical information or from the mean functional connectivity matrix within a population. Node vectors are generated from the functional connectivity matrix. (b) Metric learning with the Siamese Graph Convolutional Networks architecture is applied to the graph-represented data.

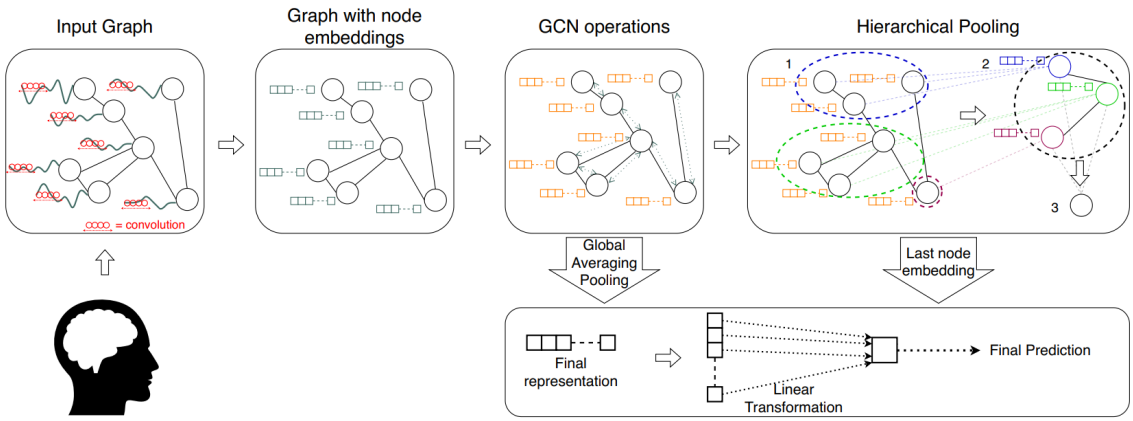


Figure 2.6: Pipeline of the spatiotemporal graph representation of fMRI data for sex prediction, reprinted from [7]. The graph is constructed with time-variant connectome sequence from independent component analysis as node values. Edges are derived from the covariance matrix and binarized with different threshold values. Node embeddings are generated by temporal convolutions of each time series. The graph-level features are extracted by Graph Convolutional Networks. The final prediction is given based on the global averaging pooling or the hierarchical pooling on graph-level features.

lutions are carried out across locations and graphs. For example, a spatio-temporal graph convolutional network has been trained to model spatial features and temporal dynamics of FC based on sub-sequences of fMRI data, as shown in Figure 2.7. Due to its computational efficiency and improved generalization, spatial GCNs have gained increasing attention in the community [171, 182].

From the volumetric representation to the surface representation, the cortical distance following the surface reflects the anatomical property of the brain instead of physical distance in 3D Euclidean space. With the graph-represented MRI data, the distance between any two regions can be evaluated by any reasonable distance metrics (e.g., structural and functional connectivity). Different from geometric deep learning with modified Euclidean operations, GCNs have been proposed with non-Euclidean convolutions on graphs.

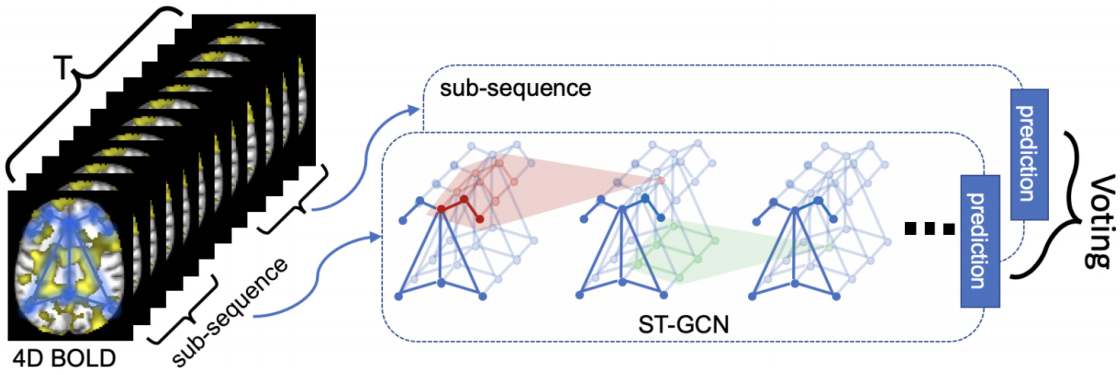


Figure 2.7: Overview of the spatio-temporal graph convolutional network for fMRI data, reprinted from [49]. Time-variant BOLD signals are taken as node values from ROI-based fMRI data. For each node, edges are obtained with top two connectomic neighbors in accordance to the functional organization of the brain. Also, another edge connects a ROI to the same ROI in the next frame of fMRI data. The graph is constructed for a sub-sequence of fMRI data, including both spatial and temporal information in one spatio-temporal graph.

Therefore, it is promising to apply deep learning architectures with non-Euclidean operations on MRI data with graph representations.

2.2.4 Deep Learning on Time Series

fMRI timeseries are put in chronological sequence such that the fluctuation of dynamic brain signals reflects the neural activity. It is interesting to study the temporal changes behind fMRI data. RNN-based deep learning architectures have been widely applied to model the temporal dynamics of brain signals. [61] successfully used RNNs to model the dynamics of brain activity with superior performance over traditional models. [43] characterized the pathophysiology of ASD with RNN. By incorporating phenotypic information, the subsequent study improved classification performance with a similar RNN architecture [41]. Our first work applied RNN to extract spatiotemporal features from an individual identification task [24], as shown in Figure 2.8. The performance was further improved by

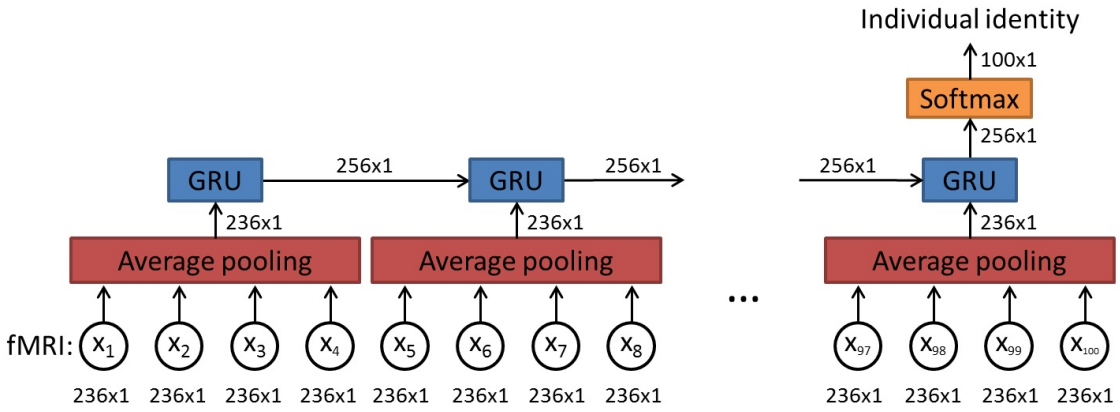


Figure 2.8: The RNN architecture on ROI-based fMRI time series for individual identification, reprinted from [24]. Spatial features are processed by the fully-connected structure of the Gated recurrent units (GRUs) layer. Temporal information between frames is leveraged based on spatial features.

using convolutional RNN architecture [167]. These studies adopted ROI-based fMRI data as inputs with reduced spatial complexity. Similarly, fMRI data can also be preprocessed by temporal ICA to reflect brain networks as independent components. [173] adopted RNN-based architecture on fMRI independent components to discriminate schizophrenia from normal controls. As shown in Figure 2.9, convolutional kernels in different scales were applied to extract spatial features, which were followed by RNN layers to extract temporal features.

The success of RNN-based architectures indicates that RNN layers effectively model the temporal dynamics behind fMRI time series. The joint training of CNN layers and RNN layers is promising to realize better performance with spatiotemporal features. Therefore, it is worthy to explore deep learning architectures with convolutions for spatial

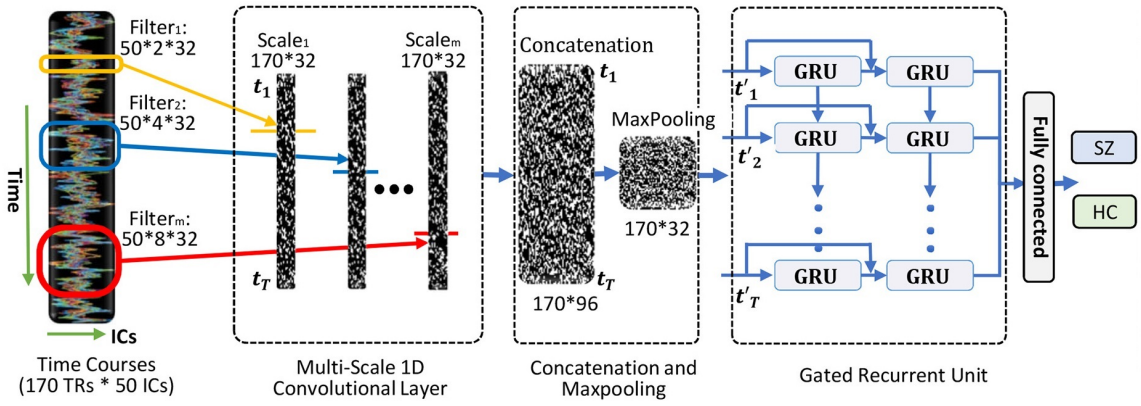


Figure 2.9: The RNN architecture on fMRI independent components to discriminate schizophrenia, reprinted from [173]. Convolutional kernels in different scales are adopted to extract spatial features between brain networks in the form of independent components, which are followed by RNN layers to extract temporal features.

features and to examine RNN layers for temporal dynamics when extracting spatiotemporal features from fMRI time series.

Chapter 3

Application of Convolutional Recurrent Neural Network for Individual Recognition Based on Resting State fMRI Data

3.1 Introduction

Mainstream fMRI studies have been focusing on deriving population consensuses using group analysis. A group analysis in neuroimaging, albeit important, commonly neglects individual-to-individual variations. The importance of individual variability in neurobiological research has drawn increasing attention [113]. Using task-fMRI, significant individual differences in brain activation were identified, reflecting alterations in cognitive

function and behavior [9]. Individual variability in FC has been successfully used to identify subjects from a large group. More specifically, static connectivity patterns throughout the brain were shown to be subject specific and distinctive across scan sessions and conditions, providing powerful features for individual identification [47]. Therefore, exploring the individual uniqueness of the brain connectivity points to a new avenue to study the brain.

Although the static FC achieved decent accuracy, it required a sufficiently long data set (600 frames, 7.2 min) and considered only the spatial pattern through the temporal correlation without taking temporal features into full account. The performance degraded with short clips of fMRI data, probably due to temporal variability (or dynamics) in the resting-state fMRI data, which leads to high variability in the FC derived from a short window. On the other hand, the dynamic information of resting state activity, if taken into account, could provide additional features for individual identification, improving the accuracy with the short time series.

With the increasing availability of large fMRI datasets (introduced in Section 2.1), deep learning architectures have been discussed with different data representations in Section 2.2. There are several limitations to model fMRI data on the perspective of image grids: 1) it is computationally intensive to deal with voxel-wise fMRI data via CNN models; 2) brain activity mostly occurs in cortical and subcortical structures, making convolutions on white matter unnecessary; 3) the fMRI time course from a single voxel is usually very noisy; and 4) spatial features may be confined to a small neighborhood in the Euclidean space, especially in shallow CNN models. Therefore, it is of great necessity to build deep learning models that are more appropriate for the organization of the brain and can efficiently extract

connectomic features beyond a single voxel and its Euclidean neighborhood. By choosing a bunch of brain regions and ignoring fMRI image grids, ROI-based analysis is efficient to unveil brain networks based on functional atlases. ROI-based fMRI data can be represented in 2D arrays, including the temporal dimension and the spatial dimension (the order of ROIs defined by the functional atlas). Some functional atlases (e.g. Power's atlas [132]) divide ROIs into communities of tightly interconnected brain regions, which helps to understand the modularity of the human brain network [14]. Though our initial study employs the traditional RNN model to extract spatiotemporal features [24], the functional configuration of ROIs is not explicitly addressed by fully-connected layers. Therefore, we propose to adopt the convolutional operation on the spatial dimension of 2D arrays to extract local features between adjacent and connectomic (according to the functional atlas) ROIs.

In the application of time sequence modeling, RNN-based architectures have shown outstanding promise in a broad range of applications, including video classification, machine translation, and biomedical image segmentation [23, 50, 147, 164]. Specifically, we have discussed RNN-based architectures to study the dynamics of brain activity on fMRI time series in Section 2.2.4. Moreover, a convolution-based RNN was introduced to make full use of features in both spatial and temporal domains, consistently outperforming fully connected RNNs [149]. Therefore, combining the local features between adjacent and connectomic ROIs by the convolutional structure and sequence modeling capability of RNN may lead to a better approach to extract spatiotemporal features for individual identification on rs-fMRI data.

In the meantime, it is also valuable to visualize the underlying features in the trained convolutional models. Although deep learning is becoming a panacea in almost every domain, it has been criticized due to its poor interpretability as being a black-box. While many attempts have been made to provide an interpretation and an intuitive understanding of trained networks, our understanding of how these networks work and what is important behind their performance have not kept up with the pace of the development of neural networks. While dedicated deep learning models have achieved amazing performance by end-to-end learning through huge volumes of data, better comprehension of the success of these models can uncover fundamental principles of deep neural networks and reveal important features within the data.

In this chapter, we adopted convolutional RNN or ConvRNN for individual identification using rs-fMRI data. The convolutional recurrent model was able to achieve individual identification with shared convolutional weights capturing local coactivation features. In-place visualization of the informative area by ConvRNN also opened up a new avenue for understanding fMRI data based on individual differences.

3.2 Materials and Methods

3.2.1 Dataset and Preprocessing

The resting-state fMRI data for 100 subjects from the Human Connectome Project (54 females, age: 22–36, and TR = 0.72 s) was used in this work. Each subject had four resting-state fMRI sessions, 1200 volumes for each session, leading to 4800 volumes per subject in total [162]. The fMRI data was preprocessed by the HCP minimal preprocessing

pipeline [52] and denoised by ICA-FIX [142], for the removal of spatial artifact/distortion and motion-related fluctuations. Surface-based registration was performed with the MSM-ALL template [138]. To decrease the computation complexity, 236 ROIs over the cerebral cortex, as shown in Figure 3.1 based on meta-analysis [132], were used for subsequent analysis. BOLD signals within each ROI (10 mm diameter sphere) were averaged spatially. We ordered our ROIs-based data in a 1D array and preserved the order of ROIs according to the Power Atlas, where ROIs having similar functional connectivities are close to each other. They were also demeaned and scaled to unit variance over the temporal axis. For each fMRI session, fMRI data with 1200 volumes was divided into twelve 100-frame clips as inputs of ConvRNN. Data from Day 1 was used as the training dataset. The two sessions from Day 2 were used as validation and testing datasets, respectively. The best model was decided based on the validation dataset and the final performance was assessed on the testing dataset.

3.2.2 Convolutional Recurrent Neural Network

The architecture of the ConvRNN is given in Figure 3.2, along with its unrolling version. In contrast to conventional RNN, convolution was applied in both the input-to-state and state-tostate transitions, in place of the Hadamard product. There were two stacked convolutional layers, with the first convolutional layer containing 8 filters and the second convolutional layer having 16 filters. The kernel size of all convolutional filters was 2. Padding was used in all convolutional layers such that the outputs from each layer had the same spatial dimension as the original input, which is very important for the subsequent

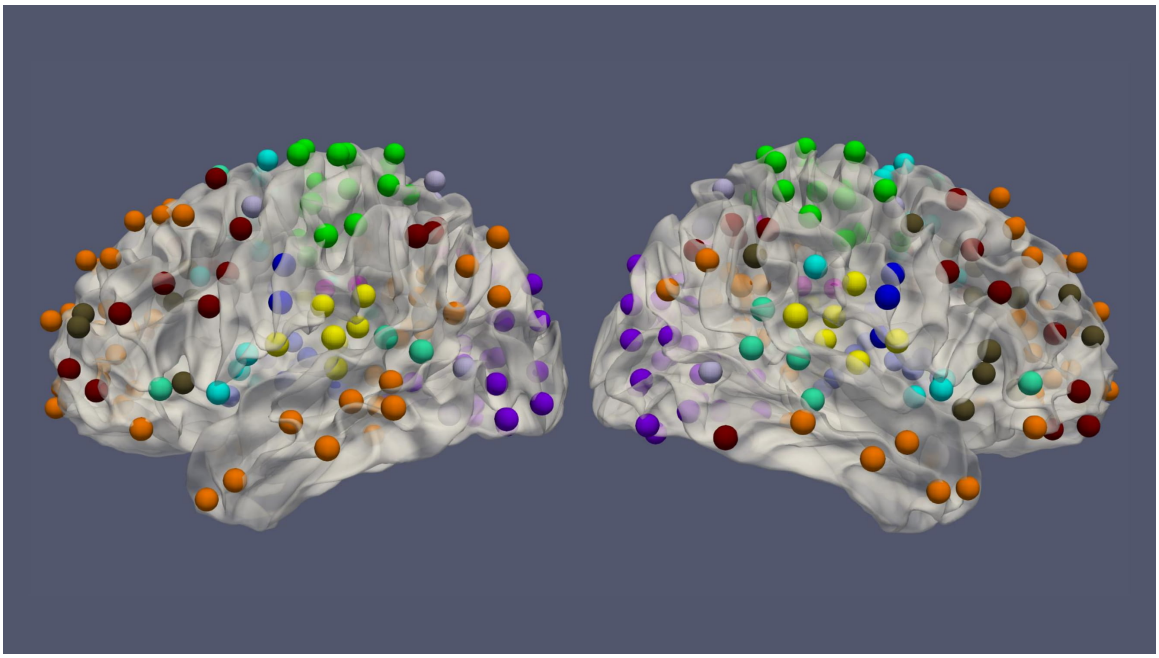


Figure 3.1: The spatial distribution of 236 ROIs over the cerebral cortex. Voxels within the 10 mm diameter sphere were averaged to get the value for each ROI.

visualization of the in-place features. Batch normalization layers were used before the non-linear activation layers of Rectified linear unit (ReLU) to reshape the distribution of convolutional layer outputs in order for better convergence and easy training [79]. The final Softmax layer with 100-category outputs was used for classification based on averaged outputs from all hidden states. No temporal or spatial pooling layer was employed to keep the spatial and temporal resolutions of the original fMRI data. All kernel weights were initialized by the Xavier uniform initializer [53], and recurrent weights were initialized as random orthogonal matrices [144].

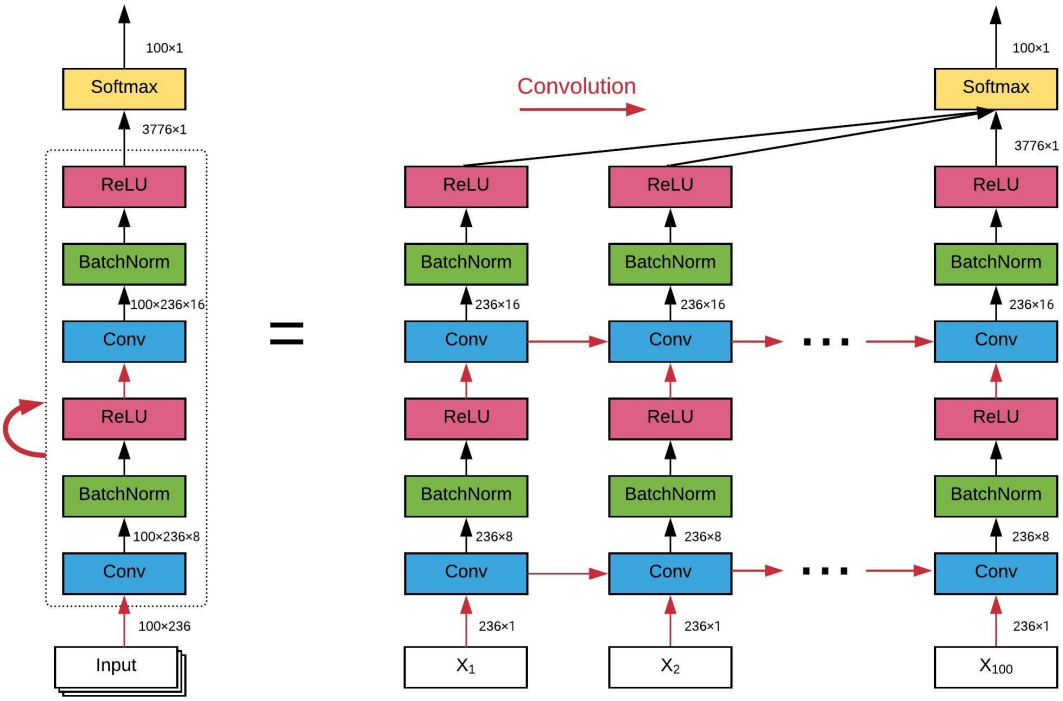


Figure 3.2: The architecture of our convolutional recurrent neural network and its unrolling version over time. All red arrows represent the convolutional operations between each input-to-state and state-to-state transitions. Batch normalization and ReLU as the non-linear activation are utilized after each convolutional layer. Final classification is based on all hidden states on average. The dimension of the data flow through the diagram is also labeled.

3.2.3 Training of The Neural Network

Our implementation of ConvRNN was carried out in Keras [30] with the Tensorflow backend [1]. Considering the limited number of frames for each subject, we chose 100 frames of fMRI data as inputs during training and validation, which is the tradeoff between the number of fMRI clips and the number of frames for each clip. Shuffled minibatches of training data as inputs were fed into the ConvRNN with the batch size of 128. Adam optimizer [89] was applied for training with the initial learning rate set to 0.001, and reduced if the validation accuracy stopped increasing. Dropout layer with 50% was utilized before the final classification to avoid overfitting only during the training [152]. After each training epoch, the model was evaluated on the validation dataset and saved only if better validation accuracy was achieved. Finally, the performance of the best model was measured on the testing dataset, which was never involved during training or validation.

It is well known that RNN is difficult to train properly, even though it is a powerful model for time series modeling. The main reasons are vanishing and exploding gradient issues of Backpropagation Through Time (BPTT) on the unrolling version of RNN [13]. Therefore, advanced architectures with gating mechanism to overcome the vanishing and exploding gradient problem, such as the Long short-term memory (LSTM) [70] and the Gated recurrent unit (GRU) [27], have gained a lot of popularity in practice to model long-term dependencies. In this work, LSTM with convolutional structure was applied. For training techniques, we used L2 regularization for recurrent weights, along with the gradient clipping strategy as a simple and computationally efficient method, effectively addressing the issue of exploding gradients (Pascanu et al., 2013). In the present work, the clipping

norm of the gradient was set to 1. Different L2 values (0.1, 0.01, 0.001, and 0.0001) on recurrent kernel weights were tested to achieve the best validation accuracy.

3.2.4 Visualization of the Individual Identification

Our ConvRNN first performed feature extraction through two convolutional recurrent layers and then fed the features into the Softmax layer for the 100-category classification. Original data was projected to a high-dimensional feature space, which was easily separated by the classification layer. In the feature space, fMRI data from the same subject are expected to be close to each other and cluster tightly. In consideration of the single classification layer, the identification accuracy of our ConvRNN relies heavily on the performance of feature extraction by convolutional recurrent layers. In order to ascertain and visualize the performance of convolutional recurrent layers in low dimensional space, t-Distributed Stochastic Neighbor Embedding (t-SNE) [160] was applied to map datapoints in high-dimensional feature space onto a two-dimensional representation.

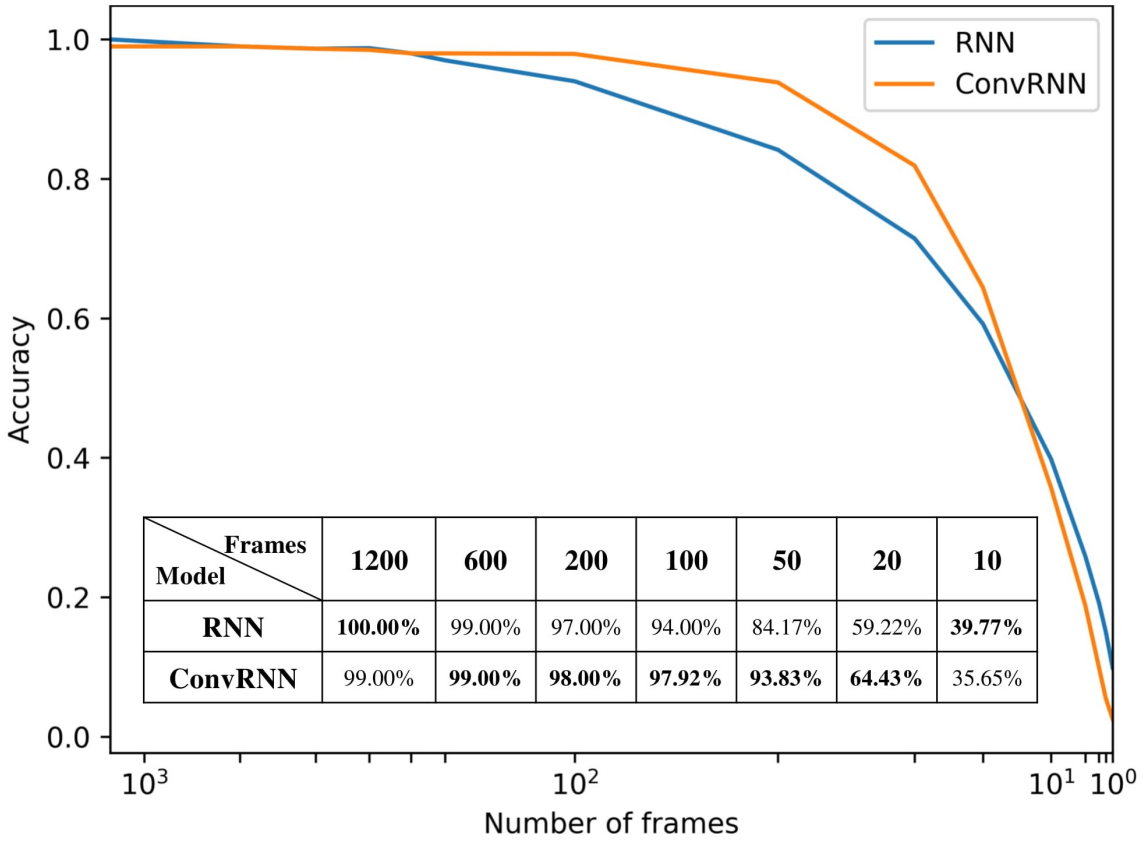
To visualize and understand informative areas related to individual identification, intermediate outputs from convolutional layers were examined. Output patterns were obtained from convolutional layers after non-linear activation and mapped onto the cortical surface (20 mm radius sphere). As all regions are considered equal in our convolutional model during the training, but they are of different importance to the final classification. We also used the occlusion method to visualize informative areas [181]. More specifically, in order to ascertain the contribution of ROIs with regard to individual identification, input of each ROI was zeroed out, and the subsequent performance decrease with the same model configuration was considered as the contribution of this ROI to the final classification.

Architecture	Number of parameters (feature extraction)	Test accuracy
RNN [24]	405K (380K)	94.43%
RNN w/o temporal pooling	405K (380K)	95.33%
ConvRNN	382K (3.8K)	**98.50%**

Table 3.1: The accuracy of different models on the testing dataset and their number of model parameters.

3.3 Results

We carried out the supervised classification task to identify each subject from a group of 100 subjects. First, the identification accuracy of different models was assessed on the testing dataset with 100 frames of fMRI data as inputs. As seen in Table 3.1, our ConvRNN model was able to achieve 98.50% accuracy on the testing dataset, where the best performance was obtained with the L2 value of 0.001 during training. The test accuracy for the traditional RNN with average temporal pooling was 94.43% [24]. In order to exclude the influence of temporal averaging, we trained another RNN without the temporal averaging and achieved an identification accuracy of 95.33%. Furthermore, we evaluated the performance of these models using different window sizes on the testing dataset. With the pre-trained models, we adopted different number of frames (from 1200 frames to single frame) as inputs from the testing dataset and evaluated the identification performance. Testing results with different number of frames are plotted in Figure 3.3. As shown in the figure, ConvRNN outperformed conventional RNN (no temporal averaging) in all cases except with 1200 frames or with less than 10 frames. In contrast, FC could achieve over



Model \ Frames	1200	600	200	100	50	20	10
RNN	**100.00%**	99.00%	97.00%	94.00%	84.17%	59.22%	**39.77%**
ConvRNN	99.00%	**99.00%**	**98.00%**	**97.92%**	**93.83%**	**64.43%**	35.65%

Figure 3.3: The relationship between identification accuracy and the window size. We evaluated pre-trained models on testing dataset. Our ConvRNN outperformed RNN except with 1200 frames or with less than 10 frames.

90% accuracy with 600 frames of fMRI data. But the individual identification accuracy drops to 70% on average when only a short period of fMRI data (100 frames) is used [47].

To visualize convolutional outputs on low-dimensional space, we applied t-SNE on intermediate outputs of our ConvRNN before the classification layer. There were 16 convolutional filters in the second convolutional recurrent layer of our ConvRNN. With proper padding, the output of the layer was made to have the same spatial dimension as the input. The feature space with 3776 dimensions was then mapped to a 2D space in Figure

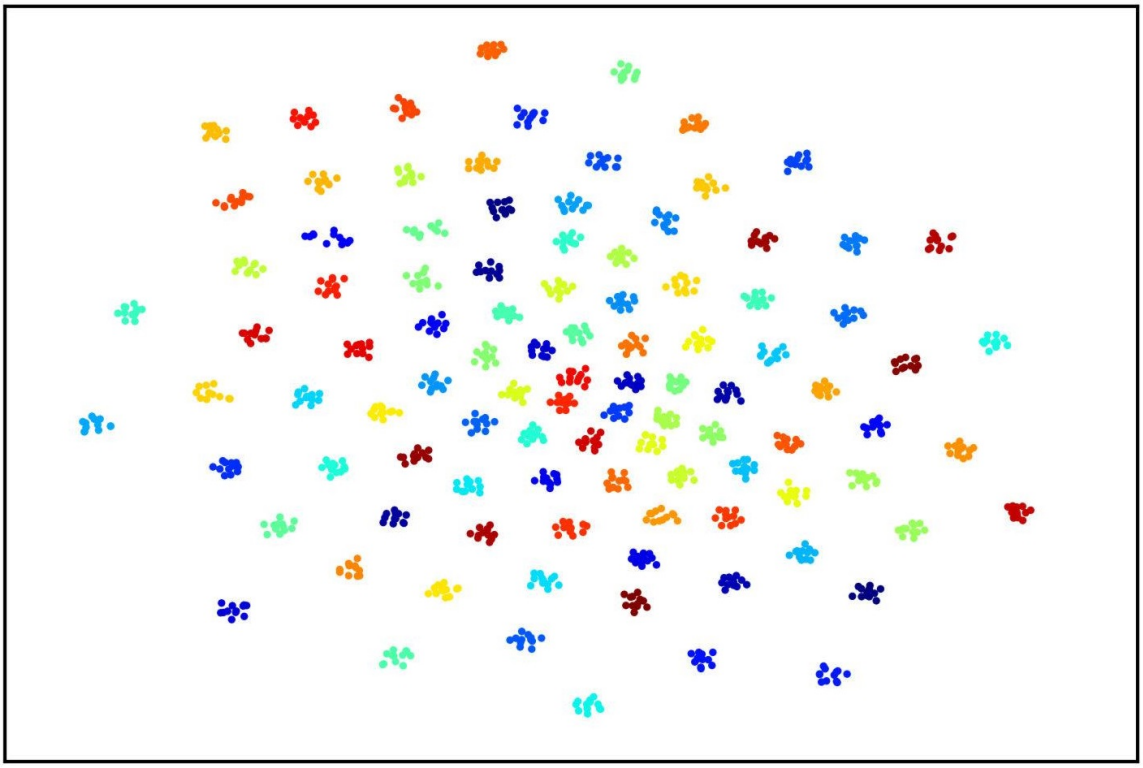


Figure 3.4: t-Distributed Stochastic Neighbor Embedding (t-SNE) visualization of 2nd convolutional recurrent layer outputs based on 100-subject testing dataset. Twelve hundred 100-frame testing data from 100 subjects were fed into ConvRNN with outputs being obtained before the classification layers and projected to 2D space by t-SNE. Projections for different subjects are in different colors.

3.4. It is clear that 100 subjects (12 clips with 100 frames for each subject) in the testing dataset appear as non-overlapping cliques in different colors with each clique representing one subject. This figure clearly indicates that spatiotemporal features, capable of individual identification, were successfully obtained by convolutional recurrent layers.

To visualize intermediate outputs of ConvRNN, average patterns from first and second convolutional layers of ConvRNN are shown in Figures 3.5 and Figure 3.6, respectively. Most patterns from the first convolutional layer were quite similar (except Filter 6) with large distinctive areas, which could be considered as the ubiquitous low-level features

from fMRI data. While highlevel patterns from the second convolutional layer had diverse informative regions, which were sparse and localized inside the area of those low-level features generated by the first convolutional layer. Meanwhile, when ROIs were individually occluded, performance degradation was observed when some ROIs were occluded, while the occlusion of some ROIs led to negligible degradation of the performance. In Figure 3.7, the absolute value of the performance degradation, normalized to reflect the contribution of each ROI is shown. It is evident that the informative area generated by alternative occlusion was similar as the patterns from first convolutional layer of ConvRNN.

3.4 Discussion

While most resting-state fMRI studies have relied on group averages, this study employed individual differences for individual identification. Unlike the first study of individual identification employing static FC [47], we incorporated both temporal and spatial features from the fMRI data. As an improvement of our previous work employing the recurrent architecture [24], we applied a convolutional recurrent neural network which led to a significant improvement in performance and a straightforward means to visualize in-place features. Figure 3.3 shows that ConvRNN is better than conventional RNN for the majority of the time windows. The performance of ConvRNN was slightly worse than conventional RNN with 1200 frames, probably due to the small number of testing data when the performance was evaluated on fMRI clips with 1200 frames. On the other hand, since our ConvRNN was trained with the fixed number of frames (i.e., 100), it is not be optimized

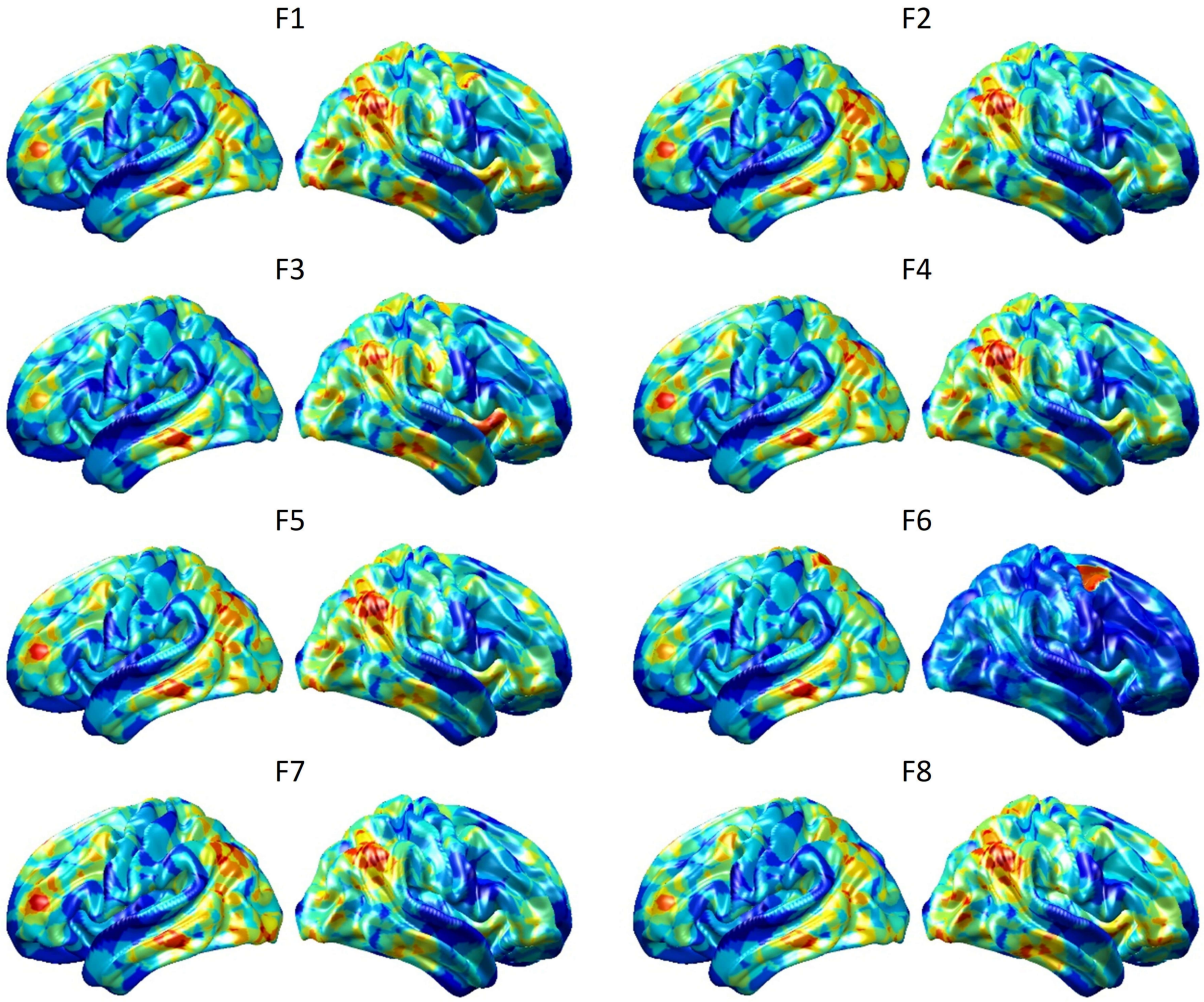


Figure 3.5: Average output patterns of the first convolutional layer with 8 convolutional kernels. Twelve hundred 100-frame testing data from 100 subjects were fed into the convolutional recurrent model with output patterns generated and averaged from the first non-linear activation layer. Red areas represent large activation values.

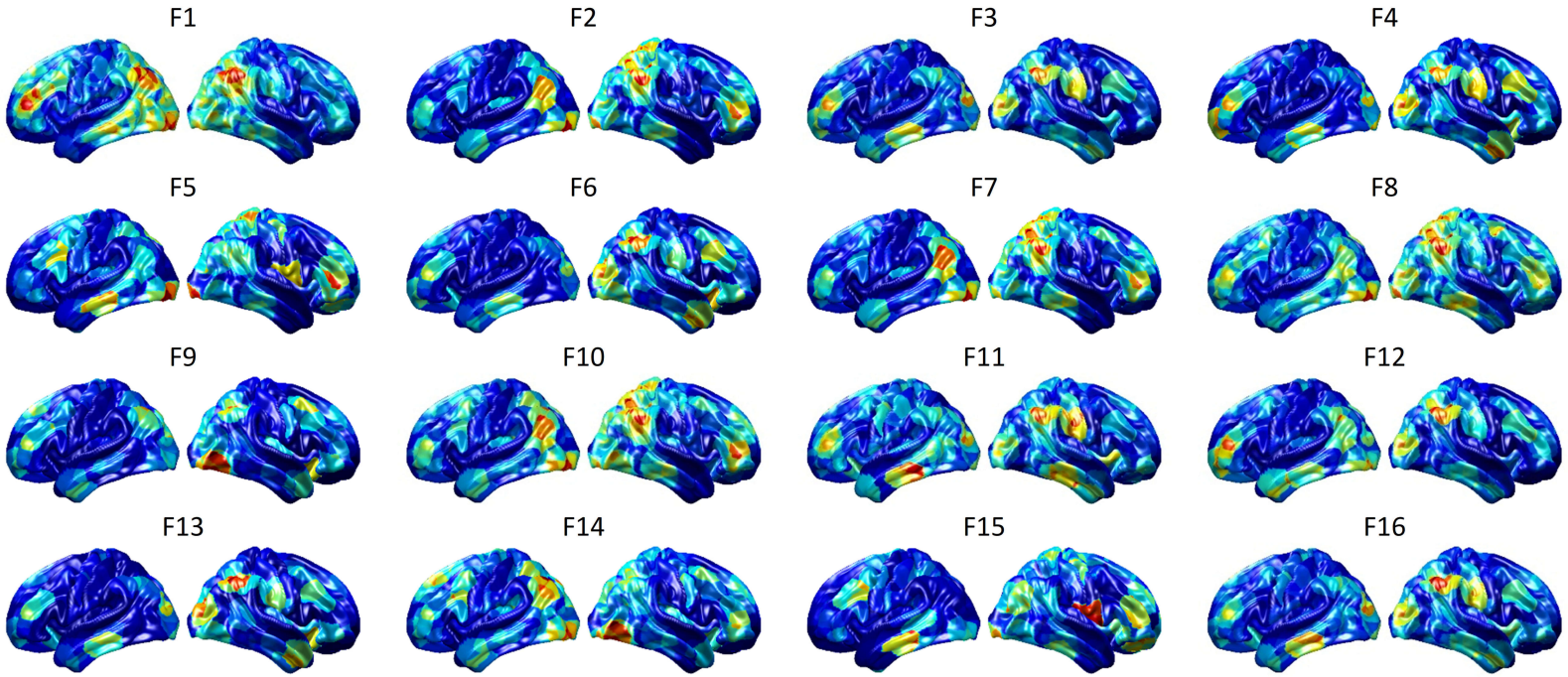


Figure 3.6: Average output patterns of the second convolutional layer with 16 convolutional kernels. Twelve hundred 100-frame testing data from 100 subjects were fed into the convolutional recurrent model with output patterns generated and averaged from the second non-linear activation layer. Red reas represent large activation values.

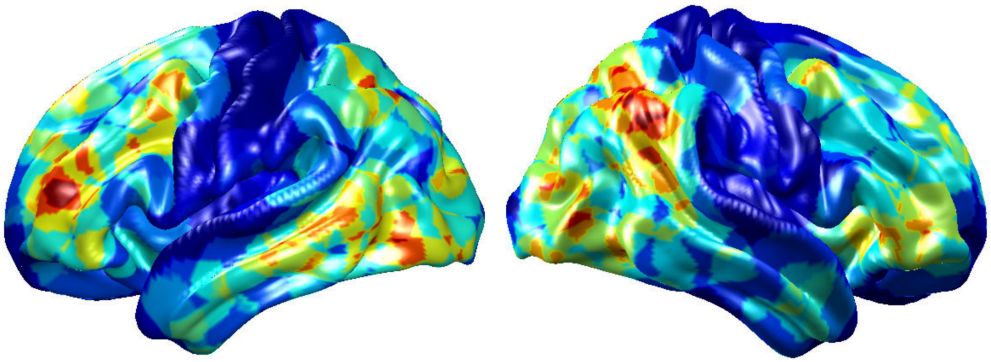


Figure 3.7: The performance degradation with occlusion. Each ROI was zeroed out separately and evaluated with the pre-trained model of ConvRNN. The performance degradation reflects the contribution of each ROI. Red region reflects large performance degradation if corresponding ROIs were occluded.

for short clips of data with less than 10 frames, and its performance with frames less than 10 is therefore worse than that of conventional RNN.

Our ConvRNN has the same number of parameters compared with the conventional RNN, indicating that both models have comparable model complexity. Apart from the different types of recurrent unit, our earlier work [24] employed a temporal pooling layer to reduce the temporal resolution. For a fair comparison with this work, another conventional RNN was applied without the temporal averaging layer. The accuracy of the conventional RNN without the temporal averaging layer was 95.33%, which was 0.9% higher than that with the temporal averaging. This improvement due to increased temporal resolution is significantly smaller than the improvement achieved with the adoption of the convolutional structure, indicating that the latter is the main contributor to the performance enhancement. Both the spatial and temporal features were fully utilized by convolutional kernels in the ConvRNN with unprecedented identification accuracy. Also, feature extraction layers of ConvRNN showed strong discriminating power for 100 subjects,

where pre-trained layers could be applied for transfer learning on new subjects or semi-supervised learning on partially labeled data. On the other hand, only one hundredths of the parameters in ConvRNN were from the convolutional recurrent layers. Convolution with shared weights in spatial and temporal spans makes it more robust against overfitting during training. Given the reduced number of trainable weights, increasing the depth and width of the model is possible without significantly increasing the model complexity, possibly capturing more sophisticated features in both spatial and temporal domains.

Furthermore, convolutional kernels with shared weights sweep across ROIs and frames. Different from the indecomposable matrix multiplication in conventional RNN, ConvRNN generates in-place features with exactly spatial correspondence as the original data. Furthermore, ConvRNN accumulates temporal information related to evolving features in the hidden state. Therefore, it is possible to examine the hidden states to have an in-place visualization and understanding of hidden features from ConvRNN. It is also clear that informative regions from two convolutional layers are totally different, in agreement with the conclusion drawn from convolutional neural networks for image classification [177]. Beside the direct visualization of the hidden state, the occlusion of ROIs served as an indirect method for visualizing significant regions under resting state for identification. Two visualization approaches came to the same conclusion of informative ROIs in term of individual identification using the resting-state fMRI data. In terms of resting state networks (RSNs) in the literature [71, 99], the informative area identified by our ConvRNN mainly contained frontoparietal network (FPN), default mode network (DMN), as well as visual network (VN). Our result is consistent with a previous study, which concluded that

the most distinguishing network was FPN, with significant improvement achieved through the combination of multiple RSNs (Finn et al., 2015). In contrast, occlusion of language network (LN) and somatosensory motor network (SMN) did not cause much reduction in performance. One possible explanation is that there was little explicit or individual specific language or motor activity during the acquisition of the resting-state fMRI datasets used here.

Although DMN and FPN stood out as the most important networks for individual identification, other networks also contributed to individual identification [46]. It is likely that contributions from most networks may be needed with more subjects to be identified. In addition, a hierarchical approach, incorporating all networks, might be the most appropriate and robust approach. Furthermore, a recent study demonstrated that task induced changes in FC provided better prediction of individuals, whereas resting state fMRI data failed to capture the full range of individual differences [56]. Such changes in FC could also be incorporated into our model to further improve the identification accuracy.

Several limitations should be noted for further work. First, the present study adopted only 236 ROIs within 10 mm diameter spheres on average, which was enough for an accurate identification for a group of 100 subjects. Inadequate power of identification could be present on new subjects beyond existing subjects. Possible reasons are the limited feature extraction of our model and the high variability of the fMRI data. It is likely that smaller and more ROIs are needed for the identification of more subjects. But more advanced models with superior performance on fMRI data should be explored. Second, current visualization of individual identification depicted in Figure 3.7 was based on group

average of ROIs' performance. While this highlights areas that are most important for individual identification, it does not explicitly depict individual features. Such features will be the focus of our future studies. Third, visualization of the spatial pattern is easy to understand, but the remarkable performance achieved by RNN suggests that a substantial amount of information is coming from temporal features. Visualizing and understanding temporal features are still necessary to gain a deeper insight into the brain dynamics. Furthermore, other popular architectures (e.g., Siamese network) and pre-trained models should be applied and compared with our approach in terms of classification performance and training efficiency in future work.

3.5 Conclusion

In this chapter, we described the application of the convolutional recurrent neural network for individual identification based on resting-state fMRI data. To explore the dynamics in the resting state fMRI data, the convolutional architecture with recurrent structure was implemented to extract and incorporate features in both spatial and temporal domains. Compared to conventional RNN model, our ConvRNN model exhibited better identification performance, with local features between neighboring ROIs being modeled by convolutional kernels. Moreover, visualization based on the ConvRNN model provides a direct understanding of the success of identification; this could lead to a promising alternative for analyzing fMRI data.

Chapter 4

Graph Convolution Network for fMRI Analysis Based on Connectivity Neighborhood

4.1 Introduction

In chapter 3, the ConvRNN architecture has been introduced on ROI data in 2D arrays. Considering the complex brain networks, the implicit graph behind ROI data can be reflected in the graph representation. Motivated by breakthroughs of deep learning on grid data, efforts have been made to extend CNNs to graphs. We have reviewed deep learning applications on graph-represented fMRI data in Section 2.2.3.

In this chapter, we present a connectivity-based Graph Convolutional Network, a spatial GCN architecture, for fMRI analysis. The graph representation was defined with

the k-nearest neighbor graph based on FC matrix. Convolutions were performed on graphs rather than image/ROI grid, which allows us to efficiently extract connectomic features of the underlying brain activity. The present model (`https://github.com/Lebo-Wang/cGCN_fMRI`) has been applied in two scenarios, and the results indicate that cGCN is effective for fMRI analysis compared with traditional deep learning architectures.

4.2 Material and Methods

4.2.1 cGCN Overview

The overview of our cGCN architecture is shown in Figure 4.1. ROIs were considered as graph nodes with the blood-oxygen-level-dependent (BOLD) signals of each frame as their attributes. Convolutions were performed within neighbors defined by the k-NN graph based on the groupwise FC matrix. Specifically, we first obtained the individual FC matrices on all data and averaged them for the group FC matrix. Based on that, a k-NN graph was generated by retaining only the top-k edges in terms of their connectivity strength (i.e., average correlation coefficient) for each node. The FC-based k-NN graph was used to guide the convolutional operations with functionally connectivity-based neighborhoods. For simplicity, the same graph was shared by all subjects and at all time frames. Although each node had a few neighbors, the convolution field of each layer was extended by stacking multiple convolutional layers in the architecture. Between convolutional layers, skip connections were added from the prior convolutional layers to the last one, providing multi-level feature fusion for classification and accelerating the model training by alleviating the vanishing gradient problem. Outputs from the convolutional layers were followed by an

RNN (or a temporal average pooling) layer to generate temporal evolutions by combining spatial representations from all frames. A Softmax layer was used at the end for the final classification.

4.2.2 Graph Construction

We represented each frame of fMRI data with a graph. A shared graph structure that reflects the intrinsic functional was derived from the group FC matrix based on all data [127]. The ROIs were considered as nodes of the graph, and the FC connections were considered as edges of the graph. To reduce the total number of edges in the graph, a k-NN graph was obtained by only keeping the top-k connectivity neighbors for each node in terms of the connectivity strength. k is the hyperparameter related to the topological structure of graphs, which controls the sparsity of the graph. To evaluate the effect of k, different values of k (3, 5, 10 and 20) were assessed in our experiments.

4.2.3 The Edge Function of cGCN

The k-NN graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ comprises nodes $\mathcal{V} = \{1, \ldots, N\}$ and edges $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$. Edge (i, j) represents the directed edge from ROI_i to ROI_j. To explicitly model the co-activation pattern between graph nodes, we chose the following edge function [168], as the convolutional operation for our cGCN convolutional layers:

$$x_i^{'} = \max_{j:(j,i)\in\mathcal{E}} h_\Theta(x_i \| x_j - x_i) \tag{4.1}$$

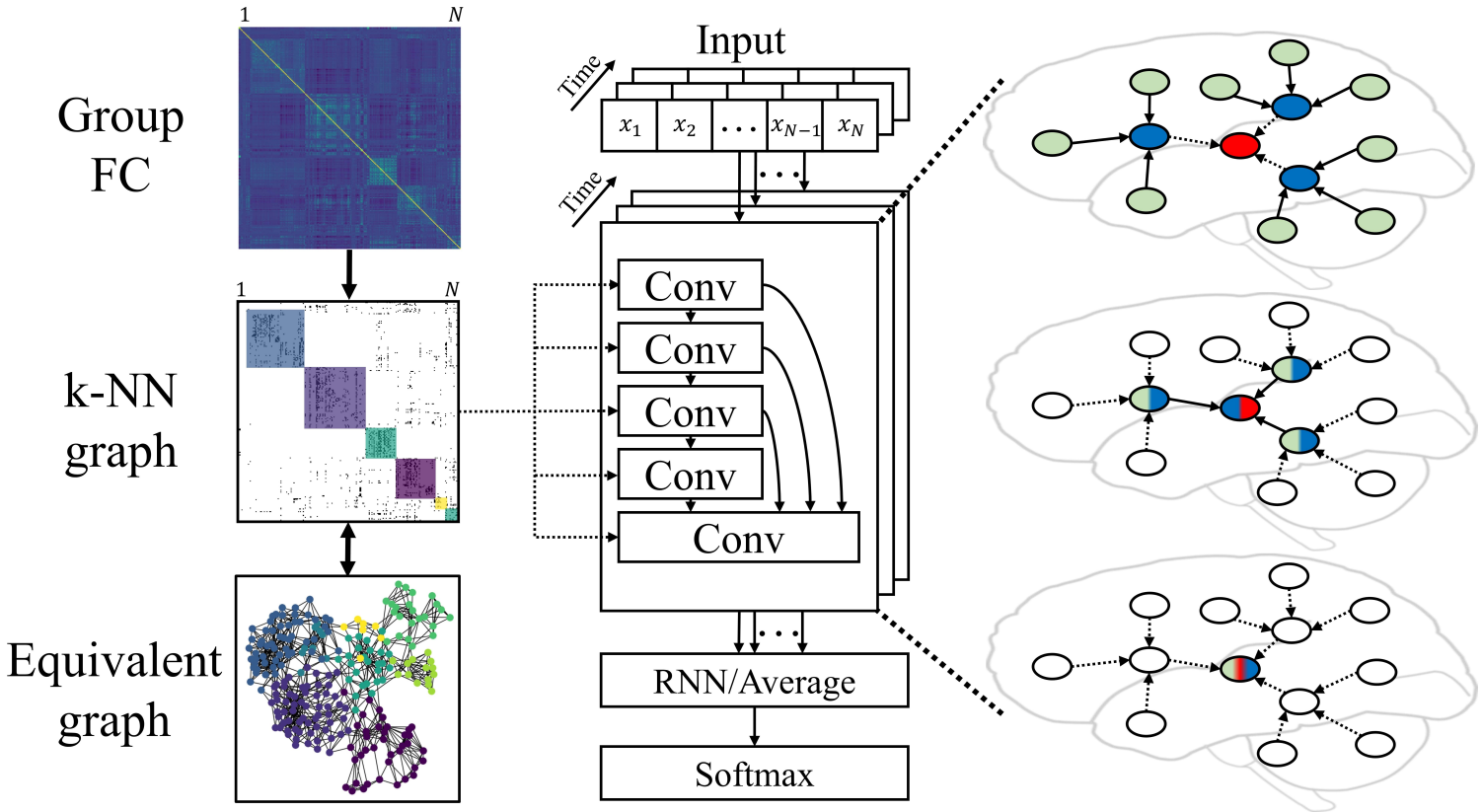


Figure 4.1: Overview of cGCN. One the left, the graph definition for cGCN was based on the group FC from all data, which can be further simplified as a k-nearest neighbors (k-NN) graph with binarized edges. In the middle, the cGCN architecture consisted of five convolutional layers. The convolutional neighborhood was defined by the shared k-NN graph across convolutional layers, time frames and subjects. The recurrent neural network (RNN) layer (or the temporal average pooling layer) obtained latent representations from all frames. The final classification was achieved by the Softmax layer. On the right, an intuitive illustration of the spatial graph convolution showed the information aggregation between neighboring nodes.

where x_i was the BOLD signal of the central node i and x_j was that of a connected neighbor node j. In addition to original features from the central node, the node value difference $x_j - x_i$ was also appended as the complementary feature. h_Θ represents trainable weights Θ with a non-linear activation function h, which was implemented by the multi-layer perceptron (MLP) with the rectified linear unit (ReLU). The final convolutional output, x_i', was the maximal activation from its k neighborhoods. In this way, spatial features between connectivity-based neighbors, including the node activity and co-activation pattern, were modeled.

4.2.4 Experiments and Settings

Two supervised classification experiments were carried out to evaluate the performance of our proposed architecture. In the first experiment, we used the "100 unrelated subjects" dataset (54 females, age: 22-36) released by the HCP and aimed to identify them based on their rs-fMRI data. Each subject had four rs-fMRI sessions (each with 1200 volumes) scanned on two days [162]. The fMRI data were preprocessed by the HCP minimal preprocessing pipeline [52] and denoised by ICA-FIX [142] to remove spatial artifacts and motion-related fluctuations. Surface-based registration was performed with the MSM-ALL method [138]. We utilized 236 ROIs based on the Power's atlas [132]. The two sessions from Day 1 were employed as the training dataset, and the two sessions from Day 2 were used as the validation dataset and test dataset, respectively. The best model was chosen according to the best validation accuracy and final identification accuracy was assessed on the test dataset. For the training and validation datasets, fMRI time courses were cut into 100-frame clips. The final identification performance was measured with different numbers

of frames (from 1200 frames to single frame) on the test dataset. To evaluate the contribution of the connectivity-based neighborhood for convolutions, we also ascertained the performance of the same cGCN architecture with random graphs (random GCNs).

In the second experiment, we used cGCN to classify ASD patients from healthy controls on the ABIDE I [38] and ABIDE II [37] datasets. There are 1057 subjects (525 ASD subjects and 532 neurotypical controls) from 17 imaging sites in ABIDE I and 1053 subjects (470 ASD subjects and 583 neurotypical controls) from 17 imaging sites in ABIDE II. All fMRI data were preprocessed by the Connectome Computation System pipeline [186], including bandpass filtering (0.01–0.1 Hz) and without global signal regression. The Craddock 200 atlas [32] was utilized to extract ROI signals. We adopted both leave-one-site-out and 10-fold cross-validations. For the leave-one-site-out cross-validation, data from each site were independently tested with the model trained on data from other sites. The leave-one-site-out cross-validation evaluated the model on heterogeneous datasets considering the site-specific variation. The stratified 10-fold cross-validation mixed all data together and split them into different folds while keeping the proportions of sites and diagnostic groups across folds. The site-specific heterogeneity was overlooked by the random partition of subsamples.

We carried out the two experiments in Keras [30] using Tensorflow as the backend [1]. Adam optimizer was applied to update model weights with adaptive learning rates [89]. Stepwise learning rate decay was also used if the validation accuracy stopped increasing, with the smallest learning rate of 1e-6. The model was evaluated on the validation dataset after each training epoch and the model parameters were saved only if better validation

accuracy was achieved. During training, the L2 regularization was used to avoid overfitting. The final performance was reported with the highest accuracy among different L2 values (0.1, 0.01, 0.001, and 0.0001).

4.2.5 Visualization

The occlusion method [181] was utilized to visualize the important ROIs for classification in both experiments. With trained model parameters, each ROI of input data was zeroed out one at a time, and the performance degradation under the same model configuration was considered as the contribution of the ROI to the classification task. The occlusion pattern was mapped onto the cortical surface for visualization. We projected ROI data to the cortical surface (each represented as a 20-mm-radius sphere) and obtained the surface renditions. For the individual identification task, the occlusion pattern was generated and averaged on the test dataset. For ASD classification, we utilized the pre-trained models from the leave-one-site-out cross-validation of ABIDE I and averaged the pattern of performance degradation on each leave-out dataset to get the visualization pattern.

4.3 Results

4.3.1 Graph Property

We compared the graph property of the k-NN graph and the FC with threshold. The k values for the k-NN graphs were set to 3, 5, 10 and 20. For a fair comparison with similar number of edges in total, the threshold values were chosen as 0.235, 0.196, 0.145 and 0.095, respectively. Intuitively, k-NN graphs have relatively sparse distribution of edges

within the same functional community, while FCs with threshold tend to have dense blobs locally as shown in Figure 4.2. We use the modularity (Q) to measure the modular quality of a graph that is divided into communities. For k-NN graphs, increasing the k value from 3 to 20 will decrease the modularity from 0.84 to 0.59. For FCs with threshold, the modularity reduces from 0.67 to 0.58 when the threshold value changes from 0.235 to 0.095. On both type of graphs, increasing the number of edges will decrease the modularity of the graph. In terms of the degree distribution for the FC with threshold, most of nodes have the degree of lower than 20 when the threshold value is over 0.196. When the threshold value decreases to 0.095, the degree distribution becomes flatten and some nodes have the degree of up to 60. Decreasing the threshold value (from 0.235 to 0.095), the number of isolated nodes reduces from 116 to 13. And the number of subgraphs drops from 122 to 14. It also means that some nodes always have relatively low FC values with other nodes. In general, the average degree of nodes will increase when the threshold value decreases, which leads to a huge diversity of the degree for each node and isolated subgraphs.

4.3.2 Performance Related to the Number of Neighbors

The hyperparameter k determines the number of edges in graphs. The classification performance was evaluated for a range of k values (3, 5, 10, and 20). In Figure 4.3, we show the performance of individual identification. cGCN with k of 5 achieved the highest identification accuracy on average. Increasing the k value beyond 5 (k=10 or 20) led to diminished performance. The performance of random GCNs was significantly lower than cGCN models, especially for small k values (k=3 or 5).

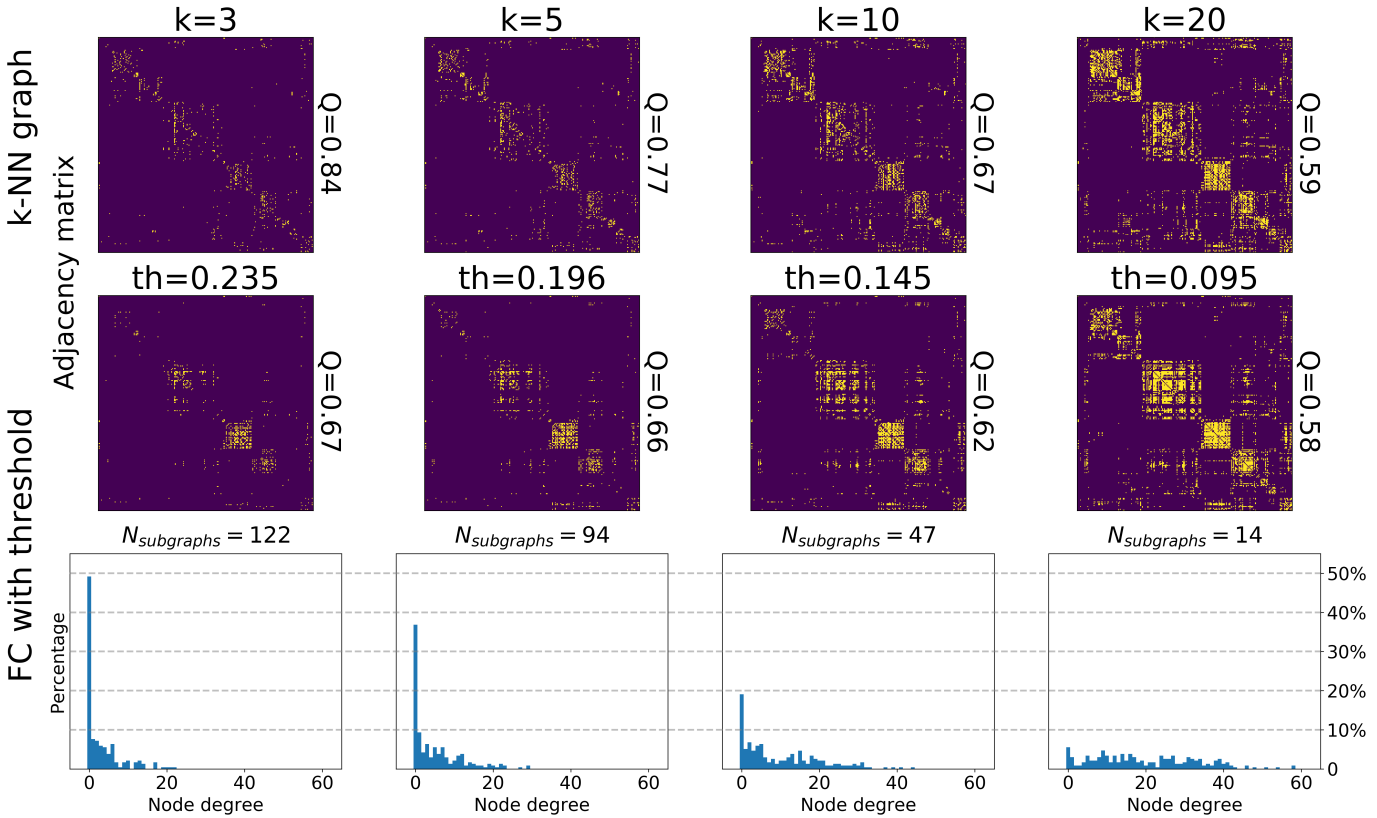


Figure 4.2: Comparison between k-NN graph and the FC with threshold. Based on the same group FC matrix, k-NN graphs in the first row are compared with FC matrices with different threshold values. The modularity value (Q) is calculated for each graph, which measures strength of division of a graph into communities. Each k-NN graph has better modularity than the corresponding FC matrix with threshold. For each k-NN graph, it is always a connected graph, and all nodes are connected with constant number of edges. For the FC matrix with threshold, we present the histograms of the node degree on the bottom. The number of subgraphs (disconnected graphs) for each FC matrix with threshold is also calculated.

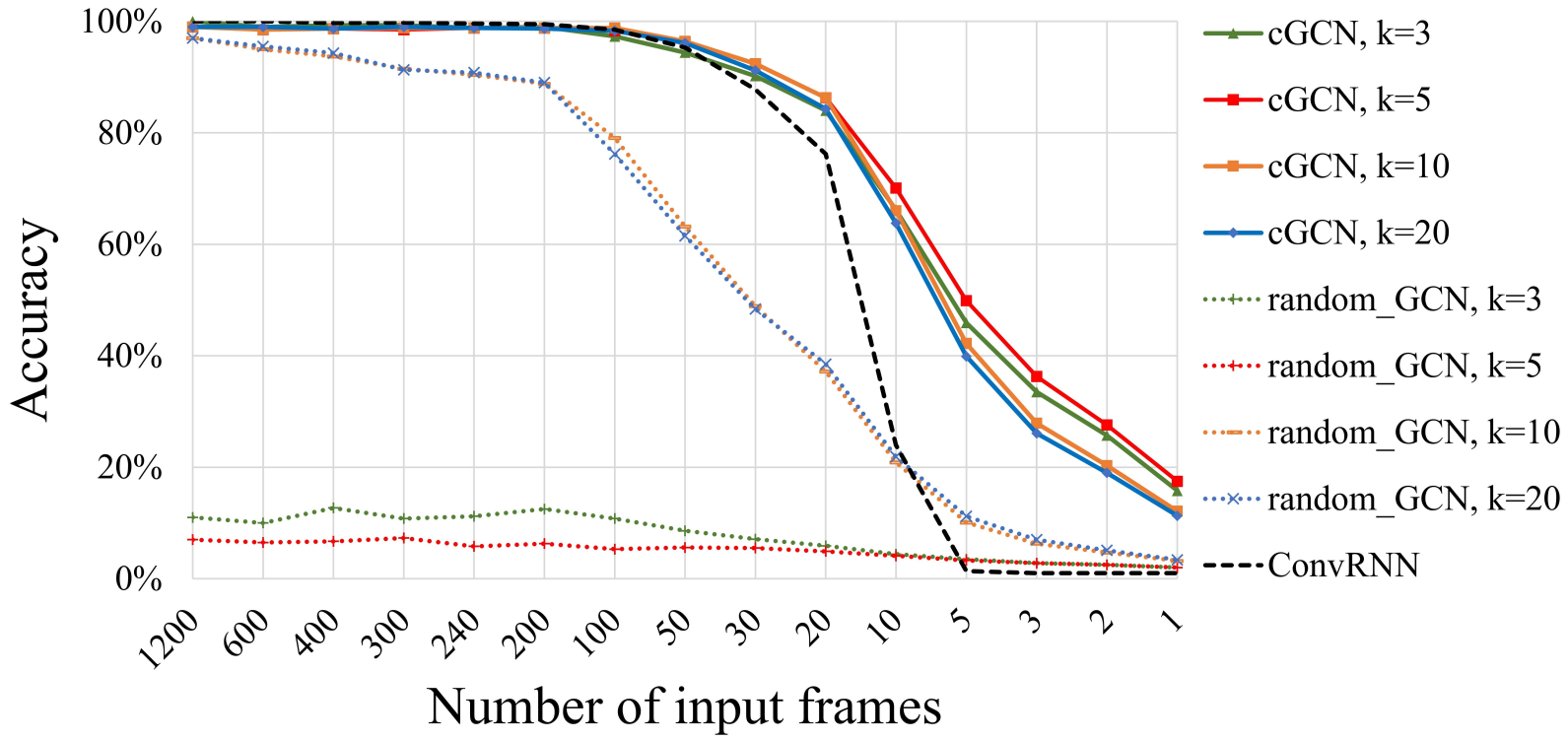


Figure 4.3: Performance on individual identification regarding the number of neighbors and the number of input frames. The highest mean classification accuracy was achieved with k=5. With 20 frames or less as input data, cGCN achieved significantly higher classification accuracy compared with random GCNs and the convolutional RNN (ConvRNN) model in our prior study [167]. Particularly with 5 frames, cGCNs obtained the identifying accuracy of over 49% with k=5, which was much higher than random GCNs and ConvRNN. It demonstrated that spatiotemporal features from each frame of fMRI data was successfully extracted by cGCNs for the individual identification task.

For ASD classification, classification results on ABIDE I and ABIDE II with leave-one-site-out and 10-fold cross validations are shown in Figure 4.4 and Figure 4.5. With leave-one-site-out cross validation, the highest mean accuracy was 71.6% (k=5) on ABIDE I and 72.1% (k=3) on ABIDE II, and the lowest mean accuracy was 70.1% (k=20) on ABIDE I and 69.6% (k=20) on ABIDE II. For the 10-fold cross-validation, the highest mean accuracy was 70.7% (k=3) on ABIDE I and 70.2% (k=5) on ABIDE II, and the lowest mean accuracy was 68.0% (k=20) on ABIDE I and 67.5% (k=20) on ABIDE II. On both cross-validations, increasing the k value for the cGCN decreased the classification accuracy.

4.3.3 Performance Related to the Number of Input Frames

For the individual identification application, we evaluated cGCNs and random GCNs with different number of frames as inputs. With cGCN models, the improved performance of individual identification was shown with increasing number of fMRI frames. As shown in Figure 4.3, cGCN achieved better performance with larger number of input frames and the performance gradually saturated at approximate 100 frames. Random GCNs exhibited a similar pattern for large k values (k=10 or 20), but with accuracy much lower than that of cGCNs. However, the performance of ConvRNN dropped dramatically with reducing frames, close to the random guess (1%) with less than 5 frames. In contrast, cGCN still obtained an identifying accuracy of >49% with 5 frames when k=5.

Similarly, for the ASD classification tasks on ABIDE I and ABIDE II shown in Figure 4.6 and Figure 4.7, better classification accuracy can be achieved with larger number

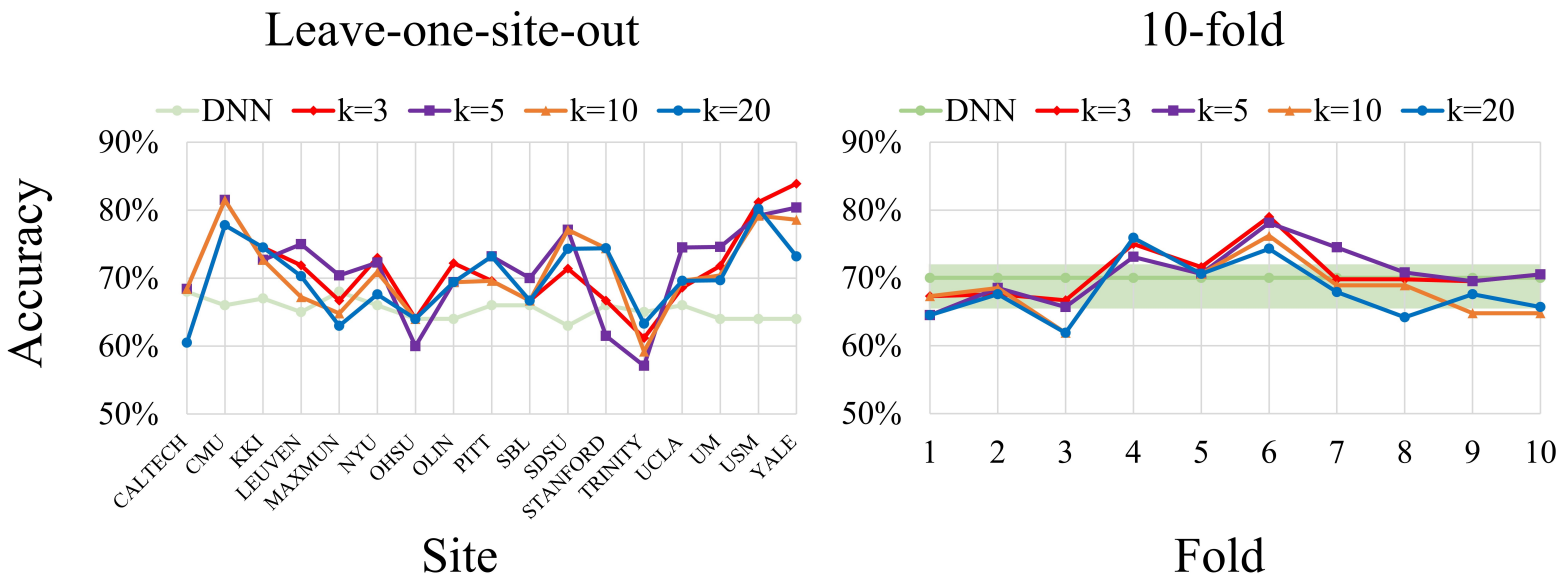


Figure 4.4: Performance on ABIDE I dataset with the leave-one-site-out cross-validation and the 10-fold cross-validation. The best average accuracy with the leave-one-site-out cross-validation was 71.6% (min: 57.1%, max: 81.5%) when k=5. As a comparison, the Deep Neural Network (DNN) model achieved mean accuracy of 65.4% (min: 63%, max: 68%) [69]. Except for some imaging sites (CALTECH, MAX_MUN, OHSU and TRINITY), cGCNs obtained distinct performance improvement compared to the previous DNN model. The best average accuracy with the 10-fold cross-validation was 70.7% (min: 66.7%, max: 79.0%) when k=3, which outperformed the DNN model of 70% (min: 66%, max: 71%).

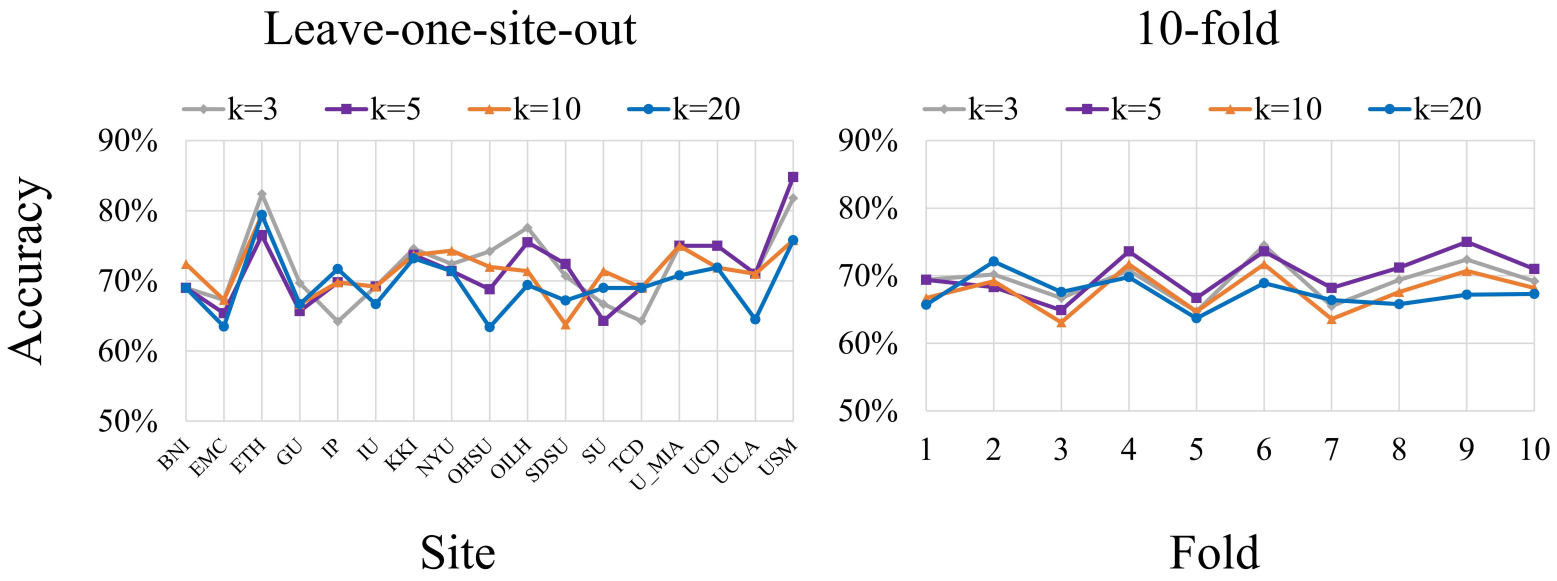


Figure 4.5: Performance on ABIDE II dataset with the leave-one-site-out cross-validation and the 10-fold cross-validation. The best average accuracy with the leave-one-site-out cross-validation was 72.1% (min: 64.2%, max: 82.4%) when k=3. The best average accuracy with the 10-fold cross-validation was 70.2% (min: 64.9%, max: 75.0%) when k=5.

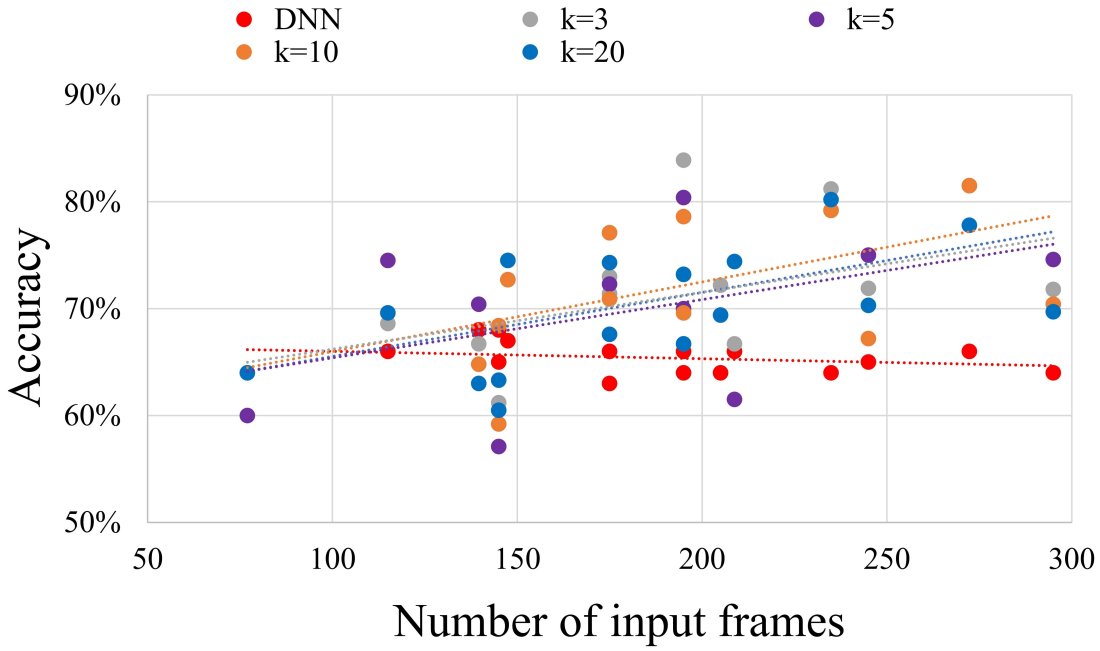


Figure 4.6: The relationship between the classification accuracy and the number of frames as inputs with the leave-one-site-out cross-validation on ABIDE I dataset. The average classification accuracy was proportional to the number of frames as inputs. While the number of frames will not affect the classification accuracy for the Deep Neural Network model [69] with FC matrices as inputs.

of input frames for different imaging sites in the leave-one-site-out cross-validation. As a comparison, the number of input frames did not significantly affect the classification accuracy of the DNN model, in which FC matrices were used as inputs without considering the temporal dimension [69].

4.3.4 Comparison

For the individual identification task, cGCN achieved an accuracy of 98.8% with k=10 when the number of input frames were fixed to 100. It outperformed several prior architectures, including the RNN model of 94.4% [23], the ConvRNN model of 98.5% [167],

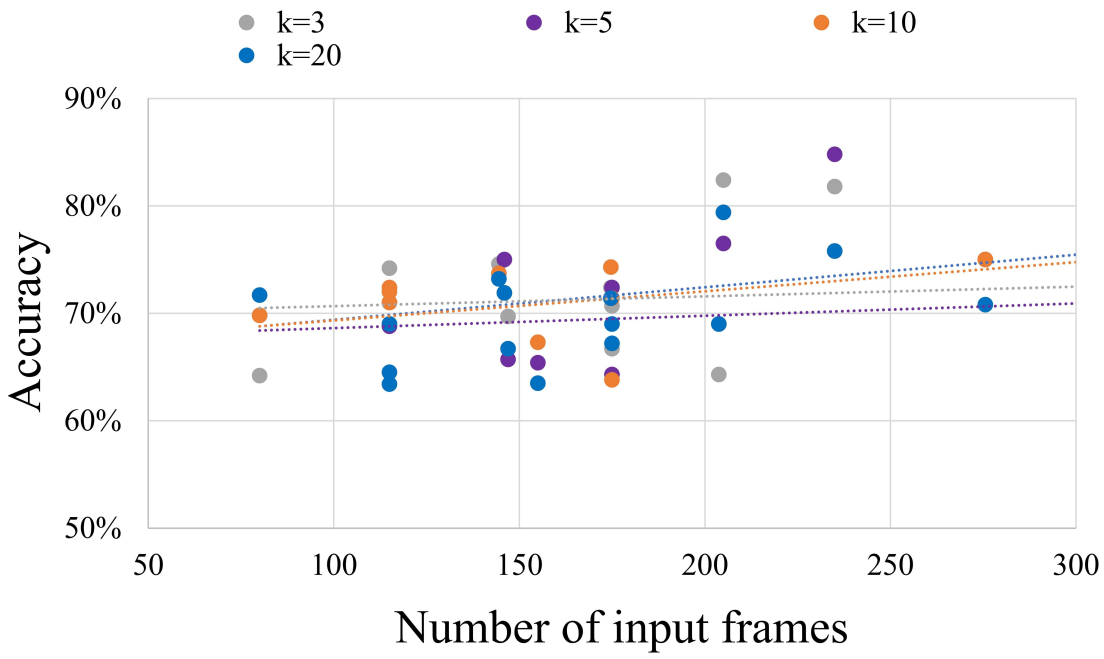


Figure 4.7: The relationship between the classification accuracy and the number of frames as inputs with the leave-one-site-out cross-validation on ABIDE II dataset. The average classification accuracy was proportional to the number of frames as inputs.

and the traditional correlation method of around 70% [47]. In particular, cGCNs showed overwhelming superiority over the traditional CNN model with fewer than 20 frames. As shown in Figure 4.3, the previous ConvRNN model with only 5 frames of input data led to an accuracy close to the random guess (1%).

For the ASD classification task with the leave-one-site-out cross-validation, the best classification accuracy was 71.6% (min: 57.1%, max: 81.5%) when k=5 as shown in Figure 4.4. Except for some imaging sites (CALTECH, MAX_MUN, OHSU and TRINITY), our cGCNs obtained significant performance improvement compared to the DNN model, whose accuracy was 65.4% on average [69]. With the 10-fold cross-validation, our best average performance was 70.7% (min: 66.7%, max: 79.0%) when k=3, which is better than the DNN model of 70% (min: 66%, max: 71%) [69] and the RNN model of 68.5% [43]. We summarized the classification performance on ABIDE dataset in Table 4.1.

4.3.5 Visualization

We applied the occlusion method to identify informative regions for two classification tasks. For the individual identification task in Figure 4.8, the smallest accuracy degradation was 0.6% when k=20 and the largest accuracy drop was 2.9% when k=3. Significant performance degradation was observed when some ROIs were individually occluded, while some regions did not suffer from any performance degradation. The salient regions for cGCN models with different k values were similar. The significant resting state networks were DMN, FPN, as well as VN.

Model	Number of subjects	Input data	Accuracy	
			Leave-one -site-out	k-fold (Chance accuracy)
RF [69]	1035	200 ROIs	-	63%
SVM [69]	1035	200 ROIs	-	65%
CNN [185]	200	whole-brain	-	70.5%
CNN [88]	774 (ABIDE I)	whole-brain	-	Best: 71.7% Ensemble: 73.3%
	393 (ABIDE II)		-	Best: 72.8% Ensemble: 71.7%
FCN [62]	110	116 ROIs	-	86.36%
FCN [69]	1035	200 ROIs	65.4%	70%
RNN [43]	1100	200 ROIs	-	68.5% (51.9%)
RNN [41]	1100	200 ROIs & Phenotypic data	-	70.1% (51.9%)
GCN [96]	871	110 ROIs	-	∼67%
GCN [122]	871	110 ROIs & Phenotypic data	-	69.5%
GCN [6]	872	110 ROIs & Phenotypic data	-	70.86%
Our cGCN	1057 (ABIDE I)	200 ROIs	71.6%	70.7% (50.5%)
	1053 (ABIDE II)		72.1%	70.2% (53.9%)

Table 4.1: Performance for the classification of Autism spectrum disorders with different deep learning models based on the Autism Brain Imaging Data Exchange (ABIDE) datasets. Some traditional machine learning methods including Random Forests (RF) and Support Vector Machine (SVM), and deep learning models including Convolutional Neural Networks (CNN), Fully-Connected Networks (FCN), Recurrent Neural Networks (RNN), Graph Convolutional Networks (GCN) are compared with our connectivity-based Graph Convolutional Network (cGCN).

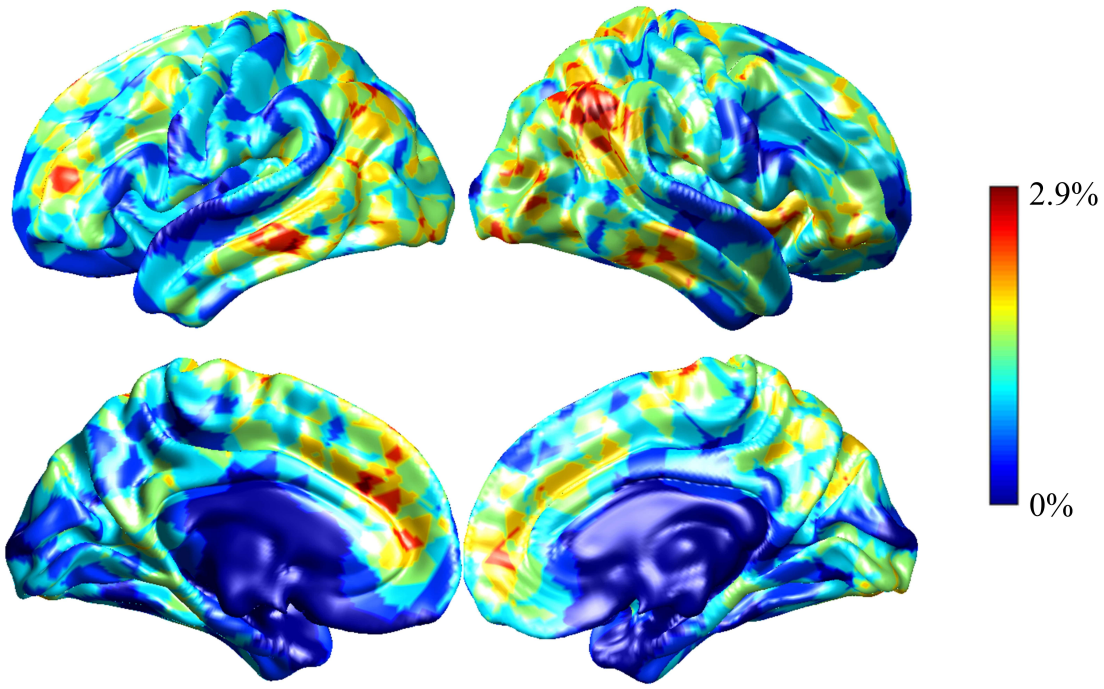


Figure 4.8: Visualization of the performance degradation through the single-ROI occlusion for the individual identification. Red region reflected large performance degradation if corresponding ROIs were occluded. The performance degradation reflects the relative contribution of each ROI. When k=20, the smallest performance drop was only 0.6%. While the largest performance drop of 2.9% was obtained with k=3. Default mode network, frontoparietal network and visual network contributed more to the individual identification.

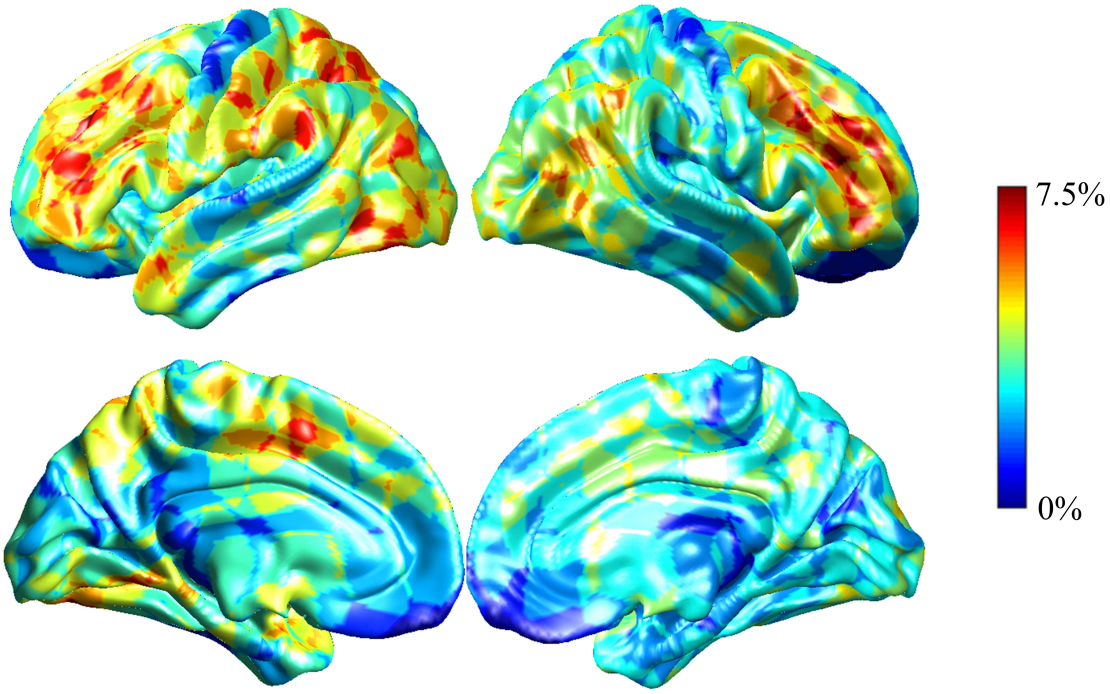


Figure 4.9: Visualization of the performance degradation through the single-ROI occlusion for the ASD classification. We averaged performance drop on all models used in the leave-one-site-out cross-validation. The maximum performance drop was 7.5% when k=3, and the smallest performance drop was 6.0% when k=20. The color of the region reflected the performance degradation if corresponding ROI was occluded. The salient regions related to the ASD classification included frontoparietal network, default mode network and ventral attention network.

For the ASD classification task in Figure 4.9, the largest performance drop was 7.5% when k=3, and the smallest performance drop was 6.0% when k=20. The salient regions for cGCN models with different k values were similar. The salient networks identified by cGCN included FPN, DMN and ventral attention network.

4.4 Discussion

We presented cGCN architecture for fMRI analysis and applied cGCN on two classification tasks with rs-fMRI data, i.e., the individual identification and classification of ASD, respectively. The superior performance compared to prior studies demonstrated that cGCN can effectively capture spatial features of fMRI data within connectomic neighbors.

A major contribution of the present work is that rather than performing convolution on structured image grids, ROI-based fMRI data are considered as graphs based on FC and cGCN is carried out between connectomic neighbors on the graph. Previous studies utilized fully-connected layers for the spatial feature extraction and ignored the brain's functional organization. In addition, it is difficult to train fully-connected models with satisfactory performance due to overfitting. Our prior paper demonstrated that it was feasible to capture spatial patterns on a small batch of ROIs from the same FC network by convolutions [167], although ROIs were artificially arranged in a pre-defined order according to the atlas. In the present work, the convolutional neighborhood was defined by the group-level FC matrix. Each ROI was involved in the convolution with its top connectivity-based neighbors, and high-level features were extracted by stacking convolutional layers hierarchically. The superior performance of cGCN over random GCN and previously used methods demonstrated that connectivity-based neighborhood for convolutions was a significant advantage for the cGCN architecture.

We adopted the k-NN graph rather than hard thresholding FC to reduce the size of the neighborhood and effects of noise in connectivity [104, 117]. The advantages of k-NN graphs lie in the following aspects. First, k-NN graphs naturally have local homogeneity

with the same number of edges originating from each node, appropriate for convolutions. Second, hierarchical feature extraction can be easily achieved with stacking layers under the guidance of the k-NN graph.

Like the convolutional kernel size in the traditional CNN models, k explicitly determines the convolutional region on graphs, as well as the computational complexity for the cGCN model. We tested the performance of our cGCN architecture with different k values. The best performance of the individual identification was achieved when k=5 while the highest classification accuracy on ASD with two types of cross-validations was achieved when k=3 or 5. For both tasks, increasing k did not improve the classification accuracy significantly. One possible reason is that convolution on too many neighbors might fail to generate local features with good generalization. The same situation also happened in the modeling of 3D objects [168]. Therefore, we suggest a k value of 5 or less for good performance, although the exact optimal value may be application dependent.

Unlike k, increasing the number of input frames improved the performance significantly for both tasks. This effect was especially obvious with a few input frames in the beginning and gradually leveled off with over 100 frames for individual identification. This result indicates that cGCN can efficiently extract spatial patterns between functionally 'adjacent' nodes from each frame of fMRI data in spite of substantial inter-subject and inter-session variability that arise due to hardware variations, and physiological variations [109]. Thus, it is beneficial to keep the temporal axis of the fMRI data and apply frame-wise feature extraction based on fMRI time course rather than the FC matrix.

There are some appealing properties by using the asymmetric edge function for convolutions according to [168]. First, the spatial feature of both global information x_i and local interactions $x_j - x_i$ is captured by the asymmetric edge function. Second, the long-distance characterization of spatial features between multi-hop neighborhood can be captured by the stacked convolutional structures, based on the topological structure of the graph.

Furthermore, the single-ROI occlusion experiments revealed the salient regions related to individual identification and classification of ASD tasks, which are in agreement with those seen in previous studies [47, 69, 119, 167]. For the individual identification task, the performance drop was much smaller than that of ConvRNN model [167], suggesting that the graph-based representation of fMRI data builds robust relationship between ROIs, and missing values of any ROI might be compensated by its functional neighbors based on the topological structure of graphs.

Some limitations should be noted for future work. First, the graph topology is invariant from layer to layer, frame to frame and subject to subject. Considering the significant variability in FC, future work shall adopt dynamic update across convolutional layers, time frames and subjects. Second, the graph in the present work only reflects the brain's organization from the functional perspective. Other image modalities, e.g., diffusion MRI, could be utilized to build graphs that reflect the brain's structural connectivity. Third, intuitive visualization of temporal features captured by cGCN is difficult. This warrants further investigation in future studies.

4.5 Conclusion

In this chapter, we describe a connectome defined neighborhood for graph convolution to extract connectomic features from rs-fMRI data for classification. Our model allows for spatial feature extraction within connectomic neighborhood rather than Euclidian ones. Significant improvement on individual identification and ASD classification tasks suggests that the cGCN model is effective in capturing connectomic features from fMRI data and is promising for fMRI analysis.

Chapter 5

Conclusion and Future Work

We have introduced two neural network architectures for fMRI analysis. Our models are RNN-based architectures to extract spatiotemporal features from fMRI sequence. Taking ROI data as inputs, we aim to extract spatial features related to functional networks based on BOLD signals from each frame of fMRI data. Our experiments demonstrated that significant performance improvement has been achieved for classification tasks. Meanwhile, interpretable visualization based on two models have provided an insightful understanding of fMRI data from a new perspective.

5.1 Similarities Between ConvRNN and cGCN

Relying on ROI-based fMRI data, it is efficient for both ConvRNN and cGCN to deploy large-scale analyses with hundreds of brain regions. Rather than performing feature extraction on the derived FC matrix, our neural network models take BOLD signals directly as inputs, and the time dimension has been explicitly kept to account for temporal

dynamics. Applying the RNN-based architecture leads to the spatial feature extraction from each frame of fMRI data and the temporal characterization from fMRI time course.

For the spatial feature extraction on fMRI data related to the functional connectivity, we assume that local interaction among connectomic brain regions reflects functional brain networks. Specifically, ConvRNN is built based on Power's atlas by putting connectomic brain regions together, in which a canonical division of ROIs into communities is adopted [14]. cGCN defines graph-represented fMRI data based on the group FC matrix, in which edges interconnect top connectomic brain regions by means of the k-NN graph. Both ConvRNN and cGCN define functionally connectomic neighborhood and acquire spatial features with the underlying configuration of brain networks.

Another similarity between ConvRNN and cGCN is that convolutions are used for the spatial feature extraction. Even though the current understanding of CNNs has been established based on the Euclidean data with structured grids (such as images and text data), the feasibility of convolutions on ROI-based fMRI data has not been fully demonstrated. Learning from the pairwise Pearson correlation and the FC matrix, both ConvRNN and cGCN have been built based on the assumption that local interactions within a small group of ROIs should exist, and can be determined by the convolutional operation. Although fully-connected layers can be applied without any assumption, the overfitting problem cannot be fully solved due to the inefficient computation. As a result, improved performance has been demonstrated with both ConvRNN and cGCN. The efficient computation of convolutional models guarantees that it is scalable to apply convolutions for large-scale fMRI analysis. Also, in-place visualization techniques based on convolutions also enable us to in-

terpret fMRI data along with classification tasks. The compatibility between convolutions and fMRI data can bring convenience when applying deep learning techniques for fMRI analysis.

5.2 Differences Between ConvRNN and cGCN

With the same fMRI time series as input data, the biggest difference between ConvRNN and cGCN is the internal data representation for each fMRI frame. ConvRNN has simply taken each frame of ROI data as a 1D array with the pre-defined order from the functional atlas. Even though the Power atlas manually groups connectomic ROIs into communities, the spatial pattern and the subsequent feature extraction are sensitive to the order of ROIs. The interpretability of extracted features is directly determined by the order of ROIs. As a comparison, cGCN considers each frame of ROI data as a graph by finding the top-connectivity neighborhood based on the group FC, which explicitly manifests interactions between two brain regions beyond the Euclidean distance. Therefore, the graph representation of fMRI data combines time-variant BOLD signals with the top-connectivity neighborhood based on the group FC matrix, which provides a powerful data representation beyond the Euclidean distance without any pre-defined order of brain regions.

In addition, ConvRNN applies traditional convolutions on ROI-based fMRI data in 1D arrays. Though high-level features are generated by stacking convolutional layers hierarchically, the modeling capacity of ConvRNN is largely confined due to the spatial configuration of 1D arrays. In contrast, the graph-represented fMRI data in cGCN defines convolutional neighbors for each ROI based on the functional connectivity, which is

equivalent to grid data with the closest neighbors taken from the structured grids. To be insensitive to the order of functional neighbors, spatial GCN has been applied with the asymmetric edge function, where both global information of ROIs and local interaction between connectomic neighbors have been taken into consideration. High-level features between multi-hop functional neighborhood based on the group FC matrix can be captured by cGCN.

5.3 Summary of Research Contributions

Our thesis introduces neural network models to perform classification tasks based on rs-fMRI data. Taking the functional configuration of the brain network into consideration, our models have achieved better performance by extracting spatiotemporal features regarding to the functional connectivity behind fMRI data. Both of our deep learning architectures employ the RNN-based architecture to generate spatiotemporal features for fMRI analysis, which is consistent with our previous study [24]. Instead of applying deep learning models with the FC matrix, we believe that spatial features should be separately extracted from each frame of fMRI data, and the dynamic features should be explored from fMRI time frames. With each time step being recurrently processed, temporal information across fMRI frames is processed by the model. Though the FC matrix depicts the spatial pattern behind the functional structure of the brain, it fails to keep the rich spatiotemporal features due to the collapsed temporal axis. The improved performance of ConvRNN and cGCN over traditional methods with FC matrices as inputs has been demonstrated.

For ConvRNN, ROI data have been arranged in 1D arrays with those ROIs from the same functional network grouped together. Spatial features related to the functional connectivity are obtained among connectomic regions. Considering the non-Euclidean property of the brain network without any pre-defined order of brain regions, the graph representation is applied for fMRI data in cGCN architecture. Specifically speaking, each frame of ROI data is considered as a graph, in which edges reflect the connectomic information between regions based on the group FC matrix. Considering the intractable complexity of the FC matrix with too many FC pairs, we assume that distinct FC-related features are derived between top connectomic regions in terms of the connectivity strength. So we further simplify the graph into a k-NN graph by keeping top FC pairs for each region and ignoring most of weak FC pairs in the original FC matrix. The linear computational complexity based on the k-NN graph can makes it scalable for large-scale analysis. Therefore, we present the graph representation for fMRI data that reflects the configuration of functional connectivity.

We also prove that convolutions are effective for the spatial feature extraction from ROI-based fMRI data, which is in accordance with main-stream deep learning techniques. For ROI data, fully-connected models have been widely applied for spatial feature extraction, considering the non-Euclidean property of the functional network. However, it is difficult to train fully-connected models properly to avoid overfitting. Our ConvRNN demonstrates that convolutions are feasible to acquire local features from ROI data with traditional data representation. The advantage of ConvRNN over fully-connected models has demonstrated convolutions are effective on ROI data. For cGCN architecture, graph

convolutions are employed on the graph-representation fMRI data, which is in line with our understanding of functional connectivity. And the improved performance has been realized with the cGCN architecture.

Along with the convolutional operation, the interpretable visualization has provided a new understanding of fMRI data along with classification tasks. Due to the in-place operation, intermediate outputs from convolutional layers of ConvRNN have been shown. Moreover, the occlusion method has been utilized on classification tasks with both ConvRNN and cGCN. The performance degradation with the occlusion is considered as an indirect measurement of the contribution of ROIs for classification tasks, which has obtained similar conclusions with previous studies. We believe visualization based on both ConvRNN and cGCN can help us to interpret fMRI data based on large-scale data-driven approaches.

5.4 Limitations and Future Work

Several limitations should be noted here for future work. First, the discriminating power and generalization of our neural network models on new data are still far from satisfactory results, although we have applied new deep learning architectures on individual identification and classification of ASD. One possible reason is that our models are not deep enough with limited modeling abilities to deal with the high variability behind fMRI data. It is probable that the model performance can be further improved by building more complex models to extract sophisticated features from fMRI data. Second, the current graph representation is proposed for each fMRI frame statically without explicitly incorporating

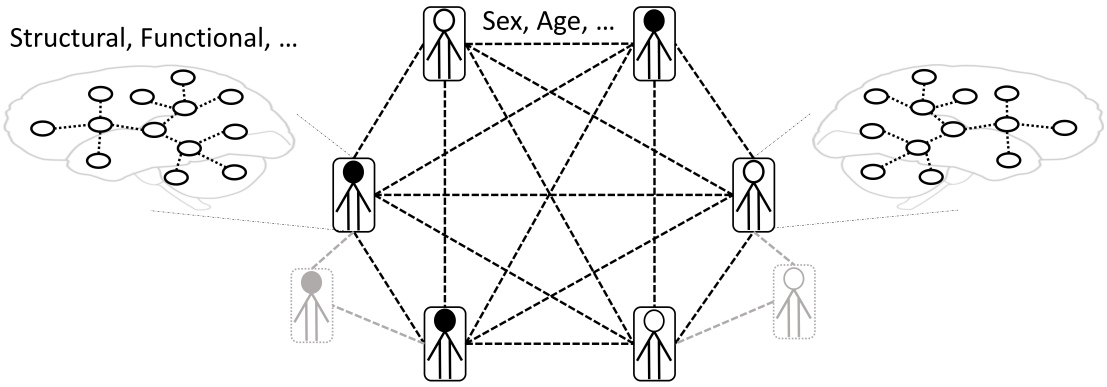


Figure 5.1: The heterogeneous graph representation for MRI classification tasks. The population graph includes nodes for different subjects and weighted edges for the similarity between two subjects based on non-image information. Each subject is also represented in a graph based on structural and functional data. Thus, the node-level classification task will be defined with the multi-level graph representation.

any temporal information of the brain networks in the graph representation. Inspired by dynamic graphs in [7, 49] and the node-level classification of the population graph [122], the graph construction for fMRI data should be revisited systematically. For example, as shown in Figure 5.1, the classification task can be configured into a heterogeneous graph, whose nodes are subjects represented in graphs based on structural and functional data. Edges in the population graph can be quantified by a combination of non-image information between subjects. Thus, a node-level classification task will be carried out on multi-level graph representations. Third, current ConvRNN and cGCN architectures have been applied without adequate optimization for fMRI analysis. The optimization of the architecture and hyperparameter tuning during the model training are of great necessity to improve classification accuracy.

Furthermore, data visualization is effective in understanding data and in interpreting neural network models intuitively. Based on deep learning models, we apply visualization approaches to determine informative regions related to the classification task, which is in agreement with previous studies. Even though current visualization methods are feasible and effective to implement, we should put more effort towards developing new visualization approaches in particular for fMRI data. Interpretable deep learning techniques are very promising to achieve accurate healthcare.

In addition to the above limitations, my current work only relies on fMRI data in single modality for classification tasks. The BOLD signals from fMRI data are considered as an indirect measurement of neural activity with great variability across sessions and subjects. The interpretation of fMRI data can not be comprehensive due to the inherent limitations of the fMRI technique. Therefore, multi-modal studies based on structural MRI data, fMRI data and phenotypic information should be integrated to have a better understanding of the brain from different aspects.

Bibliography

[1] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pages 265–283, 2016.

[2] Alexandre Abraham, Michael P. Milham, Adriana Di Martino, R. Cameron Craddock, Dimitris Samaras, Bertrand Thirion, and Gael Varoquaux. Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. 147:736–745, 2017.

[3] David Abramian and Anders Eklund. Generating fmri volumes from t1-weighted volumes using 3d cyclegan. *arXiv preprint arXiv:1907.08533*, 2019.

[4] Elena A Allen, Eswar Damaraju, Sergey M Plis, Erik B Erhardt, Tom Eichele, and Vince D Calhoun. Tracking whole-brain connectivity dynamics in the resting state. *Cerebral cortex*, 24(3):663–676, 2014.

[5] Anders H Andersen, Don M Gash, and Malcolm J Avison. Principal component analysis of the dynamic response measured by fmri: a generalized linear systems framework. *Magnetic Resonance Imaging*, 17(6):795–815, 1999.

[6] Rushil Anirudh and Jayaraman J Thiagarajan. Bootstrapping graph convolutional neural networks for autism spectrum disorder classification. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 3197–3201. IEEE, 2019.

[7] Tiago Azevedo, Luca Passamonti, Pietro Liò, and Nicola Toschi. Towards a predictive spatio-temporal representation of brain data. *arXiv preprint arXiv:2003.03290*, 2020.

[8] Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. Neural machine translation by jointly learning to align and translate. 2014.

[9] Deanna M Barch, Gregory C Burgess, Michael P Harms, Steven E Petersen, Bradley L Schlaggar, Maurizio Corbetta, Matthew F Glasser, Sandra Curtiss, Sachin Dixit, and Cindy Feldt. Function in the human connectome: Task-fMRI and individual differences in behavior. 80:169–189, 2013.

[10] Christoph Baur, Shadi Albarqouni, and Nassir Navab. Semi-supervised deep learning for fully convolutional networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 311–319. Springer, 2017.

[11] C.F. Beckmann and S.M. Smith. Tensorial extensions of independent component analysis for multisubject FMRI analysis. 25(1):294–311, 2005.

[12] Christian F Beckmann, Marilena DeLuca, Joseph T Devlin, and Stephen M Smith. Investigations into resting-state connectivity using independent component analysis. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1457):1001–1013, 2005.

[13] Yoshua Bengio, Patrice Simard, and Paolo Frasconi. Learning long-term dependencies with gradient descent is difficult. 5(2):157–166, 1994.

[14] Maxwell A. Bertolero, BT Thomas Yeo, Danielle S. Bassett, and Mark D'Esposito. A mechanistic model of connector hubs, modularity and cognition. 2(10):765, 2018.

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

[16] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. 2008(10):P10008, 2008.

[17] Joan Bruna, Wojciech Zaremba, Arthur Szlam, and Yann LeCun. Spectral networks and locally connected networks on graphs. 2013.

[18] Ed Bullmore and Olaf Sporns. Complex brain networks: Graph theoretical analysis of structural and functional systems. 10(3):186, 2009.

[19] Vince D Calhoun, Jing Sui, Kent Kiehl, Jessica A Turner, Elena A Allen, and Godfrey Pearlson. Exploring the psychosis functional connectome: aberrant intrinsic networks in schizophrenia and bipolar disorder. *Frontiers in psychiatry*, 2:75, 2012.

[20] Catie Chang and Gary H Glover. Time–frequency dynamics of resting-state brain connectivity measured with fmri. *Neuroimage*, 50(1):81–98, 2010.

[21] Colleen P Chen, Christopher L Keown, Afrooz Jahedi, Aarti Nair, Mark E Pflieger, Barbara A Bailey, and Ralph-Axel Müller. Diagnostic classification of intrinsic functional connectivity highlights somatosensory, default mode, and visual regions in autism. 8:238–245, 2015.

[22] Hao Chen, Qi Dou, Lequan Yu, and Pheng-Ann Heng. Voxresnet: Deep voxelwise residual networks for volumetric brain segmentation. *arXiv preprint arXiv:1608.05895*, 2016.

[23] Jianxu Chen, Lin Yang, Yizhe Zhang, Mark Alber, and Danny Z Chen. Combining fully convolutional and recurrent neural networks for 3d biomedical image segmentation. pages 3036–3044, 2016.

[24] Shiyang Chen and Xiaoping Hu. Individual Identification Using the Functional Brain Fingerprint Detected by the Recurrent Neural Network. 8(4):197–204, 2018.

[25] Shiyang Chen, Jason Langley, Xiangchuan Chen, and Xiaoping Hu. Spatiotemporal modeling of brain dynamics using resting-state functional magnetic resonance imaging with gaussian hidden markov model. *Brain connectivity*, 6(4):326–334, 2016.

[26] Jieyu Cheng, Adrian Dalca, Bruce Fischl, and Lilla Zollei. Cortical surface registration using unsupervised learning. *arXiv preprint arXiv:2004.04617*, 2020.

[27] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using RNN encoder-decoder for statistical machine translation. 2014.

[28] Hongyoon Choi. Functional connectivity patterns of autism spectrum disorder identified by deep feature learning. *arXiv preprint arXiv:1707.07932*, 2017.

[29] Hongyoon Choi. Functional connectivity patterns of autism spectrum disorder identified by deep feature learning. 2017.

[30] François Chollet. Keras, 2015.

[31] Cameron Craddock, Yassine Benhajali, Carlton Chu, Francois Chouinard, Alan Evans, András Jakab, Budhachandra Singh Khundrakpam, John David Lewis, Qingyang Li, and Michael Milham. The Neuro Bureau Preprocessing Initiative: Open sharing of preprocessed neuroimaging data and derivatives. 4, 2013.

[32] R. Cameron Craddock, G.Andrew James, Paul E. Holtzheimer III, Xiaoping P. Hu, and Helen S. Mayberg. A whole brain fMRI atlas generated via spatially constrained spectral clustering. 33(8):1914–1928, 2012.

[33] Antonia Creswell, Tom White, Vincent Dumoulin, Kai Arulkumaran, Biswa Sengupta, and Anil A Bharath. Generative adversarial networks: An overview. 35(1):53–65, 2018.

[34] Anders M Dale, Bruce Fischl, and Martin I Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. *Neuroimage*, 9(2):179–194, 1999.

[35] Federico De Martino, Giancarlo Valente, Noël Staeren, John Ashburner, Rainer Goebel, and Elia Formisano. Combining multivariate voxel selection and support vector machines for mapping and classification of fmri spatial patterns. *Neuroimage*, 43(1):44–58, 2008.

[36] J Demmel. Cs 267: Notes for lecture 23, april 9, 1999. graph partitioning, part 2, 1999.

[37] Adriana Di Martino, David O'connor, Bosi Chen, Kaat Alaerts, Jeffrey S Anderson, Michal Assaf, Joshua H Balsters, Leslie Baxter, Anita Beggiato, Sylvie Bernaerts, et al. Enhancing studies of the connectome in autism using the autism brain imaging data exchange ii. *Scientific data*, 4(1):1–15, 2017.

[38] Adriana Di Martino, Chao-Gan Yan, Qingyang Li, Erin Denio, Francisco X Castellanos, Kaat Alaerts, Jeffrey S Anderson, Michal Assaf, Susan Y Bookheimer, Mirella Dapretto, et al. The autism brain imaging data exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. 19(6):659, 2014.

[39] Jose Dolz, Christian Desrosiers, and Ismail Ben Ayed. 3d fully convolutional networks for subcortical segmentation in mri: A large-scale study. *NeuroImage*, 170:456–470, 2018.

[40] Jose Dolz, Christian Desrosiers, and Ismail Ben Ayed. 3D fully convolutional networks for subcortical segmentation in MRI: A large-scale study. 170:456–470, 2018.

[41] Nicha C Dvornek, Pamela Ventola, and James S Duncan. Combining phenotypic and resting-state fmri data for autism classification with recurrent neural networks. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 725–728. IEEE, 2018.

[42] Nicha C Dvornek, Pamela Ventola, and James S Duncan. Combining phenotypic and resting-state fMRI data for autism classification with recurrent neural networks. pages 725–728. IEEE, 2018.

[43] Nicha C. Dvornek, Pamela Ventola, Kevin A. Pelphrey, and James S. Duncan. Identifying autism from resting-state fMRI using long short-term memory networks. In *International Workshop on Machine Learning in Medical Imaging*, pages 362–370. Springer, 2017.

[44] Christoph Feichtenhofer, Axel Pinz, and Andrew Zisserman. Convolutional two-stream network fusion for video action recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1933–1941, 2016.

[45] Matthias Fey and Jan E. Lenssen. Fast Graph Representation Learning with PyTorch Geometric. In *ICLR Workshop on Representation Learning on Graphs and Manifolds*, 2019.

[46] Emily S Finn, Dustin Scheinost, Daniel M Finn, Xilin Shen, Xenophon Papademetris, and R Todd Constable. Can brain state be manipulated to emphasize individual differences in functional connectivity? 2017.

[47] Emily S. Finn, Xilin Shen, Dustin Scheinost, Monica D. Rosenberg, Jessica Huang, Marvin M. Chun, Xenophon Papademetris, and R. Todd Constable. Functional connectome fingerprinting: Identifying individuals using patterns of brain connectivity. 18(11):1664, 2015.

[48] Michael D Fox and Marcus E Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature reviews neuroscience*, 8(9):700–711, 2007.

[49] Soham Gadgil, Qingyu Zhao, Ehsan Adeli, Adolf Pfefferbaum, Edith V Sullivan, and Kilian M Pohl. Spatio-temporal graph convolution for functional mri analysis. *arXiv preprint arXiv:2003.10613*, 2020.

[50] Yang Gao, Jeff M. Phillips, Yan Zheng, Renqiang Min, P. Thomas Fletcher, and Guido Gerig. Fully convolutional structured LSTM networks for joint 4D medical image segmentation. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1104–1108. IEEE, 2018.

[51] Justin Gilmer, Samuel S. Schoenholz, Patrick F. Riley, Oriol Vinyals, and George E. Dahl. Neural message passing for quantum chemistry. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1263–1272. JMLR. org, 2017.

[52] Matthew F Glasser, Stamatios N Sotiropoulos, J Anthony Wilson, Timothy S Coalson, Bruce Fischl, Jesper L Andersson, Junqian Xu, Saad Jbabdi, Matthew Webster, and Jonathan R Polimeni. The minimal preprocessing pipelines for the Human Connectome Project. 80:105–124, 2013.

[53] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. pages 249–256, 2010.

[54] Karthik Gopinath, Christian Desrosiers, and Herve Lombaert. Graph convolutions on spectral embeddings for cortical surface parcellation. *Medical image analysis*, 54:297–305, 2019.

[55] Karthik Gopinath, Christian Desrosiers, and Herve Lombaert. Graph convolutions on spectral embeddings for cortical surface parcellation. 54:297–305, 2019.

[56] Abigail S Greene, Siyuan Gao, Dustin Scheinost, and R Todd Constable. Task-induced brain state manipulation improves prediction of individual traits. 9(1):2807, 2018.

[57] Hayit Greenspan, Bram Van Ginneken, and Ronald M Summers. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 35(5):1153–1159, 2016.

[58] Michael Greicius. Resting-state functional connectivity in neuropsychiatric disorders. 21(4):424–430, 2008.

[59] Michael D Greicius, Kaustubh Supekar, Vinod Menon, and Robert F Dougherty. Resting-state functional connectivity reflects structural connectivity in the default mode network. 19(1):72–78, 2009.

[60] Ludovica Griffanti, Gholamreza Salimi-Khorshidi, Christian F Beckmann, Edward J Auerbach, Gwenaëlle Douaud, Claire E Sexton, Enikő Zsoldos, Klaus P Ebmeier, Nicola Filippini, and Clare E Mackay. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. 95:232–247, 2014.

[61] Umut Güçlü and Marcel AJ van Gerven. Modeling the dynamics of human brain activity with recurrent neural networks. 11:7, 2017.

[62] Xinyu Guo, Kelli C Dominick, Ali A Minai, Hailong Li, Craig A Erickson, and Long J Lu. Diagnosing autism spectrum disorder from brain resting-state functional connectivity patterns using a deep neural network with a novel feature selection method. *Frontiers in neuroscience*, 11:460, 2017.

[63] Xinyu Guo, Kelli C Dominick, Ali A Minai, Hailong Li, Craig A Erickson, and Long J Lu. Diagnosing autism spectrum disorder from brain resting-state functional connectivity patterns using a deep neural network with a novel feature selection method. 11:460, 2017.

[64] Aric Hagberg, Pieter Swart, and Daniel S Chult. Exploring network structure, dynamics, and function using networkx. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

[65] Will Hamilton, Zhitao Ying, and Jure Leskovec. Inductive representation learning on large graphs. In *Advances in Neural Information Processing Systems*, pages 1024–1034, 2017.

[66] Michelle Hampson, Bradley S Peterson, Pawel Skudlarski, James C Gatenby, and John C Gore. Detection of functional connectivity using temporal correlations in mr images. *Human brain mapping*, 15(4):247–262, 2002.

[67] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[68] Yong He and Alan Evans. Graph theoretical modeling of brain connectivity. 23(4):341–350, 2010.

[69] Anibal Sólon Heinsfeld, Alexandre Rosa Franco, R. Cameron Craddock, Augusto Buchweitz, and Felipe Meneguzzi. Identification of autism spectrum disorder using deep learning and the ABIDE dataset. 17:16–23, 2018.

[70] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. 9(8):1735–1780, 1997.

[71] Avram J Holmes, Marisa O Hollinshead, Timothy M O'Keefe, Victor I Petrov, Gabriele R Fariello, Lawrence L Wald, Bruce Fischl, Bruce R Rosen, Ross W Mair, and Joshua L Roffman. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. 2:150031, 2011.

[72] CJ Honey, O Sporns, Leila Cammoun, Xavier Gigandet, Jean-Philippe Thiran, Reto Meuli, and Patric Hagmann. Predicting human resting-state functional connectivity from structural connectivity. 106(6):2035–2040, 2009.

[73] Yoonmi Hong, Jaeil Kim, Geng Chen, Weili Lin, Pew-Thian Yap, and Dinggang Shen. Longitudinal prediction of infant diffusion MRI data via graph convolutional adversarial networks. 38(12):2717–2725, 2019.

[74] Ehsan Hosseini-Asl, Georgy Gimel'farb, and Ayman El-Baz. Alzheimer's disease diagnostics by a deeply supervised adaptable 3d convolutional network. *arXiv preprint arXiv:1607.00556*, 2016.

[75] Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger. Densely Connected Convolutional Networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017-07.

[76] Tzu-Wei Huang, Hwann-Tzong Chen, Ryuichi Fujimoto, Koichi Ito, Kai Wu, Kazunori Sato, Yasuyuki Taki, Hiroshi Fukuda, and Takafumi Aoki. Age estimation from brain mri images using deep learning. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 849–852. IEEE, 2017.

[77] R Matthew Hutchison, Thilo Womelsdorf, Elena A Allen, Peter A Bandettini, Vince D Calhoun, Maurizio Corbetta, Stefania Della Penna, Jeff H Duyn, Gary H Glover, Javier Gonzalez-Castillo, et al. Dynamic functional connectivity: Promise, issues, and interpretations. 80:360–378, 2013.

[78] Chang Min Hyun, Hwa Pyung Kim, Sung Min Lee, Sungchul Lee, and Jin Keun Seo. Deep learning for undersampled MRI reconstruction. 63(13):135007, 2018.

[79] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. 2015.

[80] Sarah Itani and Dorina Thanou. Combining anatomical and functional networks for neuropathology identification: A case study on autism spectrum disorder. *arXiv preprint arXiv:1904.11296*, 2019.

[81] Konstantinos Kamnitsas, Christian Ledig, Virginia FJ Newcombe, Joanna P Simpson, Andrew D Kane, David K Menon, Daniel Rueckert, and Ben Glocker. Efficient multi-scale 3d cnn with fully connected crf for accurate brain lesion segmentation. *Medical image analysis*, 36:61–78, 2017.

[82] Konstantinos Kamnitsas, Christian Ledig, Virginia FJ Newcombe, Joanna P Simpson, Andrew D Kane, David K Menon, Daniel Rueckert, and Ben Glocker. Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. 36:61–78, 2017.

[83] Jeremy Kawahara, Colin J. Brown, Steven P. Miller, Brian G. Booth, Vann Chau, Ruth E. Grunau, Jill G. Zwicker, and Ghassan Hamarneh. BrainNetCNN: Convolutional neural networks for brain networks; towards predicting neurodevelopment. 146:1038–1049, 2017.

[84] Amirali Kazeminejad and Roberto C Sotero. Topological Properties of Resting-State fMRI Functional Networks Improve Machine Learning-Based Autism Classification. 12, 2018.

[85] Christopher L Keown, Michael C Datko, Colleen P Chen, Jose Omar Maximo, Afrooz Jahedi, and Ralph-Axel Müller. Network organization is globally atypical in autism: A graph theory study of intrinsic functional connectivity. 2(1):66–75, 2017.

[86] Ali Khazaee, Ata Ebrahimzadeh, and Abbas Babajani-Feremi. Identifying patients with alzheimer's disease using resting-state fmri and graph theory. *Clinical Neurophysiology*, 126(11):2132–2141, 2015.

[87] Meenakshi Khosla, Keith Jamison, Amy Kuceyeski, and Mert Sabuncu. 3D Convolutional Neural Networks for Classification of Functional Connectomes. 2018.

[88] Meenakshi Khosla, Keith Jamison, Amy Kuceyeski, and Mert R Sabuncu. 3d convolutional neural networks for classification of functional connectomes. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 137–145. Springer, 2018.

[89] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. 2014.

[90] Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. 2016.

[91] Yu Kong and Yun Fu. Human Action Recognition and Prediction: A Survey. 2018.

[92] Sergey Korolev, Amir Safiullin, Mikhail Belyaev, and Yulia Dodonova. Residual and plain convolutional neural networks for 3d brain mri classification. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 835–838. IEEE, 2017.

[93] Sergey Korolev, Amir Safiullin, Mikhail Belyaev, and Yulia Dodonova. Residual and plain convolutional neural networks for 3d brain mri classification. pages 835–838. IEEE, 2017.

[94] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. pages 1097–1105, 2012.

[95] Sofia Ira Ktena, Sarah Parisot, Enzo Ferrante, Martin Rajchl, Matthew Lee, Ben Glocker, and Daniel Rueckert. Distance metric learning using graph convolutional networks: Application to functional brain networks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 469–477. Springer, 2017.

[96] Sofia Ira Ktena, Sarah Parisot, Enzo Ferrante, Martin Rajchl, Matthew Lee, Ben Glocker, and Daniel Rueckert. Metric learning with spectral graph convolutions on brain connectivity networks. *NeuroImage*, 169:431–442, 2018.

[97] Stephen LaConte, Stephen Strother, Vladimir Cherkassky, Jon Anderson, and Xiaoping Hu. Support vector machines for temporal classification of block design fmri data. *NeuroImage*, 26(2):317–329, 2005.

[98] Megan H Lee, Christopher D Smyser, and Joshua S Shimony. Resting-state fmri: a review of methods and clinical applications. *American Journal of neuroradiology*, 34(10):1866–1872, 2013.

[99] M.H. Lee, C.D. Smyser, and J.S. Shimony. Resting-State fMRI: A Review of Methods and Clinical Applications. 34(10):1866–1872, 2013.

[100] Wenye Li and Dale Schuurmans. Modular community detection in networks. volume 22, page 1366, 2011.

[101] Xiaoxiao Li, Nicha C Dvornek, Xenophon Papademetris, Juntang Zhuang, Lawrence H Staib, Pamela Ventola, and James S Duncan. 2-channel convolutional 3d deep neural network (2cc3d) for fmri analysis: Asd classification and feature learning. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 1252–1255. IEEE, 2018.

[102] Jin Liu, Yi Pan, Min Li, Ziyue Chen, Lu Tang, Chengqian Lu, and Jianxin Wang. Applications of deep learning to mri images: A survey. *Big Data Mining and Analytics*, 1(1):1–18, 2018.

[103] Siqi Liu, Sidong Liu, Weidong Cai, Sonia Pujol, Ron Kikinis, and Dagan Feng. Early diagnosis of Alzheimer's disease with deep learning. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, pages 1015–1018. IEEE, 2014.

[104] Thomas T Liu, Alican Nalci, and Maryam Falahpour. The global signal in fMRI: Nuisance or information? 150:213–229, 2017.

[105] Xiao Liu and Jeff H Duyn. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proceedings of the National Academy of Sciences*, 110(11):4392–4397, 2013.

[106] Alexander Selvikvåg Lundervold and Arvid Lundervold. An overview of deep learning in medical imaging focusing on mri. *Zeitschrift für Medizinische Physik*, 29(2):102–127, 2019.

[107] Waqas Majeed, Matthew Magnuson, Wendy Hasenkamp, Hillary Schwarb, Eric H Schumacher, Lawrence Barsalou, and Shella D Keilholz. Spatiotemporal dynamics of low frequency bold fluctuations in rats and humans. *Neuroimage*, 54(2):1140–1150, 2011.

[108] Andrew R Mayer, David Ruhl, Flannery Merideth, Josef Ling, Faith M Hanlon, Juan Bustillo, and Jose Cañive. Functional imaging of the hemodynamic sensory gating response in schizophrenia. *Human brain mapping*, 34(9):2302–2312, 2013.

[109] David J McGonigle, Alistair M Howseman, Balwindar S Athwal, Karl J Friston, RSJ Frackowiak, and Andrew P Holmes. Variability in fMRI: An examination of intersession differences. 11(6):708–734, 2000.

[110] Martin J. McKeown and Terrence J. Sejnowski. Independent component analysis of fMRI data: Examining the assumptions. 6(5):368–372, 1998.

[111] Karla L. Miller, Fidel Alfaro-Almagro, Neal K. Bangerter, David L. Thomas, Essa Yacoub, Junqian Xu, Andreas J. Bartsch, Saad Jbabdi, Stamatios N. Sotiropoulos, and Jesper LR Andersson. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. 19(11):1523, 2016.

[112] Volodymyr Mnih, Nicolas Heess, and Alex Graves. Recurrent models of visual attention. pages 2204–2212, 2014.

[113] Peter NC Mohr and Irene E Nagel. Variability in brain activity as an individual difference measure in neuroscience. 30(23):7755–7757, 2010.

[114] Federico Monti, Davide Boscaini, Jonathan Masci, Emanuele Rodola, Jan Svoboda, and Michael M. Bronstein. Geometric deep learning on graphs and manifolds using mixture model cnns. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 5115–5124, 2017.

[115] Sophia Mueller, Danhong Wang, Michael D. Fox, BT Thomas Yeo, Jorge Sepulcre, Mert R. Sabuncu, Rebecca Shafee, Jie Lu, and Hesheng Liu. Individual variability in functional connectivity architecture of the human brain. 77(3):586–595, 2013.

[116] Brent C Munsell, Chong-Yaw Wee, Simon S Keller, Bernd Weber, Christian Elger, Laura Angelica Tomaz da Silva, Travis Nesland, Martin Styner, Dinggang Shen, and Leonardo Bonilha. Evaluation of machine learning algorithms for treatment outcome prediction in patients with epilepsy based on structural connectome data. *Neuroimage*, 118:219–230, 2015.

[117] Kevin Murphy and Michael D Fox. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. 154:169–173, 2017.

[118] Mark EJ Newman. Modularity and community structure in networks. 103(23):8577–8582, 2006.

[119] Jared A Nielsen, Brandon A Zielinski, P Thomas Fletcher, Andrew L Alexander, Nicholas Lange, Erin D Bigler, Janet E Lainhart, and Jeffrey S Anderson. Multisite functional connectivity mri classification of autism: Abide results. *Frontiers in human neuroscience*, 7:599, 2013.

[120] Heloisa Onias, Aline Viol, Fernanda Palhano-Fontes, Katia C Andrade, Marcio Sturzbecher, Gandhimohan Viswanathan, and Draulio B de Araujo. Brain complex network analysis by means of resting state fMRI and graph analysis: Will it be helpful in clinical epilepsy? 38:71–80, 2014.

[121] Graziella Orru, William Pettersson-Yeo, Andre F Marquand, Giuseppe Sartori, and Andrea Mechelli. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neuroscience & Biobehavioral Reviews*, 36(4):1140–1152, 2012.

[122] Sarah Parisot, Sofia Ira Ktena, Enzo Ferrante, Matthew Lee, Ricardo Guerrerro Moreno, Ben Glocker, and Daniel Rueckert. Spectral graph convolutions for population-based disease prediction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 177–185. Springer, 2017.

[123] Harshit S Parmar, Brian Nutter, Rodney Long, Sameer Antani, and Sunanda Mitra. Deep learning of volumetric 3d cnn for fmri in alzheimer's disease classification. In *Medical Imaging 2020: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 11317, page 113170C. International Society for Optics and Photonics, 2020.

[124] Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. On the difficulty of training recurrent neural networks. pages 1310–1318, 2013.

[125] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in PyTorch. 2017.

[126] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 701–710. ACM, 2014.

[127] Steven E Petersen and Olaf Sporns. Brain networks and cognitive architectures. *Neuron*, 88(1):207–219, 2015.

[128] Mark Plitt, Kelly Anne Barnes, and Alex Martin. Functional connectivity classification of autism identifies highly predictive brain features but falls short of biomarker standards. 7:359–366, 2015.

[129] Russell A Poldrack. Region of interest analysis for fmri. *Social cognitive and affective neuroscience*, 2(1):67–70, 2007.

[130] Russell A. Poldrack. Region of interest analysis for fMRI. 2(1):67–70, 2007.

[131] Sidney Pontes-Filho, Annelene Gulden Dahl, et al. A deep learning based tool for automatic brain extraction from functional magnetic resonance images in rodents. *arXiv preprint arXiv:1912.01359*, 2019.

[132] Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, and Bradley L Schlaggar. Functional network organization of the human brain. 72(4):665–678, 2011.

[133] Maria Giulia Preti, Thomas AW Bolton, and Dimitri Van De Ville. The dynamic functional connectome: State-of-the-art and perspectives. 160:41–54, 2017.

[134] Chen Qin, Jo Schlemper, Jose Caballero, Anthony N Price, Joseph V Hajnal, and Daniel Rueckert. Convolutional recurrent neural networks for dynamic mr image reconstruction. *IEEE transactions on medical imaging*, 38(1):280–290, 2018.

[135] Muhammad Imran Razzak, Saeeda Naz, and Ahmad Zaib. Deep learning for medical image processing: Overview, challenges and the future. In *Classification in BioApps*, pages 323–350. Springer, 2018.

[136] Ronald A Rensink. The dynamic representation of scenes. 7(1):17–42, 2000.

[137] Jonas Richiardi, Hamdi Eryilmaz, Sophie Schwartz, Patrik Vuilleumier, and Dimitri Van De Ville. Decoding brain states from fmri connectivity graphs. *Neuroimage*, 56(2):616–626, 2011.

[138] Emma C Robinson, Saad Jbabdi, Matthew F Glasser, Jesper Andersson, Gregory C Burgess, Michael P Harms, Stephen M Smith, David C Van Essen, and Mark Jenkinson. MSM: A new flexible framework for Multimodal Surface Matching. 100:414–426, 2014.

[139] Baxter P. Rogers, Victoria L. Morgan, Allen T. Newton, and John C. Gore. Assessing functional connectivity in the human brain by fMRI. 25(10):1347–1357, 2007.

[140] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241, 2015.

[141] Ünal Sakoğlu, Godfrey D Pearlson, Kent A Kiehl, Y Michelle Wang, Andrew M Michael, and Vince D Calhoun. A method for evaluating dynamic functional network connectivity and task-modulation: application to schizophrenia. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 23(5-6):351–366, 2010.

[142] Gholamreza Salimi-Khorshidi, Gwenaëlle Douaud, Christian F Beckmann, Matthew F Glasser, Ludovica Griffanti, and Stephen M Smith. Automatic denoising of functional MRI data: Combining independent component analysis and hierarchical fusion of classifiers. 90:449–468, 2014.

[143] Saman Sarraf and Ghassem Tofighi. DeepAD: Alzheimer′ s disease classification via deep convolutional neural networks using MRI and fMRI. page 070441, 2016.

[144] Andrew M Saxe, James L McClelland, and Surya Ganguli. Exact solutions to the nonlinear dynamics of learning in deep linear neural networks. 2013.

[145] William W Seeley, Vinod Menon, Alan F Schatzberg, Jennifer Keller, Gary H Glover, Heather Kenna, Allan L Reiss, and Michael D Greicius. Dissociable intrinsic connectivity networks for salience processing and executive control. *Journal of Neuroscience*, 27(9):2349–2356, 2007.

[146] Si-Baek Seong, Chongwon Pae, and Hae-Jeong Park. Geometric convolutional neural network for analyzing surface-based neuroimaging data. *Frontiers in neuroinformatics*, 12:42, 2018.

[147] Shikhar Sharma, Ryan Kiros, and Ruslan Salakhutdinov. Action recognition using visual attention. 2015.

[148] Dinggang Shen, Guorong Wu, and Heung-Il Suk. Deep learning in medical image analysis. *Annual review of biomedical engineering*, 19:221–248, 2017.

[149] Xingjian Shi, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional LSTM network: A machine learning approach for precipitation nowcasting. pages 802–810, 2015.

[150] Ohad Shitrit and Tammy Riklin Raviv. Accelerated magnetic resonance imaging by adversarial neural network. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 30–38. Springer, 2017.

[151] Stephen M Smith, Karla L Miller, Steen Moeller, Junqian Xu, Edward J Auerbach, Mark W Woolrich, Christian F Beckmann, Mark Jenkinson, Jesper Andersson, Matthew F Glasser, et al. Temporally-independent functional modes of spontaneous brain activity. *Proceedings of the National Academy of Sciences*, 109(8):3131–3136, 2012.

[152] Nitish Srivastava, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. Dropout: A simple way to prevent neural networks from overfitting. 15(1):1929–1958, 2014.

[153] Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, and Martin Landray. UK biobank: An open access resource for identifying the causes of a wide range of complex diseases of middle and old age. 12(3):e1001779, 2015.

[154] Heung-Il Suk, Chong-Yaw Wee, Seong-Whan Lee, and Dinggang Shen. State-space model with deep learning for functional dynamics estimation in resting-state fMRI. 129:292–307, 2016.

[155] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. Cvpr, 2015.

[156] Garth John Thompson, Wen-Ju Pan, Matthew Evan Magnuson, Dieter Jaeger, and Shella Dawn Keilholz. Quasi-periodic patterns (qpp): large-scale dynamics in resting state fmri that correlate with local infraslow electrical activity. *Neuroimage*, 84:1018–1031, 2014.

[157] Vincent G. van de Ven, Elia Formisano, David Prvulovic, Christian H. Roeder, and David E.J. Linden. Functional connectivity as revealed by spatial independent component analysis of fMRI measurements during rest. 22(3):165–178, 2004.

[158] Vincent G van de Ven, Elia Formisano, David Prvulovic, Christian H Roeder, and David EJ Linden. Functional connectivity as revealed by spatial independent component analysis of fmri measurements during rest. *Human brain mapping*, 22(3):165–178, 2004.

[159] Martijn P. van den Heuvel and Hilleke E. Hulshoff Pol. Exploring the brain network: A review on resting-state fMRI functional connectivity. 20(8):519–534, 2010.

[160] Laurens van der Maaten and Geoffrey Hinton. Visualizing data using t-SNE. 9:2579–2605, 2008.

[161] David C Van Essen, Heather A Drury, Sarang Joshi, and Michael I Miller. Functional and structural mapping of human cerebral cortex: solutions are in the surfaces. *Proceedings of the National Academy of Sciences*, 95(3):788–795, 1998.

[162] David C. Van Essen, Stephen M. Smith, Deanna M. Barch, Timothy EJ Behrens, Essa Yacoub, Kamil Ugurbil, and Wu-Minn HCP Consortium. The WU-Minn g: An overview. 80:62–79, 2013.

[163] DC Van Essen and HA Drury. Structural and functional analyses of human cerebral cortex using a surface-based atlas. *Journal of Neuroscience*, 17(18):7079–7102, 1997.

[164] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, pages 5998–6008, 2017.

[165] Petar Veličković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. 2017.

[166] Jinhui Wang, Xinian Zuo, and Yong He. Graph-based network analysis of resting-state functional MRI. 4:16, 2010.

[167] Lebo Wang, Kaiming Li, Xu Chen, and Xiaoping P. Hu. Application of Convolutional Recurrent Neural Network for Individual Recognition Based on Resting State fMRI Data. 13, 2019.

[168] Yue Wang, Yongbin Sun, Ziwei Liu, Sanjay E Sarma, Michael M Bronstein, and Justin M Solomon. Dynamic graph CNN for learning on point clouds. 2018.

[169] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world' networks. *nature*, 393(6684):440, 1998.

[170] Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world'networks. 393(6684):440, 1998.

[171] Zonghan Wu, Shirui Pan, Fengwen Chen, Guodong Long, Chengqi Zhang, and Philip S. Yu. A comprehensive survey on graph neural networks. 2019.

[172] Kelvin Xu, Jimmy Ba, Ryan Kiros, Kyunghyun Cho, Aaron Courville, Ruslan Salakhudinov, Rich Zemel, and Yoshua Bengio. Show, attend and tell: Neural image caption generation with visual attention. pages 2048–2057, 2015.

[173] Weizheng Yan, Vince Calhoun, Ming Song, Yue Cui, Hao Yan, Shengfeng Liu, Lingzhong Fan, Nianming Zuo, Zhengyi Yang, Kaibin Xu, et al. Discriminating schizophrenia using recurrent neural network applied on time courses of multi-site fmri data. *EBioMedicine*, 47:543–552, 2019.

[174] Guang Yang, Simiao Yu, Hao Dong, Greg Slabaugh, Pier Luigi Dragotti, Xujiong Ye, Fangde Liu, Simon Arridge, Jennifer Keegan, Yike Guo, et al. Dagan: Deep de-aliasing generative adversarial networks for fast compressed sensing mri reconstruction. *IEEE transactions on medical imaging*, 37(6):1310–1321, 2017.

[175] Guang Yang, Simiao Yu, Hao Dong, Greg Slabaugh, Pier Luigi Dragotti, Xujiong Ye, Fangde Liu, Simon Arridge, Jennifer Keegan, Yike Guo, et al. DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction. 37(6):1310–1321, 2017.

[176] Zhengshi Yang, Xiaowei Zhuang, Karthik R Screenivasan, Virendra R Mishra, and Dietmar Cordes. Robust motion regression of resting-state data using a convolutional neural network model. *Frontiers in neuroscience*, 13:169, 2019.

[177] Jason Yosinski, Jeff Clune, Anh Nguyen, Thomas Fuchs, and Hod Lipson. Understanding neural networks through deep visualization. 2015.

[178] Hongyuan You, Adam Liska, Nathan Russell, and Payel Das. Automated brain state identification using graph embedding. In *2017 International Workshop on Pattern Recognition in Neuroimaging (PRNI)*, pages 1–5. IEEE, 2017.

[179] Jure Zbontar, Florian Knoll, Anuroop Sriram, Matthew J. Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, and Hersh Chandarana. Fastmri: An open dataset and benchmarks for accelerated mri. 2018.

[180] Jure Zbontar, Florian Knoll, Anuroop Sriram, Matthew J. Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdzal, Adriana Romero, Michael Rabbat, Pascal Vincent, James Pinkerton, Duo Wang, Nafissa Yakubova, Erich Owens, C. Lawrence Zitnick, Michael P. Recht, Daniel K. Sodickson, and Yvonne W. Lui. fastMRI: An open dataset and benchmarks for accelerated MRI. 2018.

[181] Matthew D Zeiler and Rob Fergus. Visualizing and understanding convolutional networks. pages 818–833. Springer, 2014.

[182] Ziwei Zhang, Peng Cui, and Wenwu Zhu. Deep learning on graphs: A survey. *IEEE Transactions on Knowledge and Data Engineering*, 2020.

[183] Fenqiang Zhao, Shunren Xia, Zhengwang Wu, Li Wang, Zengsi Chen, Weili Lin, John H Gilmore, Dinggang Shen, and Gang Li. Spherical u-net for infant cortical surface parcellation. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 1882–1886. IEEE, 2019.

[184] Yu Zhao, Qinglin Dong, Shu Zhang, Wei Zhang, Hanbo Chen, Xi Jiang, Lei Guo, Xintao Hu, Junwei Han, and Tianming Liu. Automatic Recognition of fMRI-Derived Functional Networks Using 3-D Convolutional Neural Networks. 65(9):1975–1984, 2018.

[185] Yu Zhao, Fangfei Ge, Shu Zhang, and Tianming Liu. 3d deep convolutional neural network revealed the value of brain network overlap in differentiating autism spectrum disorder from healthy controls. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 172–180. Springer, 2018.

[186] Xi-Nian Zuo, Ting Xu, Lili Jiang, Zhi Yang, Xiao-Yan Cao, Yong He, Yu-Feng Zang, F. Xavier Castellanos, and Michael P. Milham. Toward reliable characterization of functional homogeneity in the human brain: Preprocessing, scan duration, imaging resolution and computational space. 65:374–386, 2013.