DeepFMRI: End-to-end deep learning for functional connectivity and classification of ADHD using fMRI

Atif Riaz, Muhammad Asad, Eduardo Alonso, Greg Slabaugh

City, University of London

Abstract

Background

Resting state fMRI has emerged as a popular neuroimaging method for automated recognition and classification of brain disorders. Attention Deficit Hyperactivity Disorder (ADHD) is one of the most common brain disorders affecting young children, yet its underlying mechanism is not completely understood and its diagnosis is mainly dependent on behaviour analysis.

New method

In this paper, we propose an end-to-end deep learning architecture to diagnose ADHD. Our aim is to (1) automatically classify a subject as ADHD or healthy control, and (2) demonstrate the importance of functional connectivity to increase classification accuracy and provide interpretable results. The proposed method, called DeepFMRI, is comprised of three sequential networks, namely (1) a feature extractor, (2) a functional connectivity network, and (3) a classification network. The model takes fMRI pre-processed time-series signals as input and outputs a diagnosis, and is trained end-to-end using back-propagation.

Results

Experimental results on the publicly available ADHD-200 dataset demonstrate that this innovative method outperforms previous state-of-the-art. Different imaging sites contributed the data to the ADHD-200 dataset. For the New York University imaging site, our proposed method was able to achieve classification accuracy of 73.1% (specificity 91.6%, sensitivity 65.5%).

Comparison with Existing Methods

Preprint submitted to Deep learning methods and applications in neuroimagingNovember 5, 2019

In this work, we propose a novel end-to-end deep learning method incorporating functional connectivity for the classification of ADHD. To the best of our knowledge, this has not been explored by existing studies.

Conclusions

The results suggest that the proposed end-to-end deep learning architecture achieves better performance as compared to the other state-of-the-art methods. The findings suggest that the frontal lobe contains the most discriminative power towards the classification of ADHD.

Keywords: functional MRI, Deep Learning, End-to-end network, ADHD

1. Introduction

The human brain can be perceived as a large and complex network controlling and monitoring the systems of the body. The brain network is comprised of multiple inter-connected brain regions. While supervising and executing different bodily functions, the brain regions continuously coordinate with each other in an efficient manner resulting in a complex brain connectivity pattern. Brain connectivity is a promising source for diagnosis, characterization and prediction of pathologies, which are linked to abnormal functional organization of the brain. Recently, analysis of the connectivity of brain regions has gained much research focus as it is believed that connectivity plays a key role in the cognitive processes [1].

A number of different imaging modalities have been proposed to explore the functional activity of the brain, such as Electroencephalography (EEG), Magnetoencephalography (MEG), functional Magnetic Resonance Imaging (fMRI) and Positron Emission Tomography (PET). Being non-invasive and displaying remarkable spatial resolution, fMRI is considered most suitable towards determining functional activity of the brain regions [2]. In recent years, fMRI has emerged as a popular neuroimaging modality to explore brain connectivity patterns for the classification of different neurological disorders, demonstrated in several studies [3, 4, 5] with promising outcomes by studying brain functional networks in

resting state fMRI.

fMRI data can be viewed as a 4D tensor such that the 3D volume of the brain is divided into small voxels or regions and the activity of each region is recorded for a certain duration. Two brain regions that show synchronous functional activity are assumed to be functionally connected. Functional connectivity is viewed as the relationship between the temporal activity patterns of anatomically separated brain regions, depicting the strength of functional communication between regions [1].

Recently, functional connectivity has been shown to be an important biomarker towards discrimination of different brain disorders [6, 7, 8]. Research studies have shown that brain disorders such as Alzheimer's disease, epilepsy and ADHD can alter the functional connectivity of the brain network [9]. Accurate identification of the altered functional connectivity induced by a particular disorder is considered an important task that may highlight the underlying mechanisms of the disorder. Recently, resting state fMRI has emerged as a promising neuroimaging tool to investigate functional activity of brain regions [6, 7, 10, 11, 4, 12]. In particular, fMRI has been employed to identify the connectivity alterations induced by disorders such as epilepsy [6, 7], schizophrenia [10, 11], ADHD [4, 12, 8, 13], Alzheimer's disease [14] and Parkinson's disease [15, 16].

ADHD is one of the most common neuro-developmental and mental disorders affecting 5-10% of young children [4], contributing to lifetime impairment [17], poor quality of life [18] and long-term burden on affected families [17, 18]. Like many other neurological disorders, the underlying mechanisms of ADHD are still unknown [4]. There is no single confirmed diagnostic method available for diagnosing ADHD, which may take up to several months to complete and is dependent on observations conducted by healthcare practitioners or parents.

In this work, we propose a deep learning architecture for diagnosing ADHD. The proposed method consists of an end-to-end trainable network that takes pre-processed time-series signals as input and produces predicted label as its output. The proposed architecture incorporates a functional connectivity network

which is designed to capture pair-wise region connectivity. The last component is a classifier that takes functional connectivity measures computed from pairs of brain regions as input, and produces a final prediction. The contributions of the work include 1) a deep learning architecture, trained end-to-end, for the classification of ADHD, 2) demonstration of the importance of functional connectivity for improved results and, 3) a new state-of-the-art classification accuracy on the ADHD-200 dataset.

It should be noted that a preliminary version of this work was published in a peer review conference [19]. Compared to the earlier version of this manuscript, we have extended our work by 1) performing and including additional experimental results, 2) exploring the impact of functional connectivity in an end-to-end deep network and, 3) performing anatomical analysis of our results. In this paper, we have applied our framework to ADHD data only, however, the proposed method can also be applied to other neurological disorders like schizophrenia and epilepsy.

The rest of the paper is structured as follows. Related work is introduced in Section 2. We present an overview of the fMRI data used in this work and preprocessing steps in Section 3. Our proposed multi-stage network is detailed in Section 4. Section 5 shows the experimental validation and results. The anatomical analysis is discussed in Section 6. Section 7 concludes the paper.

2. Related work

ADHD has received significant research focus, including studies employing machine learning using fMRI data to investigate functional connectivity alterations in ADHD [20, 21, 4, 9, 22]. In the following subsections, we review the related work categorised as correlation methods, dimensionality reduction methods, graph based methods, clustering based methods, and deep learning methods.

2.1. Correlation methods

Correlation is a widely used method calculating functional connectivity where the regions with high correlation are considered strongly functionally connected.

Dai et al. [23] segmented the brain into 351 Regions of Interest (ROIs) using a template provided by [24] and calculated functional connectivity using Pearson's correlation. Bohland et al. [25] applied the Automated Anatomical Labeling (AAL) atlas [26] to segment the brain into 116 ROIs and computed functional connectivity using three correlation variants: Pearson's correlation, sparse regularized inverse covariance [27] and Patel's Kappa [28]. Eloyan et al. [29] extracted five ROIs belonging to the motor network with 264 voxels as nodes and computed functional connectivity using Pearson's correlation coefficient which was later used for classification. Similarly Cheng et al. [30] employed Pearson's correlation and partial correlation to calculate functional connectivity on 90 brain regions extracted from the AAL template [26]. Multiple measures including Regional Homogeneity (ReHo), functional connectivity and fractional amplitude of low-frequency fluctuation (fALFF) were employed for classification.

Most of these studies rely on correlation-based approaches for calculation of functional connectivity. However, the correlation-based approach does not characterize the network structure of different brain regions, i.e. whether two brain regions belong to the same functional cluster or not [31]. Moreover, the network obtained by correlation is quite dense, which may degrade the performance of a classifier [31, 6]. We also note that these studies highlight the lack of consensus in using a common brain atlas for specifying brain regions in fMRI analysis.

2.2. Dimensionality reduction methods

Dimensionality reduction methods such as Independent Component Analysis (ICA) are commonly used with fMRI data. Garcia et al. [20] proposed an ICA based functional-anatomical discriminative region model for pattern classification of ADHD. This approach applied ICA to extract brain functional connectivity networks. Similarly, Tabas et al. [22] proposed a variant of ICA to characterize the differences between a healthy control group and an ADHD group. This study used 20 independent components and combined ICA and a spatial variant of Fisher's linear discriminant. ICA-based methods are considered a natural

choice for fMRI studies as these methods do not require any prior information about the spatial or temporal patterns of source signals.

In another study [32], the authors have proposed a framework for the classification of Autism and ADHD. The authors have used multiple imaging modalities, namely MRI and fMRI for classification of disorders. Also, personal characteristic data such as age, gender has been explored for the classification. In this work, a histogram of oriented gradients have been calculated from the imaging modalities and the minimum redundancy maximum relevance (MRMR) is applied as the feature selection strategy. The selected features are presented to a support vector machine classifier for the final prediction. Similarly, in another study [33], the authors have also used multiple imaging modalities (MRI and fMRI) for the classification of ADHD and Autism. In this study, the structural features from MRI are extracted through a sparse auto-encoder and principal component analysis (PCA) is applied to the fMRI data to decompose to extract principal components. Finally, an SVM classifier is used for final prediction. Both of these studies rely on multiple imaging modalities and do not incorporate functional connectivity in their work, which is an important characteristic of fMRI data. The focus in this proposed work is to utilise a single imaging modality (fMRI) and exploit functional connectivity for prediction of ADHD. We note that the proposed method produces comparable results to [32, 33], but using only a single imaging modality.

ICA-based approaches have shown success in classification tasks, however, there are limitations to these methods. First, independent components are often perceived as difficult to understand [1]. ICA is based on the assumption that components (signal sources) are independent, whether spatially or temporally. Violation of the assumption degrades performance. Moreover, selection of the number of independent components and a threshold value for the independent component maps are considered as a drawback [2], especially in the case of fMRI, where there is no prior information of the number of components available.

2.3. Graph based methods

A number of graph based approaches have been applied to model brain networks. Dey et al. [4] proposed a graph-based solution for the classification of ADHD. They modelled the brain connectivity network as a graph and represented each node of the network as a set of attributes which was termed as the signature of a node. The correlation was applied for functional network construction and a threshold was applied to construct the network. The threshold value was arbitrarily chosen and different values were employed for different imaging datasets. Similarly, Siqueira et al. [9] investigated different graph-based measures for the classification of ADHD.

2.4. Clustering based methods

Clustering is another popular approach for the evaluation of functional connectivity, where regions belonging to the same cluster are assumed to be functionally connected. Studies have shown that a clustering-based approach is more sophisticated as compared to correlation-based approaches, as the network obtained by clustering is sparse [7, 2]. Recently, a few studies have applied clustering to the ADHD-200 dataset for functional connectivity analysis [8, 3]. Zhang et al. [31] applied k-means clustering to calculate functional connectivity. However, in k-means, random initialization of clusters and priori information on the number of clusters emerges as a major drawback, as these are unknown in the case of fMRI. Hierarchical clustering can also be applied to calculate functional connectivity [34], however the selection of the thresholding and the number of clusters are not known in advance in the case of fMRI. Other studies (e.g., [7, 8, 3]) have applied affinity propagation (AP) [35] clustering for the classification of brain disorders. AP clustering does not require an initial number of clusters, which is a good choice for fMRI data. However, AP requires a similarity measure between data points as the input and output of the algorithm are affected by this measure. Selection of the distance measure presents a limitation of these methods.

2.5. Deep learning based methods

End-to-end deep learning networks have been shown to outperform existing classical machine learning models in a number of domains like image classification, image segmentation and object recognition [36]. Generally speaking, an end-to-end trainable network refers to a single learning system where the predicted label of a neural network model is predicted directly from the input, with all weights learned through back-propagation. In the domain of fMRI, there is very limited work exploring deep learning based functional connectivity for classification of neurological disorders [13].

Many studies have applied deep learning to MRI data. In [37], the authors have applied a three-dimensional convolutional neural network for the classification of Alzheimer's disease using MRI data. The method used an auto-encoder for pre-training the CNN filters. The study suggested that a CNN performs better than other classifiers for prediction of Alzheimer's disease. In [38], the authors have applied a CNN for the classification of ADHD using combined features of MRI and fMRI. Instead of applying a CNN directly on the temporal data of brain regions, the CNN was applied to the features extracted from the fMRI data. These features include regional homogeneity (ReHo), the normalized amplitude of low-frequency fluctuations (fALFF) and voxel-mirrored homotopic connectivity (VMHC). These features are calculated using conventional hand-crafted statistical measures and might not able to capture the inherent characteristics of temporal fMRI signals. Also, functional connectivity is not explored. In [39], the authors have applied an auto-encoder for the classification of Alzheimer's disease. Functional connectivity was calculated through correlation and then fed to an auto-encoder network followed by a softmax layer for the classification. The method uses a deep network, however, still relies on correlation for functional connectivity.

The use of an artificial neural network for classification of ADHD has been explored in [40]. However, this method relied on a classical machine learning model for the extraction of multiple features, which are passed to a fully connected neural network for classification. Similarly, the study in [41] addressed the

problem of classification of mild cognitive impairment (MCI) from fMRI data. The authors applied a deep autoencoder for dimensionality reduction of pre-processed fMRI signals. The representation encoded by the autoencoder was fed into a hidden Markov model to estimate the likelihood of a subject belonging to the healthy control group or the MCI group to identify its predicted label. Recently, a deep learning method, namely FCNet [13], has been proposed for the classification of ADHD from fMRI data. The method applies a convolutional neural network (CNN) to predict functional connectivity of brain regions. However, after predicting functional connectivity using deep learning, the method applies classical machine learning methods to extract discriminant features and an SVM classifier to predict classification labels. These approaches rely on classical machine learning models for prediction of a disorder. In a non peer reviewed work [42], a CNN was proposed for classification of Alzheimer's disease using fMRI data. The study applied a 2D CNN where the 4D fMRI data was converted to a stack of 2D images and classification was evaluated on the individual 2D images. The study does not incorporate the temporal information in the work, which is the most important aspect of the fMRI time-series data. Being a 2D CNN model, the prediction results are evaluated for individual images instead of per subject. For prediction of a subject, the results of individual 2D images are accumulated. A recent study [43] applied a 3D CNN for classification of autism spectrum disorder. The fMRI 3D volume was downsampled and the 3D CNN was applied on the downsampled data. The study does not incorporate functional connectivity which is an important characteristic in brain studies.

Many of the studies describe above highlight the importance of functional connectivity towards the classification of a disorder. However, most existing techniques employ classical machine learning methods for classification, typically relying solely on hand-crafted features. Discriminant features are selected and presented to a classifier for the final prediction. More importantly, all the different processing steps in a classical machine learning method are independent of each other. However, in the machine learning literature, deep learning has proved to be a powerful paradigm to simultaneously learn discriminant features

and a classifier [36].

We propose a deep neural network for diagnosis of ADHD. The network is learned end-to-end, taking pre-processed fMRI time-series as input and producing a classification (healthy or ADHD) as output. Our proposed network does not involve hand-crafted features or independent modules for feature extraction, feature selection and classification. Rather, all these modules are incorporated in a single network and the weights are learnt itself by the model during the training. To the best of our knowledge, this is the first study to propose a fully end-to-end deep learning method incorporating functional connectivity for the classification of a neurological disorder based on fMRI. Particularly, we are interested to explore if a deep network-based model can be designed for the classification of ADHD, and if it is able to outperform classical machine learning methods or a combination of deep learning and existing classical machine learning methods. The strength of deep learning comes from its representation learning capabilities, where the most discriminative features are learned during training. A deep network is composed of multiple modules, where each module learns the representation from one lower level to a higher, more abstract level.

3. Data and preprocessing

The resting state fMRI data used in this study is from the NeuroBureau ADHD-200 competition [44]. The data provided by the competition consists of MRI and resting state fMRI data as well as different phenotypic information (non-imaging data) such as age, gender and IQ, for each subject. There was a global competition held for classification of ADHD subjects, and the consortium has provided the training and independent test dataset for each imaging site. The dataset was collected and contributed by eight different imaging sites. For the development and evaluation of our proposed network, we used datasets from three imaging sites: NeuroImage (NI), New York University Medical Center (NYU), and Peking University (Peking). Our approach is designed to accept time-series signals of length 172, therefore, we discarded the imaging sites with

Table 1: Overview of the dataset used in the study.

	Train dataset		Test dataset	
	Healthy controls	ADHD	Healthy controls	ADHD
NYU	98	118	12	29
NI	23	25	14	11
Peking	61	24	24	27

signal length less than 172. All the imaging sites have a different number of subjects. Table 1 describes the overview of the data used in this study. The imaging sites have different lengths of time-series signals. The deep learning methodology employed in this work requires a fixed length of input signal and can not accept input with different input lengths. To decide the supported input length, we selected the imaging site with highest number of subjects. The site with maximum number of subjects was NYU with 226 subjects and its length of time-series signals was 172. Therefore, we designed our network to accept input length of 172. We discarded the imaging sites with length of time-series smaller than this number. Also, the time-series of length greater than 172 were truncated at the beginning to make fixed length of input signals suitable for input to the method.

The scan parameters and the equipment used were varied across different imaging sites. Some of important parameters used by different imaging sites are presented in Table 2. Additionally, the imaging sites employ different data acquisition parameters, for example NeuroImage scanned the data while subjects were asked to close their eyes, whereas in Peking, subjects were asked to keep their eyes open or close. The variations in the parameters of scan and data acquisition increase the complexity and diversity of the data.

For all our experiments, we used the pre-processed data released for the competition. This data has been pre-processed as part of the connectome project [45]. The preprocessing is performed using AFNI [46] and FSL [47] tools on

Table 2: Scan parameters per imaging sites.

	NYU	NI	Peking
Slices	33	37	33
TR (ms)	2000	1960	2000
TE (ms)	15	40	30
Thickness (mm)	4.0	3.0	3.5
FoV read (mm)	240	224	200
FoV phase (%)	80	100	100
Flip angle (degree)	90	80	90

Athena computer clusters at the Virginia Tech advanced research computing centre. The preprocessing steps include: removing of the first four time points, slice time correction, motion correction (first image taken as the reference), registration on 4×4×4 voxel resolution using the Montreal Neurological Institute (MNI) space, filtration (bandpass filter $0.009Hz < f < 0.08Hz$) and smoothing using a $6mm$ FWHM Gaussian filter. The brain is segmented into 90 regions using the well established AAL template. Although some other atlases like Craddoc et al. [24] segment the brain into 351 regions can be used here, they will yield very high dimensionality of data, so the AAL template is preferred. A number of studies [7, 3, 8, 13, 30, 25] have also employed the AAL template for brain parcellation. Interested readers may refer to the competition website for further details on the data and preprocessing [45].

4. Methods

4.1. End-to-end model

In this paper, we propose an end-to-end deep learning model for the classification of ADHD that takes pre-processed fMRI time-series signals as input and predicts a label (1 for ADHD subject and 0 for healthy control) as output. The proposed work is motivated by FCNet [13]. FCNet is used to extract functional connectivity

from fMRI time-series signals, however it combines deep learning and classical machine learning and is not trained end-to-end. For ease of understanding, our proposed architecture can be divided into three modules: 1) feature extractor network, 2) functional connectivity network, and 3) classification network. The feature extractor network is applied to a pre-processed time-series signal of individual brain region and it produces an abstracted feature as its output. These features are learned during the training. The functional connectivity network takes the abstracted features as input and produces the strength of similarity between any two brain regions. Finally, the classification network produces the final prediction label based on the functional connectivity values of all brain regions. We describe the details of each individual network below.

4.1.1. The feature extractor network

This convolutional neural network (CNN) extracts features from *individual* brain region pre-processed time-series signals and is comprised of multiple layers that are common in CNN models to learn abstract representations of data. The network is designed to accept signals of length 172 as the input and produces an abstract representation (vector of size 32). The network hyperparameters were not tuned using a validation dataset, instead, the network architecture and hyper-parameters are inspired by [48]. We use parametric ReLU and its slope is learned during the training phase. The parametric ReLU is presented as:

$$f(x) = \begin{cases} x, & x > 0 \\ ax, & x \leq 0 \end{cases}, \tag{1}$$

where a is a non-negative scalar subject to learning. Instead of three consecutive convolutional layers proposed in the network of [48], we have two convolutional layers next to each other (Layer 12 and 13). The network is presented in Figure 1a and is comprised of 15 layers (Figure 1d). All convolutional layers are one dimensional with a kernel size of 3, stride of 1 and the numbers of filters are 32, 64, 96, 64, 64 for the respective layers as presented in Figure 1. All max pooling layers pool temporally with pool length of 2 with stride 1 as proposed by [48].

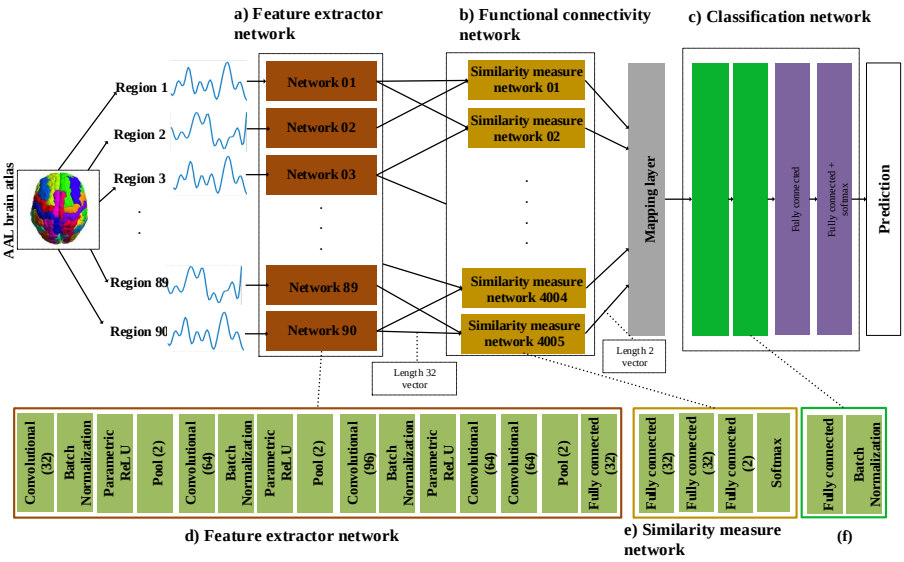


Figure 1: The DeepFMRI architecture. a) represents a set of 90 feature extractor networks where each network is applied on each individual region R. All networks share the same weights. b) represents a functional connectivity network comprising a set of 4005 similarity measure networks. Each network's input contains abstracted features of two brain regions. All networks share the same weights. c) is the classification network comprising of fully connected layers and a softmax layer. d) represents the layers in the feature extractor network, where Convolutional (c) represents a convolutional layer with c filters, Pool (p) represents a max pooling layer with pool length of p and Fully connected (f) represents a fully connected layer with f nodes. Similarly, e) represents layer architecture of similarity measure network, and f) represents the layers of an individual block in the classification network (each block has two layers, as shown above).

The last fully connected layer in the network has 32 nodes. The total number of weights in the feature extractor network are 90947.

In the proposed work, the same feature extraction steps are applied to individual brain regions. This is implemented by employing n_f feature extractor. Each feature extractor network is applied to an individual brain region ($n_f = 90$), converting individual time-series data into an abstract representation. All the feature extractor networks share the same parameters and updates are applied to these shared parameters during training.

4.1.2. The functional connectivity network

The functional connectivity network determines the functional connectivity between the brain regions and is presented in Figure 1b. The network is comprised of multiple similarity measure networks where the architecture of each similarity network is presented in Figure 1e. This Siamese-inspired similarity measure network determines the similarity between *pairs* of extracted features from two brain regions. Here, the calculated similarity measure serves as the degree of functional connectivity between the two regions. Each similarity measure network operates on two brain regions, where the input to each network are the abstracted features of the two brain regions from the feature extractor network. The neural network learns to identify functionally connected regions using a non-linear function. This function is learned from the data and is specific to this problem as compared to other commonly used generic measures like correlation. The similarity measure network is comprised of three fully connected layers, where the last layer is connected to a softmax layer with dense connections. These layers are presented in Figure 1e, where the number of nodes in these layers are 32, 32, and 2 respectively. The total number of weights in the network is 3202. The output of the similarity measure network is a length two vector, and can be interpreted as the probability the two regions are functionally connected, and the complement of the probability.

In the proposed work, all pairs of brain regions are passed through the same similarity measure network. This is implemented by employing n_s similarity measure networks. The similarity measure network is applied to all combinations of pairs of brain regions, so $n_s = 4005$ $(n_f \times (n_f - 1)/2)$. There are 4005 similarity measure networks and all the similarity measure networks are implemented with the constraint that the networks share the same parameters and updates are applied to these shared parameters. The approach is similar to a Siamese network [49].

The output of the functional connectivity network is fed to a mapping layer

using the following operation:

$$M(i) = w_1 v_1^i + w_2 v_2^i, \tag{2}$$

where v_1^i and v_2^i are the scalar outputs of the i^{th} similarity measure network, w_1 and w_2 are the weights such that $w_1 + w_2 = 1$. In order to reduce training parameters, we use $w_1 = 1$ and $w_2 = 0$. Moreover, the parameters enforce passing the functional connectivity to the classification network. The output of this network can be assumed to be the functional connectivity mapping of all the brain regions, and uses deep learning-based features from the feature extractor network. Instead of initializing weights of the feature extractor network and the similarity measure network randomly, we use weights of a pre-trained FCNet [13].

The architecture of the feature extractor network in the proposed network is the same as in the feature extractor network in FCNet [13]. Similarly, the architecture of the similarity measure network is the same in both in FCNet and the proposed network. We extracted weights of both of these networks from the pre-trained FCNet and used them to initialize the weights of the corresponding networks of the proposed network. During the training of the network, a small learning rate (10^{-5})is used for the functional connectivity network. The small learning rate allows smaller updates of weights and its advantages are two-fold: it encourages maintenance of the original characteristics of the network (i.e. to extract functional connectivity of brain regions), and at the same time adapts the network according to the end-to-end classification task. In order to evaluate whether the original characteristics of the functional connectivity network are maintained or not, we have compared the weights of the pre-trained functional connectivity network with the weights of the same network after final training of the DeepFMRI. The percentage difference in the weights is not greater than 5% for three sites, showing that the network is maintaining its original characteristics. FCNet is pre-trained using the training data and its training does not include any test data.

4.1.3. Classification network

This neural network produces the final classification results. The input to this network is the output of the mapping layer features (M) representing the functional connectivity of brain regions. The network is comprised of four fully connected layers where the last layer is connected to a softmax classifier with fully connected layers. The network is presented in Figure 1c, where the number of nodes in the fully connected layers are 100, 50, 50 and 2 respectively. The total number of weights in the network is 408602. The weights were initialized randomly.

5. Experimental settings and results

In this section, we evaluate the effectiveness of the proposed method for ADHD classification employing resting state fMRI and by comparing our results with those of the state-of-the-art methods in the literature.

5.1. Experimental settings

The proposed model is evaluated on the ADHD-200 dataset. This publicly available dataset was contributed by different imaging sites. Each imaging site provided separate training and testing datasets and we followed the split in accordance with the consortium. This also makes it possible to compare to other methods as they also tested on the independent test data. Doing cross-validation within the original training/test was not advised due to the smaller dataset. For the evaluation of our method on individual site, we train our end-to-end model on the training dataset of each imaging site and test it on the corresponding test dataset of that individual site. There are four categories of subjects in the dataset: healthy control, ADHD combined, ADHD hyperactive-impulsive and ADHD inattentive. Here, we combine all ADHD types in one category as we are interested to investigate the classification between healthy control and ADHD only.

The proposed model is created in the python programming language using the tensorflow deep learning library. The network is trained end-to-end. The

Adam optimizer [50] is used to optimize the network and the number of epochs is set to 50. After 50 epochs, the training loss converges and becomes stable.

For the initialization of the feature extractor and similarity measure networks, we use weights from a pre-trained FCNet [13] in our work, and these weights are updated through fine-tuning. The full deep network is trained the end-to-end model with the cross-entropy loss:

$$L = -\frac{1}{n}\sum_{1}^{n}[y_i log(\hat{y_i}) + (1 - y_i) log(1 - \hat{y_i})], \tag{3}$$

where n is the number of training samples, y_i is the ground truth label of the subject (1 for ADHD subject and 0 for healthy control) and $\hat{y_i}$ is the prediction by the proposed network.

As the feature extraction and similarity measure networks are initialized with a pre-trained FCNet, we employ different learning rates for i) feature extraction and similarity measure networks (10^{-5}), and ii) the classification network (10^{-4}). In our experiments, we use the Adam optimizer [50] to optimize the network.

5.2. Comparison methods

To validate the effectiveness of the proposed method, we compare it with different network architectures and state-of-the-art methods namely, an end-to-end network without functional connectivity, FCNet, a clustering method and a correlation method.

5.2.1. End-to-end model without functional connectivity

A number of studies have shown that functional connectivity plays a key role in cognitive processes of the brain [1]. Recently, studies have shown that altered functional connectivity can serve as an important biomarker towards the identification and classification of different brain disorders [6, 7, 10, 11, 4, 12, 8]. Inspired by the findings of such studies, we have integrated functional connectivity network in the proposed method architecture. In order to evaluate the importance of functional connectivity in our proposed work towards the

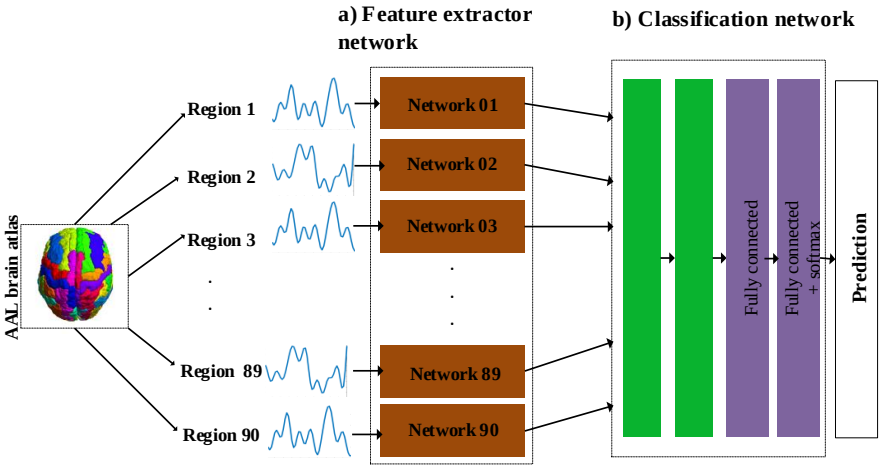


Figure 2: The end-to-end model without the functional connectivity network. a) represents a set of 90 feature extractor networks where each network is applied to each individual region R. b) is the classification network.

classification of ADHD, we have evaluated our end-to-end network without the functional connectivity network. The model without the functional connectivity network is presented in Figure 2.

In this model, the abstracted features calculated through the feature extraction network are merged and passed directly to the classification network and there is no functional connectivity network. Due to the exclusion of the functional connectivity network, there are fewer overall parameters than in the proposed model. The weights and parameters of the feature extraction network are the same as in the proposed network.

5.2.2. FCNet

The FCNet method [13] uses a CNN-based deep learning model to extract functional connectivity from the pre-processed fMRI signals. An Elastic net [21] is applied to extract the discriminant features from the calculated functional connectivity and finally an SVM classifier is applied to evaluate the classification results. This is the first method that applies a CNN-based deep learning model for the classification of ADHD.

5.2.3. Correlation method

Correlation is a popular method for calculating functional connectivity. In order to compare the proposed method with correlation, we performed correlation on pre-processed fMRI signals to calculate the functional connectivity between the brain regions. We applied an Elastic Net based feature selection to extract discriminant features. Finally, an SVM classifier was applied for classification.

5.2.4. Clustering method

A clustering-based approach for calculating functional connectivity of brain regions was used in [3]. Clustering is considered a more sophisticated technique than correlation-based techniques for calculating functional connectivity [8] as the network obtained by clustering is sparse [7, 2]. In this study, authors applied Synthetic Minority Over sampling TEchnique (SMOTE) [51] to address the problem of dataset imbalance. An Elastic Net [21] was applied to functional connectivity to extract discriminant features. Finally an SVM classifier was utilized to classify healthy vs ADHD subjects.

5.3. Feature importance of functional connectivity

A common criticism of deep networks is that they are a 'black box', mapping inputs to outputs and lacking interpretability. In a clinical context, it is of keen interest to not just produce diagnoses, but also draw some insights from network itself, particularly looking for differences between healthy control and patient groups to characterise the neurological condition. A key advantage of the proposed method is that due to the functional connectivity network, once the model is trained, we can analyse the functional connectivity of brain regions for patients and control, leading to interpretable results. As a demonstration, we carried out an experiment to rank the contribution of individual functional connectivity values towards prediction of a particular class label (in our case, class labels are healthy control and ADHD). This weighted rank can be viewed as feature importance of functional connectivity and represents the strength of the

functional connectivity of two brain regions towards calculating the assignment of a class label.

In our end-to-end network, the final prediction is calculated through the classification network. The classification network is comprised of multiple layers where it gets the functional connectivity from the mapping layer as input and produces the final prediction of the network (i.e. control or ADHD) through a softmax layer. During the training step, the network optimizes the parameters with respect to the individual class label. The network back-propagates the error from the last layer to the mapping layer (reminiscent of functional connectivity in our network) during the training phase. Thus the learned weights of this network carry important information towards determining the feature importance of the functional connectivity for each of the 4005 pairs of brain regions.

Specifically, we are interested to explore the weights assigned by the classification network to the mapping layer M in Equation 2. Deep neural networks have been applied to visualize feature importance on images [52] and videos [53]. To explore the importance of features assigned by the classification network, we carried work similar to [5]. The main idea of the approach is: given a learned neural network and a class of interest, we trace back to the original input by a backward pass with which we can determine how each input entity affects the final detection score for a specific class. In our model, we have two classes (healthy control and ADHD) and we trace back to the mapping layer values to find how each mapping layer value affects the prediction of a particular class.

Given a particular output value of mapping layer M_0, a class c and the class score function $S_c(M)$, we would like to rank the elements of M_0 based upon their influence on the score $S_c(M_0)$. Consider the linear score model for the class c:

$$S_c(M) = w_c M + b_c, \tag{4}$$

where M is the one-dimensional vector, calculated from Equation 2 and is reminiscent of the functional connectivity in our network. The w_c is the weight

and b_c is the bias of the model. Here, it is clear that the magnitude of the elements of the weight vector w_c specifies the importance of the corresponding element of M for the class c.

In the case of a deep neural network, the class score is a non-linear function of the input values, so the above assumption cannot be applied directly. However, given a vector M_0, we can approximate S_c with a linear function in the neighbourhood of M_0 by a first-order Taylor expansion [52]:

$$S_c(M) \approx wM + b. \tag{5}$$

where w is the derivative of S_c with respect to the vector M at the point M_0:

$$w = \frac{\partial S_c}{\partial M}|_{M_0}. \tag{6}$$

Another justification of the network-learned weight using the class score derivative from Equation 6 is that the magnitude of the derivative indicates which elements need to be changed the least to affect the class score the most. One can expect such elements to be more discriminative for a particular class. The derivative w in Equation 6 is calculated through back-propagation during the training of the network. We define feature importance of a node i at layer d as:

$$f_c^d(i) = \sum_{l=L-1}^{d} \sum_k w_c^{(l,l+1)} f_c^{(l+1)}(k), \tag{7}$$

where L is the total number of layers in our classification network, k is the number of nodes and f_c^L is the output of the classification network. We define I as the feature importance map for the class c, where each element is given by:

$$I_c(x) = f_c^M(x). \tag{8}$$

The I_c defines the feature importance of a particular class c.

5.4. Results

We evaluate the proposed network with the data from three imaging sites (NYU, NI and Peking) from the ADHD-200 dataset. The number of training

Table 3: Results from the proposed end-to-end network showing classification accuracy, specificity and sensitivity for individual imaging sites.

	Classification accuracy	Specificity	Sensitivity
NYU	73.1%	91.6%	65.5%
NI	67.9%	71.4%	63.6%
Peking	62.7%	79.1%	48.1%

subjects in each site is $226, 48$ and 85 respectively. ADHD-200 [44] has provided separate train and test dataset for individual imaging site. To evaluate the performance of the proposed work on the individual imaging site, the proposed end-to-end model is trained on the training dataset of each imaging site and the corresponding test dataset of the individual site is used for testing. Please note, the data used to test the method is completely independent of the data used to train. Let TP, TN, FP and FN denote true positive, true negative, false positive and false negative respectively. Sensitivity and specificity are defined as $\textit{sensitivity} = TP/(TP+FN)$ and $\textit{specificity} = TN/(TN+FP)$. The calculated results are presented in Table 3. The highest accuracy is achieved with our method for NYU with a classification accuracy of 73.1%. The classification accuracy for the NI and Peking are 67.9% and 62.7%. One concern could be that the number of subjects is very small to train a deep neural network. However, the performance of the proposed method on independent test datasets shows that the model is able to achieve generalization despite the small dataset. Recent literature [54, 5] has argued that deep neural networks ar effective to generalize well on small datasets.

5.4.1. Comparison with other methods

In order to evaluate the performance of the proposed method, we have evaluated and compared results with state-of-the-art methods as described in the previous section. The results are presented in Table 4. They show that the

proposed method outperforms the average accuracy results of the competition teams (data from the competition website [44]), the highest accuracy of competition for any individual site (from [20]), correlation-based functional connectivity results and clustering based results. Our method also performs well in comparison with the state-of-the-art FCNet method [13]. Table 1 highlights that the distribution of healthy control and ADHD class in train and test splits are different. However, in order to achieve better performance by any classifier, the training and testing data should follow a similar class-distribution. The performance of any classifier depends on the distribution of the training data. If the majority class is changed for the testing data, the classifier performance would drop badly. For the calculation of the baseline classifier accuracy, it can be assumed that a simple classifier would predict the majority class of the training dataset for all testing subjects. In the case of Peking, the majority class in the training dataset is healthy control, so the baseline accuracy for Peking on testing dataset is 47.1% $(24/(24+27))$. Similarly for NYU, with ADHD as majority class in training data set, is 70.7% $(29/(29+12))$, and for NI, with ADHD as majority class in training dataset, is 44.0% $(11/(11+14))$. The baseline accuracy for three imaging sites are presented in Table 4, where the DeepFMRI performs much better than baseline accuracy for the Peking and NYU and slightly better for the NYU.

The results show that the proposed method shows the improved results for NI and NYU and the classification accuracy is highest in all three imaging sites. For Peking, results for both the FCNet [13] and proposed method are the same.

One interesting point about the ADHD dataset is that the studies employing the dataset were not able to achieve high classification accuracy. The average and highest accuracy achieved by competing studies is presented in Table 4 where the accuracy results are around 50%. One possible reason for lower accuracy could be the heterogeneous nature of the data and the scan parameters. For example, for NI imaging site data acquisition, the subjects were asked to keep their eyes closed. No visual stimulus was presented during the scan. For NYU, the participants were asked to close their eyes, think of nothing

Table 4: Comparison of the proposed method with the baseline accuracy, average results of competition teams, highest accuracy achieved for individual site, correlation method, clustering based results [3] and state-of-the-art FCNet method [13]. The highest accuracy for NI was not quoted by [20].

	NI	Peking	NYU
Baseline accuracy	44.0%	47.1%	70.7%
Average accuracy [44]	56.9%	51.0%	35.1%
Highest accuracy [20]	–	58%	56%
Clustering method[3]	44%	58.8%	24.3%
Correlation method	52.0%	52.9%	56.1%
FCNet [13]	60.0%	**62.7%**	58.5%
Proposed method	**67.9%**	**62.7%**	**73.1%**

systematically and not fall asleep. However, a black screen was presented to them. In Peking, the participants were asked to stay still, and either keep their eyes open or closed. A black screen with a white fixation cross was displayed during the scan. Some other parameters were also not consistent across different sites, making the dataset difficult to train any single machine learning model. In order to evaluate the performance of the single model on all three imaging sites, we performed an experiment where the DeepFMRI method was trained on the combined training dataset from all three imaging sites and was evaluated on the test dataset of each individual imaging site. The results are presented in Table 5. The results show that a single model is not able to perform as well, since the data is very heterogeneous, as discussed above. The number of subjects in the NYU dataset are higher, which may account for the better accuracy for NYU, whereas the training data for Peking is very imbalanced, possibly causing lower accuracy for this site.

Table 5: Comparison of accuracies of i) trained and tested on each individual imaging site ii) trained once on the combined training data set of three imaging sites (NI, Peking and NYU) and tested individually on the three imaging sites.

Test data set	Accuracy when trained on each individual imaging site	Accuracy when trained on the combined training data set
NYU	73.1%	65.8%
NI	67.9%	60.0%
Peking	62.7%	43.1%

5.5. Performance comparison

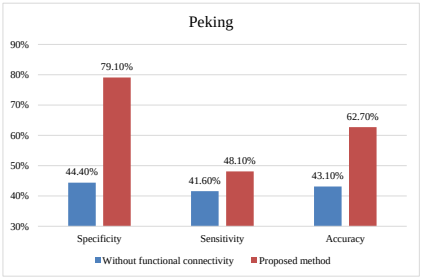
Based on the results in Tables 3 and 4, the proposed end-to-end method comprising the feature extractor, functional connectivity and the classification network to classify ADHD presents better performance than state-of-the-art methods. Although it would be helpful to conduct a statistical significance test, unfortunately, we could not conduct such a test due to very small number of available subjects in the imaging sites. However, from a methodological point of view, we are mainly interested in investigating how important each subnetwork is to produce accurate results. To this end, we additionally performed some experiments by replacing different combinations of the networks.

5.5.1. Comparison Methods

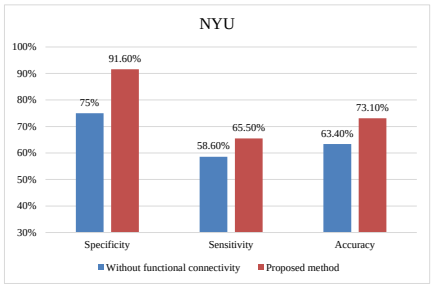
For comparison, we conducted additional experiments, namely, the effect of functional connectivity, end-to-end model without classification network, clustering + classification network and correlation + classification network which are detailed below and the results are presented in Figure 4.

Effect of functional connectivity

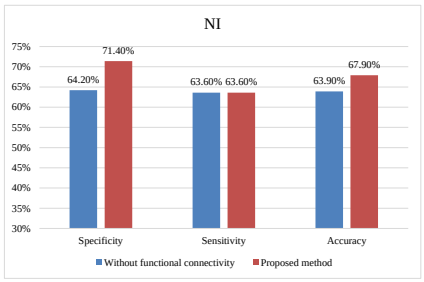
We evaluated an end-to-end model without the functional connectivity network (presented in Figure 2). The comparison of the performance of the end-to-end model with and without functional connectivity is presented in Figure 3.



(a) Peking dataset.



(b) NYU dataset.



(c) NI datatset.

Figure 3: Comparison of the performance of i) Proposed method and ii) the model without the functional connectivity network for the three imaging sites. The proposed model shows better performance as compared to the model without functional connectivity.

It is important to note that for the end-to-end model without functional connectivity, the number of parameters are less as compared to the end-to-end model with the functional connectivity network. The number of trainable parameters for end-to-end model with functional connectivity is $502,751$ vs $386,665$ for end-to-end model without functional connectivity. However, the end-to-end model with functional connectivity yields better performance than the model without functional connectivity. These findings show that functional connectivity serves as an important biomarker towards classification of ADHD.

End-to-end model without classification network

In this experiment, we are interested to determine the importance of the classification network towards diagnosis. Therefore, we use the pre-trained feature extractor and functional connectivity network to calculate functional connectivity. The proposed classification network was not used in this experiment. An Elastic Net was applied to extract discriminant features from functional connectivity and finally, an SVM classifier was applied to evaluate the classification accuracy as proposed by [13].

Clustering + classification network

We wanted also to check the importance of the functional connectivity network in our method. In this experiment, we apply clustering to calculate functional connectivity between the brain regions as proposed by [3, 8]. The calculated functional connectivity is passed to the proposed classification network to evaluate the performance of the network.

Correlation + classification network

Correlation is a popular method to calculate functional connectivity between brain regions. We employ correlation to calculate functional connectivity. Similar to the previous experiment, a classification network was employed on the calculated functional connectivity.

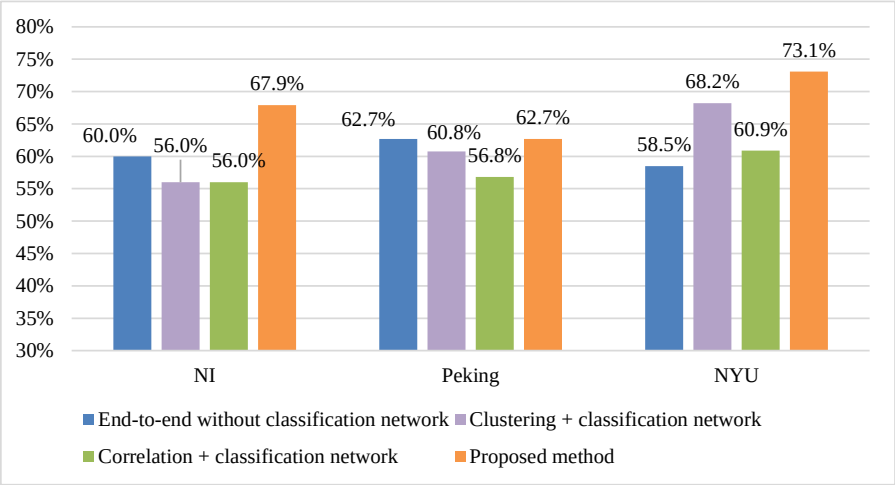


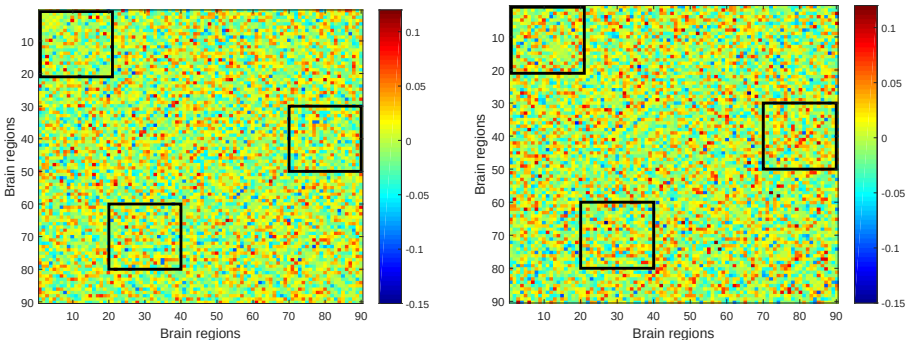
Figure 4: Comparisons of classification accuracy of different methods. The results suggest that the proposed method outperforms all other evaluated methods. The proposed method is able to achieve the highest accuracy on all three imaging sites where it outperforms in NI and NYU imaging dataset.

5.5.2. Comparison Results

We performed the comparison of these four methods and the results are presented in Figure 4. From the results, it is apparent that the proposed method outperforms all other evaluated methods or combinations. Comparison of 'clustering + classification network' and 'correlation + classification network' supports the findings of [3] that clustering is a better method to calculate functional connectivity as compared to correlation-based techniques. However, our proposed end-to-end model yields better performance.

6. Discussion

In this section, we discuss the performance comparison of networks of our proposed method and analyse the features learned by the method.



(a) $I_{healthy}$. (b) I_{ADHD}.

Figure 5: Visualization of the feature importance map (I) for the healthy and the ADHD classes for the NYU dataset. The visualization shows that a different feature importance is assigned to a particular functional connectivity. The boxes highlight some of the differences.

6.1. Analysis of learned feature importance of functional connectivity

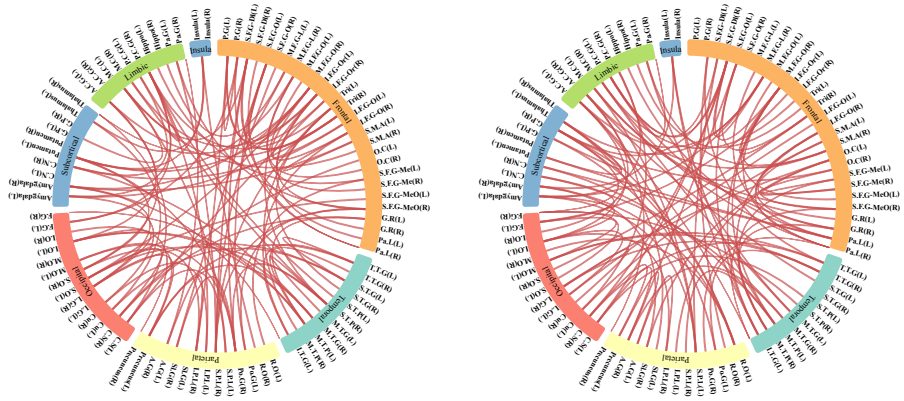
The feature importance map (I_c) from Equation 8 is a 4005 dimensional vector where each value corresponds to the importance of the respective functional connectivity value in determining a particular class. We were interested to explore the learnt feature importance values. Towards this goal, we have selected feature importance values for NYU dataset as, i) NYU has the largest number of subjects compared to other imaging sites, and ii) NYU has highest classification accuracy. We have visualized the feature importance map for both the healthy and the ADHD classes for the NYU dataset and present the result in Figure 5. It should be noted that these are the feature importance values assigned by the network to the individual features (functional connectivity) for a particular class, and it is not the functional connectivity value itself. The figure highlights the differences in feature importance learned by our method for both classes. Our method assigns different weights to an individual feature with respect to its importance towards prediction of a subject. This is in contrast to the most classical machine learning methods [8, 3], which typically employ a feature selection that assigns a single weight to a functional connectivity regardless of the class.

Next, in order to study the differences in the two classes, we have plotted the feature importance map for both classes in Figure 6. For the sake of clarity, we have plotted the top 100 feature maps for both classes. The figure highlights the differences in the feature maps of both classes.

To explore further, we have plotted the feature importance values on the brain map. The visualization of the healthy and the ADHD classes are visualized in Figure 7 and Figure 8, respectively. The figures show that in most of the cases, the importance value assigned by our network to a particular functional connectivity is different for both classes. We performed an experiment for the quantitative analysis of the feature maps of both classes. Our motivation was to compare the top 100 feature maps of both classes. The top 100 feature maps values were extracted from the healthy class and a lookup was performed in the ADHD feature maps. The result is presented in Figure 9. The figure shows that out of top 100 feature maps of healthy class, less than 10% fall in the top 500 feature maps in the ADHD class. Similarly, we extracted top 100 feature maps from the ADHD class and computed the lookup in the healthy class and the results are presented in Figure 10. As in the previous inference, out of the top 100 feature maps of the ADHD class, less than 10% fall in the top 500 feature maps in the healthy class. Our findings suggest that the altered functional connectivity between healthy control and ADHD may relate to functional brain network differences. In particular, the proposed method appears to weight different brain network structures depending on the particular class (control or ADHD).

Finally, we are interested in analyzing the learned feature importance map for both classes with respect to the inter-lobe and intra-lobe distribution. We have categorized the learned feature importance map with respect to their respective lobes and the results are visualized in Figure 11. The results suggest that for both classes, the frontal lobe carries a higher number of discriminant features in terms of both inter and intra-lobe features. The figure shows a different distribution for all of the lobes in both classes. The distribution is highlighted by the different shape of an individual lobe when comparing the two

classes. The frontal lobe is known to be involved with cognitive functioning [55]. This includes attention, the executive function that includes planning, selection, sequential organization and self-monitoring of actions, affect and mood, memory, self-awareness and personality [55]. The alterations in frontal lobe might cause abnormal behaviours in these functions including attention and mood. Studies have shown connectivity alterations in frontal, temporal, and occipital cortices locally as well as with the rest of the brain in individuals with ADHD [56]. Our findings about the frontal lobe alterations in ADHD support the results found in earlier studies [57, 58, 59].



(a) $I_{healthy}$. (b) I_{ADHD}.

Figure 6: Visualization of the learned feature importance map for a) healthy and, b) ADHD classes for the NYU dataset. For the sake of clarity, only top 100 values for an individual class are visualized. The visualization shows the differences in the feature maps of both classes.

7. Conclusions

In this paper, we have proposed an innovative end-to-end deep neural network for classification of ADHD from fMRI data. The proposed model takes pre-processed time-series signals of fMRI as input and learns to predict the classification label. We were interested to see if the classification task in fMRI can be solved by an

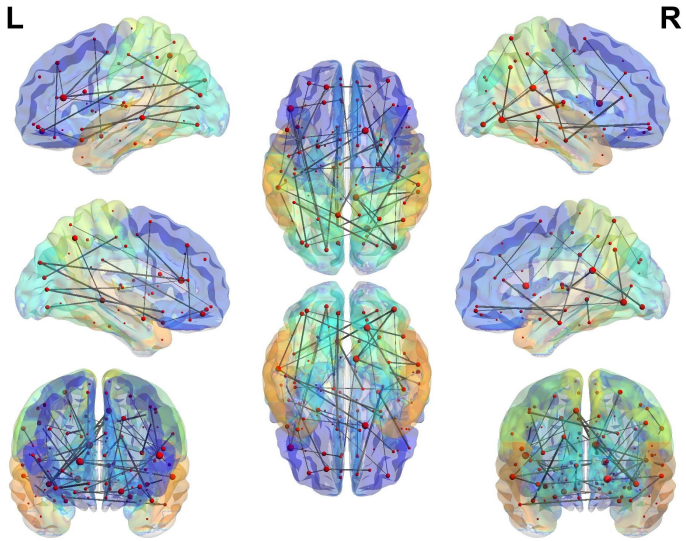


Figure 7: Visualization of the learned feature importance map for the healthy class on the brain volume. For the sake of clarity, only top 50 values are visualized. (Data visualized through the BrainNet viewer software [60]).

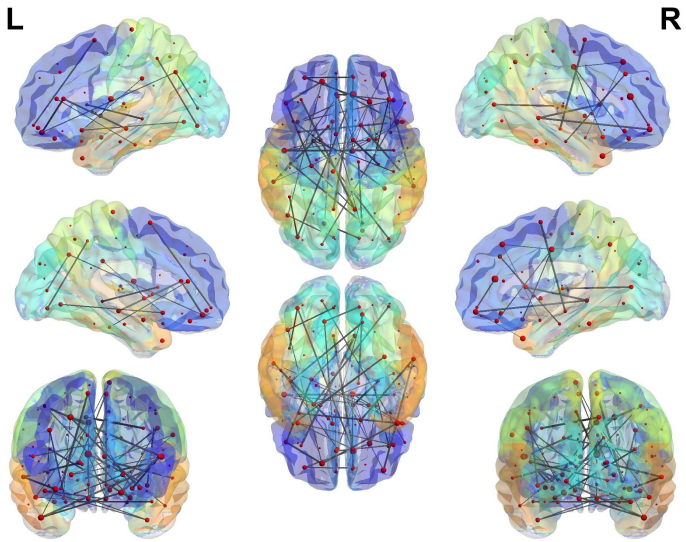


Figure 8: Visualization of the learned feature importance map for the ADHD class on the brain volume. For the sake of clarity, only top 50 values are visualized. Visualization through the BrainNet viewer software [60].

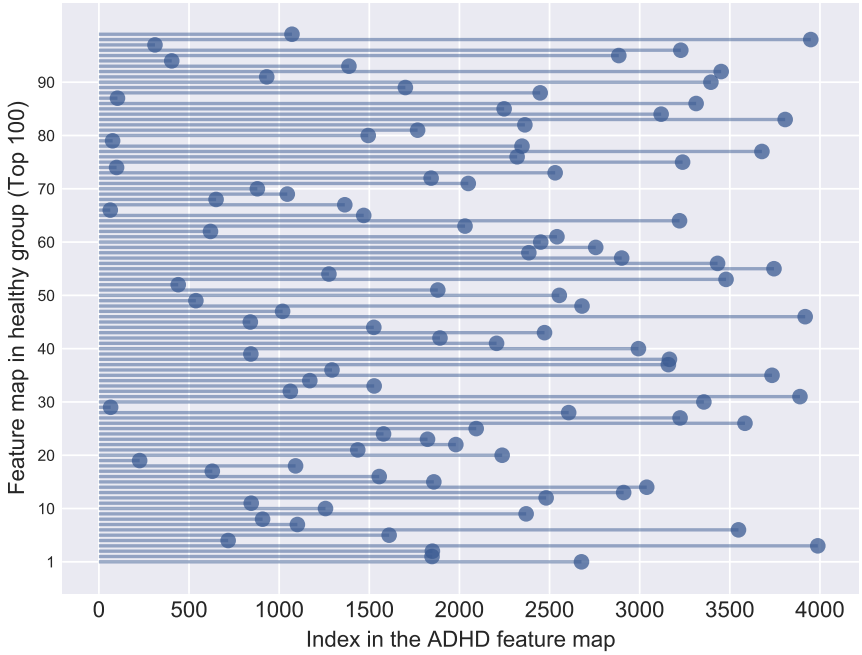


Figure 9: Plot of matching the top 100 healthy feature maps in the ADHD feature maps. The y-axis represents the top 100 feature maps in the healthy group and the x-axis represents the index of a particular healthy feature map in the ADHD feature map. The figure shows that out of top 100 feature maps of healthy class, less than 10% fall in top 500 feature maps in the ADHD class.

end-to-end network. As far as we know, this is the first attempt to apply an end-to-end network incorporating functional connectivity for classification of a neurological disorder.

We have evaluated the importance of functional connectivity in the proposed end-to-end network. Findings show that despite the large number of parameters in our method, it performs better as compared to an end-to-end network without functional connectivity with comparatively less number of trainable parameters. This result strengthens the argument that functional connectivity is an important biomarker towards the identification of a neurological disorder. Experimental results on the ADHD-200 dataset demonstrate that utilizing such a model outperforms the current state-of-the-art.

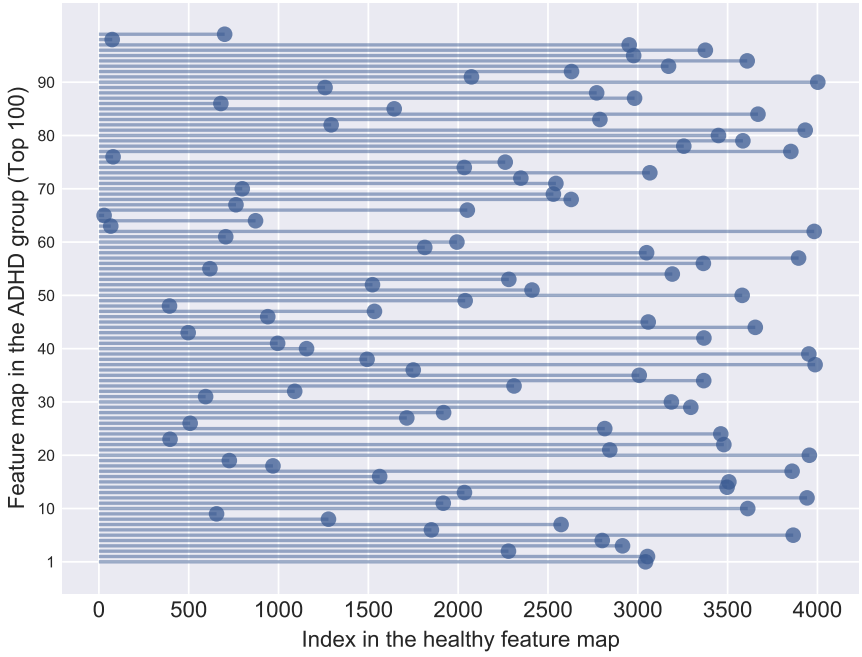
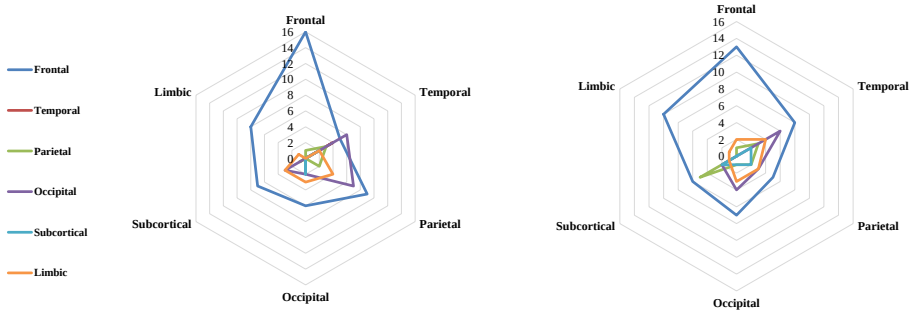


Figure 10: Plot of matching the top 100 ADHD feature maps in the healthy feature maps. The y-axis represents the top 100 feature maps in the ADHD group and the x-axis represents the index of a particular ADHD feature map in the healthy feature map. The figure shows that out of top 100 feature maps of the ADHD class, less than 10% fall in top 500 feature maps in the healthy class.



(a) Healthy class. (b) ADHD class.

Figure 11: Distribution of the top 100 features maps in the healthy and the ADHD classes.

Our proposed method is able to associate different weights to an individual functional connectivity with respect to its importance in predicting a class label (healthy control and ADHD). The proposed method appears to assign weight to different brain networks with respect to a particular class.

Our results suggest that the frontal lobe carries most discriminant power in classifying ADHD. The frontal lobe is known to be associated with cognitive functions like attention, memory, planning and mood. Our findings of the frontal lobe anomalies in ADHD concur with earlier studies results. One of the limitations of the proposed work is the small data size being evaluated. There is a smaller number of subjects in the individual imaging sites. Also, the data is very heterogeneous across different sites requiring training the network separately for each institution. The distribution of classes (healthy control and ADHD) in training and testing splits in the imaging sites is different, which makes the dataset very challenging for any classifier to achieve decent results. Due to these facts, the classification accuracy achieved by studies on this dataset is low and has room for improvement. We have compared results of the DeepFMRI and existing studies with a baseline classifier that simply chooses the majority class based on the testing data distribution. In the case of NYU, the baseline classifier accuracy is high (70.7%) and accuracy achieved by other studies are comparatively lower. However, the DeepFMRI method was able to achieve a slightly higher accuracy than the baseline, and was able to achieve much better accuracy than the baseline for the Peking and NI dataset.

In future work, we are interested to apply the proposed network to study ADHD sub-groups. The study might be based on treatment response, clinical scores, disorder outcomes etc. However, it will not require redesigning the proposed network. If the output variable is discrete, a classifier can be used as presented in this paper. If the output variable is continuous, instead a regressor can be used. We are curious to evaluate the proposed method on other disorders like epilepsy and Alzheimer's with a large number of subjects. We are also interested to explore whether the proposed method can be used to overcome variations in different imaging sites. One interesting experiment

could be to re-train the feature extractor network separately for each imaging site, keeping the parameters of the classification network and similarity measure network fixed. By this way possibly the proposed network could be adopted to different imaging sites.

References

[1] M. P. Van Den Heuvel, H. E. H. Pol, Exploring the brain network: a review on resting-state fMRI functional connectivity, European Neuropsychopharmacology 20 (8) (2010) 519–534.

[2] K. Li, L. Guo, J. Nie, G. Li, T. Liu, Review of methods for functional brain connectivity detection using fmri, Computerized Medical Imaging and Graphics 33 (2) (2009) 131–139.

[3] A. Riaz, E. Alonso, G. Slabaugh, Phenotypic integrated framework for classification of adhd using fmri, in: International Conference Image Analysis and Recognition, Springer, 2016, pp. 217–225.

[4] S. Dey, A. R. Rao, M. Shah, Attributed graph distance measure for automatic detection of attention deficit hyperactive disordered subjects, Frontiers in Neural Circuits 8.

[5] J. Kim, V. D. Calhoun, E. Shim, J.-H. Lee, Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia, NeuroImage 124 (2016) 127–146.

[6] K. Rajpoot, A. Riaz, W. Majeed, N. Rajpoot, Functional connectivity alterations in epilepsy from resting-state functional mri, PloS one 10 (8).

[7] A. Riaz, K. Rajpoot, N. Rajpoot, A connectivity difference measure for identification of functional neuroimaging markers for epilepsy, in: Neural

Engineering (NER), 2013 6th International IEEE/EMBS Conference on, IEEE, 2013, pp. 1517–1520.

[8] A. Riaz, M. Asad, E. Alonso, G. Slabaugh, Fusion of fmri and non-imaging data for adhd classification, Computerized Medical Imaging and Graphics 65 (2018) 115–128.

[9] A. dos Santos Siqueira, C. E. Biazoli Junior, W. E. Comfort, L. A. Rohde, J. R. Sato, Abnormal functional resting-state networks in ADHD: graph theory and pattern recognition analysis of fMRI data, BioMed Research International 2014.

[10] V. Kumari, E. R. Peters, D. Fannon, E. Antonova, P. Premkumar, A. P. Anilkumar, S. C. Williams, E. Kuipers, Dorsolateral prefrontal cortex activity predicts responsiveness to cognitive–behavioral therapy in schizophrenia, Biological psychiatry 66 (6) (2009) 594–602.

[11] S. P. Koch, C. Hägele, J.-D. Haynes, A. Heinz, F. Schlagenhauf, P. Sterzer, Diagnostic classification of schizophrenia patients on the basis of regional reward-related fmri signal patterns, PloS one 10 (3) (2015) e0119089.

[12] X. Wang, Y. Jiao, T. Tang, H. Wang, Z. Lu, Altered regional homogeneity patterns in adults with attention-deficit hyperactivity disorder, European journal of radiology 82 (9) (2013) 1552–1557.

[13] A. Riaz, M. Asad, S. M. R. Al-Arif, E. Alonso, D. Dima, P. Corr, G. Slabaugh, FCNet: A Convolutional Neural Network for Calculating Functional Connectivity from functional MRI, in: International Workshop on Connectomics in Neuroimaging, Springer, 2017, pp. 70–78.

[14] C.-Y. Wee, P.-T. Yap, D. Zhang, K. Denny, J. N. Browndyke, G. G. Potter, K. A. Welsh-Bohmer, L. Wang, D. Shen, Identification of MCI individuals using structural and functional connectivity networks, Neuroimage 59 (3) (2012) 2045–2056.

[15] M. Díez-Cirarda, A. P. Strafella, J. Kim, J. Peña, N. Ojeda, A. Cabrera-Zubizarreta, N. Ibarretxe-Bilbao, Dynamic functional connectivity in Parkinson's disease patients with mild cognitive impairment and normal cognition, NeuroImage: Clinical 17 (2018) 847–855.

[16] T. Wu, L. Wang, Y. Chen, C. Zhao, K. Li, P. Chan, Changes of functional connectivity of the motor network in the resting state in Parkinson's disease, Neuroscience letters 460 (1) (2009) 6–10.

[17] V. A. Harpin, The effect of adhd on the life of an individual, their family, and community from preschool to adult life, Archives of disease in childhood 90 (suppl 1) (2005) i2–i7.

[18] J. T. Nigg, Attention-deficit/hyperactivity disorder and adverse health outcomes, Clinical psychology review 33 (2) (2013) 215–228.

[19] A. Riaz, M. Asad, S. M. R. Al-Arif, E. Alonso, D. Dima, P. Corr, G. Slabaugh, Deep FMRI: An end-to-end deep network for classification of fMRI data, in: IEEE International Symposium on Biomedical Imaging (ISBI), IEEE, 2018.

[20] M. Nuñez-Garcia, S. Simpraga, M. A. Jurado, M. Garolera, R. Pueyo, L. Igual, Fadr: Functional-anatomical discriminative regions for rest fmri characterization, in: International Workshop on Machine Learning in Medical Imaging, Springer, 2015, pp. 61–68.

[21] H. Zou, T. Hastie, Regularization and variable selection via the elastic net, Journal of the Royal Statistical Society: Series B (Statistical Methodology) 67 (2) (2005) 301–320.

[22] A. Tabas, E. Balaguer-Ballester, L. Igual, Spatial discriminant ica for rs-fmri characterisation, in: Pattern Recognition in Neuroimaging, 2014 International Workshop on, IEEE, 2014, pp. 1–4.

[23] D. Dai, J. Wang, J. Hua, H. He, Classification of adhd children through multimodal magnetic resonance imaging, Frontiers in Systems Neuroscience 6 (2012) 63.

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

[25] J. W. Bohland, S. Saperstein, F. Pereira, J. Rapin, L. Grady, Network, anatomical, and non-imaging measures for the prediction of adhd diagnosis in individual subjects, Frontiers in Systems Neuroscience 6 (2012) 78.

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

[27] S. Huang, J. Li, L. Sun, J. Ye, A. Fleisher, T. Wu, K. Chen, E. Reiman, A. D. N. Initiative, et al., Learning brain connectivity of alzheimer's disease by sparse inverse covariance estimation, NeuroImage 50 (3) (2010) 935–949.

[28] R. S. Patel, F. D. Bowman, J. K. Rilling, A bayesian approach to determining connectivity of the human brain, Human brain mapping 27 (3) (2006) 267–276.

[29] A. E. et al., Automated diagnoses of attention deficit hyperactive disorder using magnetic resonance imaging, Frontiers in Systems Neuroscience 6 (2012) 61.

[30] W. Cheng, X. Ji, J. Zhang, J. Feng, Individual classification of adhd patients by integrating multiscale neuroimaging markers and advanced pattern recognition techniques, Frontiers in Systems Neuroscience 6 (2012) 58.

[31] J. Zhang, W. Cheng, Z. Wang, Z. Zhang, W. Lu, G. Lu, J. Feng, Pattern classification of large-scale functional brain networks: identification of

informative neuroimaging markers for epilepsy, PloS one 7 (5) (2012) e36733.

[32] S. Ghiassian, R. Greiner, P. Jin, M. R. Brown, Using functional or structural magnetic resonance images and personal characteristic data to identify adhd and autism, PloS one 11 (12) (2016) e0166934.

[33] B. Sen, N. C. Borle, R. Greiner, M. R. Brown, A general prediction model for the detection of adhd and autism using structural and functional mri, PloS one 13 (4) (2018) e0194856.

[34] R. Salvador, J. Suckling, M. R. Coleman, J. D. Pickard, D. Menon, E. Bullmore, Neurophysiological architecture of functional magnetic resonance images of human brain, Cerebral cortex 15 (9) (2005) 1332–1342.

[35] B. J. Frey, D. Dueck, Clustering by passing messages between data points, Science 315 (5814) (2007) 972–976.

[36] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.

[37] A. Payan, G. Montana, Predicting alzheimer's disease: a neuroimaging study with 3d convolutional neural networks, arXiv preprint arXiv:1502.02506.

[38] L. Zou, J. Zheng, C. Miao, M. J. Mckeown, Z. J. Wang, 3d cnn based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural mri, IEEE Access 5 (2017) 23626–23636.

[39] C. Hu, R. Ju, Y. Shen, P. Zhou, Q. Li, Clinical decision support for alzheimer's disease based on deep learning and brain network, in: 2016 IEEE International Conference on Communications (ICC), IEEE, 2016, pp. 1–6.

[40] G. Deshpande, P. Wang, D. Rangaprakash, B. Wilamowski, Fully connected cascade artificial neural network architecture for attention deficit

hyperactivity disorder classification from functional magnetic resonance imaging data, IEEE Transactions on Cybernetics 45 (12) (2015) 2668–2679.

[41] H.-I. Suk, C.-Y. Wee, S.-W. Lee, D. Shen, State-space model with deep learning for functional dynamics estimation in resting-state fMRI, NeuroImage 129 (2016) 292–307.

[42] S. Sarraf, G. Tofighi, Classification of alzheimer's disease using fmri data and deep learning convolutional neural networks, arXiv preprint arXiv:1603.08631.

[43] X. Li, N. C. Dvornek, X. Papademetris, J. Zhuang, L. H. Staib, P. Ventola, J. S. Duncan, 2-channel convolutional 3d deep neural network (2cc3d) for fmri analysis: Asd classification and feature learning, in: Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on, IEEE, 2018, pp. 1252–1255.

[44] Adhd-200 sample.
URL `http://fcon_1000.projects.nitrc.org/indi/adhd200/`

[45] Adhd-200 preprocessed.
URL `http://preprocessed-connectomes-project.org/adhd200/`

[46] R. W. Cox, Afni: software for analysis and visualization of functional magnetic resonance neuroimages, Computers and Biomedical research 29 (3) (1996) 162–173.

[47] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, et al., Advances in functional and structural mr image analysis and implementation as fsl, Neuroimage 23 (2004) S208–S219.

[48] X. Han, T. Leung, Y. Jia, R. Sukthankar, A. C. Berg, Matchnet: Unifying feature and metric learning for patch-based matching, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3279–3286.

[49] J. Bromley, I. Guyon, Y. LeCun, E. Säckinger, R. Shah, Signature Verification using a "Siamese" Time Delay Neural Network, in: Advances in Neural Information Processing Systems, 1994, pp. 737–744.

[50] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980.

[51] N. V. Chawla, K. W. Bowyer, L. O. Hall, W. P. Kegelmeyer, Smote: synthetic minority over-sampling technique, Journal of Artificial Intelligence Research 16 (2002) 321–357.

[52] K. Simonyan, A. Vedaldi, A. Zisserman, Deep inside convolutional networks: Visualising image classification models and saliency maps, arXiv preprint arXiv:1312.6034.

[53] C. Gan, N. Wang, Y. Yang, D.-Y. Yeung, A. G. Hauptmann, Devnet: A deep event network for multimedia event detection and evidence recounting, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 2568–2577.

[54] M. Olson, A. Wyner, R. Berk, Modern neural networks generalize on small data sets, in: Advances in Neural Information Processing Systems, 2018, pp. 3619–3628.

[55] C. Chayer, M. Freedman, Frontal lobe functions, Current neurology and neuroscience reports 1 (6) (2001) 547–552.

[56] L. Cocchi, I. E. Bramati, A. Zalesky, E. Furukawa, L. F. Fontenelle, J. Moll, G. Tripp, P. Mattos, Altered functional brain connectivity in a non-clinical sample of young adults with attention-deficit/hyperactivity disorder, Journal of Neuroscience 32 (49) (2012) 17753–17761.

[57] L. Wang, C. Zhu, Y. He, Y. Zang, Q. Cao, H. Zhang, Q. Zhong, Y. Wang, Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder, Human brain mapping 30 (2) (2009) 638–649.

[58] M.-g. Qiu, Z. Ye, Q.-y. Li, G.-j. Liu, B. Xie, J. Wang, Changes of brain structure and function in adhd children, Brain topography 24 (3-4) (2011) 243–252.

[59] M. V. Cherkasova, L. Hechtman, Neuroimaging in attention-deficit hyperactivity disorder: beyond the frontostriatal circuitry, The Canadian Journal of Psychiatry 54 (10) (2009) 651–664.

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