

UC Irvine

UC Irvine Previously Published Works

Title

Building Models of Functional Interactions Among Brain Domains that Encode Varying Information Complexity: A Schizophrenia Case Study

Permalink

<https://escholarship.org/uc/item/57r70698>

Journal

Neuroinformatics, 20(3)

ISSN

1539-2791

Authors

Batta, Ishaan
Abrol, Anees
Fu, Zening
[et al.](#)

Publication Date

2022-07-01

DOI

10.1007/s12021-022-09563-w

Peer reviewed



Published in final edited form as:

Neuroinformatics. 2022 July ; 20(3): 777–791. doi:10.1007/s12021-022-09563-w.

Building Models of Functional Interactions Among Brain Domains that Encode Varying Information Complexity: A Schizophrenia Case Study

Ishaan Batta^{*,1,3}, Anees Abrol¹, Zening Fu¹, Adrian Preda², Theo G.M. van Erp², Vince D. Calhoun^{1,3}

¹Center for Translational Research in Neuroimaging and Data Science (TReNDS): Georgia State University, Georgia Institute of Technology, and Emory University, Atlanta, USA

²Department of Psychiatry and Human Behavior, University of California Irvine, Irvine, California, USA

³Dept. of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, USA

Abstract

Revealing associations among various structural and functional patterns of the brain can yield highly informative results about the healthy and disordered brain. Studies using neuroimaging data have more recently begun to utilize the information within as well as across various functional and anatomical domains (i.e., groups of brain networks). However, most whole-brain approaches assume similar complexity of interactions throughout the brain. Here we investigate the hypothesis that interactions between brain networks capture varying amounts of complexity, and that we can better capture this information by varying the complexity of the model subspace structure based on available training data. To do this, we employ a Bayesian optimization-based framework known as the Tree Parzen Estimator (TPE) to identify, exploit and analyze patterns of variation in the information encoded by temporal information extracted from functional magnetic resonance imaging (fMRI) subdomains of the brain. Using a repeated cross-validation procedure on a schizophrenia classification task, we demonstrate evidence that interactions between specific functional subdomains are better characterized by more sophisticated model architectures compared to less complicated ones required by the others for optimally contributing towards classification and understanding the brain's functional interactions. We show that functional subdomains known to be involved in schizophrenia require more complex architectures to optimally unravel discriminatory information about the disorder. Our study points to the need for adaptive, hierarchical learning frameworks that cater differently to the features from different subdomains, not only for a better prediction but also for enabling the identification of features predicting the outcome of interest.

*Corresponding Author: Ishaan Batta, ibatta@gatech.edu.

INFORMATION SHARING STATEMENT

The datasets used in this study were published by (Keator et al., 2016) (fBIRN) and (Aine et al., 2017) (COBRE) respectively. The COBRE dataset can be directly downloaded from http://fcon_1000.projects.nitrc.org/indi/retro/cobre.html. Python codes for implementing BrainNetCNN and FNN frameworks were used from https://github.com/ThomasYeoLab/Standalone_He2019_KRDNN/. Preprocessing softwares can be found at <https://trendscenter.org/software/>

Keywords

Multilayer Perceptron; Bayesian Optimization; Hyperparameter Optimization; Schizophrenia; fMRI; Functional Connectivity; Subdomain Analysis

1 INTRODUCTION

Numerous works have studied brain disorders by employing multilayered machine learning (ML) approaches to neuroimaging data. In most cases, the main focus of these studies is to increase the accuracy with which the subjects having a certain condition can be classified from unaffected controls. However, not many studies focus on probing the predictive power of features encoded by the neuroimaging data. In addition to improving the classification capabilities of developed classifiers, it is equally important to localize the brain regions or groups of brain regions (subdomains) that are the most discriminative for a given disorder. Recently, deep learning classifiers have been applied widely in studies involving the use of neuroimaging data for classification. So far, previous work has used the desired features directly as a single input to the learning framework without any subdomaining of the feature set (Srinivasagopalan et al., 2019; Ulloa et al., 2015; Han et al., 2017). The prevailing methodologies inherently employ two implicit assumptions which may not necessarily hold, namely a homogeneity in the feature set for use as a single input set and a high uniformity in the complexity of feature interactions between subdomains, ignoring the need for flexible architectures. Mainstream architectures based upon these assumptions leave little room for interpretation because the parameters learned in multi-layered learning models are non-linear combinations of the input features. While using branched architectures can be of help with interpretation in terms of subdomains in the data, various studies which use multi-branched architectures on neuroimaging data have mainly focused only on accuracy improvement rather than interpretability and used these architectures to cater to the multimodal (Ulloa et al., 2015, 2018) and even multi-atlas scenarios (Zeng et al., 2018). Even in most of these cases, there is little flexibility in the architectures in terms of the depths of the branches. Introducing such flexibility in model complexity of multi-branched architectures allows for studying subdomains in terms of the nature of information they encode towards discriminating between two or more groups of subjects. No studies have explored the use of architectures designed to treat subdomains in the data differently and also reflect the variability with which different subdomains in the data encode predictive information. To overcome these limitations, this study introduces flexible ML architectures that take into account the variations in the complexity of interactions between subdomains in the data. Our approach demonstrates the need for as well as the benefit of disengaging the implicit assumption of feature homogeneity and uniform complexity in the nature of predictive information.

It is vital to identify interactions of functional and anatomical networks of high predictive value towards certain targeted applications. Moreover, it is equally important to study the diversity in the way this predictive information is encoded in the data with respect to the anatomical and functional subdomains of the brain. As mentioned before, the latter variation necessitates the use of architectures that are capable of incorporating inputs with

varying information complexities in the data subdomains. Given the complex manner in which interactions between various regions of the brain occur, it is reasonable that certain subdomains of the brain may need deeper architectures for better prediction, which may be indicative of a greater degree of non-linear interactions in these subdomains. In this study, we analyze the pattern in which various subdomains for functional magnetic resonance imaging (fMRI) data encode information in a schizophrenia classification task. We use a multilayer perceptron (MLP) classifier for studying the functional connectivity features associated with schizophrenia classification. Towards this goal, we use the intra-network and inter-network connections at the level of subdomains, termed as subdomain interactions (SDIs) in this paper to create separate input layers in the multi-branched architecture. The input layers in our multi-branched architecture are followed by a variable number of hidden layers in each of the branches before a late fusion step as detailed in the methods section. By optimizing over this flexible multi-branched architecture search space, we show that different subdomain interactions encode discriminatory information with variable complexity and certain subdomain interactions associated with schizophrenia consistently need more complex frameworks while others require simpler ones.

One of the potential concerns with allowing such flexibility in the depth is the vast architecture search space generated due to the multiple parameters involved. Performing optimization over this space is a hyperparameter optimization problem, a well studied field in machine learning (Shahriari et al., 2015; Luo, 2016; Bergstra et al., 2013b, 2011). It is computationally impractical to linearly traverse through exponentially huge hyper-parameter search spaces with standard methods such as random search or grid search. Bayesian optimization frameworks resolve this issue by heuristically traversing parts of the search space that are more likely to cover a solution close to the optimal one. Here we employ a Bayesian method known as the Tree Parzen Estimator (TPE) to realize the hyper-parameter optimization stage (Bergstra et al., 2011). Starting with a set of initial randomly chosen points in the search space, the TPE algorithm traverses new points in each of its iterations using a simpler (i.e., faster) surrogate function and calculating a metric for the expected improvement in classification accuracy subsection 2.4. In this way, the search space is traversed without having to compute the actual function (in this case, the validation accuracy of the architecture) and selecting a new point which would be more optimal with high probability. Analyzing the final architecture returned by the TPE procedure can reveal specific associative patterns corresponding to each subdomain. By studying the patterns in the associations of subdomains in the optimized architectures, we illustrate that allowing for and optimizing over subdomain specific variation in architectures not only enables superior prediction, but also reveals how certain subdomain interactions bear higher complexity of information while others have less complex information. Moreover, with the rapid increase in multimodal datasets for mental disorders, it is crucial to develop methodologies synthesizing features from subdomains spanning more than one modality. Our study also gives initial insights towards the need for developing such flexible frameworks for multimodal studies. Frameworks catering to and analyzing subdomain variation can be of immense use in various fields of study.

The flow of the paper, shown in Figure 1 and detailed in methods section, is as follows:

a) Spatially constrained independent component analysis (scICA) on the preprocessed

fMRI dataset is used to calculate the functional connectivity for pairs of components of interest subject-by-subject; b) Categorizing the feature set of component-component functional connectivity into subsets based on subdomains (brain networks) of the two components involved (the inter-network or inter-network connections are termed as subdomain interactions (SDIs) throughout the paper); c) Using the subdomained features as inputs to the multi-branch MLP architecture with flexible depth, with hyper-parameter optimization (TPE algorithm) on the architecture search space to determine the optimal variability in depth for each SDI (Figure 1b); d) Comparison of the performance of TPE with existing learning methods; (e) Analysis of patterns associated with certain SDIs in the validated optimal architectures returned by the TPE procedure. We demonstrate the working, performance and interpretation of the TPE algorithm in the results section. Lastly, we discuss the interpretation of these results in the discussion section.

2 METHODS AND MATERIALS

2.1 Datasets and Pre-Processing

This study uses two independent datasets, the Function Biomedical Informatics Research Network (fBIRN) dataset (Keator et al., 2016) and the Center of Biomedical Research Excellence (COBRE) dataset (Aine et al., 2017). Subjects with large head motion (> 3 deg and > 3 mm) during the scan and with functional data leading to bad full brain normalization were excluded. After applying this criteria for exclusion, the fBIRN dataset consisted of 160 healthy controls (HC) with age 19 – 59 years (mean 37.04 ± 10.86), 45/115 females/males and 151 subjects who had schizophrenia (SZ) with age 18–62 years (mean 38.77 ± 11.63), 36/115 females/males, whereas the COBRE dataset consisted of 89 healthy controls (HC) with age 18–65 years (mean 38.09 ± 11.67), 25/64 females/males and 68 SZ with age 19–65 years (mean 37.79 ± 14.45), 11/57 females/males. For avoiding confounding effects, HC and SZ subjects of both datasets were matched by age and gender (age: $p = 0.1758$ (fBIRN), 0.8874 (COBRE); gender: $p = 0.3912$ (fBIRN), 0.0794 (COBRE)).

Data collection for fBIRN dataset was done using 3-T Siemens Tim Trio scanners for six out of seven sites and 3-T General Electric Discovery MR750 scanner for one site. The same resting-state parameters were used across the scanners a standard gradient echo-planar imaging (EPI) sequence, repetition time (TR)/echo time (TE) = 2000/30 ms, voxel spacing size = $3.4375 \times 3.4375 \times 4$ mm, slice gap = 1 mm, flip angle (FA) = 77° , field of view (FOV) = 220×220 mm, number of excitations (NEX) = 1, and number of volumes = 162. Participants had their eyes closed and were instructed to rest quietly during the scanner. The COBRE data was collected on a single site using a 3-T Siemens Trio scanner. For the COBRE data, a gradient-echo EPI sequence was used to acquire T2-weighted functional images with the following parameters: TE = 29 ms, TR = 2000 ms, flip angle (FA) = 75° , slice thickness = 3.5 mm, slice gap = 1.05 mm, field of view 240 mm, matrix size = 64×64 , voxel size = $3.75 \text{ mm} \times 3.75 \text{ mm} \times 4.55 \text{ mm}$ and number of volumes = 149. For the duration of the scan, subjects were instructed to keep their eyes open and passively stare at a central cross.

For both the datasets, the statistical parametric mapping (SPM12, <http://www.fil.ion.ucl.ac.uk/spm/>) toolbox based on the Matlab 2016 platform was used to

preprocess fMRI data. To ensure equilibrium of the signal and adaptation of the subjects to scanner noise, the first five scans were removed. The SPM toolbox was used for performing slice timing correction and rigid body head motion correction. Warping of the fMRI data into the standard Montreal Neurological Institute (MNI) space was done using an echo-planar imaging (EPI) template. The data were slightly resampled to $3 \times 3 \times 3 \text{ mm}^3$ isotropic voxels. For smoothing the data, a Gaussian kernel with a full width at half maximum (FWHM) of 6 mm was used. This was followed by further feature extraction detailed in subsequent sections.

2.2 Connectivity Features and Subdomain Interactions

To estimate intrinsic connectivity networks from the fMRI data, the scICA approach described in (Du and Fan, 2013) with the Neuromark (Du et al., 2019a,b) template as reference maps was used. ICA was run for a model order of 100, out of which 53 consistent and reproducible independent non-artifactual components (ICs) were retained. The functional connectivity matrix was computed based on the static correlation between the time courses of pairs of two ICs. All the 53 ICs have been arranged into 7 distinct functional subdomains (disjoint exhaustive sets of ICs), namely the areas falling under the following subdomains: default mode network (DMN), visual (VIS), auditory (AU), cognitive control (CC), sensorimotor (SM), cerebellar (CB) and sub-cortical (SC). Based on this categorization, 28 subdomain interaction (SDI) features were created where each SDI represents the set of intra-network or inter-network connections between all possible pairs of ICs for a given pair functional subdomains or pair of functional subdomains respectively (See Figure 2). As an example, the features in the SDI named DMN-VIS refer to entries from the functional connectivity matrix (fMRI time-series correlations) corresponding to the connections between the DMN and VIS subdomain scICA components. Such connections are termed as inter-network connections. On the other hand, the intra-network connections comprise connections within the same domain, for example, the SDI named DMN-DMN would correspond to the connections where both of the scICA components belong to the DMN subdomain. Thus, connectivity features for the 28 SDIs consist of 21 sub-matrices for the inter-network connections and 7 sub-matrices for intra-network connections. Hence the SDI features are simply sub-matrices of the 53×53 functional connectivity matrix created using 53 scICA components.

2.3 Architecture Search Space

A multilayer perceptron (MLP) architecture with multiple branches was used with input layer consisting of the 28 SDI features from the 7 subdomains (sets of ICs) as described in subsection 2.2. At the input level, the architecture has 28 branches for each SDI i.e., a sub-matrix of the 53×53 static functional connectivity matrix. Each input layer is then fed into a variable number of fully connected hidden layers and a late fusion layer, finally followed by a fully connected layer (Figure 1b). In each branch, a constant factor of 0.1 was used to reduce the number of units in subsequent hidden layers, starting with the input layer. Optimizing the variation in the number of fully connected layers in the branch corresponding to each SDI can give insights into understanding the nature of information that is encoded by the component interactions for the subdomains involved in the SDI. SDIs which consistently require no or a small number of layers in an optimized architecture can

be said to contain more linear or direct information for the task at hand (schizophrenia classification), whereas the SDIs which require a higher number of fully connected layers can be said to contain more complex and indirect information. For controlling the model complexity, the number of fully connected hidden layers in each branch were limited to vary between 0 to 2. In the exponential search space thus generated with a total of 3^{28} possible configurations, a 28 length vector \mathbf{x} can be used to represent each point, with each element $x_i \in \{0, 1, 2\}$ denoting the number of fully connected layers in the branch for the i -th SDI (Figure 1b).

2.4 Tree-structured Parzen Estimator (TPE) for Hyper-Parameter Optimization

Optimizing hyper-parameters that span an exponential search space has been a well-studied problem in machine learning. In similar lines, Bergstra et al. (2011) had proposed the TPE algorithm which is used for hyper-parameter optimization in this work. TPE is a sequential model-based optimization (SMBO) framework which essentially estimates the conditional and marginal probabilities $p(x|y)$ and $p(y)$ to perform hyper-parameter optimization on hyper-parameters x and cost function y . After spanning an initial set of points selected randomly from the search space and computing the cost function for them, TPE utilizes this information to create a surrogate cost function. The surrogate function is then used to select the next set of points in the traversal sequence based on the estimates of the cost function. Towards this goal, two groups made up of the upper quartile and the lower quartile are created based on a splitting value y^* of the cost function at the randomly selected points. The two probability density functions, $g(x)$ and $l(x)$ for the upper and lower quartiles of the cost function respectively is then defined as in Equation 1:

$$p(x | y) = \begin{cases} g(x) & \text{if } y > y^* \\ l(x) & \text{if } y \leq y^* \end{cases} \quad (1)$$

Having estimated $l(x)$ and $g(x)$, the subsequent iterations of the TPE algorithm involve the optimization of the Expected Improvement function defined in Equation 2:

$$EI_{y^*}(x) = \int_{-\infty}^{y^*} (y^* - y)p(y | x)dy \propto g(x)/l(x) \quad (2)$$

The expected improvement function $EI_{y^*}(x)$ can be shown to be proportional to $l(x)/g(x)$ (Bergstra et al., 2011). This means that if points are sampled from $l(x)$ with high probability and from $g(x)$ with lower probability, then the expected improvement is maximized, thus minimizing the cost function. In each iteration, the algorithm returns the point x^* having the largest value of $EI_{y^*}(x)$ and the distributions $l(x)$, $g(x)$ are updated accordingly. An example showing how TPE works is provided in Figure 1c and indicates convergence towards the true optimal on a quadratic cost function $(x - 1)^2$ optimized for a real valued parameter x .

2.5 Analysis of Learned Models

Using features from each SDI as inputs, the MLP architecture with 28 branches as described in subsection 2.3 was optimized using the TPE implementation in the HyperOpt python library (Bergstra et al., 2013a). We used a repeated random sub-sampling cross-validation

procedure for running the TPE algorithm for 50 repetitions on the functional connectivity data, resulting in a set of optimized architecture vectors $\{\mathbf{x}^{(r)}\}_{r=1}^{50}$. The mean across repetitions of the validation accuracy has been plotted against TPE iterations in Figure 5.

The final prediction model, represented by the vector \mathbf{x}^* , was created using the most frequent number of hidden layers across repetitions for each SDI in the TPE-optimized architectures. The final architecture vector \mathbf{x}^* is defined as, $\{x_i^*\}_{i=1}^{28} = \left\{ \operatorname{argmax}_{l \in \{0, 1, 2\}} \sum_{r=1}^{50} \mathbb{1}\{x_i^{(r)} = l\} \right\}_{i=1}^{28}$, where $\mathbf{x}^{(r)}$ represents the TPE-optimized architecture for the r^{th} repetition and $\mathbb{1}\{\cdot\}$ is the indicator function. The multi-branched MLP architecture generated using the final architecture vector \mathbf{x}^* was run on a held-out test data for 50 repetitions, resulting in 50 test accuracy values. It should be noted that while the random splits for training and testing were different for each of the 50 repetitions, the same random splits were used across all methods for a given repetition to avoid any unwanted differences in results due to different training data.

2.6 Performance Comparison with Classification Methods

To study the performance of the TPE-optimized architecture, a comparison was made with baseline methods run for 50 repetitions on held-out test data with 20% of the total number of samples. These methods included standard machine learning models, neural network based models, and multi-branched architectures with as well as without flexible branch-depth. The set of methods and corresponding parameters used are detailed subsequently.

2.6.1 Standard Machine Learning Models—Standard classification models used for comparison include logistic regression (LOG), support vector machine (SVM) with radial kernel, and random forest classifier (RFC). The learning parameters like regularization strength for logistic regression and SVM were optimized using grid search with 5-fold repeated cross validation using the training data.

2.6.2 Non-Branched Neural Network Architectures—In addition to the standard models, we used neural network based architectures including the multilayer perceptron (MLP) and a feed-forward neural network with encoder-decoder architecture (FNN) (LeCun et al., 2015) as shown in Figure 3 (a),(b). Both the networks involve the vectorized resting state functional connectivity matrix as the input features. While MLP involves fully connected hidden layers of decreasing size between input and output layers, FNN uses an encoder-decoder scheme with an initial set of fully connected hidden layers of decreasing size (encoding layers) followed by a set of layers of decreasing size (decoding layers). For both MLP and FNN, the depth of the network was tuned between single, double and triple layered architectures along with tuning of other network parameters. The depths corresponding to the highest accuracy score were selected (double layered for MLP and triple layered for FNN).

Additionally, we also employed the BrainNetCNN architecture (also denoted as BCNN in this paper) recently introduced by (Kawahara et al., 2017) as a non-branched neural network based method for comparison. Unlike MLP and FNN, BrainNetCNN utilizes

the mathematical structure of the functional connectivity matrix. As shown in Figure 3 (c), BrainNetCNN feeds the functional connectivity matrix input to Edge-to-Edge (E2E), Edge-to-Node (E2N), Node-to-Graph (N2G) layers followed by a fully connected linear layer before computing the output for classification (Kawahara et al., 2017). The E2E layer connects each element from a particular filter from the incoming input matrix to an element in the outgoing matrix of the same size. The E2N layer connects each diagonal entry of the incoming matrix to an entry in the outgoing vector, while the N2G layer is a fully connected hidden layer feeding into subsequent fully connected linear layers that lead to the output layer for classification. BrainNetCNN has been employed by various studies (He et al., 2020; Pervaiz et al., 2020; Li et al., 2018) for both classification and regression problems and is known to perform while reducing the parameter space due to the convolutional layers. We used the python code provided by (He et al., 2020) for the implementation of BrainNetCNN in the context of this study.

2.6.3 Branched Neural Network Architectures without Flexibility—The standard machine learning models and non-branched architectures do not involve any subdomaining of the input features as done in the TPE procedure. For an even closer comparison, we also used three types of branched multilayer perceptron architectures without any tuning or flexibility in depths of individual branches i.e. uniform depth for each branch, unlike the TPE framework. These are denoted in the paper by UNIF0, UNIF1, UNIF2 based on the constant depths of the branches before a fully connected fusion layer (See Figure 3 (d)). Note that the MLP, FNN and BCNN architectures described previously have all the static functional connectivity features in the input layer, which are further fed to subsequent layers without any branching. While these non-branched architectures serve as a baseline to compare multi-branched vs. non-branched architecture scenarios, the UNIF architectures serve as a baseline for comparing uniform vs. flexible depth in the branch corresponding to each SDI (Figure 3). Thus, against the architectures $\mathbf{x}^{(r)}$'s returned by TPE that have variable values of layer-depth, the UNIF architectures essentially is an architecture in the same architecture search space of the TPE algorithm, but with a constant layer-depth of 1, 2 or 3 before the late-fusion step. Therefore, the depth-vector for UNIF1 is $\mathbf{u}_1^{(r)} = \{u_{1i}^{(r)}\}_{i=1}^{28}$ with $u_{1i}^{(r)} = 1 \forall 1 \leq i \leq 28$, where $u_{1i}^{(r)}$ represents the number of layers used on top of i^{th} SDI in the architecture UNIF1 for the r^{th} repetition. Thus, the depth-vector for the UNIF2 architecture is $\mathbf{u}_2^{(r)} = \{u_{2i}^{(r)}\}_{i=1}^{28}$ where $u_{2i}^{(r)} = 2 \forall 1 \leq i \leq 28$. Similarly, the architecture UNIF0 is designed to have no layer on top of each SDI branch before the late fusion step in the multi-branch MLP architecture i.e., the UNIF0 architecture's depth-vector is $\mathbf{u}_0^{(r)} = \{u_{0i}^{(r)}\}_{i=1}^{28}$ where $u_{0i}^{(r)} = 0 \forall 1 \leq i \leq 28$.

2.6.4 Branched Neural Network Architectures with Non-Uniform Branch-Depth—In addition to branching, a significant feature in the TPE framework is the optimization of over the variable branch-depth. By using a Bayesian approach for this optimization, TPE arrives at a resultant architecture in which the depths for all branches need not be uniform. However, to compare the performance with additional baseline hyper-

parameter optimization methods, we used random search (RNDS) and grid search (GRDS) as possible replacements for the TPE optimization procedure. Essentially, RNDS and GRDS tuning procedures are the most similar to TPE than the other aforementioned methods in terms of the way the architecture is structured and optimized. The performances were measured on held-out test data using architectures optimized with the same number of iterations for RNDS, GRDS, and TPE.

3 RESULTS

3.1 scICA components

After running the scICA framework on the fMRI data, 53 non-artifactual, reproducible independent components (ICs) were obtained. Table 1 shows the peak coordinates for the components along with the brain region corresponding to the component. Following this, static correlations between the time courses of these 53 components were computed resulting in a 53×53 functional connectivity matrix. The 53 ICs were assigned to 7 functional subdomains or sets of components (Table 1, Figure 4). 28 pair-wise subdomain interactions (SDIs) resulting from the 7 subdomains were constructed. The SDIs are defined by the sets of intra-network or inter-network connections between all possible pairs of subdomains; thus, there were 21 SDIs corresponding to inter-network connections and 7 for the intra-network connections (Figure 2).

3.2 Performance using TPE

Starting with an initial set of randomly selected points in the hyper-parameter search space, the TPE optimization procedure subsequently selects new points to traverse using the expected improvement metric and the surrogate cost function which is computationally faster (subsection 2.4). Figure 5 shows the validation accuracy plotted against time (iterations). The validation accuracy, which is the objective function being optimized, increases with the TPE iterations. The procedure returns an optimal architecture for each of the 50 repetitions. The most frequently occurring number of hidden layers for each SDI across repetitions are used to define the final architecture. The test accuracy values on held-out data were obtained for 50 independent repetitions by using the final architectures constructed after running the TPE optimization algorithm (Figure 6). The final TPE models reported a slightly higher prediction accuracy on held-out test data than the baseline standard machine learning methods (SVM, logistic regression and random forest classifier). Specifically, we observed a mean prediction accuracy of 0.81 for fBIRN and 0.78 for COBRE for the proposed TPE-optimized architecture. For both the datasets, TPE also performs significantly better ($p < 0.05$) than the non-branched neural network based methods (MLP, FNN, BCNN) as well as multi-branch architectures with uniform depth in each branch (UNIF0, UNIF1 and UNIF2). Interestingly, the TPE optimization procedure also results in significantly better ($p < 0.05$) test accuracy than other optimization techniques for branched architectures with variable branch-depth (RNDS, GRDS). The results have been summarized in Table 2 as well as Figure 6. The results show that treating features from certain subdomains differently in terms of the complexity of architecture required can result in superior prediction performance. In fact, this improvement over the baseline methods also

allows for a meaningful analysis of variability in the nature of information carried by the various subdomain interactions in the data.

3.3 Feature Stability

While the TPE algorithm gives the best performance, the next step would be to analyze the subdomain interactions (SDIs) requiring deeper or shallower hidden layers in the multi-branched MLP architecture. However, before checking for this characteristic variation in the SDIs, it is relevant to check whether the features learned by the optimized architecture have consistent discriminative power for each SDI in terms of their importance towards prediction. For this purpose, impurity-based feature prediction power (Louppe et al., 2013), also known as Mean Decrease Impurity (MDI), was computed by running random forest classifier on the parameters learned in the late fusion layer (Figure 1b) of the optimized TPE architecture. These were compared with the impurity-based prediction power by running random forest classifier on the features in the input layer of the network (i.e., the functional connectivity features). The comparison was made by obtaining the prediction power vectors for both the aforementioned cases and computing the cumulative prediction power for elements belonging to each subdomain interaction (SDI). The cumulative prediction power of each SDI for the prediction task, averaged over 50 repetitions of the algorithm on held-out test data is plotted in Figure 7. The results indicate that the prediction power of SDI parameters learned in the late fusion step of the final architecture is highly correlated (0.9 for fBIRN and 0.81 for COBRE) with the prediction power of the same SDI features in the functional connectivity input features. This observation suggests that the TPE-optimized architecture learned appropriate brain representations as the SDI features retained similar properties in terms of the prediction accuracy on the schizophrenia classification task as well as the importance of each SDI in the prediction. With these properties being similar, the TPE procedure that optimizes based on the depth by each SDI in the multi-branched architecture, can be said to additionally provide insights into the complexity of the information stored in the SDIs. The results from the analysis of this variation in the nature of information across the SDIs (as defined in subsection 2.5) is done in the subsequent subsection.

3.4 Analyzing Optimal Models

After having ensured that the TPE procedure for optimizing the multi-branched MLP architecture is consistent in terms of prediction accuracy on held-out test data and also shows a similar trend in the importance of SDIs, an analysis on the variation in depth required by various SDIs in the final architecture returned by TPE was done as described in subsection 2.5. This was done by considering the most frequently occurring number of fully connected layers across all repetitions of TPE for each SDI. The relationship between 7 subdomains (networks) in terms of the number of fully connected layers needed for optimal decoding of information in each pair of networks (SDI) in the final architecture is shown in Figure 8. A notable observation here is that in the final architecture of both the fBIRN and COBRE datasets, the number of fully connected layers required by each SDI is the same for 19 out of the 28 SDIs (Figure 9), which indicates a very similar common pattern across the two independent datasets in terms of complexity of information in subdomain interactions. Note that the number of such common values across two datasets can be modelled using a random variable which is distributed according to the binomial distribution $B(k, n, p)$, where

k is the number of successes (=19), n is the number of trials (=28) and p is the probability of a success, which in this case is the probability of having the same depth in both datasets for a particular SDI, given by $\frac{3}{3^2} = \frac{1}{3}$. A binomial test was done to check whether the observed number of common values (= 19) is significantly different from the expected number (= 28/3) in the case of the binomial distribution. This test indicates that the result is significant with a p -value of 2.09×10^{-4} .

3.5 Biological Interpretation

Using the TPE framework, functional connectivity features between brain regions belonging to various functional subdomains were used to analyze them in terms of the complexity involved for uncovering predictive information for schizophrenia classification. While most deep learning architectures do not allow for analyzing specific region-wise functional connectivity input features due to the mixing up in subsequent hidden layers, the TPE framework mixes the region-level features only within the branch corresponding to the functional subdomains that the regions belong to. Thus, the results, upon optimization for complexity required by various subdomain interactions, can be interpreted in terms of the functional connectivity subdomains. Figure 9 summarizes these findings in terms of the 7 functional connectivity subdomains involving brain regions belonging to default mode network (DMN), visual (VIS), auditory (AU), cognitive control (CC), sensorimotor (SM), cerebellar (CB) and sub-cortical (SC) areas.

With schizophrenia being a cognitive disorder involving audio-visual hallucinations occurring as if being sensed in real even while the subject is not actively engaged in a task, it is interesting to note that the same is reflected by the functional subdomains involved in these functions (CC, DMN, SM, VIS, AUD) requiring higher complexity models. Cognitive control deficits are well known to be prevalent in patients diagnosed with schizophrenia, with cognitive control known to regulate a lot of other cognitive systems including behavioral disorganization (Lesh et al., 2011). According to the cognitive control model, failures to retrieve episodic memories are mainly mediated through frontal areas involved in cognitive control (Ragland et al., 2009). Additionally, previous studies have revealed that impaired connectivity in the DMN is associated with reduction in capacity to effectively retrieve episodic memories and process self referential information (Dunn et al., 2014; Kim, 2010). Thus, the associations between CC and DMN domains are evident from these observations. The connectivity of the DMN regions with the SM and VIS regions is also known to be affected in schizophrenia (Wang et al., 2015). The role of connectivity between regions belonging to the SM, VIS and CC subdomains is also known in the case of schizophrenia (Kaufmann et al., 2015; Javitt and Freedman, 2015). Previous studies have found significant differences in the connectivity within the SM/VIS regions as well as their connectivity with the CC regions (Kaufmann et al., 2015; Javitt and Freedman, 2015; Butler and Javitt, 2005; Gaebler et al., 2015). Many studies have shown that inter-subdomain connectivity is affected in schizophrenia for the CC, VIS, SM, DMN and AUD areas (Javitt, 2009; Kim et al., 2009). This is in line with the observation that the inter-subdomain connections between regions from all these subdomains require more complexity (Figure 9).

4 DISCUSSION

In this work we focus on developing a new approach for modeling differing values of variation and complexity with which information is encoded in the functional subdomain interactions (SDIs). We evaluate this approach in the context of a schizophrenia classification problem. The results show that allowing for differing subspace complexity can result in improved performance of the models (Figure 6), and enables us to identify meaningful differences in the modeling of different subdomains in the data, effectively using the resulting model complexity to determine whether their data contains more or less complex information about the prediction (Figure 9), given the model performance is at par with baseline frameworks. This trend of variation in the nature of information encoded in various functional subdomain interactions is significantly consistent (p -value = 2.09×10^{-4}) across the independent datasets used in this study, our findings showed that 19 out of the 28 SDIs had the same depth in the optimized architectures for both the datasets.

Notably, it can be observed that the connections from the cognitive control (CC) to the visual (VIS) and sensorimotor (SM) subdomains require deeper models for both datasets (Figure 9(a)), indicating more complex models are needed to capture links between higher-order cognition (CC) and lower-order sensory areas (VIS, SM). Connectivity of components from SM, VIS and CC subdomains is well known to be affected in schizophrenia (Kaufmann et al., 2015; Javitt and Freedman, 2015). In fact, differences have been observed for both the focal (within SM and VIS components) as well as distal (between SM/VIS and CC components) connectivity in schizophrenia (Kaufmann et al., 2015; Javitt and Freedman, 2015; Butler and Javitt, 2005; Gaebler et al., 2015). The observation of auditory (AUD) subdomain to VIS subdomain connections requiring deeper models is interesting given previous work highlighting the disruptions in these areas in schizophrenia (Lynall et al., 2010; Calhoun et al., 2009; Gallinat et al., 2002; Rotarska-Jagiela et al., 2010; Yu et al., 2012) are connected in a complex manner and are implicated in auditory hallucinations and visual saccades, both of which are disrupted in schizophrenia, respectively. Changes in connectivity between the default mode network (DMN) and cognitive control (CC) areas are also well known, especially between the precuneus and the prefrontal cortex (Wolf et al., 2011). Interestingly, all SDIs involving the functional connectivity between components from the same network (i.e., self-connections in Figure 9) require shallower models, indicating that connections within a given network may encode less complex information compared to SDIs with inter-network connections, even if the former (intra-network connections) are known to be involved in schizophrenia (Garrity et al., 2007).

The TPE-optimized architectures perform significantly better than the corresponding architectures that do not allow for any variation in the depth required for various SDIs (UNIF0, UNIF1, UNIF2) as well as other hyperparameter optimization techniques (GRDS, RNDS) on flexible branched architectures. Moreover, we also show by comparison with baseline methods that the multi-branched architecture also learns features that are consistent in terms of their discriminative power. Unlike the standard non-branched deep learning architectures like MLP, FNN Encoder-Decoder and BrainNetCNN, it is possible to track feature importance in standard ML classifiers. However, the framework presented in this work is not only consistent with standard machine learning models in terms of feature

importance, but also additionally allows for analyzing the variation in the nature of information in feature subdomains. Results from our work further emphasize the importance of questioning the assumption that various features from data as complex as neuroimaging data would require architectures with the same complexity irrespective of the subdomains of the brain. While it is apparent that certain subdomains contain more complex information for a given application, we provide an interpretable framework that analyzes as well exploits this variability for better understanding and classifying a given condition, in this case schizophrenia. A similar analysis could be extended to other brain disorders and diseases in future work, providing biomarkers for these conditions in terms of complexity of information in the subdomains of the brain.

It should be noted that the TPE framework is less prone to over-fitting because it involves the use of branched fully connected architecture, thus requiring much fewer parameters than conventional non-branched fully connected architectures. Additionally, to avoid over-fitting we performed the whole TPE procedure is repeated with random sub-sampling, which resulted in a very low standard deviation in the test accuracy across repetitions, similar to other standard machine learning methods. Thus, TPE is able to generalize better without over-fitting. In future, work should also be focused on extending this analysis to study more parameter optimization methods as well as introducing more types of flexibility in the frameworks, while avoiding the risk of over-fitting. Approaches to automatically determine the model structure/depth and including additional measures of model complexity would be quite powerful and is not well studied. Furthermore, developing flexible architectures like these does not only give insights towards identifying relationships between brain subdomains of the same modality, but it can also be extended to create adaptable yet rigorous frameworks that exploit complex interrelationships from subdomains in a multimodal setting.

ACKNOWLEDGEMENTS

Research reported in this publication was supported by National Institute of Mental Health (HHS - NIH) of the National Institutes of Health under award numbers R01MH118695, RF1AG063153, and R01EB020407.

REFERENCES

- Aine C, Bockholt HJ, Bustillo JR, Cañive JM, Caprihan A, Gasparovic C, Hanlon FM, Houck JM, Jung RE, Lauriello J, et al. (2017). Multimodal neuroimaging in schizophrenia: description and dissemination. *Neuroinformatics*, 15(4):343–364. [PubMed: 28812221]
- Bergstra J, Yamins D, and Cox DD (2013a). Hyperopt: A python library for optimizing the hyperparameters of machine learning algorithms. In *Proceedings of the 12th Python in science conference*, pages 13–20. Citeseer.
- Bergstra J, Yamins D, and Cox DD (2013b). Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures. In *International Conference on Machine Learning*.
- Bergstra JS, Bardenet R, Bengio Y, and Kegl B (2011). Algorithms for hyper-parameter optimization. pages 2546–2554.
- Butler PD and Javitt DC (2005). Early-stage visual processing deficits in schizophrenia. *Current opinion in psychiatry*, 18(2):151. [PubMed: 16639168]
- Calhoun VD, Eichele T, and Pearlson G (2009). Functional brain networks in schizophrenia: a review. *Frontiers in human neuroscience*, 3(17).

- Du Y and Fan Y (2013). Group information guided ica for fmri data analysis. *Neuroimage*, 69:157–197. [PubMed: 23194820]
- Du Y, Fu Z, Sui J, Gao S, Xing Y, Lin D, Salman M, et al. (2019a). Neuromark: a fully automated ica method to identify effective fmri markers of brain disorders. medRxiv, 19008631.
- Du Y, Fu Z, Sui J, Gao S, Xing Y, Lin D, Salman M, Rahaman MA, Abrol A, Chen J, et al. (2019b). Neuromark: a fully automated ica method to identify effective fmri markers of brain disorders. medRxiv, page 19008631.
- Dunn CJ, Duffy SL, Hickie IB, Lagopoulos J, Lewis SJ, Naismith SL, and Shine JM (2014). Deficits in episodic memory retrieval reveal impaired default mode network connectivity in amnesic mild cognitive impairment. *NeuroImage: Clinical*, 4:473–480. [PubMed: 24634833]
- Gaebler AJ, Mathiak K, Koten JW Jr, König AA, Koush Y, Weyer D, Depner C, Matentzoglou S, Edgar JC, Willmes K, et al. (2015). Auditory mismatch impairments are characterized by core neural dysfunctions in schizophrenia. *Brain*, 138(5):1410–1423. [PubMed: 25743635]
- Gallinat J, Mulert C, Bajbouj M, Herrmann WM, Schunter J, Senkowski D, Moukhtieva R, Kronfeldt D, and Winterer G (2002). Frontal and temporal dysfunction of auditory stimulus processing in schizophrenia. *Neuroimage*, 17(1):110–127. [PubMed: 12482071]
- Garrity AG, Pearlson GD, McKiernan K, Lloyd D, Kiehl KA, and Calhoun VD (2007). Aberrant “default mode” functional connectivity in schizophrenia. *American journal of psychiatry*, 164(3):450–457. [PubMed: 17329470]
- Han S, Huang W, Zhang Y, Zhao J, and Chen H (2017). Recognition of early-onset schizophrenia using deep-learning method. In *Applied Informatics*, volume 4, pages 1–6. SpringerOpen.
- He T, Kong R, Holmes AJ, Nguyen M, Sabuncu MR, Eickhoff SB, Bzdok D, Feng J, and Yeo BT (2020). Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *NeuroImage*, 206:116276. [PubMed: 31610298]
- Javitt DC (2009). Sensory processing in schizophrenia: neither simple nor intact. *Schizophrenia bulletin*, 35(6):1059–1064. [PubMed: 19833806]
- Javitt DC and Freedman R (2015). Sensory processing dysfunction in the personal experience and neuronal machinery of schizophrenia. *American Journal of Psychiatry*, 172(1):17–31. [PubMed: 25553496]
- Kaufmann T, Skåtun KC, Alnæs D, Doan NT, Duff EP, Tønnesen S, Roussos E, Ueland T, Aminoff SR, Lagerberg TV, et al. (2015). Disintegration of sensorimotor brain networks in schizophrenia. *Schizophrenia bulletin*, 41(6):1326–1335. [PubMed: 25943122]
- Kawahara J, Brown CJ, Miller SP, Booth BG, Chau V, Grunau RE, Zwicker JG, and Hamarneh G (2017). Brainnetcnn: Convolutional neural networks for brain networks; towards predicting neurodevelopment. *NeuroImage*, 146:1038–1049. [PubMed: 27693612]
- Keator DB, van Erp TG, Turner JA, Glover GH, Mueller BA, Liu TT, Voyvodic JT, Rasmussen J, Calhoun VD, Lee HJ, and Toga AW (2016). The function biomedical informatics research network data repository. *Neuroimage*, 124:1074–1079. [PubMed: 26364863]
- Kim DI, Mathalon D, Ford J, Mannell M, Turner J, Brown G, Belger A, Gollub R, Lauriello J, Wible C, et al. (2009). Auditory oddball deficits in schizophrenia: an independent component analysis of the fmri multisite function birm study. *Schizophrenia bulletin*, 35(1):67–81. [PubMed: 19074498]
- Kim H (2010). Dissociating the roles of the default-mode, dorsal, and ventral networks in episodic memory retrieval. *Neuroimage*, 50(4):1648–1657. [PubMed: 20097295]
- LeCun Y, Bengio Y, and Hinton G (2015). Deep learning. *nature*, 521(7553):436–444. [PubMed: 26017442]
- Lesh TA, Niendam TA, Minzenberg MJ, and Carter CS (2011). Cognitive control deficits in schizophrenia: mechanisms and meaning. *Neuropsychopharmacology*, 36(1):316–338. [PubMed: 20844478]
- Li H, Satterthwaite TD, and Fan Y (2018). Brain age prediction based on resting-state functional connectivity patterns using convolutional neural networks. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pages 101–104. IEEE.
- Louppe G, Wehenkel L, Sutera A, and Geurts P (2013). Understanding variable importances in forests of randomized trees. In *Advances in neural information processing systems*, pages 431–439.

- Luo G (2016). A review of automatic selection methods for machine learning algorithms and hyperparameter values. *Network Modeling Analysis in Health Informatics and Bioinformatics*, 5:1.
- Lynall M-E, Bassett DS, Kerwin R, McKenna PJ, Kitzbichler M, Muller U, and Bullmore E (2010). Functional connectivity and brain networks in schizophrenia. *Journal of Neuroscience*, 30(28):9477–9487. [PubMed: 20631176]
- Pervaiz U, Vidaurre D, Woolrich MW, and Smith SM (2020). Optimising network modelling methods for fmri. *Neuroimage*, 211:116604. [PubMed: 32062083]
- Ragland JD, Laird AR, Ranganath C, Blumenfeld RS, Gonzales SM, and Glahn DC (2009). Prefrontal activation deficits during episodic memory in schizophrenia. *American Journal of Psychiatry*, 166(8):863–874. [PubMed: 19411370]
- Rotarska-Jagiela A, van de Ven V, Oertel-Knöchel V, Uhlhaas PJ, Vogeley K, and Linden DE (2010). Resting-state functional network correlates of psychotic symptoms in schizophrenia. *Schizophrenia research*, 117(1):21–30. [PubMed: 20097544]
- Shahriari B, Swersky K, Wang Z, Adams RP, and Freitas ND (2015). Taking the human out of the loop: A review of Bayesian optimization. In *Proceedings of the IEEE 104*, pages 148–175, no. 1.
- Srinivasagopalan S, Barry J, Gurupur V, and Thankachan S (2019). A deep learning approach for diagnosing schizophrenic patients. *Journal of Experimental & Theoretical Artificial Intelligence*, 31(6):803–816.
- Ulloa A, Plis S, and Calhoun V (2018). Improving classification rate of schizophrenia using a multimodal multi-layer perceptron model with structural and functional mr. *arXiv preprint arXiv:1804.04591*.
- Ulloa A, Plis S, Erhardt E, and Calhoun V (2015). Synthetic structural magnetic resonance image generator improves deep learning prediction of schizophrenia. In *2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–6. IEEE.
- Wang H, Zeng L-L, Chen Y, Yin H, Tan Q, and Hu D (2015). Evidence of a dissociation pattern in default mode subnetwork functional connectivity in schizophrenia. *Scientific reports*, 5(1):1–10.
- Wolf ND, Sambataro F, Vasic N, Frasch K, Schmid M, Schönfeldt-Lecuona C, Thomann PA, and Wolf RC (2011). Dysconnectivity of multiple resting-state networks in patients with schizophrenia who have persistent auditory verbal hallucinations. *Journal of psychiatry & neuroscience: JPN*, 36(6):366. [PubMed: 21791169]
- Yu Q, A Allen E, Sui J, R Arbabshirani M, Pearlson G, and D Calhoun V (2012). Brain connectivity networks in schizophrenia underlying resting state functional magnetic resonance imaging. *Current topics in medicinal chemistry*, 12(21):2415–2425. [PubMed: 23279180]
- Zeng L-L, Wang H, Hu P, Yang B, Pu W, Shen H, Chen X, Liu Z, Yin H, Tan Q, et al. (2018). Multi-site diagnostic classification of schizophrenia using discriminant deep learning with functional connectivity mri. *EBioMedicine*, 30:74–85. [PubMed: 29622496]

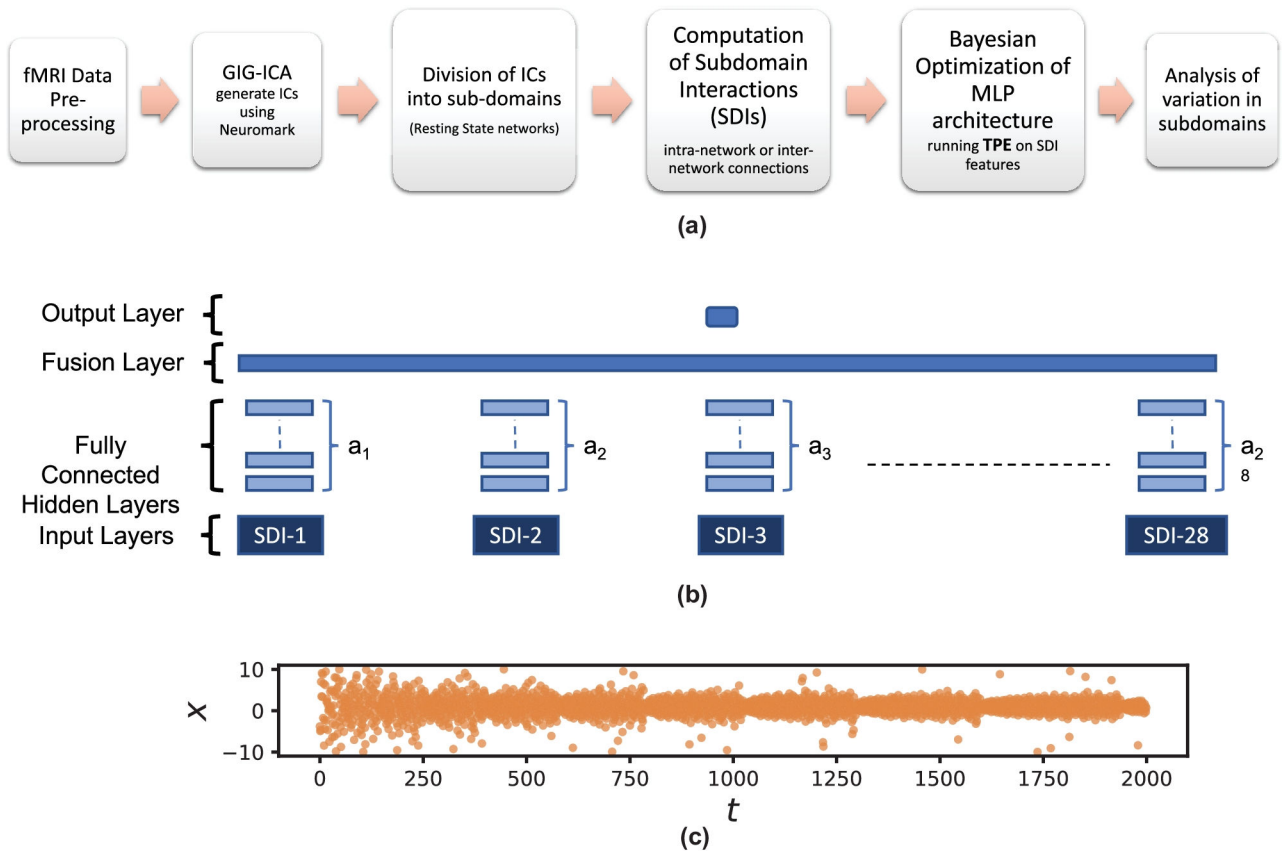


Figure 1.

(a) A step-by-step description of the whole analysis. (b) An architecture in the search space that TPE optimizes over, defined by the vector $\{x_i\}_{i=1}^{28}$, with $x_i \in \{0, 1, 2\}$ representing the number of fully connected hidden layers on top of the input node corresponding to data from the i -th subdomain interaction (SDI). Data for each SDI is a sub-matrix of the full static functional connectivity matrix with connections from participating subdomain(s). (c) TPE search space traversal on a toy example with a quadratic cost function $(x - 1)^2$ to narrow down the search closer to the optimal value.

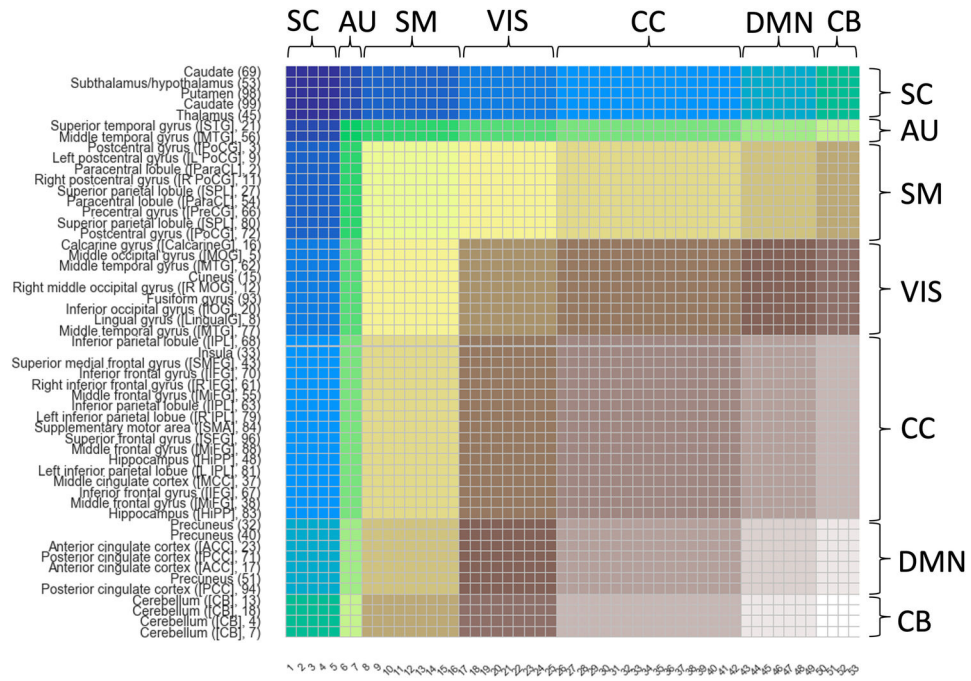
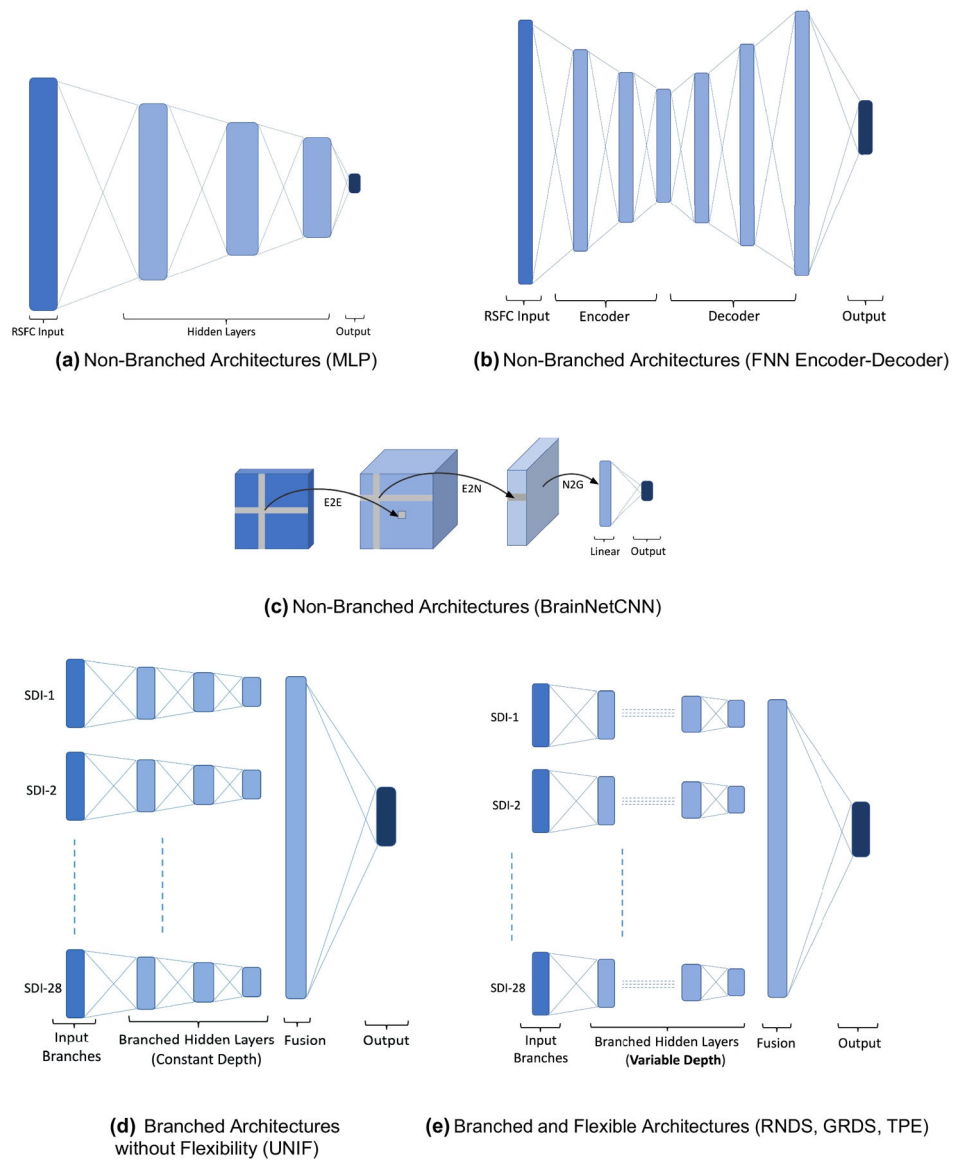


Figure 2. Illustration of how the functional connectivity matrix for the 53 components (53×53 in size) is divided into subdomain interactions (SDIs) that form the input to the branched Multilayer Perceptron (MLP) architecture. Each colored sub-matrix represents a certain SDI i.e., the set of functional connectivity values between components of a given pair of subdomains. The 7 subdomains include: default mode network (DMN), visual (VIS), auditory (AU), cognitive control (CC), sensorimotor (SM), cerebellar (CB) and sub-cortical (SC). A total of 28 SDIs, which are sub-matrices of the full functional connectivity matrix, are colored in different colors. Since the number of subdomains is 7, a total of ${}^7C_2 = 21$ out of the 28 SDIs correspond to inter-network connections (e.g. DMN-VIS) while 7 correspond to intra-network connections (e.g. DMN-DMN).

**Figure 3.**

A schematic diagram for neural network based methods used for performance comparison with the TPE based approach. In addition to standard machine learning models like SVM, logistic regression (LOG) and random forest classifier (RFC), baseline non-branched neural network architectures used were (a) multilayer perceptron (MLP), (b) feedforward neural network with encoder-decoder architecture (FNN) and (c) BrainNetCNN. Branched neural network architectures included (d) uniform architectures, UNIF0, UNIF1 and UNIF2, representing non-flexible multi-branched architectures with 0, 1 and 2 fully connected layers before the fusion step and above input layer in each SDI branch of the architecture respectively. (e) As the third baseline neural network methods, existing hyper-parameter optimization techniques including random search (RNDS) and grid search (GRDS) were used to optimize the domain of branched architectures with variable branch-depth, representing the same class as TPE.

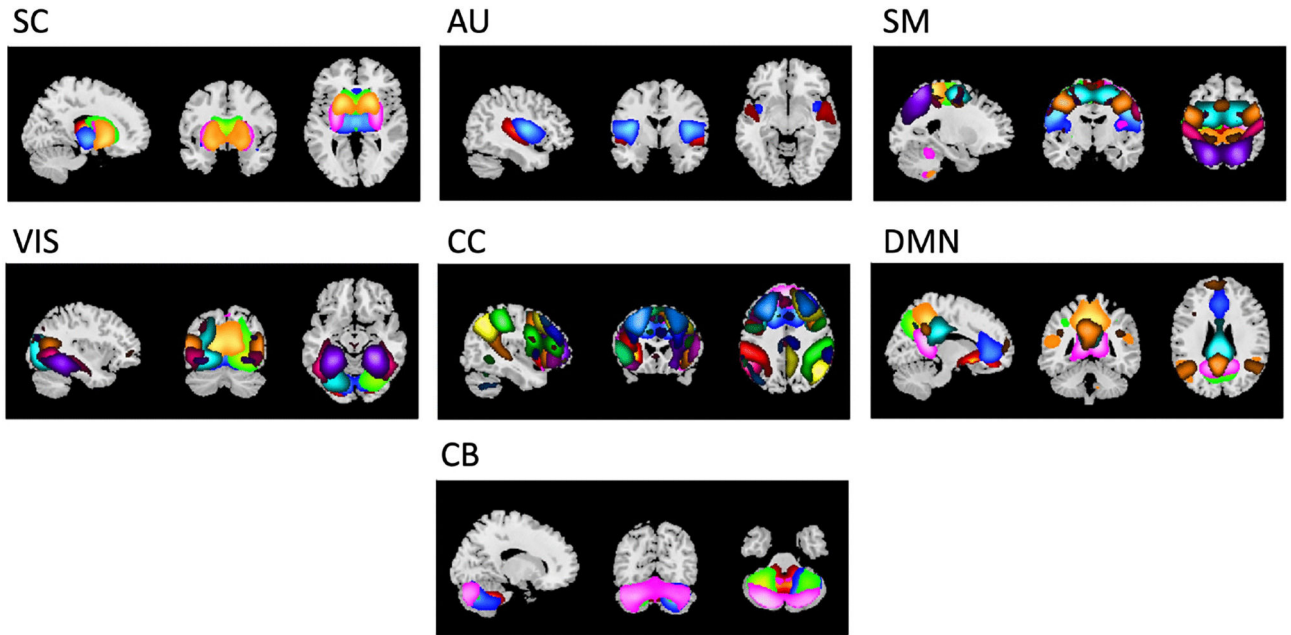


Figure 4.

The components obtained from the scICA procedure Neuromark are shown in different colors. Each map shows the components belonging to a particular subdomains i.e., networks of brain. The 7 subdomains include: default mode network (DMN), visual (VIS), auditory (AU), cognitive control (CC), sensorimotor (SM), cerebellar (CB) and sub-cortical (SC). subdomain interaction (SDI) features built using these subdomains and components were used as input to the multi-branch MLP architecture optimized for variable branch-depth by the TPE algorithm. The SDI features are comprised of functional connectivity (static time-series correlation) between pairs of components belonging the same subdomain (intra-network connections) or different subdomains (inter-network connections). The 7 subdomains shown above lead to the creation of 28 SDIs, with ${}^7C_2 = 21$ corresponding to the inter-network connections (Eg. DMN-VIS) and 7 corresponding to the intra-network connections (Eg. DMN-DMN).

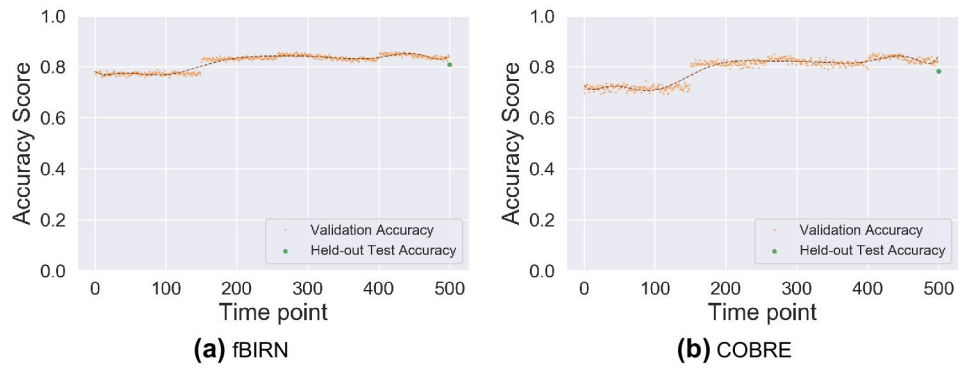


Figure 5. Mean validation accuracy vs. time point (iterations) for 50 repetitions of the TPE algorithm over the architecture search space depicted in Figure 1b. The mean test accuracy using the final architecture on held-out data for each repetition is also shown. The points traversed in the search space are selected based on expected improvement (Equation 2).

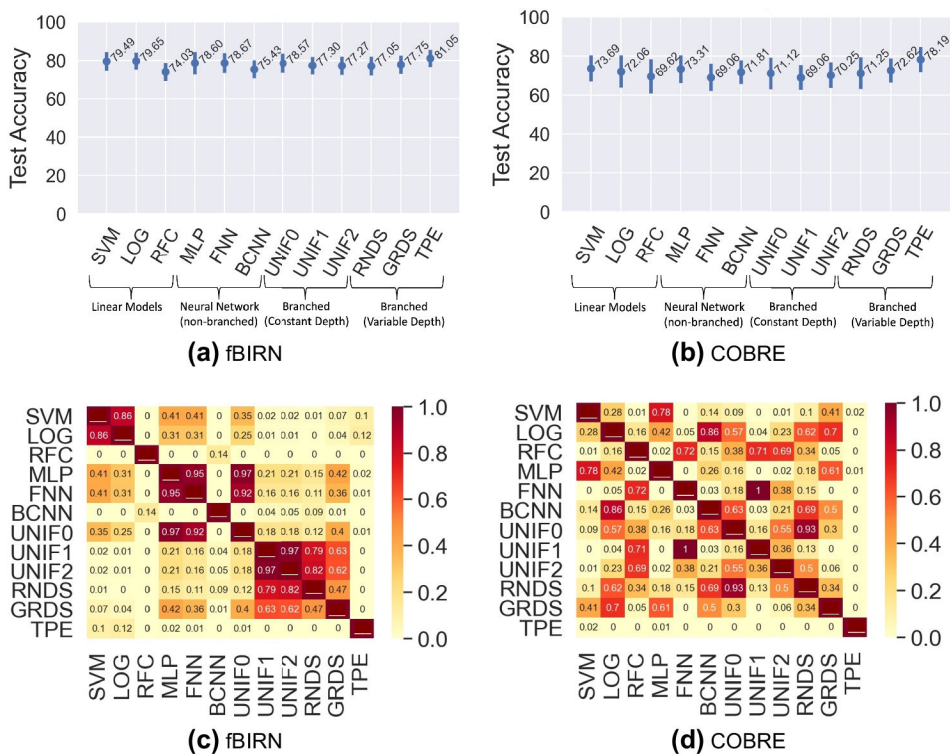


Figure 6. Mean validation accuracy with error-bar for 50 repetitions of the TPE-optimized final architecture in comparison to baseline methods for (a) fBIRN and (b) COBRE datasets along with p-values for two-sample t-test on the mean test accuracy across 50 repetitions, shown in (c),(d). Methods used for performance comparison with TPE approach include simple machine learning models (SVM, LOG, RFC), Non-Branched Neural Network Architectures (MLP, FNN, BCNN), Branched Neural Network Architectures without Flexibility (UNIF0, UNIF1, UNIF2) and also Branched Neural Network Architectures with Non-Uniform Branch-Depth (GRDS, RNDS). See Figure 3 and subsection 2.6 for visualization and detailed explanation of these methods. The architecture created from the repeated optimizations using the TPE procedure is termed as TPE in the plots. It can be noted that for both the datasets, the accuracy obtained by using the TPE-optimized architecture is significantly higher than the accuracies from uniformly branched architectures (UNIF0, UNIF1), indicating the need for flexible architectures. Moreover, the accuracy with TPE is slightly higher than the accuracy for the other baseline methods, showing the scope of interpretability in the optimized model in terms of certain subdomain interactions (SDIs) with higher complexity requiring deeper while others requiring shallow architectures.

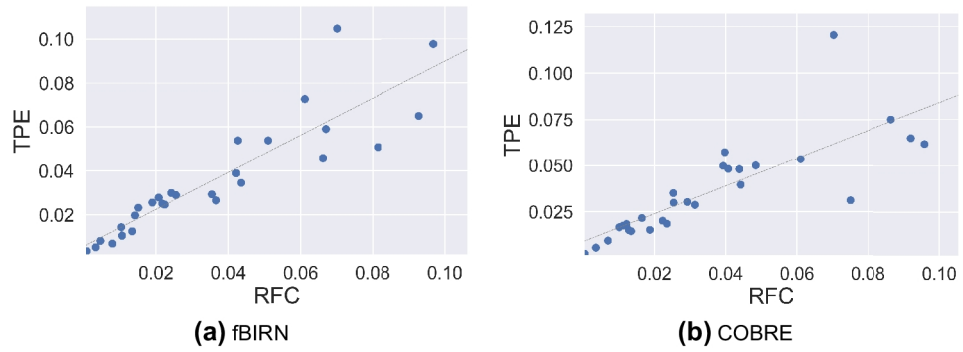


Figure 7.

To check whether the relative importance of SDI parameters learned in the fusion layer of MLP architecture is similar to the functional connectivity features of the corresponding SDI, Random Forest classifier was used to compute purity-based feature importance on the parameters in the fusion layer as well as the input layer (connectivity features). The prediction power vector obtained for both these cases was divided into 28 bins corresponding to each SDI and summed to get the cumulative prediction power of each of the SDIs. The above plots show the cumulative prediction power of SDIs, averaged over 50 repetitions, in the learned parameters inside the fusion layer (marked as TPE on the y-axis) and in the functional connectivity features in the input layer (marked as RFC on the x-axis). There was a high correlation of 0.9 and 0.81 between these prediction power values for (a) fBIRN and (b) COBRE datasets respectively. This means that in addition to being consistent in terms of the prediction accuracy, the TPE algorithm is also consistent in terms of the importance that the SDIs have for the prediction task.

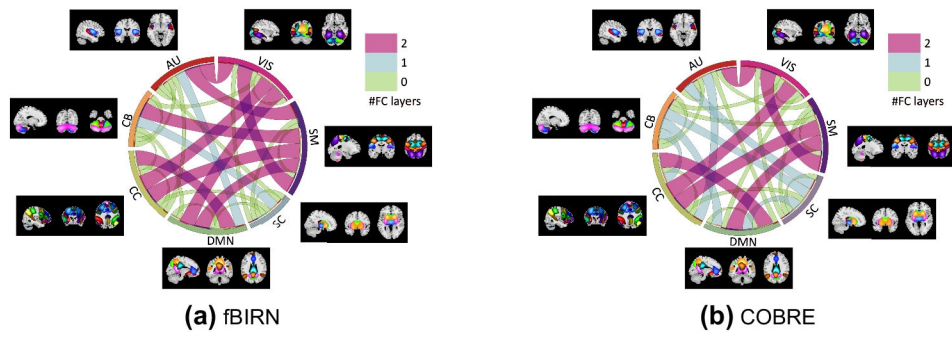


Figure 8.

(a),(b) Visualization of the depth (in terms of the number of fully connected layers) in the final optimized architecture required by various subdomain interactions (SDIs) (intra-network or inter-network functional connectivity). SDIs consistently requiring higher depth across repetitions of the algorithm can be said to carry more complex information towards the classification objective.

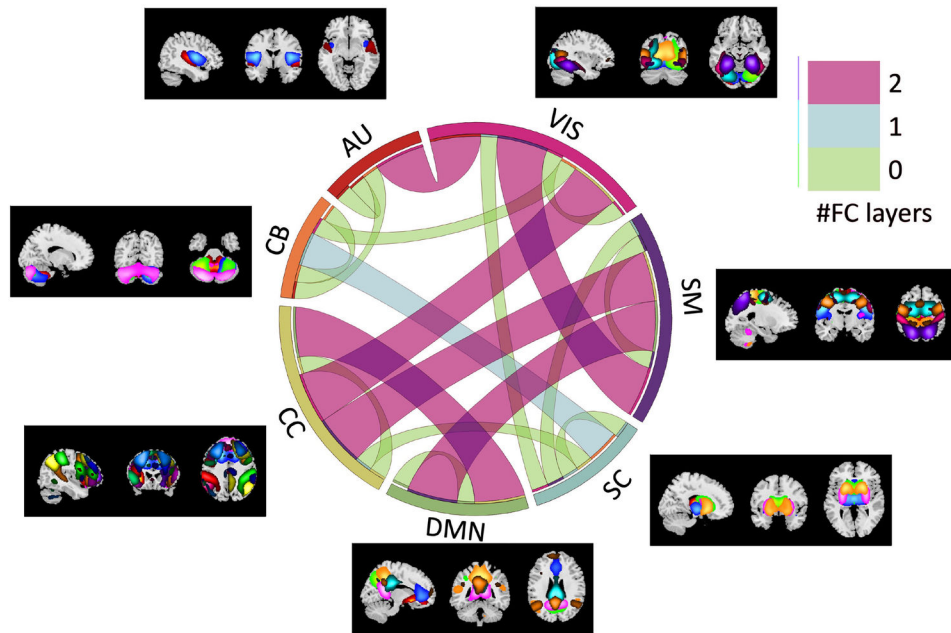


Figure 9. Connectogram showing SDIs requiring the same depth i.e., number of fully connected layers, in the TPE-optimized architectures for both COBRE and fBIRN datasets. The depth required by SDIs was the same for 19 out of 28 SDIs, indicating a common pattern across datasets in terms of certain SDIs requiring deeper models while others requiring shallower ones for a better prediction.

Table 1.

Peak coordinates and primary brain regions for the 53 components (ICs) obtained using scICA (Neuromark framework) on the fMRI time series data. The 53 components were divided into 7 subdomains (resting-state networks) as shown in different colors in the table. The time-series from every possible pair of these 53 components was used to compute the static functional connectivity (SFNC) features using Pearson correlation. The SFNC features were then divided into 28 subdomain interactions (SDIs) based on the subdomain(s) to which a given pair of components for the SFNC feature belongs (See Figure 4).

Primary region for component (IC ID)	X	Y	Z	Subdomain	Primary region for component (IC ID)	X	Y	Z	Subdomain
Caudate (69)	6.5	10.5	5.5		Inferior parietal lobule ([IPL], 68)	45.5	-61.5	43.5	
Subthalamus/hypothalamus (53)	-2.5	-13.5	-1.5		Insula (33)	-30.5	22.5	-3.5	
Putamen (98)	-26.5	1.5	-0.5	Subcortical	Superior medial frontal gyrus ([SMFG], 43)	-0.5	50.5	29.5	
Caudate (99)	21.5	10.5	-3.5		Inferior frontal gyrus ([IFG], 70)	-48.5	34.5	-0.5	
Thalamus (45)	-12.5	-18.5	11.5		Right inferior frontal gyrus ([R IFG], 61)	53.5	22.5	13.5	
Superior temporal gyrus ([STG], 21)	62.5	-22.5	7.5	Auditory	Middle frontal gyrus ([MiFG], 55)	-41.5	19.5	26.5	
Middle temporal gyrus ([MTG], 56)	-42.5	-6.5	10.5		Inferior parietal lobule ([IPL], 63)	-53.5	-49.5	43.5	
Postcentral gyrus ([PoCG], 3)	56.5	-4.5	28.5		Left inferior parietal lobue ([R IPL], 79)	44.5	-34.5	46.5	
Left postcentral gyrus ([L PoCG], 9)	-38.5	-22.5	56.5		Supplementary motor area ([SMA], 84)	-6.5	13.5	64.5	Cognitive Control
Paracentral lobule ([ParaCL], 2)	0.5	-22.5	65.5		Superior frontal gyrus ([SFG], 96)	-24.5	26.5	49.5	
Right postcentral gyrus ([R PoCG], 11)	38.5	-19.5	55.5		Middle frontal gyrus ([MiFG], 88)	30.5	41.5	28.5	
Superior parietal lobule ([SPL], 27)	-18.5	-43.5	65.5	Sensorimotor	Hippocampus ([HiPP], 48)	23.5	-9.5	-16.5	
Paracentral lobule ([ParaCL], 54)	-18.5	-9.5	56.5		Left inferior parietal lobue ([L IPL], 81)	45.5	-61.5	43.5	
Precentral gyrus ([PreCG], 66)	-42.5	-7.5	46.5		Middle cingulate cortex ([MCC], 37)	-15.5	20.5	37.5	
Superior parietal lobule ([SPL], 80)	20.5	-63.5	58.5		Inferior frontal gyrus ([IFG], 67)	39.5	44.5	-0.5	
Postcentral gyrus ([PoCG], 72)	-47.5	-27.5	43.5		Middle frontal gyrus ([MiFG], 38)	-26.5	47.5	5.5	
Calcarine gyrus ([CalcarineG], 16)	-12.5	-66.5	8.5		Hippocampus ([HiPP], 83)	-24.5	-36.5	1.5	
Middle occipital gyrus ([MOG], 5)	-23.5	-93.5	-0.5		Precuneus(32)	-8.5	-66.5	35.5	
Middle temporal gyrus ([MTG], 62)	48.5	-60.5	10.5		Precuneus(40)	-12.5	-54.5	14.5	
Cuneus(15)	15.5	-91.5	22.5		Anterior cingulate cortex ([ACC], 23)	-2.5	35.5	2.5	
Rightmiddle occipital gyrus ([R MOG], 12)	38.5	-73.5	6.5	Visual	Posterior cingulate cortex ([PCC], 71)	-5.5	-28.5	26.5	Default mode

Primary region for component (IC ID)	X	Y	Z	Subdomain	Primary region for component (IC ID)	X	Y	Z	Subdomain
Fusiform gyrus (93)	29.5	-42.5	-12.5		Anterior cingulate cortex ([ACC], 17)	-9.5	46.5	-10.5	
Inferior occipital gyrus ([IOG], 20)	-36.5	-76.5	-4.5		Precuneus (51)	-0.5	-48.5	49.5	
Lingual gyrus ([LingualG], 8)	-8.5	-81.5	-4.5		Posterior cingulate cortex ([PCC], 94)	-2.5	54.5	31.5	
Middle temporal gyrus ([MTG], 77)	-44.5	-57.5	-7.5		Cerebellum ([CB], 13)	-30.5	-54.5	-42.5	
					Cerebellum ([CB], 18)	-32.5	-79.5	-37.5	Cerebellum
					Cerebellum ([CB], 4)	20.5	-48.5	-40.5	
					Cerebellum ([CB], 7)	30.5	-63.5	-40.5	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Test accuracy scores for all the classification methods used on fBIRN and COBRE datasets. See Figure 3 and subsection 2.6 for visualization and detailed explanation of these methods respectively.

Method	Architecture Category	Test Accuracy (fBIRN) (mean \pm std)	Test Accuracy (COBRE) (mean \pm std)
SVM	Linear	79.492 \pm 4.886	73.688 \pm 6.646
LOG	Linear	79.651 \pm 4.430	72.062 \pm 8.254
RFC	Linear	74.032 \pm 4.674	69.625 \pm 8.818
MLP	Neural Net	78.603 \pm 5.813	73.312 \pm 7.132
FNN	Neural Net	78.667 \pm 5.189	69.062 \pm 6.967
BCNN	Neural Net	75.429 \pm 4.701	71.812 \pm 6.011
UNIF0	Branched	78.571 \pm 4.924	71.125 \pm 8.138
UNIF1	Branched	77.302 \pm 4.515	69.062 \pm 6.351
UNIF2	Branched	77.270 \pm 4.767	70.250 \pm 6.502
RNDS	Branched + Flexible	77.048 \pm 4.910	71.250 \pm 8.077
GRDS	Branched + Flexible	77.746 \pm 4.747	72.625 \pm 6.142
TPE	Branched + Flexible	81.052 \pm 4.515	78.188 \pm 6.468