Joint functional brain network atlas estimation and feature selection for neurological disorder diagnosis with application to autism

Mhiri, Islem; Rekik, Islem

Published in:
Medical Image Analysis

DOI:
10.1016/j.media.2019.101596

Publication date:
2020

Document Version
Peer reviewed version

Joint Functional Brain Network Atlas Estimation and Feature Selection for Neurological Disorder Diagnosis With Application to Autism

Islem Mhiri[a,b], Islem Rekik[a,c,*]

[a] *BASIRA lab, Faculty of Computer and Informatics, Istanbul Technical University, Istanbul, Turkey*
[b] *National Engineering School of Sousse (ENISo), Sousse, Tunisia*
[c] *School of Science and Engineering, Computing, University of Dundee, UK*

Abstract

Image-based brain maps, generally coined as 'intensity or image atlases', have led the field of brain mapping in health and disease for decades, while investigating a wide spectrum of neurological disorders. Estimating representative brain atlases constitute a fundamental step in several MRI-based neurological disorder mapping, diagnosis, and prognosis. However, these are strikingly lacking in the field of brain connectomics, where connectional brain atlases derived from functional MRI (fRMI) or diffusion MRI (dMRI) are almost absent. On the other hand, conventional connectomic-based classification methods traditionally resort to feature selection methods to decrease the high-dimensionality of connectomic data for learning how to diagnose new patients. However, these are generally limited by high computational cost and a large variability in performance across different datasets, which might hinder the identification of reproducible biomarkers. To address both limitations, we unprecedentedly propose a brain network atlas-guided feature selection (NAG-FS) method to disentangle the healthy from the disordered connectome. To this aim, given a population of brain connectomes, we propose to learn how estimate a centered and representative functional brain network atlas (i.e., a population center) to reliably map the functional connectome and its variability across training individuals, thereby capturing their shared traits (i.e., connectional fingerprint of a population). Essentially,

*Corresponding author; Dr Islem Rekik, http://basira-lab.com/, GitHub code: https://github.com/basiralab/NAGFS

we first learn the pairwise similarities between connectomes in the population to map them into different subspaces. Next, we non-linearly diffuse and fuse connectomes living in each subspace, respectively. By integrating the produced subspace-specific network atlases we ultimately estimate the population network atlas. Last, we compute the difference between healthy and disordered network atlases to identify the most discriminative features, which are then used to train a predictive learner. Our method boosted the classification performance by 6% in comparison to state-of-the-art FS methods when classifying autistic and healthy subjects.

Keywords: functional network atlas estimation, brain network fusion, connectomic feature selection, multi-kernel network manifold learning, discriminative biomarker identification, brain connectome, autism spectrum disorder, classification

---

1. Introduction

Autism spectrum disease (ASD) is the fastest growing neurodevelopmental disorder characterized by varied impairments in cognitive function, restricted interests, and behavioral challenges (Amaral et al., 2008; Ecker et al., 2010; Association et al., 2013; Wee et al., 2016). According to (Autism and Investigators, 2014), an estimate of 1 in 68 American children under 8 years old has ASD. (Leigh and Du, 2015) mentioned that the direct cost of autism was $268 billion in 2015 in the United States only, and it is expected to be $461 billion by 2025. To efficiently identify biomarkers that can improve ASD diagnosis, non-invasive measures derived from resting-state functional magnetic resonance imaging (rfMRI) have been used in combination with advanced machine learning techniques (Brown and Hamarneh, 2016; Minshew and Keller, 2010; Uddin et al., 2013; Abraham et al., 2017; Guo et al., 2017; Li et al., 2018). One of the most commonly used measures is functional connectivity (FC), quantifying the correlation of blood oxygen level-dependent (BOLD) signals between pairs of anatomical regions of interest (ROIs) (Minshew and Keller, 2010; Guo et al., 2017). FC is conventionally encoded in a symmetric functional connectivity matrix (i.e., connectome) of size $r \times r$, where r denotes the number of ROIs (Yap et al., 2010; Van Den Heuvel and Pol, 2010).

Although promising, existing machine learning techniques suffer from the curse of *'high-dimensional-feature-low-sample-size'*, where the number of fea-

tures largely exceeds the number of training samples –particularly in medical datasets. To address this fundamental issue, several machine learning pipelines integrated dimension reduction techniques (Khalid et al., 2014), which can be categorized into feature extraction and feature extraction methods. As for feature extraction methods such as principal component analysis (PCA) (Shams and Rahman, 2011) or graph embedding techniques (Morris and Rekik, 2017), they are not suitable for neurological biomarker discovery since there is no biological meaning for the new extracted features. Furthermore, following the projection onto a low-dimensional space, the original neurological features (e.g., brain connectivity) become intractable, hindering biomarker discovery. On the other hand, feature selection methods explicitly select the most relevant features from the original data (Kosmicki et al., 2015; Guo et al., 2017) without generally resorting to any intractable feature transformation. Besides, they are more interpretable and readable than feature extraction methods (Tang et al., 2014; Xue et al., 2016), hence their common use for ASD diagnosis and biomarker identification using FC data (Zhao et al., 2018; Dryburgh et al., 2019). However, despite the major advances in the vibrant field of feature selection methods, there remains key challenges in using these methods for identifying discriminative brain features for neurological disorder diagnosis including ASD.

The first challenge is generalizability and reproducibility. A recent work (Abraham et al., 2017) investigated the performance of different FS methods in identifying functional connectomic features distinguishing between normal controls (NC) and ASD subjects. They found out that the performance of FS methods is not consistent across datasets, which makes the task of choosing the best feature selection method for a particular dataset difficult and time-consuming. Also, they pointed out other challenges for FS methods such as data heterogeneity and the computational cost including time and space complexities. In a more recent study, (Georges et al., 2018) addressed the problem of identifying reproducible connectomic features associated with a particular brain disorder for a connectomic dataset of interest. This was highlighted as one of the grand challenges in connectomic data analysis since the performance of a particular FS method in training a typical classifier highly fluctuates depending on the input data (Georges et al., 2018).

The second challenge is scalability (Xue et al., 2016). Due to the large scale of the available datasets, the number of features dramatically increased, which caused an increase in the computational cost and a decrease in the classification performance. Furthermore, the performance of FS methods

might lack stability when scaling up the training dataset (Abraham et al., 2017; Georges et al., 2018).

The third challenge is data heterogeneity. Given that connectomic data might vary in number of features and distribution, the problem of heterogeneity is present. Typical FS methods might fail to handle the heterogeneity of the connectomic data distribution (Abraham et al., 2017; Georges et al., 2018) and identify connectional features that are reproducible and stable across heterogeneous datasets.

The fourth challenge is computational time. Devising an FS method with low time and space complexities, yet accurate and fast remains a compelling problem to solve. Overall, existing FS methods which were demonstrated to perform well on diverse datasets and outperform several state-of-the-art methods for feature selection have a non-constant time complexity. For instance, the time complexity of a landmark feature selection method (infinite feature selection –IFS) (Roffo et al., 2015) is quadratic, which cannot handle big data with a large number of samples.

To address all these challenges, we resort to *brain maps* (also called atlases) to design an efficient feature selection method for boosting neurological disorder diagnosis using connectomic data. Brain maps are widely used to derive discriminative features (or biomarkers) for diagnosing different brain disorders such as mild cognitively impaired patients and Alzheimer's disease patients (Liu et al., 2015; Adler et al., 2018), and autistic patients (Dhifallah et al., 2018). **We hypothesize that *neurological connectional biomarkers* can be effectively identified by first *learning* then *comparing* well-representative and centered brain network atlases of populations of healthy and disordered brains, respectively (Figure 1).** To do so, we *first* aim to efficiently produce a unified normalized connectional representation of a population of brain networks. *Second*, by comparing the difference between the learned healthy and disordered brain atlases, we can easily spot discriminative brain connectivities, which will be used to train a predictive learner for accurate and fast diagnosis.

For the last few decades, neuroscientific and neuroimaging studies have relied heavily on the use of anatomical brain atlases for brain mapping, normalization and comparison across individuals and populations (Toga and Thompson, 2001; Gholipour et al., 2007). However, the connectional aspect of the brain, captured by the wiring of its functional and structural neural connections, was overlooked in the field of building *population-based brain maps*. A few recent landmark works (Rekik et al., 2017, 2018b,a; Dhifallah

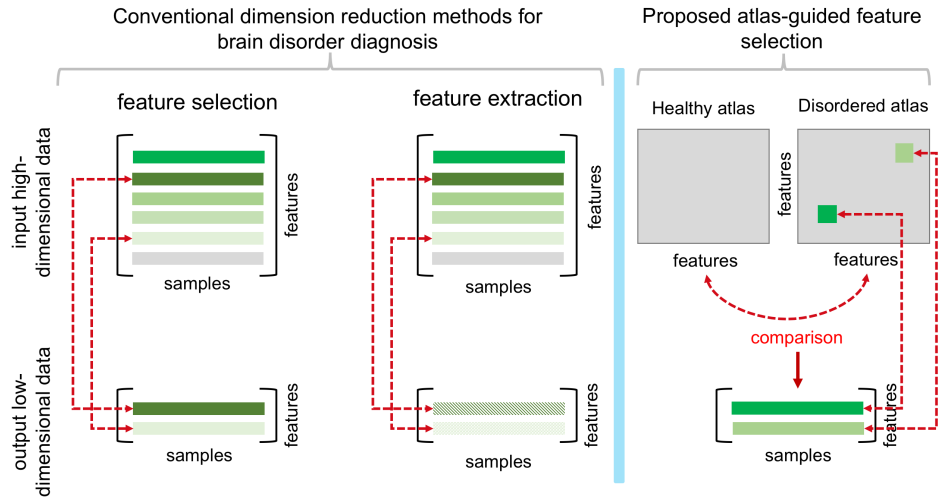

Figure 1: *Conventional dimensional reduction methods for neurological biomarker discovery and proposed network atlas-guided feature selection method.* While typical feature selection (FS) methods aim to identify the most discriminative features in the *original* feature space for the target classification task, feature extraction (FE) methods cannot track the original features as they extract new discriminative features via projection. Hence, FS methods are more convenient for clinical applications for biomarker discovery. However, existing FS methods are generally challenged by space, time, scalability, and reproducibility. To address these issues, we design a simple but effective feature selection method, which identifies the most discriminative features by comparing healthy and disordered brain network atlases to learn.

et al., 2018) aimed to fill this gap in connectional brain mapping by introducing and developing the concept of a *network atlas)* estimated from a population of brain networks. Although pioneering, the majority of these methods were not evaluated on functional brain networks. More importantly, the *discriminative potential* of the estimated connectional brain atlases for discovering neurological biomarkers that can reliably distinguish between typical and disordered brain networks remains unexplored.

To meet the four aforementioned challenges and address this gap, we unprecedentedly propose a novel brain network atlas estimation technique for fast and efficient feature selection method to boost neurological disorder diagnosis. Specifically, we propose a novel network atlas-guided feature selection (NAG-FS) method, which is *reproducible*, *scalable*, *computationally efficient*, and can handle *data heterogeneity*. By *learning* how to cluster similar functional brain networks into non-overlapping subspaces using multiple kernels, we can capture potential data distribution heterogeneity with different bandwidths. Next, we leverage the network diffusion and fusion technique introduced in (Wang et al., 2014) to nonlinearly fuse networks lying in the same subspace, hence creating a cluster-specific network atlas. Last, we obtain the population functional network atlas by non-linearly diffusing and fusing network atlases. This produces a centered and well-representative network atlas for a population of brain networks. Ultimately, to identify the top N_f most discriminative connections, we simply compute the absolute difference between the healthy network atlas, estimated using healthy individuals, and the disordered network atlas, estimated using disordered individuals. This time-constant operation $O(1)$ pins down the most distant connections in both brain templates, which will be used to train a linear classifier. By cleaving each functional network into its positive and negative parts, we also define a *positive functional network atlas* and a *negative functional network atlas*. We evaluate NAG-FS using negative, positive, and whole functional connectomes for efficiently and accurately classifying healthy and autistic brains.

2. Proposed Network Atlas-Guided Feature Selection Method

2.1. NAG-FS overview

Figure 2 provides an overview of the key steps of the proposed network atlas-guided feature selection (NAG-FS) for a fast and accurate classification of normal controls (NC) and ASD subjects. In the first stage, we learn how to estimate a centered and representative network atlas for positive (when

the activity in a brain region increases, there is the same increase in activity in the correlated region), negative (when the activity in a brain region increases, there is the opposite effect in activity in the correlated region (Chen et al., 2011; McGrath et al., 2013; Parente and Colosimo, 2016)), and whole functional brain connectivities (Smith et al., 2015), respectively. To this aim, we first divide our population into NC and ASD groups. Next, to estimate a representative and centered average functional network for each group, we learn multiple local network atlases, each occupying the center of a subpopulation in the target group as illustrated in **Figure 3**. To tease apart the different clusters within each group, thereby capturing the heterogeneous distribution of the data with most likely varying bandwidths, we leverage Single Cell Interpretation via Multikernel Learning (SIMLR) framework for clustering (Wang et al., 2018). SIMLR efficiently learns sample-to-sample similarity measure that best fits the structure of the data by combining multiple kernels. Next, we use the learned similarity matrix to cluster the functional networks within each group into different subpopulations. We then estimate the average local network at the center of each subpopulation by utilizing similarity network fusion (SNF) technique (Wang et al., 2014; Rekik et al., 2017). SNF non-linearly integrates the input networks into a single network using a local diffusion and global fusion processes. SNF is an efficient network fusion algorithm in the sense that (i) it can capture both on common patterns and complementary information across samples, (ii) it can derive useful information even from a small number of samples, (iii) is robust to noise and data heterogeneity, and (iv) scales to a large number of features. To estimate the final population-specific network atlas, we apply SNF to fuse all local network atlases estimated for each subpopulation. The last step in the proposed NAG-FS is to utilize the estimated ASD and NC network atlases to devise an accurate but fast feature selection technique. Specifically, by computing the residual network atlas (absolute difference between both network atlases), we select the non-zero features with the highest discrepancy and use those to train a linear support vector machine (SVM) classifier in a leave-one-out fashion.

2.2. Construction of a centered and representative functional network atlas

In this section, we detail the first step of NAG-FS framework, which aims to estimate centered and representative network atlases using the healthy and disordered populations respectively. Throughout the paper, matrices are denoted by boldface capital letters, e.g., $\mathbf{X}$, vectors by boldface lower-

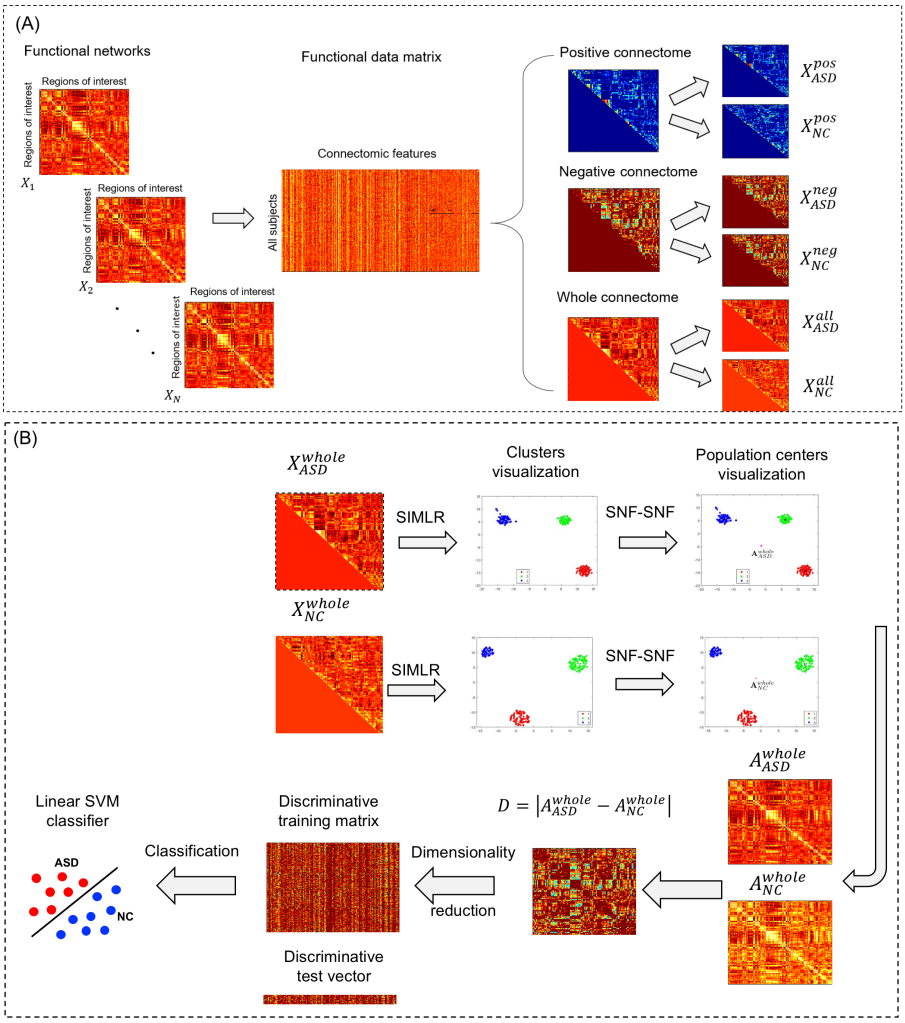

Figure 2: *Illustration of the proposed network atlas-guided feature selection (NAG-FS) framework for a fast and accurate classification of normal controls (NC) and subjects diagnosed with autism spectrum disorder (ASD).* (**A**) Given a population of N functional brain networks, each individual i is encoded in a matrix $\mathbf{X}_i$ of size $r \times r$. Given the symmetry of each $\mathbf{X}_i$, we extract features from the upper off-diagonal triangular part to define a connectivity feature vector $\mathbf{c}_i$, then concatenate them all together into a functional data matrix of size $N \times \frac{r \times (r-1)}{2}$. In addition to whole functional network, we define positive and negative functional networks by cleaving each functional network into its positive and negative parts for both ASD and NC groups. (**B**) In order to disentangle the heterogeneous distribution of the functional networks, we adapt the multi-kernel clustering method (Single Cell Interpretation via Multikernel Learning (SIMLR) (Wang et al., 2018)), which can effectively capture the inherent data distribution. For each cluster, we non-linearly diffuse and fuse all networks into a local centered network atlas using similarity network fusion (SNF (Wang et al., 2014)), then we merge all these local atlases into a global population-centered network atlas. Next, by computing the residual network atlas which is the absolute difference between ASD and NC network atlases $\mathbf{A}_{ASD}^{whole}$ and $\mathbf{A}_{NC}^{whole}$, we select the features with the highest discrepancy and use those to train a linear support vector machine (SVM) classifier within a leave-one-out cross-validation scheme. In this figure, we illustrate the steps for whole connections. These are also applied to negative and positive connectomes, respectively.

Table 1: *Major mathematical notations used in this paper.*

Mathematical notation	Definition
N	total number of subjects in the population
r	number of anatomical regions of interest
$\mathbf{X}_i$	brain network in $\mathbb{R}^{r\times r}$ of the i-th subject
$\mathbf{c}_i$	feature vector of a subject i
$\mathbf{k}_k$	k-th learning kernel in $\mathbb{R}^{N\times N}$
$\mathbf{S}$	similarity matrix in $\mathbb{R}^{N\times N}$
N_c	number of clusters
n_k	number of kernels
$\mathbf{w}_k$	weighting vector of the kernels in $\mathbb{R}^{n_k}$
$\mathbf{L}$	latent matrix in $\mathbb{R}^{N\times N_c}$
$\mathbf{I_N}$	identity matrix in $\mathbb{R}^{N\times N}$
$\mathbf{G}_i = (\mathbf{V}_i, \mathbf{E}_i)$	brain network graph of a single subject (connectome)
$\mathbf{V}_i$	brain ROIs (nodes)
$\mathbf{E}_i$	edges connecting brain regions
$\mathbf{Q}_i$	kernel similarity matrix of the i-th subject
$\mathbf{W}_i$	connectivity matrix of the i-th subject
$\mathbf{P}_i$	full kernel matrix of the i-th subject
$\mathbf{A}_c$	subpopulation network atlas
$\mathbf{A}$	population network atlas
N_f	number of selected features
$\mathbf{D}$	difference absolute matrix between $\mathbf{A}_{ASD}$ and $\mathbf{A}_{NC}$

case letters, e.g., $\mathbf{x}$, and scalars are denoted by lowercase letters, e.g., x. $\mathbf{X}^T$ denotes the transpose operator. For easy reference and enhancing the readability, we have summarized the major mathematical notations in (**Table** 1).

Given a population of N functional networks, each network i is encoded in a matrix $\mathbf{X}_i$ of size $r \times r$, where r denotes the number of anatomical regions of interest (ROIs). The diagonal of the matrix is zeroed since self-connectivity is not considered in this context. An element x_{kl} of network $\mathbf{X}_i$ of the i^{th} individual represents the Pearson correlation between the mean blood oxygen-level dependent (BOLD) signal between ROIs k and l, which falls in the $[-1\ 1]$ range, allowing for positive and negative connectivities between brain regions. Since each matrix $\mathbf{X}_i$ is symmetric, we vectorize the upper off-diagonal triangular part to define a connectivity feature vector $\mathbf{c}_i$ for subject i of size $\frac{r\times(r-1)}{2}$ (**Fig.** 2–A).

Next, we display in **Fig.** 2–B the fundamental stages for estimating the target population connectional center. In the first stage, we aim to disentangle the heterogeneous distribution of the functional networks by leveraging multiple kernels to *learn* a connectomic manifold for healthy and disordered networks, respectively. Multiple kernels have been shown to correspond to different informative representations of biological data and often are more flexible than a single kernel (Gönen and Alpaydın, 2011). In particular,

we adapt the recently developed SIMLR method presented in (Wang et al., 2018) to our aim as it has shown significant outperformance in comparison with clustering methods that used pre-defined similarity measures such as Euclidean similarity and Pearson correlation, instead of learning it in a data-driven manner. In the second stage, we estimate a local network atlas for each cluster, then we merge all local atlases into a global population-centered network atlas.

Multiple kernel-based clustering of brain networks. Fig. 3 provides an overview of the key steps for brain network data clustering taking as input the functional data matrix of size $N \times \frac{r \times (r-1)}{2}$ where each row vector $\mathbf{c}_i$ denotes the feature vector of a subject i. Specifically, (Wang et al., 2018) proposes to learn the network manifold by learning the weights $\{w_k\}_{k=1}^{n_k}$ associated with a set of Gaussian kernels $\{\mathbf{K}_k\}_{k=1}^{n_k}$ with different bandwidths that can capture the diverse statistical characteristics of the input data. The learned manifold also addresses the challenge of high levels of dropout events by employing a rank constraint in the learned cell-to-cell similarity.

Each Gaussian kernel is defined as: $\mathbf{K}(\mathbf{c}^i, \mathbf{c}^j) = \frac{1}{\epsilon_{ij}\sqrt{2\pi}} e^{(-\frac{|\mathbf{c}^i - \mathbf{c}^j|^2}{2\epsilon_{ij}^2})}$, where $\mathbf{c}^i$ and $\mathbf{c}^j$ denote the feature vectors of the i-th and j-th network atlas respectively and ϵ_{ij} is defined as: $\epsilon_{ij} = \sigma(\mu_i + \mu_j)/2$, where σ is a tuning parameter and $\mu_i = \frac{\sum_{l \in KNN(\mathbf{c}^i)} |\mathbf{c}^i - \mathbf{c}^j|}{k}$, where $KNN(\mathbf{c}^i)$ (K-nearest neighbors) represents the top k neighboring subjects of subject i. The weighted kernels are then averaged to produce the target similarity matrix $\mathbf{S}$. These are estimated along with an $N \times N_c$ latent matrix $\mathbf{L}$ capturing N_c inherent distributions of the data by solving the following optimization problem:

$$\min_{\mathbf{S},\mathbf{L},\mathbf{w}} \sum_{i,j,k} -w_k \mathbf{K}_k(\mathbf{c}^i, \mathbf{c}^j)\mathbf{S}_{ij} + \beta||\mathbf{S}||_F^2 + \eta \mathbf{tr}(\mathbf{L}^T(\mathbf{I}_N - \mathbf{S})\mathbf{L}) + \rho \sum_k w_k log w_k \quad (1)$$

Subject to: $\sum_k w_k = 1$, $w_k \geq 0$, $\mathbf{L}^T\mathbf{L} = \mathbf{I}_{N_c}$, $\sum_j \mathbf{S}_{ij} = 1$, and $\mathbf{S}_{ij} \geq 0$ for all (i, j).

The first term refers to the relation between the similarity and the kernel distance with weights w_k between two networks. The second term denotes a regularization term that avoids over-fitting the model to network data. The learned similarity $\mathbf{S}$ should be small if the distance between a pair of networks is large. The matrix $(\mathbf{I}_N - \mathbf{S})$ denotes the graph Laplacian. The last term imposes constraints on the kernel weights to avoid selection of a single

kernel. An alternating convex optimization is adopted where each variable is optimized while fixing the other variables until convergence (Wang et al., 2018). Once the similarity matrix $\mathbf{S}$ is computed, we can perform subpopulation identification by extracting the N_c blocks where similar functional networks group together (**Fig.** 3).

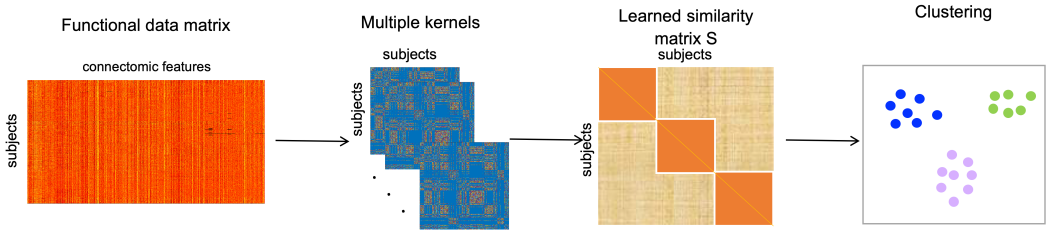

Figure 3: *Illustration of brain network data clustering using SIMLR (Wang et al., 2018).* We input the functional data matrix to SIMLR to group samples into N_c clusters. Next, we learn the proper weights associated with multiple Gaussian kernels that can capture the inherent statistical distributions of the input data. By leveraging the weighted kernels, we construct the target similarity matrix $\mathbf{S}$ which contains N_c blocks where similar functional networks group together.

Proposed population-based network atlas estimation strategy. The proposed network atlas estimation framework lies in first computing local centers at the heart of subpopulations unraveled by the previous network data clustering step, then fusing all these network centers to produce the population network atlas (**Fig.** 4). Note that the network atlas for the healthy population is estimated independently of that of the disordered (ASD in the current study) population. For each individual i in the N_c^{th} subpopulation (or cluster) composed of $\sharp N_c$ networks, we define a graph $\mathbf{G}_i(\mathbf{V}_i, \mathbf{E}_i)$ where each vertex in $\mathbf{v}_k$ denotes an ROI and each edge in E connecting two ROIs k and l denotes the strength of their correlation. Prior to applying SNF method, we first normalize each subject-specific feature vector $\mathbf{c}$ as follows: $\tilde{\mathbf{c}} = \frac{\mathbf{c}-\mathbf{E}(\mathbf{c})}{\sqrt{var(\mathbf{c})}}$, where $\tilde{\mathbf{c}}$ presents the corresponding normalized feature vector. $\mathbf{E}(\mathbf{c})$ denotes the empirical mean of $\mathbf{c}$ and $var(\mathbf{c})$ represents the variance of $\mathbf{c}$.

Next, we define a kernel similarity matrix $\mathbf{Q}_i$ for each individual i in the population, which encodes its local structure by computing the similarity between each of its elements ROI k and its nearest ROIs l as follows:

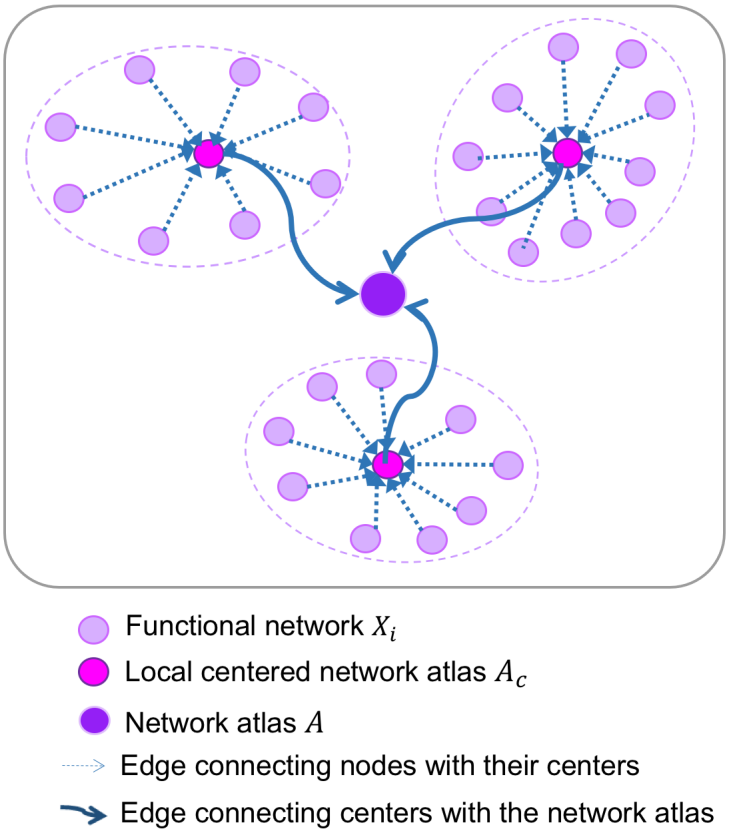

Figure 4: *Proposed SNF-SNF network diffusion and fusion strategy to estimate the population network atlas.* The basic idea lies in first learning how to estimate the local center of each cluster produced by SIMLR, then fusing local centers into a global one. To do so, we use similarity network fusion (Wang et al., 2014) to fuse networks within each cluster by diffusing their local structure across one another.

$$\mathbf{Q}_i(k,l) = \begin{cases} \frac{\mathbf{W}_i(k,l)}{\sum_{p \in N_k} \mathbf{W}_i(k,p)} & l \in N_k \\ 0, & otherwise \end{cases} \tag{2}$$

N_k represents the set of q neighbors of ROI k identified using KNN algorithm.

$\mathbf{W}_i$ represents the connectivity matrix where element $\mathbf{W}_i(k,l)$ denotes the connectivity between ROIs k and l. A scaled exponential similarity kernel is used to determine the weight of each edge. For each subject i, $\mathbf{W}_i$ is computed as follows:

$$\mathbf{W}_i(k,l) = \exp(-\frac{\rho^2(k,l)}{\mu\varepsilon_{k,l}}) \tag{3}$$

Where $\rho(k,l)$ denotes the Euclidean distance between ROIs k and l. μ is a hyperparameter and $\varepsilon_{k,l} = \frac{mean(\rho(k,l)+\rho(k,N_k)+\rho(l,N_l)}{3}$ is used to solve scaling problem (Wang et al., 2014).

To capture the global structure of each network i, we define a full kernel matrix $\mathbf{P}$, carrying the full information about the similarity of each ROI to all other ROIs as follows:

$$\mathbf{P}_i(k,l) = \begin{cases} \frac{\mathbf{W}_i(k,l)}{2\sum_{p \neq k} \mathbf{W}_i(k,p)} & l \neq k \\ 1/2, \ l = k \end{cases} \tag{4}$$

We note that SNF is robust to noise thanks to $\mathbf{Q}$, which can reduce noise between instances. In order to integrate the different networks into a single network, the status matrices $\mathbf{P}_i$ are iteratively updated for each individual by diffusing the the global structure $\mathbf{P}_j$ of $\sharp N_c - 1$ networks ($j \neq i$) along the local structure $\mathbf{Q}_i$ of subject i as follows:

$$\mathbf{P}_i = \mathbf{Q}_i \times \left(\frac{\sum_{j \neq i} \mathbf{P}_j}{\sharp N_c - 1} \right) \times \mathbf{Q}_i^T, \ j \in \{1, \ldots, \sharp N_c\} \tag{5}$$

Where $\frac{\sum_{j \neq i} \mathbf{P}_j}{\sharp N_c - 1}$ denotes the diffusion structure computed as the mean global structures of all other individuals in the subpopulation. We iterate operation 5 N_t times to progressively update each network in relation to other networks using this diffusion process. Ultimately, following N_t iterations, we produce the subpopulation network atlas by averaging the diffused status matrices $\mathbf{P}_i$ at the final iteration N_t:

$$\mathbf{A}_c = \frac{\sum_{i=1}^{\sharp N_c} \mathbf{P}_i^{N_t}}{\sharp N_c} \tag{6}$$

Ultimately, the population network atlas $\mathbf{A}$ is computed by nonlinearly fusing the subpopulation atlases using SNF (Wang et al., 2014). Using the proposed strategy, we estimate a network atlas $\mathbf{A}_{ASD}$ for ASD subjects and $\mathbf{A}_{NC}$ normal controls.

2.3. Discriminative connectional biomarker identification

Due to the high dimensionality of the extracted connectomic features of the order of r^2 with r denoting the number of ROIs used to construct each connectome in our population, we propose a novel feature selection strategy (NAG-FS) that selects the most discriminative features distinguishing between two populations based on the learned population-specific network atlases (**Fig.** 6). Using leave-one-out cross-validation (LOO-CV) strategy, we first use the training population to estimate $\mathbf{A}_{ASD}$ and $\mathbf{A}_{NC}$ network atlases. Next, we compute the absolute distance between both estimated training network atlas matrices $\mathbf{A}_{ASD}$ and $\mathbf{A}_{NC}$ as follows: $\mathbf{D}(NC, ASD) = |\mathbf{A}_{ASD} - \mathbf{A}_{NC}|$. The intuition behind computing a residual network (i.e., absolute different network) is very simple: we hypothesize that the distance between healthy and disordered network atlas is much larger than the difference between healthy network atlas and healthy individual network. The elements in the difference (residual network $\mathbf{D}$) network with the highest

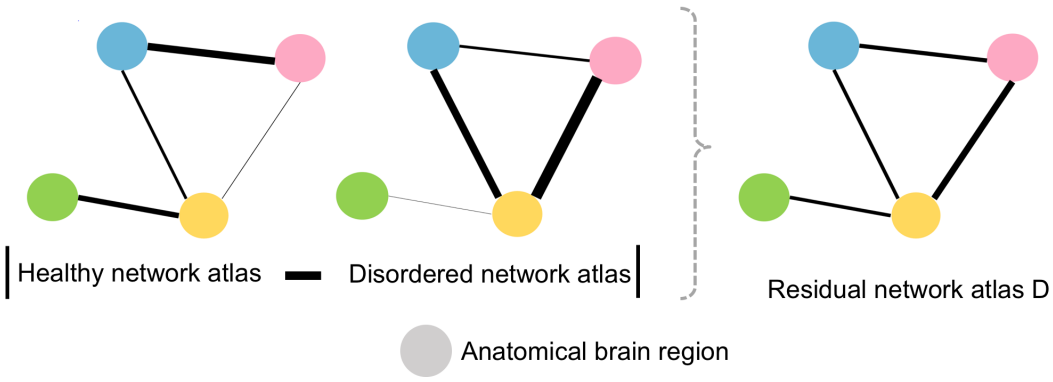

Figure 5: *Construction of the residual network atlas* $\mathbf{D}$ *using the absolute difference between both estimated healthy and disordered network atlases for easy and accurate identification of the most discriminative connectivities (i.e. connectional features).*

values are identified as those most discriminative (**Fig.** 5). By taking all elements in the upper off-diagonal part of $\mathbf{D}(NC, ASD)$, we select the top N_f features with the largest non-zero distance. By extracting the top N_f from all training networks, we train a linear SVM classifier to learn the mapping from the selected feature space to the label space $\{-1, 1\}$ (-1 for ASD and 1 for NC subjects). In the testing stage, we extract the same features from the testing functional network, then pass the produced N_f-dimensional feature vector to the trained classifier for predicting the testing subject label.

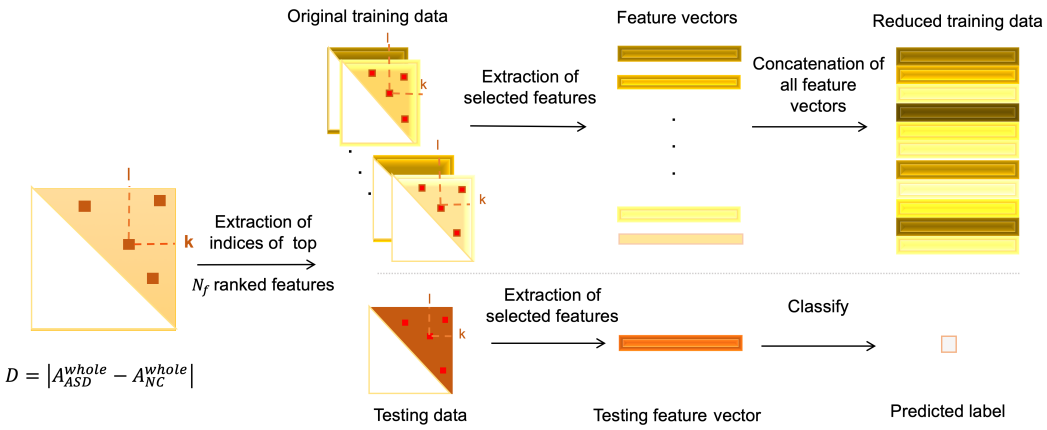

Figure 6: *Proposed network atlas-guided feature selection (NAG-FS) framework.* Given the upper off-diagonal part of $\mathbf{D}$, we select the indices of top N_f ranked features. Next, we extract the top N_f discriminative training features to train a linear classifier. In the testing stage, we extract the same ranked features from the testing functional network, then we produce a feature vector of size N_f, which is fed into SVM to predict the subject label.

3. Experimental results

Evaluation dataset and preprocessing pipeline. We evaluated our proposed framework on 517 subjects (245 ASD and 272 NC) from ABIDE preprocessed dataset [1]. Several preprocessing steps were implemented by the data processing assistant for resting-state fMRI (DPARSF) pipeline, which is based on statistical parametric maps (SPM) and resting-state fMRI data

[1] http://preprocessed-connectomes-project.org/abide/

analysis toolkit (REST). First, to ensure a steady signal, the first 10 volumes of rs-fMRI images were discarded. Based on a six-parameter (rigid body), all images were slice timing corrected and were realigned to the middle to cut down on inter-scan head motion (Tang et al., 2018). Then, the functional data were registered in montreal neurological institute (MNI) space with a resolution of $3 \times 3 \times 3\ mm^3$. To improve signal to noise ratio, spatial smoothing were then applied with a Gaussian kernel of 6 mm. Finally, band-pass filtering (0.01-0.1 Hz) was performed on the time series of each voxel (Price et al., 2014; Huang et al., 2017). These steps are detailed in this link: `http://preprocessed-connectomes-project.org/abide/`. Each brain rfMRI was partitioned into 116 ROIs.

Method parameters. For SIMLR parameters, we empirically set the number of cluster $N_c = 3$ and the number of nearest neighbors $k = 20$. For SNF parameters, we also set the number of nearest neighbors to $q = 20$, the number of iterations $N_t = 20$ as recommended in (Wang et al., 2014) for convergence. Concerning the number of nearest neighbors, we tuned them empirically for both SNF and SIMLR (q and $k = \{10 : 10 : 100\}$). We have remarked that the variation of these parameters did not influence the performance. Indeed, changes in evaluation measures was negligible when varying the number of neighbors. Hence, we opted for setting q and k to 20 which produced the best performance, although the improvement was minimal. To investigate the discriminative power of positive and negative brain connectivities in ASD diagnosis, we evaluated NAG-FS on positive, negative, and whole functional network populations, respectively using LOO-CV as explained in Section 2.3.

Source code. NAG-FS source code is available at `https://github.com/basiralab/NAGFS`.

NAG-FS evaluation using different clustering techniques. We compared SIMLR with the widely used hierarchical clustering (HC) algorithm (Stevens et al., 2000). For instance, HC was previously used to reveal the hierarchical structure of the brain function (Marui et al., 2018). To evaluate the centeredness and representativeness of the network atlas produced by our method, we computed the mean Frobenius distance between estimated network atlas and all individual networks in the population using SIMLR clustering algorithm and HC algorithm with different numbers of clusters (**Fig.** 7) for both ASD and NC classes. Clearly, SIMLR consistently achieved the minimum distance across different numbers of clusters N_c for positive, negative and whole functional network data. In particular, $N_c = 3$

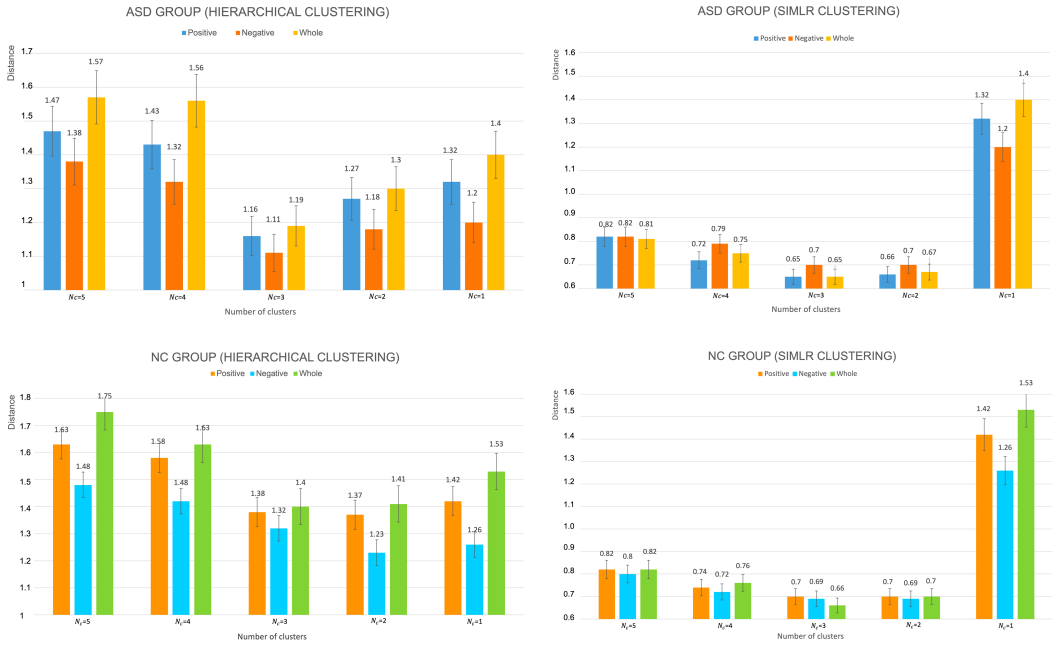

Figure 7: *Evaluation of the estimated network atlas for NC and ASD populations using different clustering techniques.* We display the mean Frobenius distance between estimated network atlas and all individual network atlases in the population using hierarchical clustering (HC) and SIMLR with different numbers of clusters. Notably, SIMLR achieves the minimum distance with $N_c = 3$ for positive, negative and whole data in both ASD and NC groups.

represented the best number of clusters for SIMLR for both ASD and NC groups while $N_c = 2$ represented the best number of clusters for HC for NC group and $N_c = 3$ for ASD group. For $N_c = 5$, HC produced the highest distance for positive $\{1.63, 1.47\}$, negative $\{1.48, 1.38\}$ and whole data $\{1.75, 1.57\}$ for ASD and NC groups respectively. We also notice that the mean Frobenius distance highly increases when $N_c = 1$ (without clustering), which indicates that clustering well captures data heterogeneity. Besides, SIMLR reduced weak similarities and clustered subpopulations more accurately in comparison with HC algorithm.

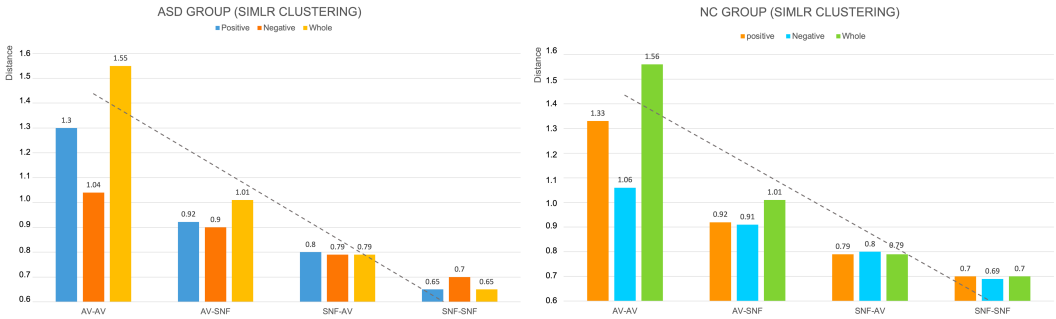

Figure 8: *Evaluation of the estimated network atlas for NC and ASD populations using different fusion strategies.* We display the mean Frobenius distance between estimated network atlas and all individual networks in the population using AV-AV, AV-SNF, SNF-AV and SNF-SNF. Clearly, SNF-SNF achieves the minimum distance for positive, negative and whole data in both ASD and NC groups.

NAG-FS evaluation using different fusion strategies. For network atlas estimation, we benchmarked our method against four network fusion strategies. Particularly, we used AV-AV algorithm which first averages networks in each cluster, then averages the cluster-specific atlas networks to generate the final network. The same process is repeated using AV-SNF, SNF-AV and SNF-SNF. Then, we computed the mean Frobenius distance defined as $d_F(\mathbf{A}, \mathbf{B}) = \sqrt{\sum_i \sum_j |a_{ij} - b_{ij}|^2}$ between the population network atlas and all individual networks in the population. We observe that the best result was given by the SNF-SNF method for positive, negative and whole data followed by SNF-AV, AV-SNF and SNF-SNF, respectively. We also note that linear fusion using AV-AV produced the highest distance for positive {1.30, 1.33}, negative {1.04, 1.06} and whole data {1.50, 1.55} in ASD and NC groups, respectively (**Fig.** 8).

Rank	First region	Second region
1	Right inferior frontal gyrus, pars triangularis (Frontal lobe)	Left globus pallidus (Basal Gonglia lobe)
2	Left inferior frontal gyrus, pars orbitalis (Frontal lobe)	Left gyrus rectus (Frontal lobe)
3	Right inferior frontal gyrus, pars opercularis (Frontal lobe)	Left thalamus (Basal Gonglia lobe)
4	Right inferior frontal gyrus, pars opercularis (Frontal lobe)	Right putamen (Basal Gonglia lobe)
5	Left inferior frontal gyrus, pars opercularis (Frontal lobe)	Left superior temporal gyrus (Temporal lobe)

Table 2: *Top 5 discriminative connections using* ***positive*** *connectomic data.*

Comparison with different feature selection methods. First, we compared NAG-FS with four baseline methods: (1) without using any feature selection method, (2) recursive feature elimination with random forest (RFE-RF) (Granitto et al., 2006), (3) local learning-based clustering feature selection (LLCFS) method (Zeng and Cheung, 2010), and ultimately (4) using a landmark feature selection method (supervised IFS) (Roffo et al., 2015), which was evaluated on 13 datasets and outperformed 8 feature selection methods. Using leave-one-out, we trained an SVM classifier using the top N_f most discriminative features identified by each of the following feature selection methods: (i) RFE-RF, (ii) LLCFS, (iii) IFS, and (iv) the proposed NAG-FS. Clearly, our method achieved the best classification accuracy for the whole connectome (65.03%), positive connectome (59.69%) and negative connectome (63.22%) (**Fig.** 9) with the lowest computational time compared to using IFS (Roffo et al., 2015) and LLCFS (Zeng and Cheung, 2010) (**Fig.** 11), respectively. Although SVM achieved the lowest computational time as it was not combined with any feature selection method (**Fig.** 11), its classification performance was the lowest (**Fig.** 9). The computational time displayed for the proposed and comparison methods was computed using the difference between the time at the end of LOO-CV and the time right before it started.

4. Discussion

The brain network atlas is a recent stimulating discovery in the field of neuroscience connectomics. Estimating representative and reliable brain atlases constitute a fundamental step in several MRI-based neurological disorder mapping, diagnosis, and prognosis. In this paper, we proposed the first network atlas-guided feature selection framework to disentangle the healthy from

Rank	First region	Second region
1	Left superior frontal gyrus (Frontal lobe)	Left superior frontal gyrus, orbital part (Frontal lobe)
2	Right superior frontal gyrus (Frontal lobe)	Right inferior frontal gyrus, pars opercularis (Frontal lobe)
3	Right superior frontal gyrus (Frontal lobe)	Left middle frontal gyrus, orbital part (Frontal lobe)
4	Right superior frontal gyrus (Frontal lobe)	Left inferior frontal gyrus, pars opercularis (Frontal lobe)
5	Left superior frontal gyrus (Frontal lobe)	Right caudate nucleus (Cerebellum lobe)

Table 3: *Top 5 discriminative connections using* ***negative*** *connectomic data.*

Rank	First region	Second region
1	Right posterior cingulate gyrus (Cingulate lobe)	Right crus I of cerebellar hemisphere (Cerebellum lobe)
2	Left posterior cingulate gyrus (Cingulate lobe)	Left caudate nucleus (Basal Ganglia lobe)
3	Right inferior frontal gyrus, pars orbitalis (Frontal lobe)	Left medial orbitofrontal cortex (Frontal lobe)
4	Left supplementary motor areas (Frontal lobe)	Right medial orbitofrontal cortex) (Frontal lobe)
5	Left inferior frontal gyrus, pars orbitalis (Frontal lobe)	Left paracentral lobule (Parietal lobe)

Table 4: *Top 5 discriminative connections using* ***whole*** *connectomic data.*

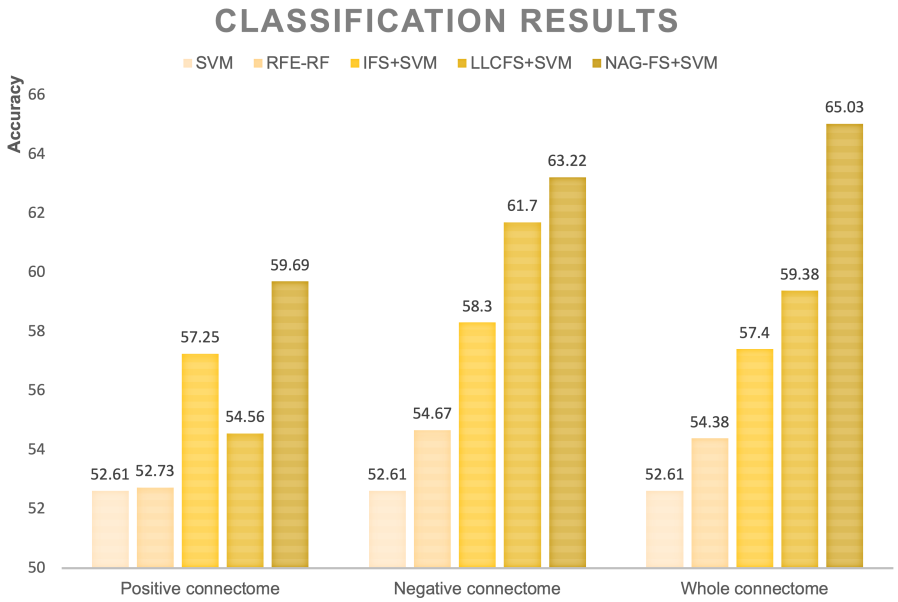

Figure 9: *ASD identification accuracy using our method (NAG-FS + SVM), IFS+SVM (Roffo et al., 2015), LLCFS+SVM (Zeng and Cheung, 2010), RFE-RF+SVM (Granitto et al., 2006), and SVM methods.* The best performance was achieved respectively using positive, negative and whole data by our method.

the disordered connectome while easily pinning down the disorder-specific connectional fingerprint (or features). Our learning-based framework comprises two steps. In the training stage, for each group of individuals, we first learn how to estimate a centered and representative functional network atlas to reliably map the functional connectome and its variability across training individuals, thereby capturing their shared traits (connectional fingerprint of a population). Second, instead of resorting to feature selection methods for network dimensionality reduction, we propose a simple but effective feature selection method called network atlas-guided feature selection (NAG-FS), shifting the way we usually design feature selection methods for biomarker discovery. Our approach has two compelling strengths: (1) the estimated network atlas was most centered and representative in comparison with network atlases produced by baseline methods, and (2) the proposed NAG-FS method largely boosted the classification of ASD patients in comparison with several feature selection methods including the landmark IFS technique.

Insights into baseline methods. *Clustering step.* As shown in **Fig.** 7, hierarchical clustering produced the highest atlas-to-individuals distance for both ASD and NC classes using positive, negative and whole correlations compared to SIMLR. These results can be explained by the fact that the hierarchical clustering is not effective in unraveling the inherent data distribution. For this reason, we adopted SIMLR, a learning-based clustering technique, which learns a pairwise network similarities prior to clustering (Wang et al., 2018). Notably, SIMLR consistently outperformed HC in both NC and ASD populations.

Fusion step. To evaluate the centerdness of the estimated network atlases, we computed the average Frobenius distance between the estimated network atlas and individual networks in the population. As shown in **Fig.** 8, the SNF-SNF method produced the most centered atlas for both ASD and NC populations. This might indicate that alternative network fusion methods simply average all networks without capturing their nonlinear relationship. Therefore, these results demonstrate the effectiveness of SNF in capturing both common patterns and complementary information across samples and producing a representative fused network (Wang et al., 2014).

Classification step. To evaluate the performance of the proposed NAG-FS, we benchmarked it against a landmark IFS (Roffo et al., 2015), which outperformed several feature selection methods in the state-of-the-art, and other feature selection methods including RFE-RF (Granitto et al., 2006), and LLCFS method (Zeng and Cheung, 2010). Clearly, our method achieved

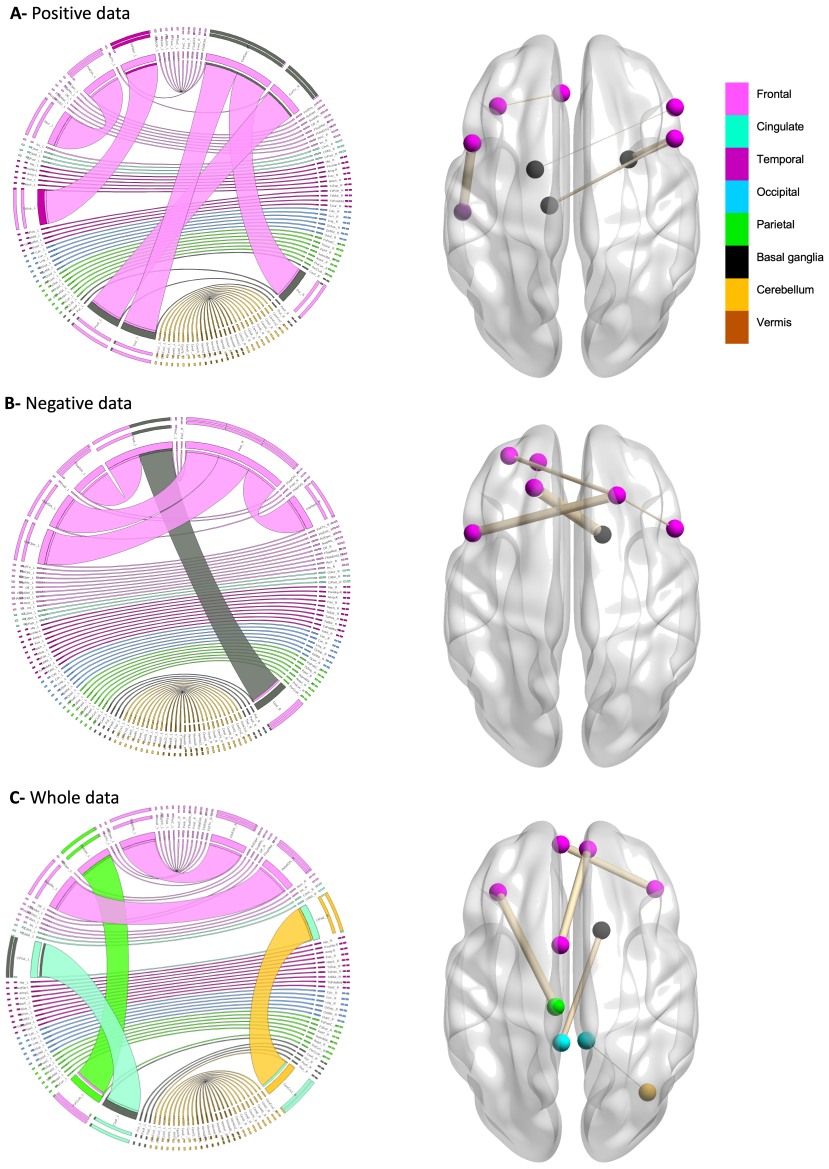

Figure 10: *Strongest connections present the 5 most discriminative network connections between ASD and NC classes using positive, negative, and whole functional connectomes.* The circular graphs were generated using Circos table viewer (Krzywinski et al., 2009). We used BrainNet Viewer Software (Xia et al., 2013) to visualize the regions of interest involving the most discriminative connectivities.

the best results as shown in **Fig.** 9 and **Fig.** 11 in terms of classification accuracy as well as computational time.

Clinical interest. ASD is a progressive neurodegenerative disease. All ASD patients displayed severe social isolation, language impairment, motor stereotypes and insistence on sameness (Caronna et al., 2008). Although autism is a lifelong condition, researchers showed that an early intervention creates a big difference to the development of children especially babies outstanding to improved outcomes for ASD children, including the increased daily and social living skills, higher intelligence (IQ) (Prior et al., 2011; Magiati et al., 2012). Thus, different to previous studies that focused on selecting features for classification without paying attention to the inter-population variabilities and intra-population differences which are much more of a clinical interest, estimating a well-representative brain network atlas for healthy and autistic groups might help clinicians better interpret the altered brain connections.

Insights into most discriminative connectional features. In **Fig.** 10, we display the top 5 most discriminative functional connectivities identified by NAG-FS. We found that the frontal lobe has the highest discriminative power (**Table** 2, **Table** 3, **Table** 4). In particular, for positive functional connectomic data, we notice that all second most discriminative functional brain connectivities involved the frontal lobe. In **Table** 3, we found that 80% of the second most discriminative connectivities also involved the frontal lobe. Furthermore, we identified other discriminative connectivities between different regions such as the parietal cortex, temporal lobe, basal ganglia, cingulate and cerebellum, which were reported in previous studies (Chandana et al., 2005; Sundaram et al., 2008; Ha et al., 2015). The frontal lobe has a major role in speech and language production (Alexander et al., 1989), understanding and reacting to others, forming memories (Curran et al., 1997) and making decisions (Collins and Koechlin, 2012), which might explain the prevalence of altered brain connectivities in this brain region. (Ha et al., 2015) also reported that the frontal, temporal, parietal cortex and basal ganglia mediate clinical phenotypes of ASD. (Balsters et al., 2016) demonstrated that the anterior cingulate cortex activates very weakly in ASD patients. Additionally, the development of the brain during childhood in ASD seems to be predominated by an enlarged brain volume of the frontal lobe (Ha et al., 2015). In the functional brain imaging research, the left inferior and middle frontal gyri have showed hypo-activation in adolescent with ASD compared with NC (Kana et al., 2006). (Kim et al., 2010; Adolphs, 2001) showed that

the mediate impairments of social behaviors caused by the abnormalities in the frontal lobe, superior temporal cortex and parietal cortex. All these studies support the discriminative functional connectivities revealed by our estimated functional network atlases.

Multi-site data heterogeneity. As mentioned above, we evaluated our method using the worldwide multi-site (ABIDE) dataset. ABIDE dataset consists of 1112 subjects with different age, gender and handedness. Hence, all these factors make the data highly heterogeneous. So, to limit this issue, first, we have included 50% of the available data (only 517 subjects). Second, to design a robust method that can handle multi-site datasets in a fully unsupervised way, we have leveraged SIMLR framework in order to disentangle the heterogeneous distribution of the data. Besides, we have used the SNF network fusion algorithm which is robust to data heterogeneity. All these steps help our method better handle data heterogeneity in an unsupervised and generic way, which is blind to data sites/sources.

Computational cost. Most efficient feature selection methods are computationally expensive (Duch et al., 2012; Wong et al., 2018). Devising a fast and efficient feature selection technique with low computationally cost that can meet the formidable challenge of exploding large datasets remains an active area of research. In this work, we demonstrate that brain network atlases can be used to select discriminative and reliable features in a constant computational time $O(1)$ as it simply relies on computing the difference between two matrices. Although RFE-RF method achieved lower computational time than our method, our method boosted the classification performance by more than 11% in comparison to RFE-RF (**Fig.** 11). We also note that the RFE-RF may not scale up to high-dimensional data which makes it less effective in handling connectomic datasets (Darst et al., 2018). Notably, NAG-FS is less computationally expensive than IFS and LLCFS methods. As demonstrated in (Roffo et al., 2015), IFS is less computationally expensive than 7 state-of-the-art methods. This shows that NAG-FS is not only accurate, but also practical and efficient.

Remark 1. We note that the whole pipeline (**Fig.** 2) is in $O(n^3)$. Yet, in our work we would like to accentuate in how to leverage well-estimated brain network atlases to effectively select the most relevant features without resorting to training complex FS methods with high computational time and memory space, whose performance generally largely fluctuates when varying the dataset they are trained on (Georges et al., 2018). Hence, once the network atlases are estimated, we easily and reliably identify the most dis-

criminative features in O(1) time complexity.

Remark 2. In **Fig.** 11 we report the computational time including the network atlas estimation step. However, once the network atlases are estimated, the identification of the most discriminative feature is performed in almost a constant time $O(1)$. In a real-world clinical setting, pre-estimated network atlases can be directly utilized for diagnosis in a constant time.

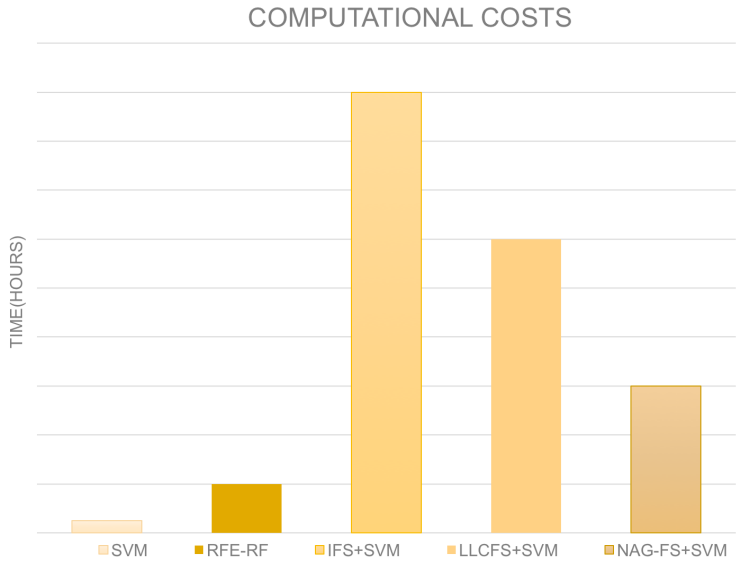

Figure 11: *Computational time.* Run time comparison between the proposed (NAG-FS + SVM) method, (RFE-RF+SVM), (LLCFS+SVM), (IFS+SVM), and SVM methods in ASD/NC classification. Notably, SVM does not use any feature selection technique, hence it is the least time-consuming. All our experiments were performed using a Windows machine with 4 Go RAM and i5 Cores.

Limitations and future directions. Although our proposed method achieved promising results, there are several improvements that can be explored as future work.

Integrating other brain network representations. Our method was evaluated using only functional brain networks derived from rfMRI. We intend to explore the discriminative power of brain network atlases derived from other brain modalities such as structural brain networks estimated from diffusion MRI and morphological brain networks estimated from conventional T1-weighted MRI (Morris and Rekik, 2017; Soussia and Rekik, 2017; Lisowska et al., 2018; Mahjoub et al., 2018; Soussia and Rekik, 2018; Nebli and Rekik, 2019; Seidlitz et al., 2018). Integrating multi-modal connectomic networks

to estimate a holistic network atlas of the brain may provide complementary information to improve ASD/NC diagnosis (Liu et al., 2017).

Atlas parcellation. The produced brain network atlases were derived from a single brain parcellation. Leveraging multiple parcellations will provide a more coarse-to-fine landscape of the effect of a neurological disorder on functional brain connectivities. One can also utilize a multi-modal brain parcellation of the brain to estimate multimodal brain network atlases (Glasser et al., 2016).

Integrating multi-view brain networks. In addition to using an algorithm which integrates *multi-view* brain networks, we can use multi-view clustering techniques to take advantage of the multi-view aspect of the data and create more homogenous clusters across views (Yang and Wang, 2018).

Integrating multi-view clustering techniques. Add to integrating multi-view brain networks, we can use multi-view clustering techniques to take advantage of the multi-view information and unleash the power of knowledge (yang et al. 2018). So, we can simultaneously carry out the clustering using each view of data samples to exploit the complementary information.

Deep learning models as a potential future research direction. At present, most investigators (Guo et al., 2017; Heinsfeld et al., 2018) focused on developing deep learning algorithms thanks to its efficiency in classification task. Nevertheless, these techniques have not been broadly adopted by the neuroimaging studies and still immature (Li et al., 2018). Also, they used modest sample size. And as it is known deep learning algorithms are data hungry which needs massive samples. Consequently, the use of small or modest data size make the reliability and the reproducibility of these approaches debatable. (Wang et al., 2019) proposed another technique called "multi-site adaption framework via low-rank representation decomposition (maLRR)" which addresses deep learning algorithms issues. However, from a neuroscience and treatment perspectives, there is a key limitation that occurs both in deep learning and maLRR methods. These techniques do not allow to track the most discriminative features (i.e., connectivities) which may help to discover the disease-relevant markers, but, they learn new feature representations, which might boost the classification results but unfortunately loose data interpretability. With the increase of the public pool of connectomic datasets, designing *connectivity-based* deep learning models for estimating centered and well-representative network atlases without loss of interpretability and tractability would be of high utility in mapping the human brain and diagnosing neurological disorders.

5. Conclusion

In this paper, we proposed how to estimate representative and centered brain network atlases, which can be leveraged to identify discriminative brain connectivities between healthy and disordered populations. The proposed network atlas-guided feature selection (NAG-FS) achieved the best results in terms of classification accuracy and overall computational time in comparison to state-of-the-art feature selection methods. We also used NAG-FS to investigate the discriminative power of positive, negative, and whole functional connectomes in classifying ASD and NC subjects, thereby identifying disordered brain connectivities fingerprinting autism spectrum disorder. To sum up, the estimation of centered brain network atlases provides a new and exciting venue for better understanding a wide spectrum neurological disorders and easily and effectively spotting reliable biomarkers for improving diagnosis and prognosis. In our future work, we will further investigate the discriminative power of network atlases in the diagnosis of neurodegenerative disorders including dementia.

Abraham, A., Milham, M.P., Di Martino, A., Craddock, R.C., Samaras, D., Thirion, B., Varoquaux, G., 2017. Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example. NeuroImage 147, 736–745.

Adler, D.H., Wisse, L.E., Ittyerah, R., Pluta, J.B., Ding, S.L., Xie, L., Wang, J., Kadivar, S., Robinson, J.L., Schuck, T., et al., 2018. Characterizing the human hippocampus in aging and Alzheimers disease using a computational atlas derived from ex vivo MRI and histology. Proceedings of the National Academy of Sciences 115, 4252–4257.

Adolphs, R., 2001. The neurobiology of social cognition. Current opinion in neurobiology 11, 231–239.

Alexander, M.P., Benson, D.F., Stuss, D.T., 1989. Frontal lobes and language. Brain and language 37, 656–691.

Amaral, D.G., Schumann, C.M., Nordahl, C.W., 2008. Neuroanatomy of autism. Trends in neurosciences 31, 137–145.

Association, A.P., et al., 2013. Diagnostic and statistical manual of mental disorders (DSM-5) .

Autism, Investigators, D.D.M.N.S.Y..P., 2014. Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, united states, 2010. Morbidity and Mortality Weekly Report: Surveillance Summaries 63, 1–21.

Balsters, J.H., Apps, M.A., Bolis, D., Lehner, R., Gallagher, L., Wenderoth, N., 2016. Disrupted prediction errors index social deficits in autism spectrum disorder. Brain 140, 235–246.

Brown, C.J., Hamarneh, G., 2016. Machine learning on human connectome data from MRI. arXiv preprint arXiv:1611.08699 .

Caronna, E.B., Milunsky, J.M., Tager-Flusberg, H., 2008. Autism spectrum disorders: clinical and research frontiers. Archives of Disease in Childhood 93, 518–523.

Chandana, S.R., Behen, M.E., Juhász, C., Muzik, O., Rothermel, R.D., Mangner, T.J., Chakraborty, P.K., Chugani, H.T., Chugani, D.C., 2005.

Significance of abnormalities in developmental trajectory and asymmetry of cortical serotonin synthesis in autism. International Journal of Developmental Neuroscience 23, 171–182.

Chen, G., Chen, G., Xie, C., Li, S.J., 2011. Negative functional connectivity and its dependence on the shortest path length of positive network in the resting-state human brain. Brain connectivity 1, 195–206.

Collins, A., Koechlin, E., 2012. Reasoning, learning, and creativity: frontal lobe function and human decision-making. PLoS biology 10, e1001293.

Curran, T., Schacter, D.L., Norman, K.A., Galluccio, L., 1997. False recognition after a right frontal lobe infarction: Memory for general and specific information. Neuropsychologia 35, 1035–1049.

Darst, B.F., Malecki, K.C., Engelman, C.D., 2018. Using recursive feature elimination in random forest to account for correlated variables in high dimensional data. BMC genetics 19, 65.

Dhifallah, S., Rekik, I., Alzheimer's Disease Neuroimaging Initiative, 2018. Clustering-based multi-view network fusion for estimating brain network atlases of healthy and disordered populations. Journal of neuroscience methods .

Dryburgh, E., McKenna, S., Rekik, I., 2019. Predicting full-scale and verbal intelligence scores from functional connectomic data in individuals with autism spectrum disorder. Brain imaging and behavior , 1–10.

Duch, W., Nowak, W., Meller, J., Osiński, G., Dobosz, K., Mikołajewski, D., Wójcik, G.M., 2012. Computational approach to understanding autism spectrum disorders. Computer Science 13, 47–61.

Ecker, C., Marquand, A., Mourão-Miranda, J., Johnston, P., Daly, E.M., Brammer, M.J., Maltezos, S., Murphy, C.M., Robertson, D., Williams, S.C., et al., 2010. Describing the brain in autism in five dimensions-magnetic resonance imaging-assisted diagnosis of autism spectrum disorder using a multiparameter classification approach. Journal of Neuroscience 30, 10612–10623.

Georges, N., Rekik, I., et al., 2018. Data-specific feature selection method identification for most reproducible connectomic feature discovery fingerprinting brain states. International Workshop on Connectomics in Neuroimaging , 99–106.

Gholipour, A., Kehtarnavaz, N., Briggs, R., Devous, M., Gopinath, K., 2007. Brain functional localization: a survey of image registration techniques. IEEE transactions on medical imaging 26, 427–451.

Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., et al., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178.

Gönen, M., Alpaydın, E., 2011. Multiple kernel learning algorithms. Journal of machine learning research 12, 2211–2268.

Granitto, P.M., Furlanello, C., Biasioli, F., Gasperi, F., 2006. Recursive feature elimination with random forest for ptr-ms analysis of agroindustrial products. Chemometrics and Intelligent Laboratory Systems 83, 83–90.

Guo, X., Dominick, K.C., Minai, A.A., Li, H., Erickson, C.A., Lu, L.J., 2017. Diagnosing autism spectrum disorder from brain resting-state functional connectivity patterns using a deep neural network with a novel feature selection method. Frontiers in neuroscience 11, 460.

Ha, S., Sohn, I.J., Kim, N., Sim, H.J., Cheon, K.A., 2015. Characteristics of brains in autism spectrum disorder: structure, function and connectivity across the lifespan. Experimental neurobiology 24, 273–284.

Heinsfeld, A.S., Franco, A.R., Craddock, R.C., Buchweitz, A., Meneguzzi, F., 2018. Identification of autism spectrum disorder using deep learning and the abide dataset. NeuroImage: Clinical 17, 16–23.

Huang, X.H., Song, Y.Q., Liao, D.A., Lu, H., 2017. Detecting community structure based on optimized modularity by genetic algorithm in resting-state fMRI. International Symposium on Neural Networks , 457–464.

Kana, R.K., Keller, T.A., Cherkassky, V.L., Minshew, N.J., Just, M.A., 2006. Sentence comprehension in autism: thinking in pictures with decreased functional connectivity. Brain 129, 2484–2493.

Khalid, S., Khalil, T., Nasreen, S., 2014. A survey of feature selection and feature extraction techniques in machine learning. Science and Information Conference (SAI), 2014 , 372–378.

Kim, J.E., Lyoo, I.K., Estes, A.M., Renshaw, P.F., Shaw, D.W., Friedman, S.D., Kim, D.J., Yoon, S.J., Hwang, J., Dager, S.R., 2010. Laterobasal amygdalar enlargement in 6-to 7-year-old children with autism spectrum disorder. Archives of general psychiatry 67, 1187–1197.

Kosmicki, J., Sochat, V., Duda, M., Wall, D., 2015. Searching for a minimal set of behaviors for autism detection through feature selection-based machine learning. Translational psychiatry 5, e514.

Krzywinski, M.I., Schein, J.E., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., Marra, M.A., 2009. Circos: An information aesthetic for comparative genomics. Genome Research http://genome.cshlp.org/content/early/2009/06/15/gr.092759.109.full.pdf+html.

Leigh, J.P., Du, J., 2015. Brief report: Forecasting the economic burden of autism in 2015 and 2025 in the United States. Journal of Autism and Developmental Disorders 45, 4135–4139.

Li, H., Parikh, N.A., He, L., 2018. A novel transfer learning approach to enhance deep neural network classification of brain functional connectomes. Frontiers in neuroscience 12, 491.

Lisowska, A., Rekik, I., Alzheimer's Disease Neuroimaging Initiative, 2018. Joint pairing and structured mapping of convolutional brain morphological multiplexes for early dementia diagnosis. Brain connectivity .

Liu, M., Zhang, D., Shen, D., Initiative, A.D.N., 2015. View-centralized multi-atlas classification for alzheimer's disease diagnosis. Human brain mapping 36, 1847–1865.

Liu, M., Zhang, J., Yap, P.T., Shen, D., 2017. View-aligned hypergraph learning for Alzheimers disease diagnosis with incomplete multi-modality data. Medical image analysis 36, 123–134.

Magiati, I., Tay, X.W., Howlin, P., 2012. Early comprehensive behaviorally based interventions for children with autism spectrum disorders: a summary of findings from recent reviews and meta-analyses. Neuropsychiatry 2, 543.

Mahjoub, I., Mahjoub, M.A., Rekik, I., 2018. Brain multiplexes reveal morphological connectional biomarkers fingerprinting late brain dementia states. Scientific reports 8, 4103.

Marui, W., Kan, S., Nii, M., Shibata, M., Kobashi, S., 2018. Md-lms algorithm based brain functional connectivity analysis in resting state fMRI. 2018 World Automation Congress (WAC) , 1–5.

McGrath, J., Johnson, K., O'Hanlon, E., Garavan, H., Leemans, A., Gallagher, L., 2013. Abnormal functional connectivity during visuospatial processing is associated with disrupted organisation of white matter in autism. Frontiers in human neuroscience 7, 434.

Minshew, N.J., Keller, T.A., 2010. The nature of brain dysfunction in autism: functional brain imaging studies. Current opinion in neurology 23, 124.

Morris, C., Rekik, I., 2017. Autism spectrum disorder diagnosis using sparse graph embedding of morphological brain networks. Graphs in Biomedical Image Analysis, Computational Anatomy and Imaging Genetics , 12–20.

Nebli, A., Rekik, I., 2019. Gender differences in cortical morphological networks. Brain imaging and behavior , 1–9.

Parente, F., Colosimo, A., 2016. The role of negative links in brain networks. Biophysics and Bioengineering Letters 9.

Price, T., Wee, C.Y., Gao, W., Shen, D., 2014. Multiple-network classification of childhood autism using functional connectivity dynamics. International Conference on Medical Image Computing and Computer-Assisted Intervention , 177–184.

Prior, M., Roberts, J.M., Rodger, S., Williams, K., Sutherland, R., 2011. A review of the research to identify the most effective models of practice in early intervention for children with autism spectrum disorders. Australian Government Department of Families, housing, Community Services and Indigenous Affairs (FaHCSIA). Australia .

Rekik, I., Li, G., Lin, W., Shen, D., 2017. Estimation of brain network atlases using diffusive-shrinking graphs: Application to developing brains. International Conference on Information Processing in Medical Imaging , 385–397.

Rekik, I., Li, G., Lin, W., Shen, D., 2018a. Do baby brain cortices that look alike at birth grow alike during the first year of postnatal development? International Conference on Medical Image Computing and Computer-Assisted Intervention , 566–574.

Rekik, I., Li, G., Lin, W., Shen, D., 2018b. Estimation of shape and growth brain network atlases for connectomic brain mapping in developing infants. Biomedical Imaging (ISBI 2018), 2018 IEEE 15th International Symposium on , 985–989.

Roffo, G., Melzi, S., Cristani, M., 2015. Infinite feature selection. Proceedings of the IEEE International Conference on Computer Vision , 4202–4210.

Seidlitz, J., Váša, F., Shinn, M., Romero-Garcia, R., Whitaker, K.J., Vértes, P.E., Wagstyl, K., Reardon, P.K., Clasen, L., Liu, S., et al., 2018. Morphometric similarity networks detect microscale cortical organization and predict inter-individual cognitive variation. Neuron 97, 231–247.

Shams, W.K., Rahman, A.W.A., 2011. Characterizing autistic disorder based on principle component analysis. Industrial Electronics and Applications (ISIEA), 2011 IEEE Symposium on , 653–657.

Smith, S.M., Nichols, T.E., Vidaurre, D., Winkler, A.M., Behrens, T.E., Glasser, M.F., Ugurbil, K., Barch, D.M., Van Essen, D.C., Miller, K.L., 2015. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nature neuroscience 18, 1565.

Soussia, M., Rekik, I., 2017. High-order connectomic manifold learning for autistic brain state identification. International Workshop on Connectomics in Neuroimaging , 51–59.

Soussia, M., Rekik, I., 2018. Unsupervised manifold learning using high-order morphological brain networks derived from T1-w MRI for autism diagnosis. Frontiers in Neuroinformatics 12.

Stevens, M.C., Fein, D.A., Dunn, M., Allen, D., Waterhouse, L.H., Feinstein, C., Rapin, I., 2000. Subgroups of children with autism by cluster analysis: A longitudinal examination. Journal of the American Academy of Child & Adolescent Psychiatry 39, 346–352.

Sundaram, S.K., Kumar, A., Makki, M.I., Behen, M.E., Chugani, H.T., Chugani, D.C., 2008. Diffusion tensor imaging of frontal lobe in autism spectrum disorder. Cerebral cortex 18, 2659–2665.

Tang, C., Wei, Y., Zhao, J., Nie, J., 2018. The dynamic measurements of regional brain activity for resting-state fMRI: d-ALFF, d-fALFF and d-ReHo. International Conference on Medical Image Computing and Computer-Assisted Intervention , 190–197.

Tang, J., Alelyani, S., Liu, H., 2014. Feature selection for classification: A review. Data classification: Algorithms and applications , 37.

Toga, A.W., Thompson, P.M., 2001. The role of image registration in brain mapping. Image and vision computing 19, 3–24.

Uddin, L.Q., Supekar, K., Lynch, C.J., Khouzam, A., Phillips, J., Feinstein, C., Ryali, S., Menon, V., 2013. Salience network–based classification and prediction of symptom severity in children with autism. JAMA psychiatry 70, 869–879.

Van Den Heuvel, M.P., Pol, H.E.H., 2010. Exploring the brain network: a review on resting-state fMRI functional connectivity. European neuropsychopharmacology 20, 519–534.

Wang, B., Mezlini, A.M., Demir, F., Fiume, M., Tu, Z., Brudno, M., Haibe-Kains, B., Goldenberg, A., 2014. Similarity network fusion for aggregating data types on a genomic scale. Nature methods 11, 333.

Wang, B., Ramazzotti, D., De Sano, L., Zhu, J., Pierson, E., Batzoglou, S., 2018. SIMLR: A Tool for Large-Scale Genomic Analyses by Multi-Kernel Learning. Proteomics 18, 1700232.

Wang, M., Zhang, D., Huang, J., Yap, P.T., Shen, D., Liu, M., 2019. Identifying autism spectrum disorder with multi-site fmri via low-rank domain adaptation. IEEE Transactions on Medical Imaging .

Wee, C.Y., Yap, P.T., Shen, D., 2016. Diagnosis of autism spectrum disorders using temporally distinct resting-state functional connectivity networks. CNS neuroscience & therapeutics 22, 212–219.

Wong, E., Anderson, J.S., Zielinski, B.A., Fletcher, P.T., 2018. Riemannian regression and classification models of brain networks applied to autism. International Workshop on Connectomics in Neuroimaging , 78–87.

Xia, M., Wang, J., He, Y., 2013. BrainNet Viewer: a network visualization tool for human brain connectomics. PloS one 8, e68910.

Xue, B., Zhang, M., Browne, W.N., Yao, X., 2016. A survey on evolutionary computation approaches to feature selection. IEEE Transactions on Evolutionary Computation 20, 606–626.

Yang, Y., Wang, H., 2018. Multi-view clustering: a survey. Big Data Mining and Analytics 1, 83–107.

Yap, P.T., Wu, G., Shen, D., 2010. Human brain connectomics: networks, techniques, and applications. IEEE Signal Processing Magazine 27, 131–134.

Zeng, H., Cheung, Y.m., 2010. Feature selection and kernel learning for local learning-based clustering. IEEE transactions on pattern analysis and machine intelligence 33, 1532–1547.

Zhao, F., Zhang, H., Rekik, I., An, Z., Shen, D., 2018. Diagnosis of autism spectrum disorders using multi-level high-order functional networks derived from resting-state functional MRI. Frontiers in human neuroscience 12.