High-order Connectomic Manifold Learning for Autistic Brain State Identification

Soussia, Maysa ; Rekik, Islem

Published in:
Connectomics in NeuroImaging

DOI:
10.1007/978-3-319-67159-8_7

Publication date:
2017

Document Version
Peer reviewed version

The final publication is available at Springer via http://dx.doi.org/10.1007/978-3-319-67159-8

High-order Connectomic Manifold Learning for Autistic Brain State Identification

Mayssa Soussia[1,2] and Islem Rekik[1⋆]

[1]BASIRA lab, CVIP group, School of Science and Engineering, Computing, University of Dundee, UK
[2]The National Engineering School of Tunis (ENIT), Department of Electrical Engineering, Tunis, Tunisia

Abstract. Previous studies have identified *disordered* functional (from fMRI) and structural (from diffusion MRI) brain connectivities in Autism Spectrum Disorder (ASD). However, 'shape connections' between brain regions were rarely investigated in ASD –e.g., how morphological attributes of a specific brain region (e.g., sulcal depth) change in relation to morphological attributes in other regions. In this paper, we use conventional T1-w MRI to define morphological connectivity networks, each quantifying shape similarity between different cortical regions for a specific cortical attribute at both *low-order* and *high-order* levels. For ASD identification, we present a connectomic manifold learning framework, which learns multiple kernels to estimate a similarity measure between ASD and normal controls (NC) connectomic features, to perform dimensionality reduction for clustering ASD and NC subjects. We benchmark our ASD identification method against supervised and unsupervised state-of-the-art methods, while depicting the most discriminative high- and low-order relationships between morphological regions in the left and right hemispheres.

1 Introduction

Autism spectrum disease (ASD) is a neurodevelopmental disorder characterized by altered cognitive functions, specifically difficulties in learning and impairment in social interaction. The centers for Disease Control and Prevention (CDC) estimates autism's prevalence as 1 in 68 children in the United States. Recently, interest in understanding how ASD alters connectivity between different brain regions has grown with the development of important technological and methodological neuroimaging tools. Most of connectomic studies on ASD in the literature [1,2,3] have mainly focused on structural and functional connectivity (FC) derived from diffusion tensor imaging (DTI) and functional magnetic resonance imaging (fMRI), respectively. For example, [4] used functional MRI to quantify consistent spatial temporal FC patterns to distinguish between ASD subjects and normal controls (NC). Another work [5] applied a unified connectivity framework

⋆ Corresponding author: irekik@dundee.ac.uk, www.basira-lab.com

on DTI that learns the underlying patterns of ASD pathology through projective non-negative decomposition into sets of discriminative, developmental and reconstructive components. However, one of the potential limitations of solely relying on fMRI or DTI are: (1) fMRI has low signal-to-noise ratio possibly induced by non-neural noise, hence its derived functional connectivity strength between pairs of ROIs can be spurious or noisy, and (2) fiber tractography methods can produce largely variable and somewhat biased structural brain networks [6]. As an alternative, we propose a different type of brain network: a *morphological* network *solely* constructed from T1-w MRI. Since recent research on ASD reported brain cortical thickness changes in autism during early childhood [7], we specifically propose to build different morphological networks based on the morphology of the cortical surface, where each network is associated with a unique low-order cortical attribute such as sulcal depth or cortical thickness. However, since simply concatenating morphological brain networks overlooks how their relationship might be affected at a *higher-order* level by autism, we introduce *morphological high-order brain networks* for autism identification. Unlike functional high-order networks which model the dynamic brain activity within a time-window [8], our high-order network (HON) is able to characterize more complex interaction patterns among brain regions *in morphology*.

On the other hand, a very limited number of works used machine-learning methods on human connectome data from MRI for ASD/NC classification [9], in a supervised manner. For instance, [10] used a functional network estimated from resting state fMRI to distinguish between ASD subjects and healthy controls. [11], which adopted a network regularized support vector machine, used DTI to identify faulty sub-networks associated with ASD. However, while the majority of supervised machine-learning techniques are somewhat limited in terms of scalability as they require reliable and accurate labeling of medical data, unsupervised learning techniques can provide decision support for early intervention, help develop data-driven guidelines for care plan management, and help group patients by similar non-semantic features (i.e., latent representation of brain disorder group or subgroup), to enable comparative effectiveness research (e.g., of medications) [12]. From a connectomic perspective, very few studies applied unsupervised learning methods for brain disease applications [9]. For instance, [13] computed spectral graph clustering to identify significant connectome modules for different brain disorder groups (Alzheimer's disease (AD) and Significant Memory Concern (SMC)). Another work [14] used a multi-view spectral clustering to group functional and structural brain networks of traumatic brain injury (TBI) patients. However, to the best of our knowledge, no previous unsupervised learning methods were used to distinguish between autistic and healthy brains [9].

To fill this gap, we propose a high-order morphological connectomic manifold learning for ASD identification using a novel unsupervised data clustering method called single-cell interpretation via multikernel learning (SIMLR) [15]. SIMLR has many appealing aspects. First, it inputs high-order networks and efficiently learns a similarity matrix between networks by combining multiple

kernels which provides a better fit to the inherent statistical distribution of the HON data. Second, it is scalable and separates subpopulations more accurately than conventional methods (e.g., PCA or t-SNE). Third, it improves weak similarities between samples through graph diffusion, which adds transitive similarities between dissimilar regions that have many similar neighboring regions. We compare our framework with both supervised and unsupervised disease identification techniques. To the best of our knowledge, this is the first work that: (1) defines high-order *morphological* brain networks, (2) jointly integrates multiple cortical morphological brain networks for autism identification, and (3) utilizes unsupervised SIMLR technique on ASD connectomic data.

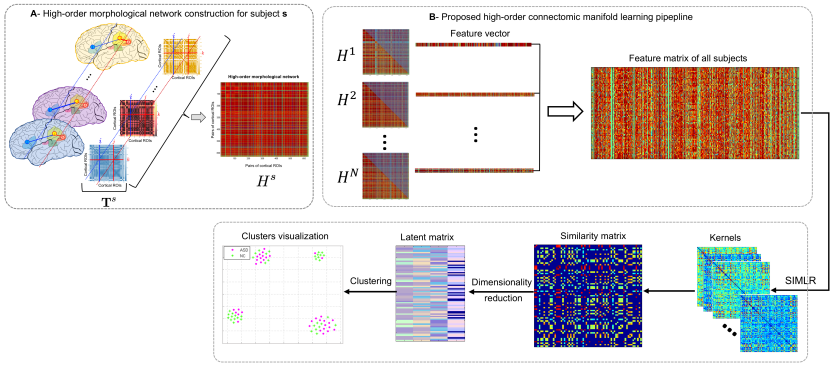

Fig. 1: *Illustration of the proposed high-order connectomic manifold learning for autistic brain state identification.* (A) High-order morphological network construction using multiple brain networks, each measuring a unique cortical attribute (e.g., thickness) on the cortical surface. These are stacked together to form a morphological brain tensor $\mathcal{T}^{\mathbf{s}}$ for subject s. (B) Given the high-order feature matrix of all subjects, we used SIMLR [15] to learn proper weights for multiple kernels, which measure different distances between subjects. Next, we use the learned kernels to construct a symmetric similarity matrix $\mathbf{S}$ between subjects. SIMLR imposes a low-rank constraint on $\mathbf{S}$ such that different populations of the input data will be embedded into independent block-diagonal structure that clusters similar samples. This outputs a latent data representation in a low-dimensional space, which is inputted to a clustering algorithm. Each point in the 2D scatter plot represents an ASD or NC subject, and the corresponding colors represent the true labels in each cluster.

2 High-order Connectomic Manifold Learning for Unsupervised Clustering of Autistic and Healthy Brains

In this section, we present the high-order connectomic manifold learning for ASD identification using multiple kernels based on SIMLR technique introduced

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

Mathematical notation	Definition
$\mathcal{T}^{\mathbf{s}}$	brain tensor of subject s in $\mathbb{R}^{n_r \times n_r \times n_v}$
$\mathbf{X}^k$	brain network in $\mathbb{R}^{n_r \times n_r}$ denoting the k-th frontal-view of tensor $\mathcal{T}$
$\mathbf{H}^s$	high-order morphological brain network for subject s
$\mathbf{h}_s$	high-order feature vector extracted from the upper triangular part of $\mathbf{H}^s$
$\mathbf{K}_l$	l-th learning kernel in $\mathbb{R}^{n \times n}$
n	number of subjects
$\mathbf{S}$	similarity matrix in $\mathbb{R}^{n \times n}$ for connectomic manifold learning
$\mathbf{L}$	latent matrix in $\mathbb{R}^{n \times c}$
c	number of clusters
m	number of kernels
$\mathbf{w}$	weighting vector of the kernels in $\mathbb{R}^m$
$\mathbf{I}_n$	identity matrix in $\mathbb{R}^{n \times n}$

in [15]. We denote tensors by boldface Euler script letters, e.g., $\mathcal{X}$. Matrices are denoted by boldface capital letters, e.g., $\mathbf{X}$, and scalars are denoted by lowercase letters, e.g., x. For easy reference and enhancing the readability, we have summarized the major mathematical notations in **Table** 1. **Fig.** 1 illustrates the proposed pipeline for ASD/NC identification in four major steps. 1) construction of low-order morphological network 2) construction of high-order morphological network 3) feature extraction 4) connectomic manifold learning using SIMLR.

Low-order morphological network construction. For each subject s, we construct a brain tensor $\mathcal{T}^{\mathbf{s}}$ of size $\mathbb{R}^{n_r \times n_r \times n_v}$ for each cortical hemisphere, where n_r is the number of cortical regions of interest (ROIs) and n_v is the number of the tensor frontal views. Basically, for each cortical attribute (e.g., thickness), we construct a morphological brain network that constitutes a frontal view in $\mathcal{T}^{\mathbf{s}}$. Let x_i^k and x_j^k denote the mean of a cortical attribute of the i-th ROI and the j-th ROI in the k-th frontal view respectively. We then compute the absolute difference between x_i^k and x_j^k which depicts the connectivity strength between ROIs i and j. An element in the i-th row and j-th column of the k-th frontal view $\mathbf{X}^k$ is defined as: $\mathbf{X}_{ij}^k = |x_i^k - x_j^k|$.

High-order morphological network construction (HON). As the low-order network is unable to reveal the intrinsic similarities between more than a pair of ROIs, we propose to construct a high-order morphological network based on Pearson correlation to detect more complex interaction patterns between multiple brain regions. In addition to maintaining the pairwise relationship between ROIs in the same morphological view, the morphological HON underlines the relationship between ROIs across different views. Let y_{ij}^s denote the vector of the s-th subject corresponding to the connectivity strength between the i-th and j-th ROIs across all views. Each row in the high-order network $\mathbf{H}^s$ represents a pair of ROIs (i, j) and each column denotes a pair of ROIs (p, q). For a subject s, an element in H^s is defined using the Pearson's correlation coefficient as $\mathbf{H}_{ij,pq}^s = corr(y_{ij}^s, y_{pq}^s)$. We note that the entries $\mathbf{H}_{ij,pq}^s$ of the HON matrix

indicate the influence of the connectivity strength between the i-th and j-th ROI on the connectivity strength between the p-th and q-th ROI. Thus, it underlines the higher order relationship between multiple ROIs.

Feature Extraction. For each subject, features are extracted in a naive way. Due to their symmetry, we concatenate the upper triangle elements of the HON matrix for subject s into a long feature vector $\mathbf{h}^s$. The weights on the diagonal are set to zero to avoid self-connectedness. Next, using K-fold data partition scheme, the extracted features of all ASD and NC subjects, while excluding the K-th fold, are fed into SIMLR.

High-order connectomic manifold learning. In this section, we briefly present the framework introduced in [15] and how we extended it to our aim. The main idea of SIMLR is to learn a pairwise similarity matrix of size $n \times n$ from an input matrix of size $n \times d$ where n is the number of subjects and d is the dimension of their associated feature vectors. This allows to learn the connectomic manifold where all HON features $\{\mathbf{h}^1, \ldots, \mathbf{h}^n\}$ are nested. Instead of using one predefined distance metric which may fail to capture the nonlinear relationship in the data, we use multiple Gaussian kernels with learned weights to better explore in depth the similarity patterns among ASD and NC HONs. In other words, adopting multiple kernels allows to better fit the true underlying statistical distribution of the input matrix of high-order features. Additionally, constraints are imposed on kernel weights to avoid a single kernel selection [15]. The Gaussian kernel is expressed as follows: $\mathbf{K}(\mathbf{h}^i, \mathbf{h}^j) = \frac{1}{\epsilon_{ij}\sqrt{2\pi}} e^{(-\frac{|\mathbf{h}^i - \mathbf{h}^j|^2}{2\epsilon_{ij}^2})}$, where $\mathbf{h}^i$ and $\mathbf{h}^j$ denote the feature vectors of the i-th and j-th subjects respectively and ϵ_{ij} is defined as: $\epsilon_{ij} = \sigma(\mu_i + \mu_j)/2$, where σ is a tuning parameter and $\mu_i = \frac{\sum_{l \in KNN(\mathbf{h}^i)} |\mathbf{h}^i - \mathbf{h}^j|}{k}$, where $KNN(\mathbf{h}^i)$ represents the top k neighboring subjects of subject i. The computed kernels are then averaged to further learn the similarity matrix $\mathbf{S}$ through an optimization framework formulated as follows:

$$\min_{\mathbf{S},\mathbf{L},\mathbf{w}} \sum_{i,j} -w_l \mathbf{K}_l(\mathbf{h}^i, \mathbf{h}^j)\mathbf{S}_{ij} + \beta||\mathbf{S}||_F^2 + \gamma \mathbf{tr}(\mathbf{L}^T(\mathbf{I}_n - \mathbf{S})\mathbf{L}) + \rho \sum_l w_l log w_l \quad (1)$$

Subject to: $\sum_l w_l = 1$, $w_l \geq 0$, $\mathbf{L}^T\mathbf{L} = \mathbf{I}_c$, $\sum_j \mathbf{S}_{ij} = 1$, and $\mathbf{S}_{ij} \geq 0$ for all (i, j), where:

1. $\sum_{i,j} -w_l \mathbf{K}_l(\mathbf{h}^i, \mathbf{h}^j)\mathbf{S}_{ij}$ refers to the relation between the similarity and the kernel distance with weights w_l between two subjects. The learned similarity should be small if the distance between a pair of subjects is large.
2. $\beta||\mathbf{S}||_F^2$ denotes a regularization term that avoids over-fitting the model to the data.
3. $\gamma \mathbf{tr}(\mathbf{L}^T(\mathbf{I}_n - \mathbf{S})\mathbf{L})$: $\mathbf{L}$ is the latent matrix of size $n \times c$ where n is the number of subjects and c is the number of clusters. The matrix $(\mathbf{I}_n - \mathbf{S})$ denotes the graph Laplacian.
4. $\rho \sum_l w_l log w_l$ imposes constraints on the kernel weights to avoid selection of a single kernel.

Fig. 2: *ASD identification accuracy using our method and comparison supervised and unsupervised methods.* We evaluated each of these methods on i) the concatenated low-order morphological networks (i.e., 4 views) that we term with CON, and ii) the high-order morphological networks (HON).

An alternating convex optimization is adopted where each variable is optimized while fixing the other variables until convergence [15]. Once, the similarity matrix $\mathbf{S}$ is obtained, a dimensionality reduction is performed on $\mathbf{S}$ using t-SNE [16]. In other words, the data is projected onto a lower dimension that preserves the similarity depicted in $\mathbf{S}$ resulting in an $n \times c$ latent matrix $\mathbf{L}$. For visualization, the same algorithm is used to create an embedding of $\mathbf{S}$ in a 2D space. A K-means clustering is then applied to the latent matrix $\mathbf{L}$ to cluster similar subjects and assess the concordance with the true labels (**Fig.** 1). It should be noted that the true labels were only used in the form of distinct colors to intuitively visualize the groups in (**Fig.** 1).

3 Results and Discussion

Evaluation dataset and method parameters. We evaluated the proposed clustering framework on 80 subjects (40 ASD and 40 NC) from Autism Brain

Imaging Data Exchange (ABIDE I)[1] public dataset, each with structural T1-w MR image [17]. We used FREESURFER to reconstruct both right and left cortical hemispheres for each subject from T1-w MRI. Then we parcellated each cortical hemisphere into 35 cortical regions using Desikan-Killiany Atlas. For each subject, we generated $n_v = 4$ cortical morphological networks: $\mathbf{X}^1$ denotes the maximum principal curvature brain view, $\mathbf{X}^2$ denotes the mean cortical thickness brain view, $\mathbf{X}^3$ denotes the mean sulcal depth brain view, and $\mathbf{X}^4$ denotes the mean of average curvature. For SIMLR parameters, using a nested grid search, we set the number of clusters to $c = 4$. We used $m = 21$ kernels where each kernel is determined by a set of hyperparameters ($\sigma = 1 : 0.25 : 2.5$, number of top KNN neighbors in $\{10, 12, 14\}$), where σ is the variance parameter of the Gaussian function.

Method evaluation and comparison methods. To evaluate the reproducibility of our high-order connectomic manifold learning and clustering framework, we used two k-fold cross-validation schemes ($k = 5$ and $k = 10$) using randomized partitioning of data samples. The process was repeated 30 times and the average classification performance reported as final result for all comparison methods. We first compared our ASD/NC clustering method with Ward's linkage clustering [18], a widely used hierarchical clustering algorithm which optimizes a Euclidean objective function as a criterion for merging a pair of clusters at each step. This method was previously used for clustering functional HON networks for Alzheimer's disease diagnosis in [8]. Second, we compared the ASD/NC segregation efficiency of our method with two classification frameworks based on supervised linear support vector machine (SVM) classifier. Specifically, the first supervised method learns a single SVM using training connectomic features. To further evaluate SIMLR in a supervised manner, we propose a SIMLR-based ensemble classifier learning framework, where we use SIMLR to cluster the training data into different clusters, and then train an SVM classifier for each training cluster. In the testing stage, we use label majority voting by all trained SVM classifiers to label an input testing subject. Each of these methods was evaluated on i) the concatenated low-order morphological networks (i.e., 4 views) that we term with CON, and ii) the high-order morphological networks (HON). **Fig.** 2 displays ASD identification accuracies of all methods.

For the left hemisphere (LH), our method (**Fig.** 2–unsupervised SIMLR HON) had the best performance in distinguishing between ASD/NC subjects among all methods using both 5-fold and 10-fold cross-validation schemes, with an average performance of 63.64%. We note that the accuracies increased in average with HON across all methods, which might indicate that LH has more discriminative regions at a higher-order morphological level, except for the supervised SIMLR based ensemble SVM which scored better with CON. The low performance of supervised SIMLR-based ensemble SVM can be explained by the fact that SIMLR tends to produce more homogenous clusters, hence creating a non-balanced data samples for SVM training. This points to the imbalanced data issue for training supervised methods. On the other hand, results for the right

[1] http://fcon_1000.projects.nitrc.org/indi/abide/

hemisphere (RH) were better in average with unsupervised SIMLR CON. The performance of other methods also peaked when using CON features, except for supervised SIMLR ensemble based SVM. This might indicate that morphological changes due to ASD in RH regions occur at a low-order morphological connectivity level rather than a higher order level. In other words, the RH pairwise connectivity strength between regions in the same view depicts better the changes associated with autism than the high order relationship between regions of different views. Still, the unsupervised methods scored better in performance than supervised methods and the best discriminative power was obtained when using the LH. For our best performing methods, we identified the top 2 discriminative high-order relationships for LH: (1) (fusiform gyrus, parahyppocampal gyrus) and (Lingual gyrus, pericalcarine cortex), and (2)(entorhinal cortex, transverse temporal gyrus) and (fusiform gyrus, posterior cingulate cortex); along with the top 2 discriminative low-order regions for RH: (1) entorhinal cortex and posterior singulate cortex, and (2) precuneus cortex and postcentral gyrus.

4 Conclusion

In this paper, we presented the first work on a high-order connectomic manifold learning using morphological brain networks for autism identification. Our framework outperformed both supervised and unsupervised baseline methods and was able to further identify the most discriminative relationships between *pairs* of morphological brain connectivities. Noting that ASD classification is a challenging problem, achieving 65.62% is quite promising based on solely T1-w MR images. To improve the connectomic manifold learning for a more accurate ASD/NC segregation, we will evaluate our method on the whole ABIDE dataset, which allow more powerful statistical analysis of our results. Further, we will extend our unsupervised learning method to spatiotemporal connectomic data for monitoring and predicting ASD progression.

References

1. T.Price, Wee, C.Y., Gao, W., Shen, D.: Multiple-network classification of childhood autism using functional connectivity dynamics. **17** (2014) 177–84 1
2. H. Koshino, PA . Carpenter, .M..C..K..J..: Functional connectivity in an fMRI working memory task in high-functioning autism. Neuroimage **24** (2005) 810–21 1
3. Uddin, L.Q., Supekar, K., Lynch, C.J., Khouzam, A., J.Phillips, Feinstein, C., Ryali, S., Menon, V.: Salience network-based classification and prediction of symptom severity in children with autism. JAMA psychiatry **24** (2013) 1
4. Zhu, Y., Zhu, X., Zhang, H., Gao, W., Shen, D., Wu, G.: Reveal consistent spatial-temporal patterns from dynamic functional connectivity for autism spectrum disorder identification. (2017) 1
5. Ghanbari, Y., Smith, A.R., Schultz, R.T., Verma, R.: Connectivity subnetwork learning for pathology and developmental variations. **16** (2013) 90–7 1
6. Jbabdi, S., Johansen-Berg, H.: Tractography: Where do we go from here? Brain Connect **1** (2012) 169–183 2

7. E, S., A, T., D, G., C, F., S, S., J, G., A, R.: Cortical thickness change in autism during early childhood. Hum Brain Mapp **37** (2016) 2616–29 2
8. Chen, X., Zhang, H., Shen, D.: Ensemble hierarchical high-order functional connectivity networks for MCI classification. MICCAI (2016) 18–25 2, 7
9. Brown, C., Hamarneh, G.: Machine learning on human connectome data from MRI. arXiv:1611.08699v1 (2016) 2
10. Iidaka, T.: Resting state functional magnetic resonance imaging and neural network classified autism and control. Elsevier **63** (2014) 55–67 2
11. Li, H., Xue, Z., Ellmore, T.M., Frye, R.E., Wong, S.T.: Identification of faulty DTI-based sub-networks in autism using network regularized SVM. IEEE ISBI (2012) 2
12. Wang, X., Sontag, D., Wang, F.: Unsupervised learning of disease progression models. KDD'14 Proceedings (2014) 85–94 2
13. H.Gao, C.Cai, J.Yan, L.Yan, Cortes, J., Y.Wang, F.Nie, J.West, A.Saykin, Shen, L., H.Huang: Identifying connectome module patterns via new balanced multigraph normalized cut. Springer **9350** (2015) 169–176 2
14. H.Chen, A.Iraji, X.Jiang, J.Lv, Z.Kou, T.Liu: Longitudinal analysis of brain recovery after mild traumatic brain injury based on groupwiseconsistent brain network clusters. Springer **9350** (2015) 194–201 2
15. Wang, B., Zhu, J., Pierson, E., Ramazzotti, D., Batzoglou, S.: Visualization and analysis of single-cell rna-seq data by kernel-based similarity learning. Nature **70** (2017) 869–79 2, 3, 4, 5, 6
16. Maaten, L., Hinton, G.: Visualizing data using t-sne. Journal of Machine Learning Research **9** (2008) 2579–2605 6
17. Mueller, S.G., Weiner, M.W., Thal, L.J., Petersen, R.C., Jack, C., Jagust, W., Trojanowski, J.Q., Toga, A.W., Beckett, L.: The Alzheimer's Disease Neuroimaging Initiative. Neuroimaging Clinics of North America **10** (2005) 869–877 7
18. Joe, H., Ward, J.: Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association **58** (1963) 236–244 7