BRAIN GEOMETRY PERSISTENT HOMOLOGY MARKER FOR PARKINSON'S DISEASE

Amanmeet Garg, Donghuan Lu, Karteek Popuri, Mirza Faisal Beg

School of Engineering Science, Simon Fraser University, Canada

ABSTRACT

The geometry of the human brain changes due to age and neurodegeneration. The brain geometry is expected to undergo a similar change in shape with a normal aging, however such change may differ in patients suffering from neurodegenerative disorders. In the novel framework proposed in this work, we model the brain geometry as a 3D point cloud and study the algebraic topology features of this point cloud. Specifically, we compute the persistence timelines of a simplicial complex in a multiscale simplicial homology of the underlying topology space. Further, persistence landscape summary features are obtained from the timelines and studied for their difference between the two groups. The statistical significance obtained in a permutation testing experiments highlights the ability of the persistence landscape features to differentiate between the PD and healthy control brain geometry.

Index Terms— Persistence homology, Persistence landscape, Brain geometry, Parkinson's disease.

1. INTRODUCTION

Brain shrinkage is consistently observed with normal aging and with neurodegeneration in Parkinson's disease (PD) [1]. The brain geometry change attributed to the brain shrinkage is often neglected as a potential differentiating signature between healthy individuals and patients. In this work we present a novel framework to quantify the geometric arrangement of brain regions to study its alterations in the presence of a neurodegenerative disease such as PD.

The current study of brain networks [2], commonly as binary undirected graphs is based on the simplifying assumption of a pairwise (dyadic) linkage between the nodes of the network. Such an assumption limits the ability to investigate higher order interactions (polyadic, many-to-many) between the nodes of the network. The recent advances in Topological Data Analysis have enabled the study of such polyadic interactions with principles rooted in persistent homology of simplicial complexes [3, 4]. Such approaches have been recently applied to study networks of brain function, however, have been limited to study of group level homology features [5]. In this work, we present a framework to quantify and characterize the geometrical arrangement of the cortical surface patches via a morphological network for an individual. Further, we compute the algebraic topology features to capture the multiway interactions between networks nodes.

Our geometry network framework encodes the geometric arrangement of the cortical anatomical patches with an inter-patch Euclidean distance in the 3D space. Further, we induce a filtration of the Vietoris-Rips complex on the distance matrices to obtain the persistence information (persistence landscapes (PL) [6]) of the k-simplices in the filtration (explained in section 2.3). The PL features are then tested for their ability to differentiate between the two groups (PD vs. CN). The key contributions of our work are two fold, 1) a study of the geometrical arrangement of cortical brain regions as affected by Parkinson's disease and 2) a novel framework to characterize the polyadic interactions between brain regions to obtain discriminative features.

2. METHODS

Our interest in geometry networks is based on the premise that the inter-regional geometrical arrangement between the cortical areas changes with neurodegeneration. The method we present here focuses on the alterations in cortical surface geometry approximated as a set of anatomically defined patches, whose centroids are representative of their geometric localization in the 3D space. Further, the topological analysis takes the inter-regional (node-to-node) Euclidean distance information and induces filtration of simplicial complexes to obtain features of persistent homology for the brain point cloud. Below we describe our method in further detail.

2.1. Geometry Network

The central aim of our work is to quantify the geometrical arrangement of the cortical structures in the brain. Thus, their localization in the 3D space is the signal of interest and is captured in the polyadic (many-to-many) interaction via the Euclidean distance between the patch centroids. Such $n \times n$ inter-patch distance matrices for n nodes enable us to study the topology of the geometrical arrangement and compare them across groups. The method to obtain the cortical parcellation is outlined below in the section 2.5.

The inter-patch distance for each subject varies depending on their respective brain size. As a result, in addition to the disease related change, the scale of the overall brain size in-

978-1-5090-1172-8/17/$31.00

fluences the graph filtration generated based solely on the raw distance values. Thus, to overcome such brain size difference, we normalize the values of the inter-patch distance for each individual to the range of $[0, 1]$ prior to the generation of the topological features. This enables a comparison of the network and topology features across subjects and groups by potentially reducing the affect of scale variation of the brain size.

2.2. Persistent Homology

The central idea behind the theory of persistent homology (PH) is to build a sequence of nested subsets on a space of simplicial complexes, studied at different resolutions. For our work, the Vietoris-Rips (VR) complex completely defined by the underlying 1-skeleton is induced on a symmetric distance matrix of pairwise distances between points in a point cloud.

A VR complex is defined on a metric space M for a specific distance value γ by forming a k-simplex for every finite set of $k+1$ points that has diameter at most γ. For a set of k nodes in the point cloud, the VR complex has at most $(k-1)$ simplices, enabling the geometry networks to obtain higher dimensional interactions limited in binary networks to 1-dimensional simplices (edges). Monotonically increasing values of the scale parameter ε_k lead to a VR filtration where $VR_{\varepsilon_1} \subseteq VR_{\varepsilon_2} \cdots \subseteq VR_{\varepsilon_k} \cdots \subseteq VR_{\varepsilon_n}$. For each filtered persistence module of the VR complex we obtain the tuples (b_i, d_i), with $b_i < d_i$ commonly known as a *birth-death pairs* of a k-dimensional simplex in the filtration. The length $d_i - b_i$ provides information of persistence of a k-simplex where long persistence times are suggestive of signal and short persistence times indicate towards noise. Each (b_i, d_i) tuple represented on a 2D plot provides a *persistence diagram* which can then be used to obtain topological summaries and perform subsequent statistical analyses.

2.3. Persistence Landscapes

A persistence landscape for each $\{(b_i, d_i)\}_{i=1}^n$ is a sequence of functions $\lambda_k : \mathbb{R} \to [0, \infty], k = 1, 2, 3, \ldots$ where $\lambda_k(x)$ is the k-th largest value of $\{f_{b_i,d_i}(x)\}_{i=1}^n$ [6]. For every birth-death pair (b, d) we define a piecewise linear function $f_{(b,d)} : \mathbb{R} \to [0, \infty]$ such as:

$$f_{(b,d)} = \begin{cases} 0, & if x \notin (b, d), \\ x - b & if x \in (b, \frac{b+d}{2}], \\ -x + d & if x \in (\frac{b+d}{2}, b). \end{cases}$$

For a set of persistence landscapes $\lambda^1, \ldots, \lambda^N$ we compute the average landscape as $\bar{\lambda} = \sum_{j=1}^N \frac{1}{N} \lambda^j$.

The distance between two persistence landscapes $\mathbb{L} = \{\mathbb{L}_k\}$ and $\mathbb{L}' = \{\mathbb{L}'_k\}$ can be obtained as the L^p norms for $1 < p < \infty$ which is defined as,

$$\|\mathbb{L}_k - \mathbb{L}'_k\|_p = \left[\sum_{k=1}^K \int \|\mathbb{L}_k - \mathbb{L}'_k\|_p^p \right]^{\frac{1}{p}} \quad (1)$$

and L^∞ can be obtained as $\|\mathbb{L}_k - \mathbb{L}'_k\|_\infty = \max_k \|\mathbb{L}_k - \mathbb{L}'_k\|$

2.4. Permutation testing

For the two groups of persistence landscapes $\mathbb{L}^1, \ldots, \mathbb{L}^N$ and $\mathbb{L}^1, \ldots, \mathbb{L}^M$, let δ be the true L_p distance between their average PLs, $\overline{\mathbb{L}_N}$ and $\overline{\mathbb{L}_M}$. We permute the group labels and compute the group level average landscapes $\overline{\mathbb{L}_N}$ and $\overline{\mathbb{L}_M}$ and find the corresponding L_p distance between them. The p-value of the statistical test equals the proportion of random permutations in which the distance between $\overline{\mathbb{L}_N}$ and $\overline{\mathbb{L}_M}$ is greater than the true difference.

Computational tools: We input the distance matrices for the geometry networks in the package *Perseus* with the parameter (*distmat*) to compute the PH of the Vietoris-Rips complex for each brain point cloud from the inter-patch distance matrices [4]. We obtain birth-death pairs (b_i, d_i) for the k-dimensional simplices for $k = 0, 1, 2, 3$. A recent *persistence landscapes* toolbox was released for research use by [6] enabling computation and statistical inference of the persistence landscapes. The number of landscapes λ^i varies dependent upon the underlying persistence diagrams which inherently depend on the birth-death pairs (b_i, d_i). Further, we perform permutation testing on the persistence landscapes in the two groups for 2000 random group assignments.

Comparison with classical network features: We computed the classical network features of nodal degree, local efficiency, clustering coefficient and betweeness centrality for the graphs in the filtration for each subject. Further, we test the group level difference between the features in the two groups in a Hotelling's T2 test in a permutation testing experiment (10000 permutations) after PCA dimensionality reduction.

2.5. Image processing

Cortical segmentation: The outer surface boundary for the cortical gray matter (pial surface) was obtained from the T1 structural MRI images via the FreeSurfer 5.1 method [7]. The surfaces were quality controlled for segmentation accuracy and, inaccurate segmentations were corrected manually where possible or removed from further analysis. The final cortical surface model for each subject is represented as a triangulated mesh with vertices and faces.

Adaptive Cortical Parcellation: As the first step in our method we normalize the cortical surface of the individual target subjects to an average template to obtain vertex wise correspondence in the cortical mesh of all subjects across the database. The cortical surface thus obtained was further subdivided into patches constrained to be within the anatomical parcellations obtained from the Freesurfer method [7] along with a constraint on the minimum number of vertices per patch (*mvcpp*). The cortical patch-wise parcellation from the template surface was transferred to individual target subjects in their respective spaces with spherical daemons registration [7]. The coordinates of the centroid for each patch were obtained as the average of the coordinates of the vertices within each patch.

Head size normalization: In order to study the the geometrical arrangement of the brain regions, it is vital to reduce the variability across subjects due to head size. The first step of normalization uses a 12 degree of freedom affine transformation between the 'cranial vault shapes' as an anatomically-meaningful way to normalize the geometry based on the estimated contour of each brain in its adult disease-free state. The affine transformations estimate potentially different scaling in the x, y and z directions based on registering the cranial-vault shapes which could be important when the goal is to study the relative geometrical arrangement of brain regions, and not just the volumes of brain regions. The vertices on the surface of the cortical gray matter (pial surface) were then transformed with this affine transformation to account for the effects of head size-related variability across subjects.

2.6. Materials

Imaging data for this work was obtained from the publicly available database provided under the Parkinson's Progressive Markers Initiative (*PPMI*). Detailed protocol for image acquisition and quality control for the study is available at the website *www.ppmi-info.org*. The two groups with De Novo PD patients (n = 339, age = 61.28$\pm$9.84, 209M/130F) and healthy controls (CN) (n = 150, age = 59.60$\pm$11.33, 93M/57F) were selected and analyzed through the method as described above.

3. RESULTS

A template parcellation into patches with *mvcpp* = 2000 was transferred to individual subject cortical surfaces. This resulted in 70 patches per hemisphere for a total of $N = 140$ patches per brain. Computing the centroid coordinates of each patch and the Euclidean distance between the patch centroids resulted in 489 (339 PD, 150 CN) distance matrices of size $N \times N = 140 \times 140$.

The persistence landscape features from the VR filtration from each brain showed a variation across the two groups (Figure 1). The PD group average persistence landscapes showed consistently higher values of the function in comparison to CN group for all three dimensions ($k = 0, 1, 2$)(Figure 1). The difference in the average persistence landscapes was further tested in a permutation testing experiment (2000 randomized group assignments) and showed strong statistically significant difference ($p < 0.05$). All three dimensions ($k = 1, 2, 3$) reached significance for the L_1, L_2 and L_∞ distance functions suggesting a true difference in the original groups.

In the permutation testing experiment, the network features showed no statistically significant difference between the two groups (nodal degree (p=0.242), local efficiency (p=0.182), clustering coefficient (p=0.341) and betweeness centrality (p=0.481)).

Table 1. Results for the permutation testing experiment of persistent landscapes of the Parkinson's disease and healthy control group geometry networks. The persistence landscapes for k-dimensional simplices were tested for the group level difference. The numbers present the p-value of the testing for L_p distance with $p = 1, 2, \infty$. All results show significant group level difference with p-value $<$0.05.

k	$p = 1$	$p = 2$	$p = \infty$
0	0	0	0
1	0.008	0.008	0.007
2	0.023	0.023	0.022

4. DISCUSSION

The current work presents a novel *Geometry Networks* framework as a non-invasive imaging marker for neurodegenerative disorders and supporting statistical evidence of its ability to differentiate between patient (Parkinson's disease) and healthy groups. These networks capture the 3D geometrical arrangement of the anatomically defined cortical (gray matter) regions and the change in this geometry with disease. Persistent homology features of these data were obtained for the distance matrices as they encapsulate the polyadic interactions between the points in the 3D point cloud. Statistical evidence presented in our work (section 3, Table 1) highlights, the ability of the homology features to capture the subtle changes in the geometry of the human cortex with the affects of disease, whereas the classical network features are unable to capture the geometry change in the disease group. Further, the evidence suggests that the geometrical arrangement of the cortical regions is influenced by PD related neurodegeneration.

It is important to note that the persistence landscapes in all three dimensions ($k = 1, 2, 3$) showed strong ($p < 0.05$) statistical difference in the L_1, L_2 and L_∞ distance metrics. On the contrary, the network features (section 3) showed no such difference ($p > 0.05$) in the permutation testing experiment. Further, the PLs for PD show larger values than the healthy group (figure 1) suggesting a longer range interaction

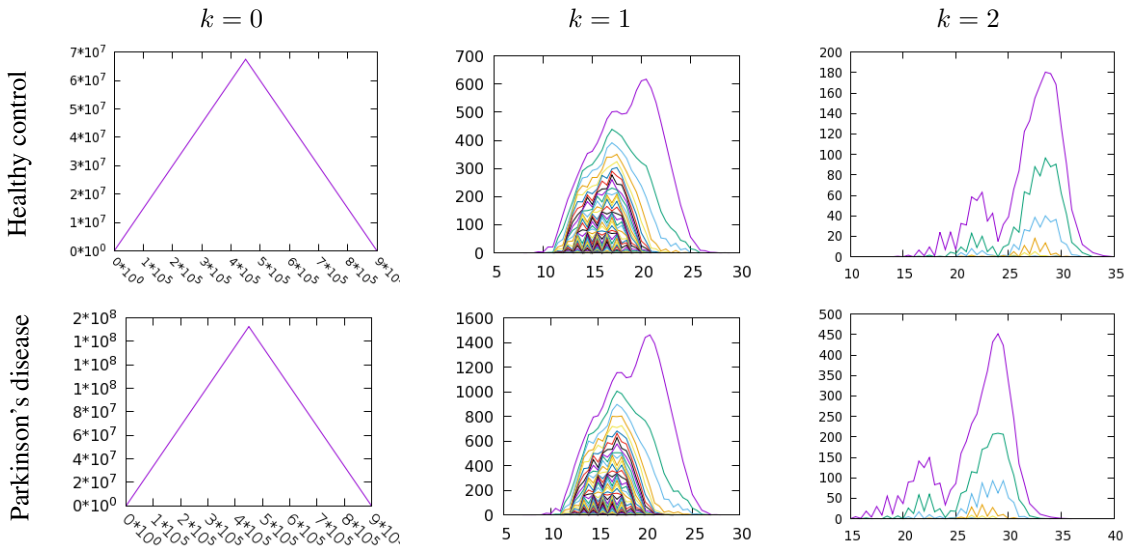

Fig. 1. Visualization of the average persistence landscapes in the healthy control (top row) and Parkinson's disease (bottom row) groups. The landscapes are for $k = 0$ (first column), $k = 1$ (third column) and $k = 2$ (third column).

between regions, versus, the short range (small (b_i, d_i)) in the healthy controls. Thus, indicative of need for longer range interactions with localized neurodegeneration in PD subjects.

In this work we selected a *mvcpp* = 2000 for adaptive parcellation of the cortical surface. This parameter choice was guided by the computational limitation of time and memory requirements for computing homology features for larger distance matrices. Further research into algorithmic development is needed to enable computation of homology of larger matrices thus, enabling more precise geometrical assessment of the cortical surface.

As geometry and size of the brain change with age and development, the framework presented in this work provides with the opportunity to study the geometric arrangement in different age groups, its change with age and potentially detrimental affects of disease. The results of this work lead to the future work of extending the framework to consider a combination of cortical anatomy with subcortical and white matter anatomical regions to build a whole brain geometrical network. Additionally, the features can be included in a machine learning framework to enable diagnosis of individual subjects.

Acknowledgment

Data used in the preparation of this article were obtained from the Parkinsons Progression Markers Initiative (PPMI) database (http://www.ppmi-info.org/data). For up-to-date information on the study, visit www.ppmi-info.org.

5. REFERENCES

[1] Tracy R Melzer, Richard Watts, Michael R MacAskill, Toni L Pitcher, Leslie Livingston, Ross J Keenan, John C Dalrymple-Alford, and Tim J Anderson, "Grey matter atrophy in cognitively impaired Parkinson's disease.," *Journal of neurology, neurosurgery, and psychiatry*, vol. 83, no. 2, pp. 188–94, feb 2012.

[2] Mikail Rubinov and Olaf Sporns, "Complex network measures of brain connectivity: Uses and interpretations.," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, sep 2010.

[3] Gunnar Carlsson, *Topology and data*, vol. 46, jan 2009.

[4] Konstantin Mischaikow and Vidit Nanda, "Morse Theory for Filtrations and Efficient Computation of Persistent Homology," *Discrete and Computational Geometry*, vol. 50, no. 2, pp. 330–353, 2013.

[5] Hyekyoung Lee, Moo K Chung, Hyejin Kang, Boongnyun Kim, and Dong Soo Lee, "Computing the shape of brain networks using graph filtration and Gromov-Hausdorff Metric," in *MICCAI 2011, Part 2, LNCS 6892*, 2011, pp. 302–309.

[6] Peter Bubenik, "Statistical topological data analysis using persistence landscapes," *Journal of Machine Learning Research*, vol. 16, pp. 25, 2015.

[7] Bruce Fischl, Martin I MI I Sereno, and Anders M AM M Dale, "Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system.," *Neuroimage*, vol. 9, no. 2, pp. 195–207, feb 1999.