Topology-based Clustering of Functional Brain Networks in an Alzheimer's Disease Cohort

Frederick H. Xu, BS[1], Michael Gao[1], Jiong Chen, MS[1], Sumita Garai, PhD[1], Duy Anh Duong-Tran, PhD[2], Yize Zhao, PhD[3], Li Shen, PhD[1]
[1]University of Pennsylvania, Philadelphia, PA, USA; [2]United States Naval Academy, Annapolis, MD, USA; [3]Yale University, New Haven, CT, USA

Abstract

Alzheimer's disease is a progressive neurodegenerative disease with many identifying biomarkers for diagnosis. However, whole-brain phenomena, particularly in functional MRI modalities, are not fully understood nor characterized. Here we employ the novel application of topological data analysis (TDA)-based methods of persistent homology to functional brain networks from ADNI-3 cohort to perform a subtyping experiment using unsupervised clustering techniques. We then investigate variations in QT-PAD challenge features across the identified clusters. Using a Wasserstein distance kernel with a variety of clustering algorithms, we found that the 0^{th}-homology Wasserstein distance kernel and spectral clustering yielded clusters with significant differences in whole brain and medial temporal lobe (MTL) volume, thus demonstrating an intrinsic link between whole brain functional topology and brain morphometric structure. These findings demonstrate the importance of MTL in functional connectivity and the efficacy of using TDA-based machine learning methods in network neuroscience and neurodegenerative disease subtyping.

Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disease characterized by a loss of memory recall and decreased higher-order cognitive function. While ample evidence is present regarding AD's relationship with local grey matter atrophy and build-up of amyloid-β and tau proteins,[1] less is understood about the systematic effect of degeneration on brain function as the disease progresses.[2,3] From a functional perspective, cognition in healthy human adults is, from a brain connectivity perspective, expressed through highly putative subnetworks of regions of interest (ROIs) with significantly correlated patterns of activity at rest (also known as resting state networks).[4] In AD, although reductions in connectivity are frequently observed in the default mode network (DMN),[5] the primary origin of pathogenesis, there is additional evidence of dysfunction related to AD in functional subnetworks outside of the DMN, thus implicating a whole-brain phenomenon.[6] As such, disruptions to functional connectivity in AD at a whole-brain level are not fully understood or characterized, leaving room for further investigation of how the disease's development may be related to whole-brain functional alterations.

Functional connectivity using fMRI can be represented through a functional brain network, where nodes in the network are defined by ROIs and the edges between nodes in the network are correlations between ROIs' aggregate fMRI BOLD signal time series. Topological properties of functional brain networks can be quantified using graph theoretical measures, which capture notions of whole-brain integration and segregation in a computationally efficient and intuitive manner.[7] However, these measures are sensitive to network construction methods such as the selection of edge thresholding value[8,9] and potentially leading to diverging interpretations of network topology in the context of AD.[10–12] As such, threshold-free methodology via topological data analysis (TDA) has been introduced to the field of network neuroscience to offer whole-brain descriptions of network topology without the need to threshold the network for a priori edge selection.[13,14] The most widely-used TDA tool is persistent homology, where a graph filtration is conducted to investigate the persistence of topological invariants throughout all thresholds of the graph. The outputs of persistent homology may be used in various machine learning approaches, including kernel representations via well-defined distance metrics between homological profiles.[15]

In the present study, we perform an AD subtyping experiment using unsupervised clustering strategies on functional brain network data from the Alzheimer's Disease Neuroimaging Initiative Phase 3 (ADNI3) cohort. We use a Wasserstein distance kernel for both 0th- and 1st-order persistent homology profiles, obtained via Rips filtration, and benchmark the performance across several popular clustering algorithms. For the purposes of subtyping comparison, we investigate the diagnostic status and QT-PAD challenge features and compare the differences between the learned

clusters. Additionally, we will compare the results from TDA methodology against similar methods using Euclidean distance between vectors of graph theoretical measures computed with the Brain Connectivity Toolbox (BCT).[7] Lastly, we investigate the association between functional network topology and AD-related phenotypes within the clusters.

Materials and Methodology

Data Acquisition

Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD).[16,17] All participants provided written informed consent and study protocols were approved by each participating site's Institutional Review Board (IRB). Up-to-date information about the ADNI is available at `www.adni-info.org`. Resting-state functional (rs-fMRI) and structural MRI (sMRI) imaging data were acquired from the ADNI Phase 3 cohort. Scans obtained were either from the baseline or year 1 time point as many subjects do not have an fMRI scan during screening for ADNI enrollment. Subjects were subsequently grouped based on diagnostic status at the time of the fMRI scan as healthy control (HC), MCI, or dementia (AD). In total, 417 subjects had usable network data after image processing with 197 CN, 158 MCI, 58 AD subjects, and 4 subjects with no diagnostic entry at the time of the scan. Subjects' demographics data were also obtained from ADNI and were filtered based on the closest date to their fMRI scan's date. Data were then organized into the Brain Imaging Data Structure (BIDS) format.

Image Processing

Image processing was conducted using the ConnectomeMapper3 pipeline built in Nipype. The pipeline consolidates several open-source packages for linear registration, segmentation, and rs-fMRI processing. Further details of packages may be found in Tourbier et al. 2022[18] and related literature. The sMRI volumes are first segmented using FreeSurfer's reconall function before being parcellated using the Lausanne 2018 parcellation.[19] We note that the lowest scale of Lausanne 2018 is equivalent to the native FreeSurfer space. For the rs-fMRI images, despiking and slice timing correction were done using AFNI and motion-corrected and distortion-corrected using FSL. The images were then registered to the b0 sMRI using the FLIRT toolbox in FSL. The BOLD time signals of each ROI are then bandpass filtered and detrended using a linear regression. Global signal regression (GSR) was not applied for this study. Functional brain networks were constructed for each subject by computing the Pearson correlation between the BOLD time signals of pairs of ROIs, and saved as an adjacency matrix (Figure 1a). Each entry of the weighted adjacency matrix thus corresponds to an edge defined as $w(i,j) = corr(BOLD_i, BOLD_j)$.

Brain connectivity graph-theoretical measures

Brain Connectivity Toolbox (BCT)[7] is widely used by researchers for complex brain-network analysis. Here we computed four network-level graph theoretical measures using BCT. For global measures, we computed global efficiency and global clustering coefficient. For node level information, we used mean clustering coefficient and average nodal betweenness centrality. It is to be noted that global efficiency and average betweenness centrality are integrative measures whereas transitivity and mean clustering coefficients are segregative measures, thus quantifying different properties network topology. For distance kernel computation, we computed the Euclidean distance between vectors of BCT features to obtain a distance kernel for the population of subjects.

Persistent Homology

The field of topological data analysis was recently introduced to the field of network neuroscience.[13] Algebraic topology involves the quantification of geometric properties of shapes classified as topological invariants that do not change under transformations of said shapes. The most commonly used procedure to detect such features, persistent

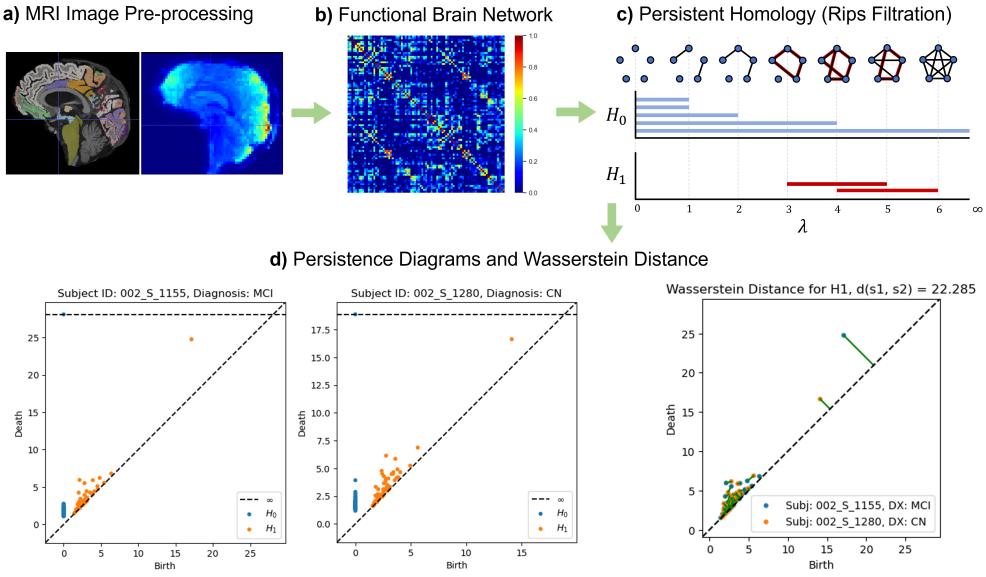

Figure 1: The schematic describes how we arrive at the Wasserstein distance between homologies from an MRI image. The images are a) first processed using the pre-processing pipeline and b) a functional brain network is constructed via correlation between ROI BOLD signal timeseries. c) Persistent homology is conducted. The diagram shows a simple example to provide a depiction of the Rips Filtration process that is used for calculating persistent homology. d) The results are then represented as a persistence diagram. From here, the Wasserstein distance can be computed in this coordinate space between subjects' persistence diagrams.

homology, quantifies these topological invariants across the structural hierarchy of a given object.

We computed the 0^{th}- and 1^{st}-order persistent homologies (H_0 and H_1 respectively) for each subject's functional brain network. The process involves first converting each brain network into a distance matrix by inverting each edge weight $D(i,j) = 1/W(i,j)$ and performing a Rips filtration using the distance matrix, starting with a trivial graph with only nodes and no edges, and progressively introducing edges in ascending order of distance based on a moving threshold λ. As the filtration is conducted, the persistence of topological invariants are computed. In H_0, the topological invariant is a connected component and in H_1 the topological invariant is a one-dimensional hole that can be represented by a cycle. The birth threshold λ_b is defined as the value of λ where the topological invariant first appears, while the death threshold λ_d is defined as the value of λ where the topological invariant disappears. For H_0, the death of a connected component occurs when it is eventually merged with another connected component, while in H_1 the death of a cycle occurs when it becomes comprised of cliques or triangles (Figure 1c).[13,20]. Rips filtration was performed using the Ripser Python package from scikit-tda[21]. The output of the Ripser function is a list of birth-death tuples (λ_b, λ_d) corresponding to the persistence lifetime of each equivalence class found during the graph filtration.[13] The result can be be displayed as a persistence diagram with each component's birth-death tuple represented as a coordinate (λ_b, λ_d) in the birth-death space (Figure 1c,d), which can be used for further kernel computation (Figure 1d).

Wasserstein Distance on Persistence Diagrams

The space of persistence diagrams has well-defined distance metrics, including the p-Wasserstein distance.[15,22] The application of the p-Wasserstein distance to the study of topology is derived from its definition in probability theory and optimal transport,[22]

$$W_p(X;Y) = \left(\inf_{\phi: X \to Y} \sum_{x \in X} ||x - \phi(x)||_\infty^p \right)^{1/p}, \tag{1}$$

which provides the distance between two multisets X and Y that may contain multiple instances of any element. It is assumed that a bijection ϕ exists such that there is a one-to-one matching of elements from X to the elements of Y. For every element $x \in X$, the difference from the bijection $\phi(x) = y \in Y$ can be quantified using the L_∞ norm of the p-th degree, thus obtaining a difference between the sets X and Y under the bijection $\phi : X \to Y$. However, it is noted that a strictly one-to-one bijection may not exist between two given multisets. In these instances, "virtual points" of infinite multiplicity may be considered to allow for a bijection $\phi : X \to Y$ to be found for any multisets X and Y with no impact on the final value of the cost function.[23]

In the context of persistence diagrams, let P_h^i and P_h^j be the persistence diagrams representing the results of two h^{th}-order homologies from subjects (i,j). Persistence diagrams may be viewed as two multisets containing the birth-death threshold pairs (λ_d, λ_b) for each topological invariant. Thus, the L_∞ norm for birth-death pairs is defined as

$$\left|\left|(\lambda_b^i, \lambda_d^i) - \left(\lambda_b^j, \lambda_d^j\right)\right|\right|_\infty = \max\left(\left|\lambda_b^i - \lambda_b^j\right|, \left|\lambda_d^i - \lambda_d^j\right|\right). \tag{2}$$

Substituting the definition of the L_∞ norm for persistence diagrams (Eq. 2) into the definition of the p-Wasserstein distance (Eq. 1) yields the desired definition of the p-Wasserstein distance on the space of persistence diagrams. Details of the full definition and exploration of the metric properties may be found in Tsizh et al. 2023.[22]

For this study, we will use the Persim implementation of the Wasserstein distance, which uses the Hungarian algorithm to perform a linear sum assignment to obtain the bijection ϕ. The implementation computes the 2-norm, thus resulting in a 2-Wasserstein distance. The Persim library is part of the wider scikit-tda library used to compute the Rips filtration in the previous section.[21] A visualization of the Wasserstein distance between two persistence diagrams can be seen in (Figure 1d). Using the Wasserstein distance, we compute a distance kernel for the subject space for the h^{th}-order homology, with each entry in the kernel corresponding to the Wasserstein distance between two subjects' h^{th}-order persistence diagrams $W^h(i,j) = W_2(P_h^i; P_h^i)$.

Clustering Strategies

For both the BCT Euclidean distance kernel and TDA Wasserstein Distance kernels, we use the same bevy of clustering algorithms with similar hyperparameter settings for a consistent comparison. To maintain a common computing environment, we used the sklearn library in Python for the following clustering algorithms.[24] The first algorithm used was agglomerative clustering with complete linkage. To select a distance threshold cut-off to resolve clusters, we constructed a dendrogram for each run and selected a cut-off that yields a reasonable number of clusters ($n < 5$). The second algorithm tested was Density-based Spatial Clustering of Applications with Noise (DBSCAN). The algorithm takes 2 hyperparameters: ϵ representing the maximum distance between two samples for them to be considered as neighbors and *min samples* representing the minimum number of neighbors for a sample to be considered as a core point. ϵ was first tuned using k-Nearest Neighbors with $k = 10$ and the kneedle algorithm. *min samples* was left at the default recommendation of 5. The third algorithm tested was Ordering Points To Identify the Clustering Structure (OPTICS), which is closely related to DBSCAN. The *min samples* parameter was set to 5 to match the DBSCAN hyperparameter choice. The OPTICS algorithm also takes an additional parameter, *xi* which denotes the minimum slope in a reachability plot that leads to the creation of a new cluster. The *max eps* parameter was left at ∞ to search across all possible scales of clusters, and *xi* was set to the default recommendation of 0.05. The fourth algorithm tested was spectral clustering. For this study, we opted to construct an affinity matrix using a k-nearest neighbors algorithm with $k = 8$, which was then used to compute a graph laplacian and its corresponding eigenvalues and eigenvectors. The eigenvectors of the lowest n eigenvalues are selected, where n is decided by an elbow plot of the ranked eigenvalues. Based on initial testing computations with elbow plots, we set $n = 3$. Subsequently, k-Means clustering with $k = 3$ is performed to obtain clusters.

Clustering Results Evaluation and Interpretation

The clustering results are evaluated using unsupervised and supervised metrics. For unsupervised metrics, the quality of the clusters are evaluated using silhouette score, which evaluates the goodness of a clustering technique based on

how clearly distinguished the clusters are in the distance space. For supervised metrics, the Rand index and mutual information between the cluster assignments and the diagnostic labels were computed. These metrics measure how closely the clusters recapture the diagnostic label assignments. For the training process, 100 training subsets were obtained by random sampling 80% of the data with replacement. Each algorithm was fit for each training subset to obtain a mean for each quality metric, per algorithm. 95% confidence intervals (CI) of these mean quality metrics were obtained via bootstrapping on the results and subsequently converted to and reported as margin-of-error.

To interpret the cluster groupings and identify disease phenotypes, we evaluated the distribution of Quantitative Templates for the Progression of AD (QT-PAD) challenge features from clusters obtained when fitting the algorithms on the entire dataset.[25] Fifteen of the 16 QT-PAD measures were obtained from the ADNI3 project demographics data, as amyloid-PET was missing from the dataset. For CSF-ABETA and CSF-TAU, indicators of > 1700 and < 8 were changed to values of 1700 and 8, respectively. It is to be noted that the brain imaging quantitative traits of FreeSurfer volume measures, cerebral spinal fluid measures, and FDG-PET were taken from baseline scans. For each clustering result, Kruskal-Wallis was conducted to test for significance ($P < 0.05$) of differences among median QT-PAD measures of clusters. A false discovery rate (FDR) correction was applied to correct for multiple comparisons. For clustering results that exhibited significance, a Dunn posthoc was conducted to identify pairwise differences between clusters.

For followup analyses on topological properties, we investigated two measures connectivity using properties of the The first measure is the integrated Betti number for the 0$^{\text{th}}$ homology (β_0 AUC), which quantifies the area under the Betti curve.[13] The second measure we compute is persistent entropy, derived from Shannon entropy, which measures the degree of order in the persistence lifetimes ($\lambda_d - \lambda_b$). A graph that has topological invariants with very different persistence lifetimes, such as some invariants having short lifetimes while others have long lifetimes, will possess a lower persistent entropy measure, while a graph that contains topological invariants with approximately similar persistence lifetimes will have a higher persistent entropy measure.[26] These measures are computed using the scipy and Persim libraries in Python.

Results

Clustering Results Based on Quality Metrics

The clustering results were first evaluated using clustering quality metrics (Table 1). For the unsupervised metric of silhouette score, we found that the BCT features showed the best silhouette scores when using agglomerative clustering ($SS_{BCT,Agg} = 0.825 \pm 0.012$), while the H_0 and H_1 Wasserstein kernels showed the best silhouette scores when using DBSCAN ($SS_{H_0,DBS} = 0.498 \pm 0.006$, $SS_{H_1,DBS} = 0.425 \pm 0.004$). When evaluating the clustering quality based on supervised metrics comparing the cluster assignments to the original diagnostic labels, different kernel-clustering techniques exhibit better performance. For BCT features, OPTICS clustering yielded the best supervised metrics performance with the highest Rand index ($RI_{BCT,OPT} = 0.542 \pm 0.003$) and mutual information ($MI_{BCT,OPT} = 0.0160 \pm 0.0013$). For the H_0 Wasserstein kernel, spectral clustering performed the best in Rand Index ($RI_{H_0,Spec} = 0.517 \pm 0.001$), while OPTICS clustering performed the best in mutual information ($MI_{H_0,OPT} = 0.137 \pm 0.0014$). Agglomerative clustering showed the best Rand index ($RI_{H_1,Agg} = 0.482 \pm 0.005$) and mutual information ($MI_{H_1,Agg} = 0.0088 \pm 0.0007$) for the H_1 Wasserstein kernel.

Overall, we note that these metrics in combination do not indicate an overall best kernel-clustering method. While BCT with OPTICS performs the best on supervised metrics, it is the worst-performing on silhouette score among the clustering algorithms used with BCT features (Table 1). A similar observation is made for H_1 and agglomerative clustering, where it exhibited the best performance on supervised metrics among the H_1 Wasserstein kernel techniques, but is the second-worst performing on silhouette score. For H_0 Wasserstein, there is no definitive best choice based on a combination of metrics. Additionally, performance on mutual information is poor, with all kernel-clustering combinations scoring close to 0, indicating that the diagnostic labels cannot be adequately recovered in the cluster assignments. As such, we tested individual QT-PAD measures instead to observe finer-grained variations in the population demographics that may be recaptured by the unsupervised methods.

Table 1: Clustering Quality Metrics

Method	Sil. Score	Rand Ind.	M.I.
BCT Agglom.	**0.825 ± 0.012**	0.409 ± 0.002	0.0157 ± 0.0008
BCT DBSCAN	0.793 ± 0.026	0.414 ± 0.001	0.0069 ± 0.0008
BCT OPTICS	-0.047 ± 0.018	**0.542 ± 0.003**	**0.0160 ± 0.0013**
BCT Spectral	0.185 ± 0.041	0.489 ± 0.005	0.0031 ± 0.0003
H0 Agglom.	0.375 ± 0.006	0.497 ± 0.004	0.0074 ± 0.0008
H0 DBSCAN	**0.498 ± 0.006**	0.404 ± 0.001	0.0066 ± 0.0003
H0 OPTICS	-0.422 ± 0.029	0.429 ± 0.003	**0.0137 ± 0.0014**
H0 Spectral	0.338 ± 0.004	**0.517 ± 0.001**	0.0052 ± 0.0006
H1 Agglom.	0.294 ± 0.010	**0.482 ± 0.005**	**0.0088 ± 0.0007**
H1 DBSCAN	**0.425 ± 0.004**	0.416 ± 0.001	0.0008 ± 0.0002
H1 OPTICS	-0.182 ± 0.022	0.409 ± 0.001	0.0087 ± 0.0007
H1 Spectral	-0.017 ± 0.003	0.427 ± 0.001	0.0045 ± 0.0006

*Boldface numbers are best performance in each metric per graph analytical technique. Margin-of-error is derived from the 95% CI of the mean.

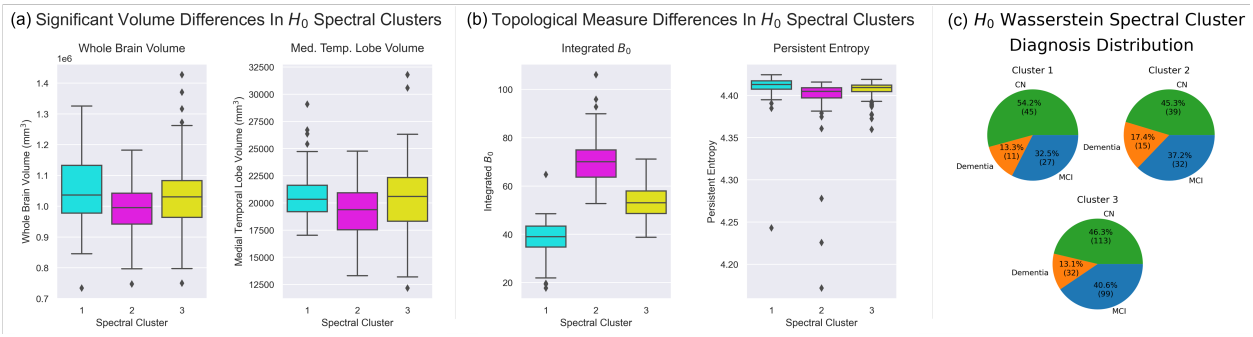

Figure 2: H_0 Wasserstein spectral cluster measures were computed to investigate the significant differences in these clusters. We identified that cluster 2 had (a) lower median WB volume and (b) lower median MTL volume. This corresponded to increased functional network segregation via (c) higher median integrated β_0. While persistent entropy was significantly different under statistical testing, the (d) median persistent entropy values were only marginally lower.

QT-PAD Measures

To evaluate more subtle variations in the population, we tested the significance of differences in QT-PAD challenge features among cluster results. Each clustering method was fit on the entire dataset to obtain a single clustering result per method for subsequent analyses. When comparing differences in QT-PAD features among the clusters observed with each method, it was found that the BCT feature clustering methods and H_1 Wasserstein clustering methods did not yield any significance at a significance level of $\alpha = 0.05$ after FDR correction was applied. For H_0, two morphologic QT-PAD measures of whole brain volume ($P_{WBV} = 0.015$) and medial temporal lobe (MTL) volume ($P_{MTLV} = 0.015$) showed significant differences ($\alpha < 0.05$) level post-FDR correction. We focus on these significant results for further investigation (Figure 3). To identify pairwise differences among the H_0 Wasserstein spectral clusters, we performed a Dunn posthoc test with a significance level of $\alpha = 0.05$ (Table 2). We found that median whole brain volume of cluster 2 was significantly different from that of clusters 1 and 3 ($P_{1,2} = 0.0019$, $P_{2,3} = 0.0098$). We note a similar observation in the medial temporal lobe (MTL) volume, where the median MTL volume of cluster 2 was significantly different from that of the other clusters ($P_{1,2} = 0.0092$, $P_{2,3} = 0.0061$). When investigating the median values of each cluster, we find that these two volume measures are lower in cluster 2 compared to the volumes in cluster 1 and 3 (Table 3, Figure 2a). These findings lead to the conclusion that differences in functional connectivity as observed through persistent homology are linked to structural differences in whole brain volume and MTL volume.

To further investigate the noted association between structure and function, we compute two topological measures of integrated-β_0 and persistent entropy to quantify connectivity, and statistically compared the measures across the H_0 Wasserstein spectral clusters. Kruskal-wallis revealed that the clusters had significant differences ($\alpha = 0.05$) in both integrate-β_0 (H-statistic $= 270.445$, $P = 1.877 \times 10^{-59}$) and persistent entropy (H-statistic $= 54.752$, $P = 1.291 \times 10^{-12}$). Upon further testing using Dunn post-hoc and comparing median metric values, we found

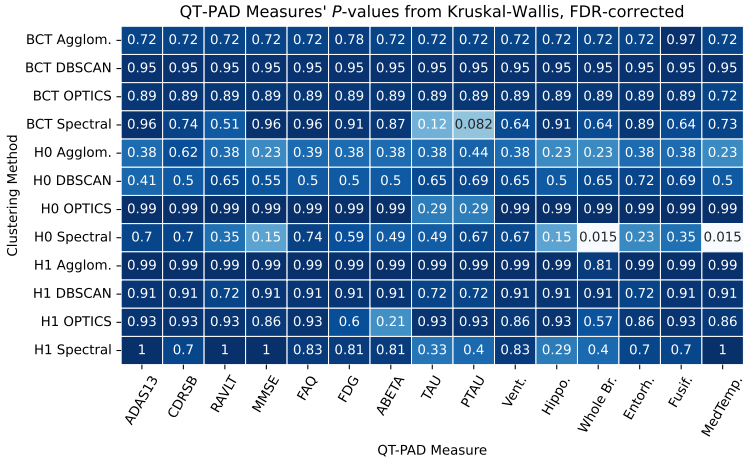

Figure 3: The P-values were obtained from Kruskal-Wallis tests for group differences in QT-PAD measures across obtained clusters for each clustering method. False discovery rate correction was used to correct for multiple comparisons. The heatmap color uses a logarithmic color scale to better depict the range of P-values present across magnitudes. Of the methods and measures tested, it is observed that the H_0-Wasserstein Spectral Clustering method yielded clusters that differed significantly at the $\alpha = 0.05$ level in whole-brain volume ($P = 0.015$) and medial temporal lobe volume ($P = 0.015$), after FDR correction was applied.

that cluster 2's integrated-β_0 was significantly higher than that of the other two cluster's ($P_{1,2} = 2.807 \times 10^{-60}$, $P_{2,3} = 1.332 \times 10^{-22}$, Table 2, Figure 2b). Since the Rips filtration is performed in a distance space, weaker edges correspond to higher values of λ. As such, in this definition, a higher λ_d indicates that weaker edges are required to achieve a minimally connected graph, leading to a right-shift in the Betti curve and higher values for integrated-β_0. From this, we conclude that cluster 2 exhibits weaker connectivity and thus increased segregation of the brain network, which is associated with decreased whole brain and MTL volumes. For persistent entropy, significant differences were also found ($\alpha = 0.05$) with cluster 2 showing lower median persistent entropy values (Figure 2b). However, we note that although the difference significant under statistical testing, the magnitude of difference is marginal (Table 3).

Table 2: Dunn Posthoc P-Values for Differences in H_0 Wasserstein Spectral Clusters

Measure	1 vs. 2	2 vs. 3	1 vs. 3
Whole Br. Vol.	**0.0019**	**0.0098**	0.2379
Med. Temp. Lobe Vol.	**0.0092**	**0.0061**	0.6385
Integrated β_0	$\mathbf{2.807 \times 10^{-60}}$	$\mathbf{1.332 \times 10^{-22}}$	$\mathbf{2.405 \times 10^{-24}}$
Persistent Entropy	$\mathbf{4.203 \times 10^{-13}}$	$\mathbf{6.262 \times 10^{-6}}$	$\mathbf{1.366 \times 10^{-05}}$

*Boldface indicates significance at the $\alpha = 0.05$ level.

Discussion

Alzheimer's disease is a progressive neurodegenerative disease that has relationships with whole-brain functional network alterations that are not fully characterized.[2] In the present study, we employ novel topological data analysis techniques via Wasserstein distance-based clustering of homological profiles in an unsupervised learning task to investigate the relationship between functional brain network topology and QT-PAD challenge features that are used to identify biomarkers of AD. Of the clustering methods tested, the H_0 Wasserstein distance kernel with spectral clustering yields clusters with significant differences in two physiological QT-PAD measures of whole brain and MTL volume. Additionally, the cluster that exhibits the lowest of these two volumetric measures also possesses the highest measure of integrated-β_0, indicating that functional topological segregation is associated with decreased whole brain and MTL volume. These findings point towards an important relationship between structural brain atrophy and functional brain network connectivity in neurodegenerative disease, particularly that of the MTL.

The MTL is a key brain region that exhibits biomarkers in the progression of Alzheimer's disease. It has been im-

Table 3: Median Feature Values in H_0 Wasserstein Spectral Clusters

Measure	Cluster 1	Cluster 2	Cluster 3
Whole Br. Vol. (10^5 mm^3)	10.361 ± 7.770	9.995 ± 4.991	10.303 ± 5.997
Med. Temp. Lobe Vol. (10^3 mm^3)	20.343 ± 1.218	19.377 ± 1.703	20.599 ± 2.012
Integrated β_0	39.058 ± 4.361	70.050 ± 5.629	53.112 ± 4.669
Persistent Entropy	4.413 ± 0.005	4.405 ± 0.006	4.409 ± 0.004

plicated in early progression of the disease, where atrophy that occurs in the MTL spreads to the anterior-temporal and posterial-medial systems.[27] Additionally, although MTL atrophy typically occurs with aging, it has been found to adversely affect those with MCI and has served as a predictor of an MCI patient transitioning into AD.[28,29] From a functional perspective, the MTL has been shown to play a key role in learning and memory systems, where abnormal functional connectivity is associated with cognitive deficits in late-stage MCI and AD subjects.[30] The present study further confirms the MTL's importance in large-scale brain systems, where functional brain network segregation is associated with atrophy in the MTL. In addendum to the the findings regarding MTL, the same clusters that exhibit low MTL volume also exhibit low whole-brain volume, which aligns with previous findings in literature where MTL atrophy propagates and spreads to wider brain systems. However, we note that the clusters are heterogeneous in their distribution of diagnostic statuses, thus requiring further investigation to confirm how the relationship between brain morphologic features and whole-brain functional network topology contributes to the progression of AD.

From a methodological perspective, the present study demonstrates the viability of TDA-based clustering methods as applied to functional brain network topology. The H_0 Wasserstein kernel with spectral clustering yields clusters with significant differences in QT-PAD challenge features after FDR-correction and thus performs notably better than BCT features for identifying variations in the population. It has been found in previous studies that graph theoretical measures such as those computed with BCT are highly sensitive to network construction methods, with significant differences found among AD cohorts fluctuating based on choice of parcellation scheme[11,12] and an overall inconsistency in the directionality of topological changes in relation to AD among studies that use different network construction protocols.[10,31] These noted inconsistencies in AD brain topology research reflect the overall findings in network neuroscience where graph theoretical measures have a noted sensitivity to edge thresholding procedures[9] and selection of parcellation scheme.[8] We surmise that the BCT features are unable to yield clusters with signficant biomarker differences due to these sensitivities and that perhaps a different network construction protocol could produce significant findings, with the caveat that the methodology may not be robust. The overall conclusion is that advancements in more robust methodology are required to garner a stable depiction of brain network topology. TDA offers a means of avoiding the edge threshold problem by searching across the entire edge hierarchy of a network. However, the application of TDA to network neuroscience in human connectome mapping has been limited thus far, with only a small number of studies demonstrating success in human brain networks. Our work here shows the potential for TDA to yield significant findings in neurodegenerative disease, linking whole-brain functional network topology to morphologic biomarkers. Moreover, the methodology employed here demonstrates success on a larger cohort size using data mining techniques of unsupervised clustering, thus providing an avenue for further integration of larger dementia datasets from the UKBioBank and OASIS challenge to verify the methodology's robustness and yield potential new findings on neurodegenerative disease's relationship with brain network topology. Lastly, further advancement of TDA in network neuroscience, which is in turn derived from neuroimaging, has the potential to offer novel and robust non-invasive biomarkers that can aid in the early detection of AD, a crucial step to identify opportunities for early intervention in the preclinical stages of the disease.[32]

The study has a few key limitations. Data used for this study were obtained from the ADNI-3 cohort, which uses a strict enrollment and screening process for subjects with pre-clinical and clinical AD.[17] As such, the data are largely homogeneous in disease phenotype, which may not yield as many insights in an AD subtyping experiment. The experiment here can be further improved if multiple cohorts, such as the UK BioBank and Penn Frontal Lobe Degeneration cohort, are integrated together for subtyping and investigating the potentially heterogeneous nature of AD. For the machine learning methods, we acknowledge that the present study can be further improved with parameter tuning, as each clustering method possesses numerous variations in parameter settings that may be tuned via grid search. However, the goal of the present study is to demonstrate the viability of TDA-based unsupervised learning methods across a suite of clustering algorithms, leading to increased computational load. As such, we note that the findings in this study are

observed under a limited scope of parameter tuning, and findings may be improved for significance or interpretability when further testing is conducted for each of the individual clustering algorithms. Lastly, persistent homology yields global topological features without locality. Further investigation is required to identify exactly which regions and connections contribute to topological differences in brain networks.

Conclusion

In this study, we have tested the novel application TDA-based clustering methods using a Wasserstein distance kernel to an AD unsupervised learning problem on the ADNI3 cohort. We have found a significant link between global brain network functional topology and structural brain volume, where increased functional network segregation is associated with lowered whole brain volume and MTL volume. The findings here provide additional insights into the relationship between disease physiology and whole-brain functional connectivity in the context of dementia, as well as further demonstrating the efficacy of TDA in an unsupervised machine learning task in brain networks. Future work can seek to expand upon the methodology, investigating the robustness of persistent homology across different network constructions, quantify localized regional contributions to overall brain network topology, and further develop brain network topology as a potential biomarker for neurodegenerative disease.

Acknowledgements

This work was supported in part by the National Institutes of Health (NIH) grants T32 AG076411, RF1 AG068191, U01 AG066833, U01 AG068057, and R01 AG071470, and the National Science Foundation grant IIS 1837964. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (NIH Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012).

References

1. Sneham Tiwari, Venkata Atluri, Ajeet Kaushik, Adriana Yndart, and Madhavan Nair. Alzheimer's disease: pathogenesis, diagnostics, and therapeutics. *International Journal of Nanomedicine*, Volume 14:5541–5554, 2019.
2. Matthew R. Brier, Jewell B. Thomas, and Beau M. Ances. Network dysfunction in alzheimer's disease: Refining the disconnection hypothesis. *Brain Connectivity*, 4(5):299–311, 2014.
3. Jessica S Damoiseaux. Resting-state fmri as a biomarker for alzheimer's disease? *Alzheimer's Research amp; Therapy*, 4(3):8, 2012.
4. Thomas B. T. Yeo, Fenna M. Krienen, Jorge Sepulcre, and et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106(3):1125–1165, 2011.
5. Goran Simic, Mirjana Babic, Fran Borovecki, and Patrick R. Hof. Early failure of the default-mode network and the pathogenesis of alzheimer's disease. *CNS Neuroscience amp; Therapeutics*, 20(7):692–698, 2014.
6. M. R. Brier, J. B. Thomas, A. Z. Snyder, and et al. Loss of intranetwork and internetwork resting state functional connections with alzheimer's disease progression. *Journal of Neuroscience*, 32(26):8890–8899, 2012.
7. Mikail Rubinov and Olaf Sporns. Complex network measures of brain connectivity: Uses and interpretations. *NeuroImage*, 52(3):1059–1069, 2010.
8. Andrew Zalesky, Alex Fornito, Ian H. Harding, Luca Cocchi, Murat Yücel, Christos Pantelis, and Edward T. Bullmore. Whole-brain anatomical networks: Does the choice of nodes matter? *NeuroImage*, 50(3):970–983, 2010.
9. Mark Drakesmith, Karen Caeyenberghs, Anirban Dutt, Glyn H. Lewis, Anthony S. David, and Derek K. Jones. Overcoming the effects of false positives and threshold bias in graph theoretical analyses of neuroimaging data. *NeuroImage*, 118:313–333, 2015.
10. Betty M. Tijms, Alle Meije Wink, Willem De Haan, and et al. Alzheimers disease: connecting findings from graph theoretical studies of brain networks. *Neurobiology of Aging*, 34(8):2023–2036, 2013.
11. Frederick Xu, Sumita Garai, Duy Duong-Tran, Andrew J. Saykin, Yize Zhao, and Li Shen. Consistency of graph theoretical measurements of alzheimer's disease fiber density connectomes across multiple parcellation scales. *(BIBM) 2022 IEEE International Conference on Bioinformatics and Biomedicine IEEE*, Regular Paper, 2022.

12. Zhanxiong Wu, Dong Xu, Thomas Potter, and Yingchun Zhang. Effects of brain parcellation on the characterization of topological deterioration in alzheimers disease. *Frontiers in Aging Neuroscience*, 11, 2019.

13. Ann E. Sizemore, Jennifer E. Phillips-Cremins, Robert Ghrist, and Danielle S. Bassett. The importance of the whole: Topological data analysis for the network neuroscientist. *Network Neuroscience*, 3(3):656–673, 2019.

14. T. Songdechakraiwut, L. Shen, and M. Chung. Topological learning and its application to multimodal brain network integration. *Med Image Comput Comput Assist Interv*, 12902:166–176, 2021.

15. Danielle Barnes, Luis Polanco, and Jose A. Perea. A comparative study of machine learning methods for persistence diagrams. *Frontiers in Artificial Intelligence*, 4, 2021.

16. M. W. Weiner, D. P. Veitch, P. S. Aisen, et al. The alzheimer's disease neuroimaging initiative: a review of papers published since its inception. *Alzheimers Dement*, 9(5):e111–94, 2013.

17. M. W. Weiner, D. P. Veitch, P. S. Aisen, et al. Recent publications from the Alzheimer's disease neuroimaging initiative: Reviewing progress toward improved ad clinical trials. *Alzheimer's & Dementia*, 13(4):e1–e85, 2017.

18. Sebastien Tourbier, Joan Rue-Queralt, Katharina Glomb, and et al. Connectome mapper 3: A flexible and open-sourcepipeline software for multiscale multimodal human connectomemapping. *Journal of Open Source Software*, 7(74):4248, 2022.

19. Leila Cammoun, Xavier Gigandet, Djalel Meskaldji, and et al. Mapping the human connectome at multiple scales with diffusion spectrum mri. *Journal of Neuroscience Methods*, 203(2):386–397, 2012.

20. Ann E. Sizemore, Chad Giusti, Ari Kahn, Jean M. Vettel, Richard F. Betzel, and Danielle S. Bassett. Cliques and cavities in the human connectome. *Journal of Computational Neuroscience*, 44(1):115–145, 2017.

21. Nathaniel Saul and Chris Tralie. Scikit-tda: Topological data analysis for python, 2019.

22. Maksym Tsizh, Vitalii Tymchyshyn, and Franco Vazza. Wasserstein distance as a new tool for discriminating cosmologies through the topology of large-scale structure. *Monthly Notices of the Royal Astronomical Society*, 522(2):2697–2706, 2023.

23. Herbert Edelsbrunner and John L. Harer. Computational topology: An introduction. *Miscellaneous Books*, 69, 2010.

24. F. Pedregosa, G. Varoquaux, A. Gramfort, and et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

25. Alzheimer's disease modelling challenge: Modelling the progression of alzheimer's disease. *Portland Institute for Computational Science*, 2012.

26. Nieves Atienza, Rocio Gonzalez-Díaz, and Manuel Soriano-Trigueros. On the stability of persistent entropy and new summary functions for topological data analysis. *Pattern Recognition*, 107:107509, 2020.

27. Robin de Flores, Sandhitsu R. Das, Long Xie, and et al. Medial temporal lobe networks in alzheimer's disease: Structural and molecular vulnerabilities. *The Journal of Neuroscience*, 42(10):2131–2141, 2022.

28. Léa Chauveau, Elizabeth Kuhn, Cassandre Palix, and et al. Medial temporal lobe subregional atrophy in aging and alzheimer's disease: A longitudinal study. *Frontiers in Aging Neuroscience*, 13, 2021.

29. P. J. Visser, F. R. J. Verhey, P. A. M. Hofman, P. Scheltens, and J. Jolles. Medial temporal lobe atrophy predicts alzheimer's disease in patients with minor cognitive impairment. *Journal of Neurology, Neurosurgery & Psychiatry*, 72(4):491–497, 2002.

30. Bradford C. Dickerson and Reisa A. Sperling. Functional abnormalities of the medial temporal lobe memory system in mild cognitive impairment and alzheimer's disease: Insights from functional mri studies. *Neuropsychologia*, 46(6):1624–1635, 2008.

31. Jon delEtoile and Hojjat Adeli. Graph theory and brain connectivity in alzheimer's disease. *The Neuroscientist*, 23(6):616–626, 2017.

32. Jaime Gomez-Ramirez and Jinglong Wu. Network-based biomarkers in alzheimer's disease: Review and future directions. *Frontiers in Aging Neuroscience*, 6, 2014.