Topological Descriptors for Quantitative Prostate Cancer Morphology Analysis

Lawson PJ[1], Berry E[2], Brown JQ[1], Fasy BT[2], Sholl, AB[1], Wenk C[1]
[1]Tulane University, New Orleans, LA, [2]Montana State University, Bozeman, MT

INTRODUCTION

This project aims to discover new quantitative image-based prognostic biomarkers for prostate cancer, focusing on an investigation of novel concepts from persistent homology applied to prostate cancer glandular architecture. In this project we propose new topological descriptors that capture architectural features of prostate glands in histopathology images.

PERSISTENT HOMOLOGY

Persistent homology can be used to capture the topology (connectivity of a space) of a given set of nuclei positions at different scales, by continuously expanding a ball over time t and recording cycles that are formed as the balls connect, and cycles that are destroyed as the balls merge. This set of nested topological spaces is a filtration, as shown in Figure 1. The nuclei can be described as a sublevel set $d_N^{-1}((-\infty, t])$ of the distance function to the nuclei $d_N: S \to \mathbb{R}$, where the domain of the function S is the H&E slide and $d_N(x)$ is the distance from $x \in S$ to the closest nuclei. As where t is increased from $-\infty$ and the connectivity of the sublevel set changes either at a local minimum where a new component is added, or at a local maximum where two components are merged as seen in Figure 1. The time δ_b at which a component, or cycle, is formed is its birth, and the time δ_d, at which it is destroyed, is its death. A cycle persists for $\delta_d - \delta_b$ and the multiset of all birth-death pairs is plotted as a persistence diagram as shown in Figure 3. representing three circles, their distance function, and persistence diagram.

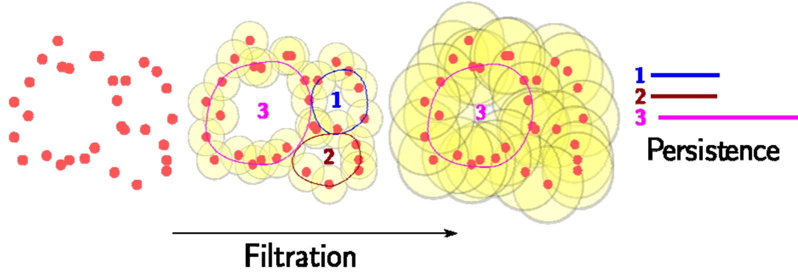


Figure. 1 Filtration and Corresponding Persistence [2]

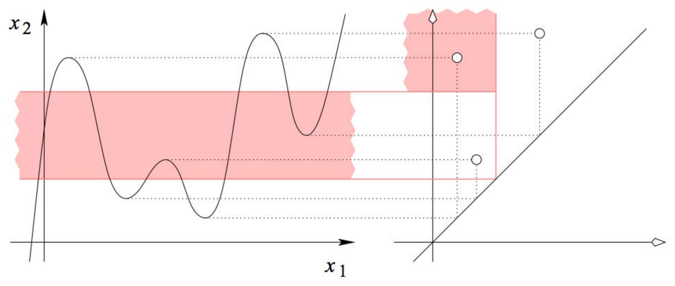


Figure 2. Persistence From Sublevel Set [1]

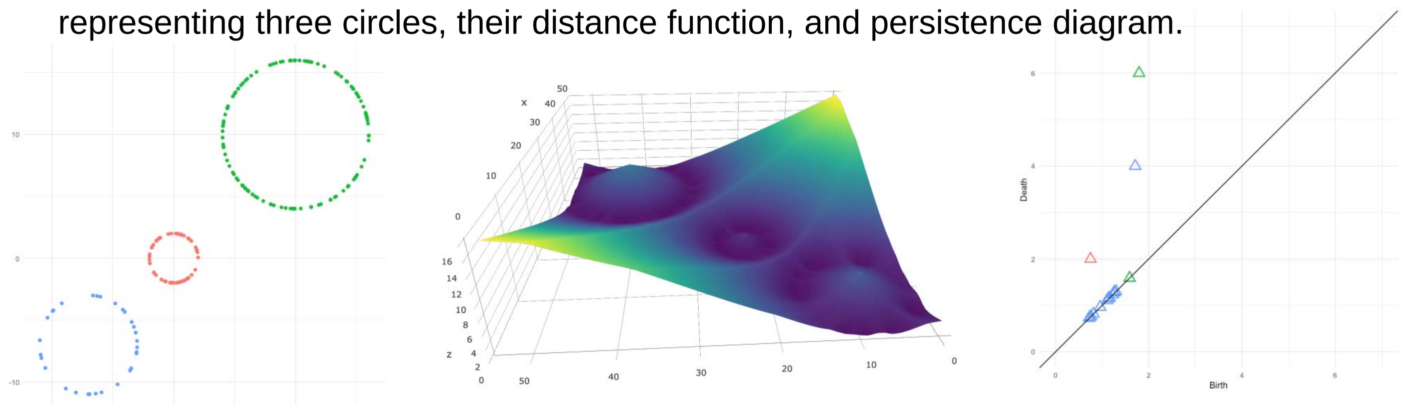


Figure 3. Example Sublevel Set and Corresponding Persistence Diagram

APPLICATION TO GLEASON GRADING

- Gleason grade of prostatic adenocarcinoma is an established prognostic indicator based on the architectural pattern.
- The qualitative nature of Gleason grade assessment results In high inter and intra observer variability.
- Due to the clearly differentiated morphology between Gleason grades the use of topological data analysis to generate quantitative descriptors is a natural application.
- The arrangement of nuclei, corresponding to different architectural features of tumors, can be captured by persistence homology, as seen in figures 4 and 5.

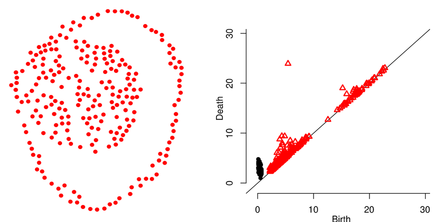


Figure 4. Simulated Glomeruloid Gland

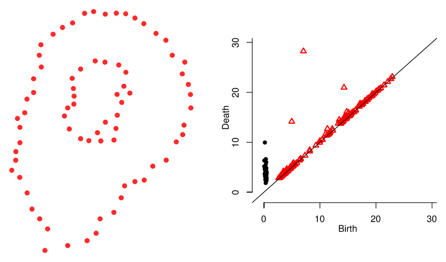


Figure 5. Simulated Telescoping Gland

METHODS

PREPROCESSING

- H&E specimens were digitized at 20X with an Aperio CS2 whole slidescanner .
- Nuclei stained by hematoxylin were isolated by color [3].
- Nuclear segmentation was performed using maximum correlation thresholding.
- Stromal nuclei were filtered based on thresholding of nuclei eccentricity.

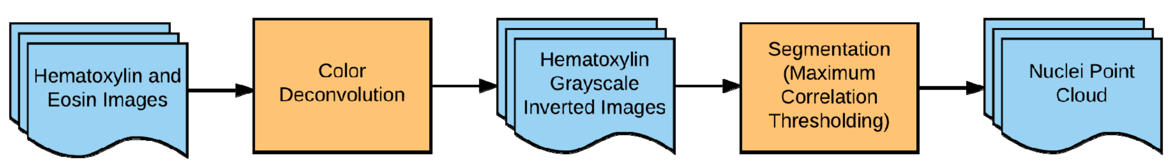


Figure 6. Preprocessing Workflow

PERSISTENT HOMOLOGY

The point clouds generated from selected ROIs had a distance to measure function computed over them, defined as $\hat{d}_{m0}(y) = \left(\frac{1}{k}\sum_{x_i \in N_k(y)} \|x_i - y\|^r\right)^{1/r}$, where $k = [m0 * n]$ and $N_k(y)$ is the set containing the k nearest neighbors of y among $x_1, \ldots, x_n$. Finally persistence diagrams were computed over the function generated for each ROI.

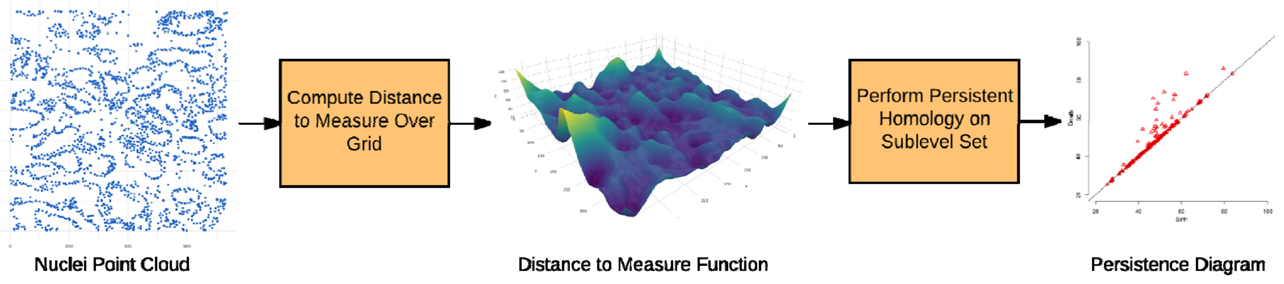


Figure 7. Computing Persistence Diagrams and Corresponding Measures

Functional and vectorized summaries of persistence diagrams were used to to facilitate the application of nonparametric statistics:

- **Persistence Landscapes** (Figure 8), a series of piecewise linear functions that encode a persistence diagram.
- **Persistence Intensity Functions** (Figure 9), a smoothed two-dimensional histogram using the points in a persistence diagram.
- **Bottleneck Distance**, the infimum of the supremum of all possible bijections between two persistence diagrams.
- **Wasserstein Distance (Earthmover Distance),** the distance between probability distributions between two persistence diagrams.

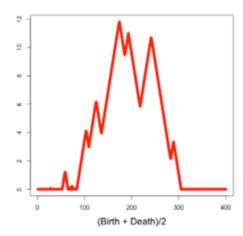


Figure 8. Landscape

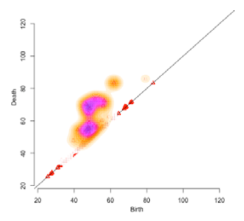
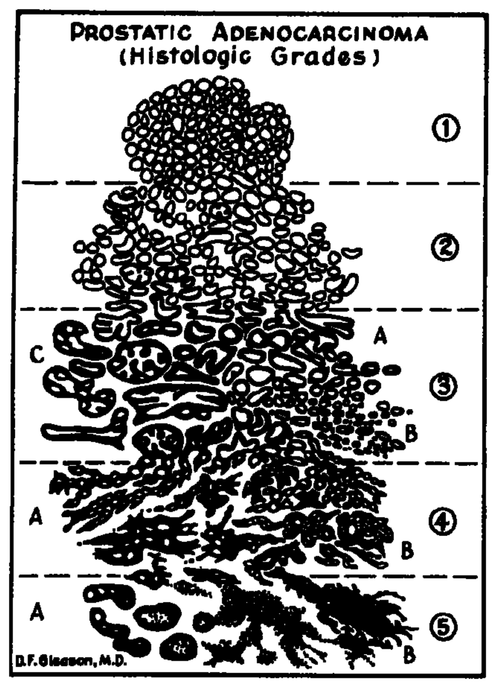


Figure 9. Intensity Plot

RESULTS

APPLICATION TO GLEASON DRAWING

By applying persistent homology to the original Donald Gleason sketch of the Gleason grading system, individually by grade, a clear pattern emerges. The increase in architectural heterogeneity proceeding from Gleason 1 to Gleason 5 is captured in the the widening distribution of persistence in the corresponding persistence diagrams.



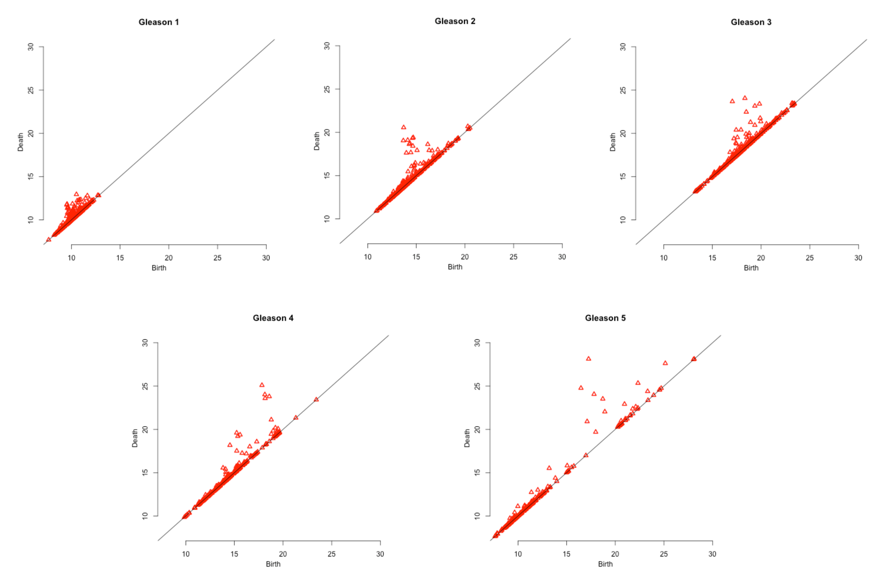


Figure 10. Persistence Diagrams of Gleason Sketch [4]

APPLICATION TO PATIENT DATA

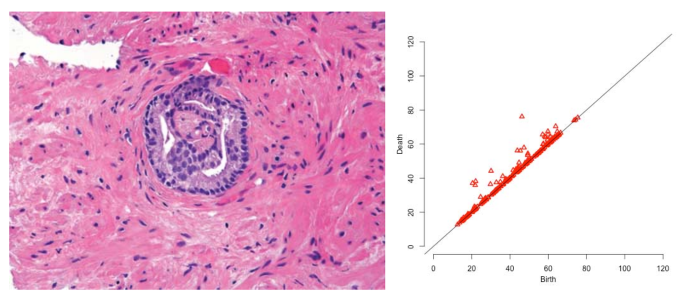


Figure 11. Perineural Invasion [5]

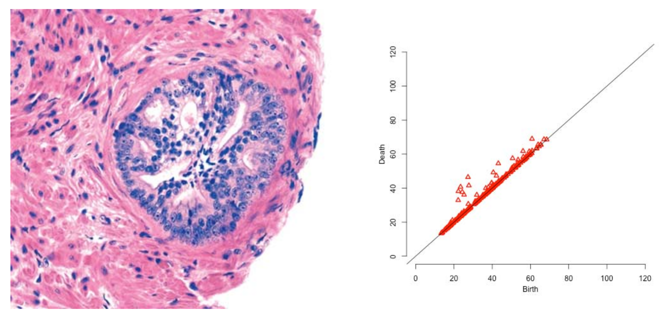


Figure 12. Single Cribriform Gland [5]

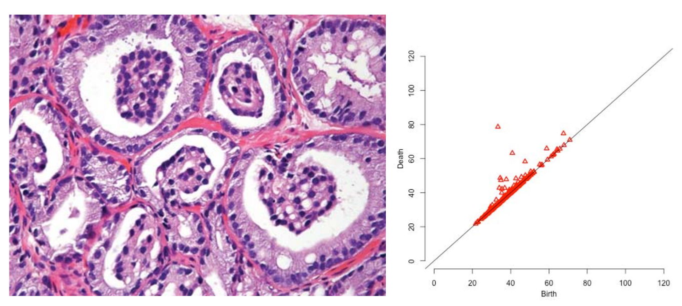


Figure 13. Small Glomeruloid Glands [5]

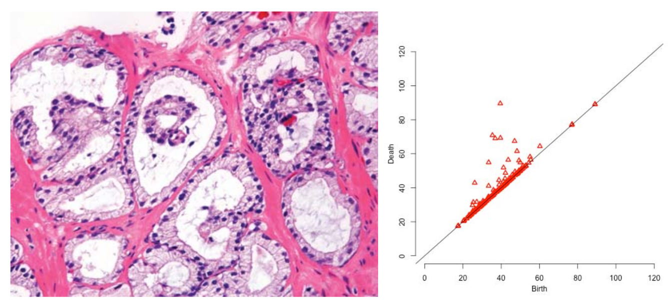


Figure 14. Telescoping Glands [5]

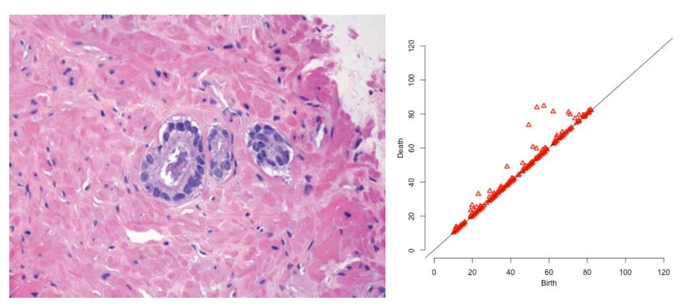


Figure 15. Gleason 3 Pattern [5]

When applied to unique architectural features in prostatic adenocarcinoma, such as the telescoping glands (glands within glands) seen in Figure 14, distinct persistence patterns arise. Architecturally similar features bear similar persistence diagrams, whereas unrelated features are distinct.

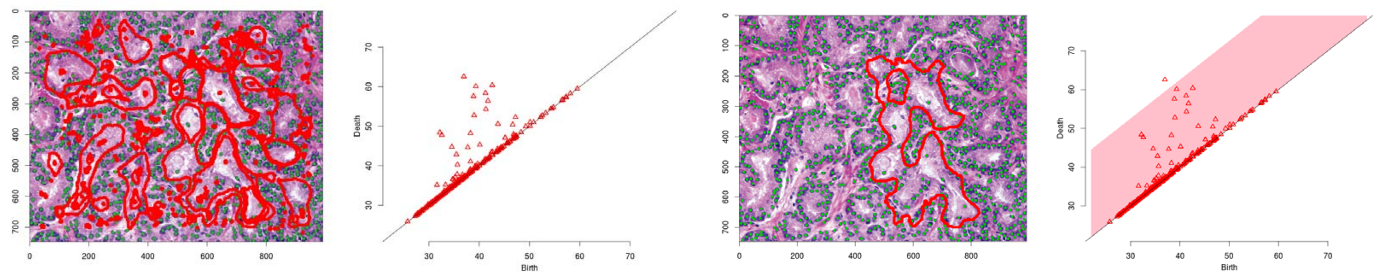


Figure 16. Cycles Over Plotted Over Gleason 3

Figure 16 shows a purely graded Gleason 3 ROI and corresponding persistence diagram, with the cycles identified plotted over the original image. The far right represents the same H&E image as above except with a 95% confidence band generated by bootstrapping subsampled persistence diagrams, with only the topological features considered significant plotted (Those outside the pink banded region).

The image to the right, Figure 17, shows persistence and the corresponding landscapes being computed over larger, purely graded, ROIs. A clear pattern emerges as the heterogeneous distribution of glandular sizes in Gleason 3 approaches the sheet of cells in the Gleason 5. The transition from heterogeneous to homogenous can be reflected in the narrowing of the persistence range and decrease in the full width half maximum of the landscape.

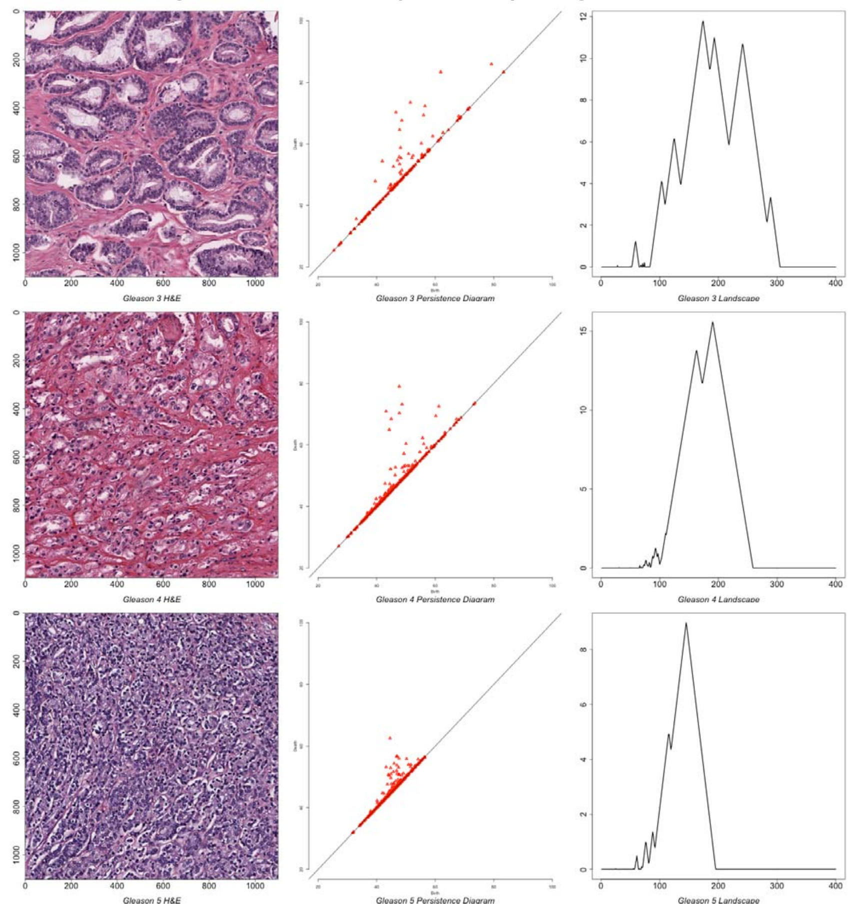


Figure 17. Gleason Grade 3 to 5 Persistence and Landscape

QUANTIFYING PERSISTENCE

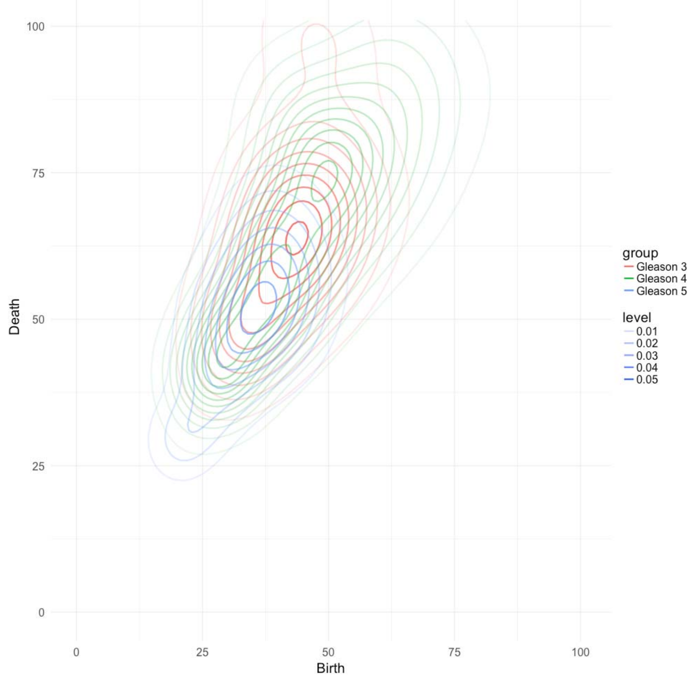


Figure 18. Contour plot of averaged intensity of purely-graded ROIs

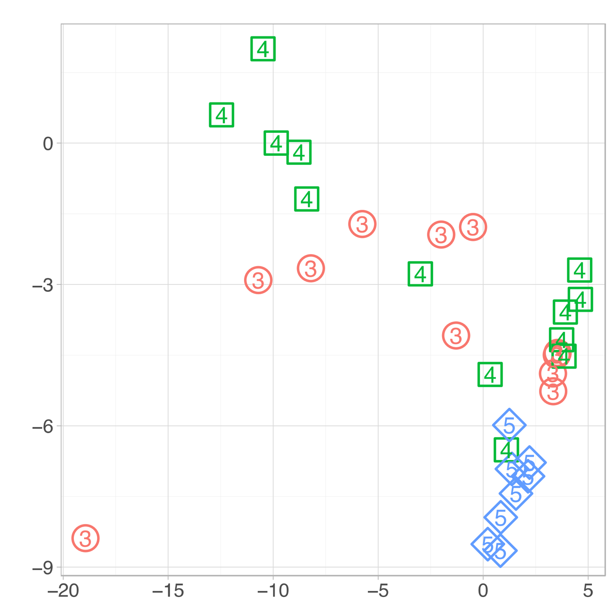


Figure 19. Multidimensional scaling applied to the bottleneck distance between a set of purely-graded ROIs.

Figure 18 represents an averaging of intensity functions generated for purely graded ROIs, with some separation by grade. Figure 19 corresponds to multidimensional scaling performed on a dissimilariy matrix of bottleneck distances of Gleason grades 3 to 5 and shows that Gleason grades can be differentiated by their persistence diagrams. While tentative, this work shows that persistence diagrams can eventually be quantified and leveraged for discriminatory capacity.

CONCLUSIONS

Persistent homology shows a clear capacity to distinguish architectural features of prostatic adenocarcinoma. We hope to leverage the discriminatory capacity of persistent homology to generate an intuitive, quantitative descriptor for histopathology slides that at minimum recapitulates Gleason grading, and ideally extends beyond to it provide finer granularity descriptors to prostatic adenocarcinoma architectural features.

REFERENCES

[1] Singh, "Topological Descriptors of Histology Images," 2014.
[2] H. Edelsbrunner and J. Harer, "Persistent homology - a survey," Contemporary mathematics, vol. 453, pp. 257–282, 2008.
[3] A. C. Ruifrok and D. A. Johnston. Quantification of histochemical staining by color deconvolution. Analytical and Quantitative Cytology and Histology, 23(4):291–299, August 2001.
[4] Peter A. Humphrey. Gleason grading and prognostic factors in carcinoma of the prostate. Modern Pathology, 17(3):292–306, February 2004.
[5] Epstein JI, editor. The Gleason grading system. A complete Guide for Pathologists and Clinicians. Philadelphia: Wolters Kluver Health | Lipincott Williams & Wilkins, 2013.

ACKNOWLEDGEMENTS: Research reported in this poster was supported by the Division of Mathematical Sciences of the National Science Foundation under award number 1557750