Atlas of Science
another view on science
http://atlasofscience.org

Computational topology for biomedical image segmentation in pathomics

Topology concerns the study of fundamental and intrinsic properties of spaces which are invariant to continuous deformations, extrinsic evaluations or changes in magnitude e.g. scale, shape, distance, intensity, etc. Topology became an influential field of research by the middle of the 20th century. For most of its history topology has not been applied to the scientific domain due to its abstract and qualitative nature. However, during the last ten years it has emerged as one of the most effective representation models for exploring massive and complex data sets gaining all sorts of attention in the fields of computational science and big data analytics.

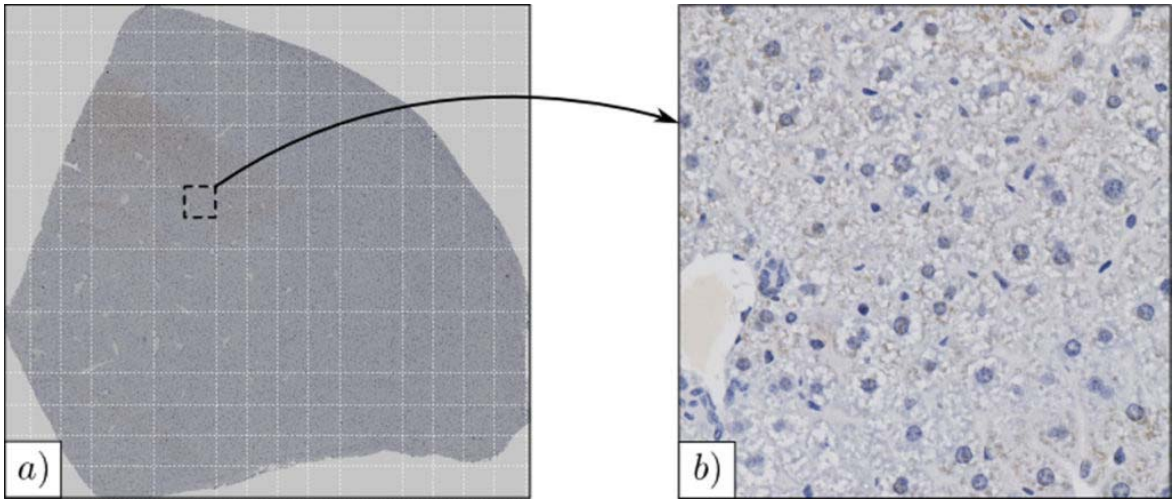

Fig. 1. Sample images acquired from histological section of liver tissue used in the analysis: (a) whole slide image, and (b) cropped square section.

Topological invariants have great potential in the image processing field since they are related to structural properties, resulting in less sensitivity regarding changes in scale and intensity, thus simplifying tasks such as image segmentation. One topological invariant used in image and data analysis is the Betti number. A standard approach for obtaining Betti numbers relies on embedding the images in a Euclidean space; pixels become vertices. Two nearby vertices span an edge, three nearby vertices span a triangle comprised of three edges, and four nearby vertices span a tetrahedron formed by four triangles and six edges. These configurations are called geometric simplicial complexes. Many algorithms can compute Betti numbers in this tessellated space. The number of complexes spanned by k neighboring vertices is $(2^{k+1} - 1)$. Thus, a small increment in k will introduce a substantial increase of simplicial complexes that would make unattainable to compute the already computationally expensive Betti numbers. This computational issue is tackled by dropping the number of vertices drastically; this is achieved by splitting the image into connected components at different scales in a process called image dismantling which is a kind of data coarsening. Betti numbers are computed from the inclusion relations between the connected components. Instead of geometric simplicial complexes, abstract simplicial complexes are spanned which is a combinatorial

version based on sets of sets. From this abstract representation, Betti numbers can be efficiently obtained.

The segmentation of cell nuclei is an important step towards the automated analysis of histological images. The presence of numerous nuclei in whole-slide images necessitates methods that are computationally tractable in addition to being effective. The automated segmentation of cell nuclei is a well-studied topic for which many algorithms are available in the literature and newer methods continue to be investigated. Nevertheless, a major limitation of many segmentation methods is the under/over-segmentation of cell nuclei. The aim is to develop a method for the robust segmentation of cell nuclei in histological images based on the principles of persistent homology. Essentially, the approach deals with the persistence of disconnected sets in the image, thus identifying salient regions that express patterns of persistence. By introducing an image representation based on topological features, the task of segmentation is less dependent on variations of color or texture. This results in a novel approach that generalizes well and provides stable performance.

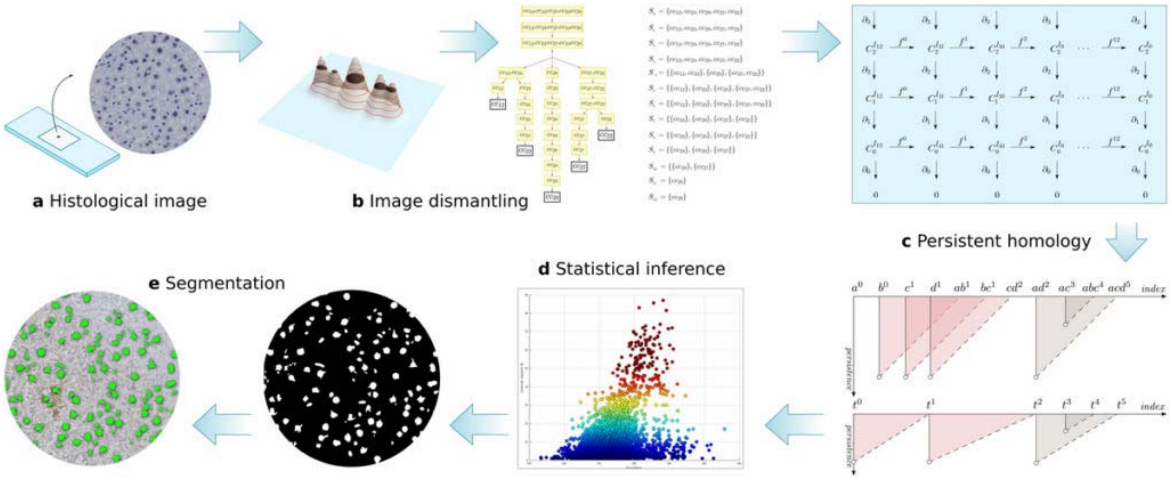

Fig. 2. Overview of automated workflow based on the principles of persistent homology for the segmentation of cell nuclei in whole-slide images.

Images acquired from histological sections of liver tissue are used as a case study; the histological landscape consists of hepatocytes and non-parenchymal cells (Fig. 1). The accuracy of the proposed methodology (Fig. 2) is verified against an automated workflow created by the output of a conventional filter bank (validated by experts) and the supervised training of a random forest classifier. The proposed workflow successfully detected both hepatocyte and non-parenchymal cell nuclei with an accuracy of 84.6%, and hepatocyte cell nuclei only with an accuracy of 86.2%. The proposed method is useful for obtaining unsupervised robust initial segmentations that can be further integrated in image/data processing and management pipelines.

Rodrigo Rojas Moraleda [1,5]***, Wei Xiong*** [2]***, Niels Halama*** [3]***, Katja Breitkopf-Heinlein*** [4]***, Steven Dooley*** [4]***, Luis Salinas*** [5]***, Dieter W. Heermann*** [2]***, Nektarios A. Valous*** [1]

[1]*Applied Tumor Immunity Clinical Cooperation Unit, National Center for Tumor Diseases, German Cancer Research Center, Heidelberg, Germany*
[2]*Statistical Physics and Theoretical Biophysics Group, Institute for Theoretical Physics, Heidelberg University, Heidelberg, Germany*
[3]*Department of Medical Oncology, National Center for Tumor Diseases, University Hospital Heidelberg, Heidelberg, Germany*
[4]*Department of Medicine II, Faculty of Medicine at Mannheim, Heidelberg University, Mannheim, Germany*
[5]*Department of Informatics, Technical University Federico Santa María, Valparaiso, Chile*

Publication

Robust detection and segmentation of cell nuclei in biomedical images based on a computational topology framework.
Rojas-Moraleda R, Xiong W, Halama N, Breitkopf-Heinlein K, Dooley S, Salinas L, Heermann DW, Valous NA
Med Image Anal. 2017 May