World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:14, No:12, 2020

# Identification of Clinical Characteristics from Persistent Homology Applied to Tumor Imaging

**Authors :** Eashwar V. Somasundaram, Raoul R. Wadhwa, Jacob G. Scott

**Abstract :** The use of radiomics in measuring geometric properties of tumor images such as size, surface area, and volume has been invaluable in assessing cancer diagnosis, treatment, and prognosis. In addition to analyzing geometric properties, radiomics would benefit from measuring topological properties using persistent homology. Intuitively, features uncovered by persistent homology may correlate to tumor structural features. One example is necrotic cavities (corresponding to 2D topological features), which are markers of very aggressive tumors. We develop a data pipeline in R that clusters tumors images based on persistent homology is used to identify meaningful clinical distinctions between tumors and possibly new relationships not captured by established clinical categorizations. A preliminary analysis was performed on 16 Magnetic Resonance Imaging (MRI) breast tissue segments downloaded from the 'Investigation of Serial Studies to Predict Your Therapeutic Response with Imaging and Molecular Analysis' (I-SPY TRIAL or ISPY1) collection in The Cancer Imaging Archive. Each segment represents a patient's breast tumor prior to treatment. The ISPY1 dataset also provided the estrogen receptor (ER), progesterone receptor (PR), and human epidermal growth factor receptor 2 (HER2) status data. A persistent homology matrix up to 2-dimensional features was calculated for each of the MRI segmentation. Wasserstein distances were then calculated between all pairwise tumor image persistent homology matrices to create a distance matrix for each feature dimension. Since Wasserstein distances were calculated for 0, 1, and 2-dimensional features, three hierarchal clusters were constructed. The adjusted Rand Index was used to see how well the clusters corresponded to the ER/PR/HER2 status of the tumors. Triple-negative cancers (negative status for all three receptors) significantly clustered together in the 2-dimensional features dendrogram (Adjusted Rand Index of .35, $p = .031$). It is known that having a triple-negative breast tumor is associated with aggressive tumor growth and poor prognosis when compared to non-triple negative breast tumors. The aggressive tumor growth associated with triple-negative tumors may have a unique structure in an MRI segmentation, which persistent homology is able to identify. This preliminary analysis shows promising results in the use of persistent homology on tumor imaging to assess the severity of breast tumors. The next step is to apply this pipeline to other tumor segment images from The Cancer Imaging Archive at different sites such as the lung, kidney, and brain. In addition, whether other clinical parameters, such as overall survival, tumor stage, and tumor genotype data are captured well in persistent homology clusters will be assessed. If analyzing tumor MRI segments using persistent homology consistently identifies clinical relationships, this could enable clinicians to use persistent homology data as a noninvasive way to inform clinical decision making in oncology.

**Keywords :** cancer biology, oncology, persistent homology, radiomics, topological data analysis, tumor imaging
**Conference Title :** ICCTTDA 2020 : International Conference on Computational Topology and Topological Data Analysis
**Conference Location :** London, United Kingdom
**Conference Dates :** December 10-11, 2020