Fast and Accurate Tumor Segmentation of Histology Images using Persistent Homology and Deep Convolutional Features

Talha Qaiser[a], Yee-Wah Tsang[b], Daiki Taniyama[c], Naoya Sakamoto[c], Kazuaki Nakane[d], David Epstein[e], Nasir Rajpoot[a,b,f,*]

[a]*Department of Computer Science, University of Warwick, UK*
[b]*Department of Pathology, University Hospitals Coventry and Warwickshire, UK*
[c]*Department of Molecular Pathology, Hiroshima University Institute of Biomedical and Health Sciences, Japan*
[d]*Graduate School of Medicine, Division of Health Science, Osaka University, Japan*
[e]*Mathematics Institute, University of Warwick, UK.*
[f]*The Alan Turing Institute, UK*

Abstract

Tumor segmentation in whole-slide images of histology slides is an important step towards computer-assisted diagnosis. In this work, we propose a tumor segmentation framework based on the novel concept of *persistent homology profiles* (PHPs). For a given image patch, the homology profiles are derived by efficient computation of persistent homology, which is an algebraic tool from homology theory. We propose an efficient way of computing topological persistence of an image, alternative to simplicial homology. The PHPs are devised to distinguish tumor regions from their normal counterparts by modeling the atypical characteristics of tumor nuclei. We propose two variants of our method for tumor segmentation: one that targets speed without compromising accuracy and the other that targets higher accuracy. The fast version is based on a selection of exemplar image patches from a convolution neural network (CNN) and patch classification by quantifying the divergence between the PHPs of exemplars and the input image patch. Detailed comparative evaluation shows that the proposed algorithm is significantly faster than competing algorithms while

---

[*]Corresponding author
 Email address: `N.M.Rajpoot@warwick.ac.uk` ()

achieving comparable results. The accurate version combines the PHPs and high-level CNN features and employs a multi-stage ensemble strategy for image patch labeling. Experimental results demonstrate that the combination of PHPs and CNN features outperform competing algorithms. This study is performed on two independently collected colorectal datasets containing adenoma, adenocarcinoma, signet, and healthy cases. Collectively, the accurate tumor segmentation produces the highest average patch-level F1-score, as compared with competing algorithms, on malignant and healthy cases from both the datasets. Overall the proposed framework highlights the utility of persistent homology for histopathology image analysis.

Keywords: Tumor Segmentation, Persistent Homology, Deep Learning, Histology Image Analysis, Computational Pathology, Colorectal (colon) Cancer.

2018 MSC:

1. Introduction

Colorectal cancer (CRC), also known as colon or bowel cancer, originates in the colon or the rectum, due to the abnormal growth pattern of cells. CRC is the third most commonly diagnosed cancer in males and the second most in females Ferlay et al. (2010), with an estimated 1.4 million cases and 693,000 deaths occurring in 2012 Torre et al. (2015). According to the National Cancer Institute (NCI), around 4.3% of the human population will be diagnosed with CRC during their lifetime, based on 2012-2014 data[1]. CRC results from excessive growth of malignant (cancer) cells in the colon or the rectum. The most common form of CRC is adenocarcinoma (found in up to 95% of CRC cases) which develops in epithelial gland cells— these are the glands that secrete mucus, which lubricates the colorectal region.

In a routine diagnostic process, a pathologist analyzes tissue sections on

---

[1]https://seer.cancer.gov/statfacts/html/colorect.html

glass slides under the microscope to observe morphological features and variability of nuclear morphology. However, careful visual examination of tissue slides is difficult when workloads are high, and the subjective nature of the task inevitably leads to inter and even intra-observer variability Viray et al. (2013); Smits et al. (2014). In contrast, automated algorithms can analyze the data with reproducible results. The reliability of the results of an algorithm can be objectively measured (for example against a patient's subsequent clinical progress) and then improved against an objective standard. This is now possible with the advent of digital slides scanners, as hundreds of glass tissue slides can now be scanned in a single run of the scanner. A whole-slide image (WSI) is a multi-resolution gigapixel image typically stored in a pyramid structure, formed by scanning a conventional glass slide at microscopic resolution. In view of the increasing number of CRC cases and shortcomings of the conventional diagnosis system, it is imperative to develop fast and reliable algorithms that can assist the histopathologists in their diagnosis of cancer.

Localization of malignant tumor regions in Hematoxylin and Eosin (H&E) stained slides is an important first task for a pathologist while diagnosing CRC. Manual segmentation of tumor regions from glass slides is a challenging and time consuming task. Therefore, automated localization of tumor-rich areas is a vital step towards a computer-assisted diagnosis and quantitative image analysis. Accurate segmentation of tumor-rich areas may also assist pathologists in understanding disease aggressiveness and selection of high power fields for tumor proliferation grading and scoring. In a recent study Dunne et al. (2016), it has been shown that precise localization of tumor epithelial regions in CRC can overcome the association of nonmalignant stroma regions in gene expression profiling, which also provides substantial prognostic information for individual cases. Hence, automated tumor segmentation of CRC tissue slides could potentially speed up the diagnostic process and overcome the inter-observer variability of conventional methods Litjens et al. (2016); Qaiser et al. (2018).

Tumor regions can be distinguished from normal regions using the appearance of cell nuclei Baba and Câtoi (2007); Clevert et al. (2015). In tumor

3

regions, epithelial nuclei have atypical characteristics — relatively large nuclei, with heterogeneous chromatin texture and irregularities in their shape and size. Due to uncontrolled cell division, tumor nuclei sometimes form clusters filling inter-cellular regions, as shown in Figure 1. In some cases with moderately and poorly differentiated grades, the structure of individual nuclei is difficult to discern. In contrast, nuclei retain their structure and morphological appearance in normal regions including stroma, lymphocytes, normal mucosa, and adipose tissue regions.

In this paper, we show that important morphological differences between normal and cancer nuclei can be measured using persistent homology, a mathematical tool explained in Section 3. We propose two persistent homology methods for tumor segmentation of H&E stained WSIs including *a)* homology based fast and reliable tumor segmentation *b)* accurate tumor segmentation by combining the homological and deep convolutional features to enhance the classification accuracy of a deep convolutional neural network (CNN).

We validate the proposed methods with relatively large datasets containing both malignant and healthy cases from two independent institutions. We repeat results of weak and strong cross-validation on the two datasets. Generally, in a clinical setup, a WSI scanner processes 500 to 1000 glass slides each day and so analyzing data from a single scanner may require the processing of several terabytes of new data each day. We, therefore, make a particular point in this paper of quoting run times, and show that our fast and reliable tumor segmentation algorithm is significantly faster than a conventional convolution neural network (CNN) and other competing approaches.

1.1. Our Contributions

We start with an algorithm to derive *persistent homology profiles* which effectively transform an input patch into two 1-D statistical distributions. These distributions capture the degree of nuclear connectivity in a given patch of a WSI. We propose two approaches for tumor segmentation based on the PHPs.
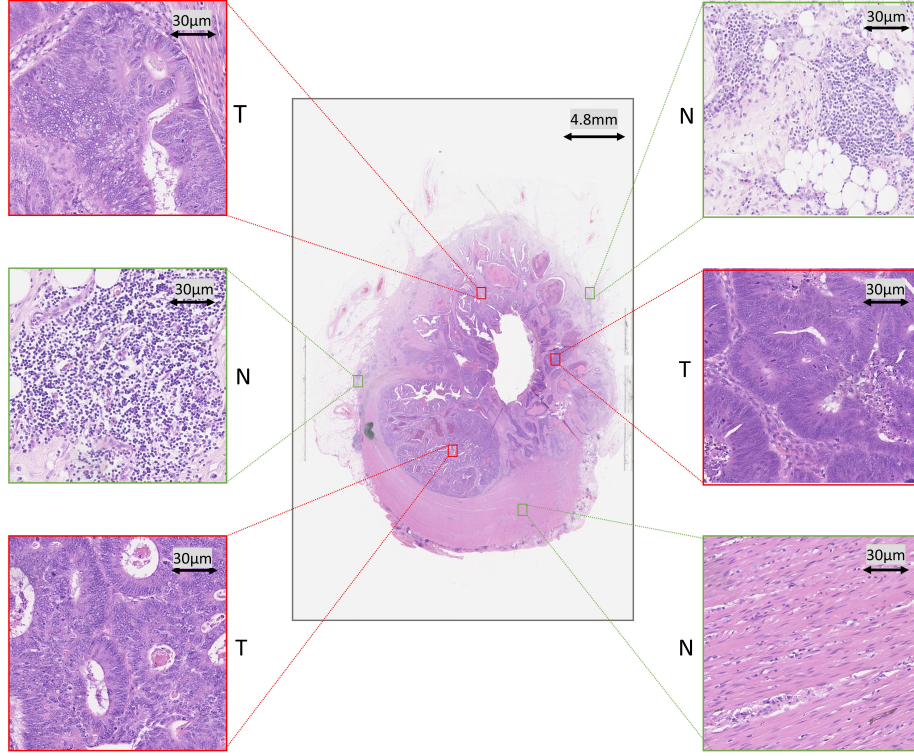
4

Figure 1: An example of a whole slide image with 6 regions-of-interest (ROIs) to illustrate the degree of connectivity between nuclei in tumor and normal regions. The zoomed-in regions are of size $140.8 \times 140.8 \ \mu m^2$, which is equivalent to $20\times$ magnification. ROIs with green rectangles (with label N) shows non-tumor whereas subregion with red rectangles (with label T) shows tumor areas.

In the first instance, we seek to develop a fast and reliable tumor segmentation algorithm. Presented with an image patch from a WSI, the algorithm computes the PHP, and then compares it with PHPs derived from tumor and normal *exemplar patches* using the symmetrized Kullback-Leibler divergence (KLD). The exemplar patches are precomputed, using CNN activations but the CNN itself is not used again after precomputation. Finally, we employ a novel variant of k-NN classifier to predict the outcome of a given patch as tumor or non-tumor. The fast algorithm, compared with the state-of-the-art, is significantly faster while competitive in accuracy.

Our second algorithm for accurate tumor segmentation is a combination of deep convolutional and topological features (PHPs). In contrast to the fast algorithm, this algorithm continues to use a CNN and extracts a set of learned features. We compute PHPs of the training dataset and separately employ a Random Forest regression model for both sets of features. We propose a multi-stage ensemble strategy to fuse the output from regression models. Experimental results on both the datasets demonstrate that the combination of topological and CNN features produces high segmentation accuracy.

Preliminary versions of these approaches were presented in Qaiser et al. (2016, 2017). In this paper, we provide a comprehensive literature review on tumor segmentation of histology WSIs, an intuitive explanation of persistent homology and our proposed algorithms by providing more details on the underlying principles. We report the results of our methods on two independent datasets comprising a range of malignant CRC cases including adenoma, adenocarcinoma, signet as well as healthy cases. This paper also contains a more robust approach for selection of exemplar patches from a CNN, discussion and comparative results on alternative options for selection of exemplar patches, runtime analysis of tumor segmentation approaches and finally, a detailed analysis on the robustness of the proposed methods.

1.2. Overview

The rest of the article is organized as follows: In Section 2, we briefly review some notable tumor segmentation algorithms in the literature. A brief introduction to persistent homology is given in Section 3. In Section 4, we describe the proposed method for computing the PHPs and our proposed algorithms for segmenting tumor regions in a WSI. A comprehensive evaluation, comparative analysis of the proposed framework followed by robustness analysis of the proposed methods on an unseen dataset is discussed in Section 5. Finally, we conclude with a summary and future directions in Section 6.

2. Related Work

Existing literature on tumor segmentation in histology images can be broadly classified into two categories: 1) hand-crafted feature based methods and 2) data-driven deep feature based algorithms. In this section, we review previous work regarding the tumor segmentation on images of H&E stained slides. Literature review on other related methods is covered in relevant sections, where appropriate.

A wide range of studies have been published on the use of texture, morphological and color features for tumor segmentation. Perception-based features Bianconi et al. (2015), local binary patterns (LBP) along with contrast measure features Linder et al. (2012), color graphs Altunbay et al. (2010) , Gabor and histogram features Khan et al. (2013); Kather et al. (2016), bags-of-superpixels pyramid Akbar et al. (2015), have been used to segment tumor rich areas. Weakly supervised multiple clustered instance approaches Xu et al. (2014, 2012) have also been proposed for segmentation of tumor in tissue micro arrays (TMAs) of colon cancer images, whereby bags of selected patches are generated to learn the model in a multiple instance framework. The multiple clustered instance model learns from a set of general features like the L*a*b color histogram, LBP, multiwavelet transforms, and scale invariant feature transforms. However, the selection of an optimal set of features for fully or weakly supervised learning is an onerous task and poses the risk of over-emphasizing some particular features of a dataset. Moreover, a major shortcoming of the above algorithms is the fact that their scope is mainly limited to hand-picked visual fields or TMAs. In a clinical setup, a tumor segmentation solution should be capable of scaling the results to the WSI level.

Deep learning has recently produced exceptional performance on tasks in computer vision Krizhevsky et al. (2012) and in medical image processing Bejnordi et al. (2017); Chen et al. (2017); Sirinukunwattana et al. (2017). One of the well-known methods for segmentation is to learn a set of hierarchical features by employing a combination of down and up sampling convolution layers,

7

such as U-net Ronneberger et al. (2015). These approaches perform reasonably well for pixel level segmentation but are computationally expensive and may encounter the vanishing gradients problem while training. Cruz-Roa et al. (2017) presented a CNN framework for tumor detection in breast histology images. Our CNN architecture for tumor segmentation has some basic similarities with the proposed framework in Cruz-Roa et al. (2017). However, our CNN model is relatively deep, enabling the model to learn a set of features at various levels of abstraction. In a supervised learning environment, one can think of deep learning features as a set of data-driven features, learnt by using back propagation (BP). The BP algorithm penalizes the kernel maps by forcing them to learn from their mistakes on the training dataset. Instead of relying on a set of handcrafted features, deep learning models learn the set of optimal features without human intervention.

Each of the several physical processes involved in creating a WSI, starting with the original biopsy or resection and ending with laying a 3µm thick section on a glass slide, is random with respect to orientation. Therefore the textural and geometric features in a WSI will have a random orientation, though the orientations of different features may well be correlated with each other. However, deep learning models and especially CNNs find difficulty in learning the rotationally invariant characteristics of an input image Sabour et al. (2017). In contrast, our biologically interpretable PHPs not only capture the degree of connectedness among nuclei but are also invariant to rotational transformations.

3. Introduction to Persistent Homology

Persistent homology is an algebraic tool, whereby, given a topological space, certain *algebraic invariants* are computed using the structure of that space. It is a fairly recent concept of homology theory, with a wide range of applications in different domains of data analytics including protein structure Xia and Wei (2014); Cang and Wei (2017), robotics Bhattacharya et al. (2015); Pokorny et al. (2016), neuroscience Curto (2017), shape modelling Turner et al. (2014),

8

analyzing brain arteries Bendich et al. (2016), classification of endoscopy images Dunaeva et al. (2016), mutational profile and survival analysis Wubie (2016), video surveillance Lamar-Leon et al. (2016), time series modelling Otter et al. (2017) and natural language processing Zhu (2013). The concept of persistent homology is relatively new for medical image analysis in general and for histology image analysis in particular.

Persistent Homology Theory is the study of the homology of a filtered space, by which we mean a sequence

$$\{\emptyset = X_0 \subseteq X_1 \subseteq X_2 \subseteq \cdots \subseteq X_k = \mathbf{X}\} \tag{1}$$

This is referred to as a *filtration* of the topological space X. Readers are referred to Edelsbrunner and Harer (2008); Carlsson (2009); Ghrist (2008) for a description of the general theory.

In our case, we are analyzing 2D greyscale images, and each X_i is the union of closed pixels in a single image. $X_k = X$ is the entire image. These subspaces are so special, and with such nice properties, that dramatic simplifications are possible in computing the persistent homology. The only homology groups that are non-zero are in dimensions 0 and 1. We have no need to consider what is often a significant aspect of persistent homology, namely the birth and death of homology classes. It is sufficient for our purposes to consider only the Betti numbers $\beta_0(X_i)$ and $\beta_1(X_i)$ for $1 \leq i < k$. Moreover, these Betti numbers can be computed using basic topological ideas, namely a count of connected components, which is a simple and rapid computational procedure.

To explain how we generate the filtration of a given greyscale image, we suppose for definiteness that the intensity of each pixel is an integer in the range $[0, 255]$. We then select a sequence of integers $0 = t_0 < t_1 < \cdots < t_{k-1} < t_k = 256$. These integers are various threshold levels, at which the image is binarized. We define X_i to be the union of closed pixels p, with intensity $I(p) < t_i$, so that $X_0 = \emptyset$; and $X_k = X$. Each grayscale image gives rise to a single filtration. A

9

careful choice[2] of $t_1, t_2, \ldots, t_{k-1}$ balances density in $[0, 255]$ to give all (or nearly all) the information needed from the original grayscale image, against sparsity to make the algorithm run fast. We compute two relevant homology groups $H_0(X_i)$ and $H_1(X_i)$ for $1 \leq i < k$ as follows. Instead of using computationally expensive constructs of simplicial topology, as is normal in the literature using persistent homology Nakane et al. (2015, 2013), we compute $H_0(X_i)$ by counting the connected components of X_i, giving the Betti number $\beta_0(X_i)$ and $H_1(X_i)$ by counting the components of $X \setminus X_i$, giving the Betti number $\beta_1(X_i)$. For our purposes, it is therefore sufficient to calculate, for each i with $1 \leq i < k$, these two numbers, giving a total of $2k - 2$ length of feature vector.

A given small patch as in Figure 2 represents a filtered space extracted from a WSI. For a range of t it gives a list of subspaces such that all pixels of previous subspace X_{i-1} are present in X_i. To construct PHPs we recorded the rank of homology groups $(H_0(X_i), H_1(X_i))$ for an entire range of t, such as $0 < t_1 < t_2 < \cdots < t_{k-1} < t_k = 256$. We trade a considerable improvement in speed of computation for a negligible loss in information, as explained in Footnote 2. The algorithm for computing PHPs is the main foundation for fast tumor segmentation. In this case, the *algebraic invariants* turn out to be nothing more complicated than whole numbers (ranks of homology groups, which are 0th and 1st Betti numbers). Hence we do not need to build computationally expensive simplical complexes in order to compute persistent features. This provides an alternative approach to simplicial homology for 2D images.

4. The Proposed Approaches

In this section, we present two proposed approaches based on PHPs. We first describe the workflow of the fast tumor segmentation and then explain a variant for accurate tumor segmentation. For a given WSI, we divide it into patches of

---

[2]We empirically decided on choosing every third possible value in the range 1..255, giving 84 levels instead of 255.
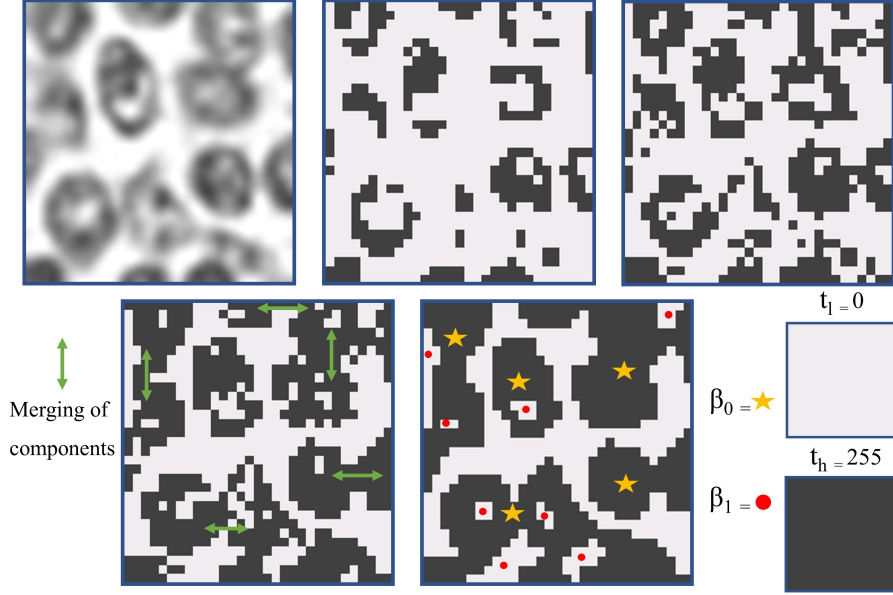
Figure 2: An illustration of the filtered space associated to an image. We first show an input image at $40\times$ magnification, and then four images showing growing sequence of subspaces, obtained by thresholding at increasing values of threshold. β_0 (yellow-stars) represents the number of connected components and, similarly, β_1 (red-dots) shows the number of one-dimensional voids. ~~(red arrows only show few of them).~~

256×256 at $20\times$ magnification. The problem then reduces to classification of each patch as either tumor or non-tumor.

4.1. Fast Tumor Segmentation

The algorithm for fast tumor segmentation is established on three pivotal steps: *1)* an efficient way of computing PHPs, *2)* selection of representative images from the activation maps of a convolutional network, and *3)* a novel algorithm for patch classification.

4.1.1. Persistent Homology Profiles

As mentioned in Section 1, tumor nuclei carry atypical characteristics and exhibit chromatin texture allowing us to distinguish them from non-tumor nuclei. Here we characterize these phenomena with the help of persistent homology

for tumor regions in CRC histology images. Our topological features provide a global description of image patch I by finding the relationships among data points (pixels), contrary to textural and geometrical features where precise distances, angles and spatial arrangement are important. For a given patch I, we derive two statistical distributions by computing the ranks of their 0th and 1st dimensional homology groups, denoted by β_0 and β_1. We refer to these statistical distributions as *persistent homology profiles*.

Highly variable colour of histology specimens is mainly due to non standardization in staining protocols. To overcome this problem we first perform stain deconvolution Ruifrok et al. (2001) separating an RGB image patch into three channels, Hematoxylin, Eosin, and background. For follow up analysis, we only use the Hematoxylin channel to improve consistency in intensity appearances. To extract the topological features, we binarize the Hematoxylin channel to record the corresponding Betti numbers (β_0, β_1). In our case, the topological features were only recorded for one third of possible threshold values – see Footnote 2. These limited threshold values retain the discriminative characteristics of PHPs by using three times fewer parameters as compared to using all threshold values Qaiser et al. (2017). For each t, the Betti numbers are computed by counting the connected components and one-dimensional voids, as explained in Section 3 and Figure 2. Rather than relying on hand-picked threshold values, we observe the topological features at $0 = t_0 < t_1 < \cdots < t_{k-1} < 256$. Later, we convert each statistical curve into a discrete probability distribution, scaling the values so that the area under each curve is one.

Differences between tumor and non-tumor regions are reflected in their homology invariants. This can be seen in Figure 3 and 4, which show the curves representing median of our PHPs for selected exemplar patches from a CNN (explained in Section 4.1.2) for both tumor and normal classes, with first (Q_1) and third (Q_3) quartile. The green dotted line shows the PHPs (β_0, β_1) for image patches as shown in the first column. It is worth mentioning here the magnitude of derived PHPs is less relevant instead, the pivotal aspect is the noticeable trend in growth of homology classes. As we start increasing the
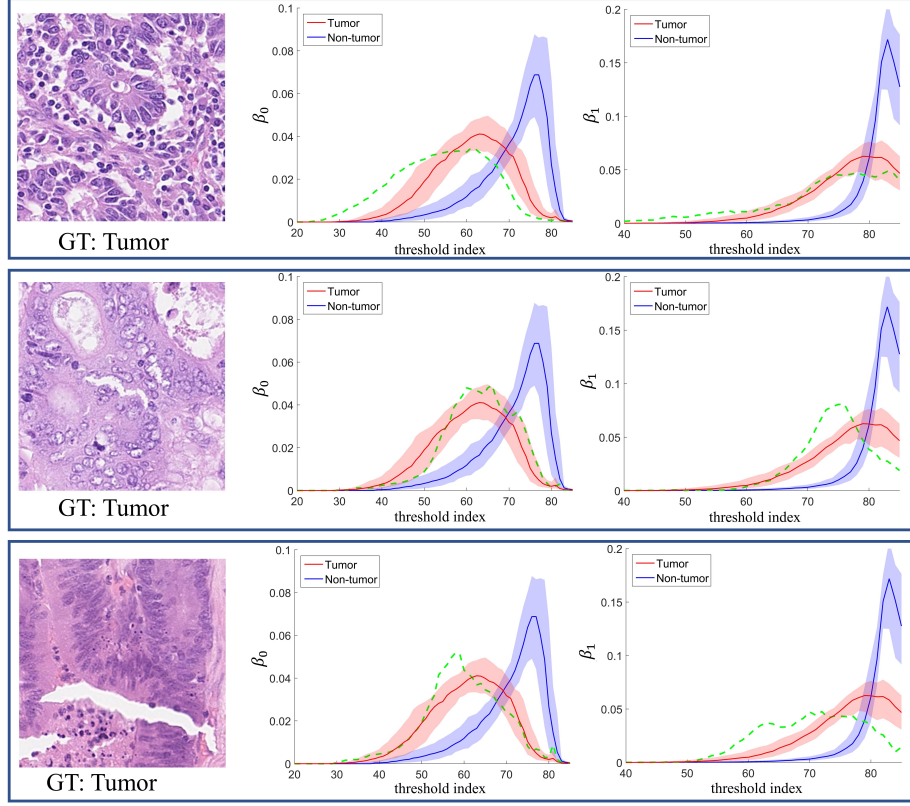
12

Figure 3: An example of persistent homology profiles (PHP) for the selected tumor patches (left): original images with ground truth (GT), (middle): PHP for β_0, (right): PHP for β_1. The shaded regions in β_0 and β_1 show the first and third quartile of the exemplar patches, whereas the green dotted line shows the PHP of selected patch.

threshold t from the lower limit (t_0) to upper limit (t_{k-1}), the filtering subspace propagates from an empty set to the entire topological space. Since tumor regions carry more irregularity in their shape, size and tumor nuclei lie relatively close to each other filling the inter-cellular cytoplasmic space, their homology ranks (β_0, β_1) do not show rapid change while merging and forming into new classes as compared to those for non-tumor regions.
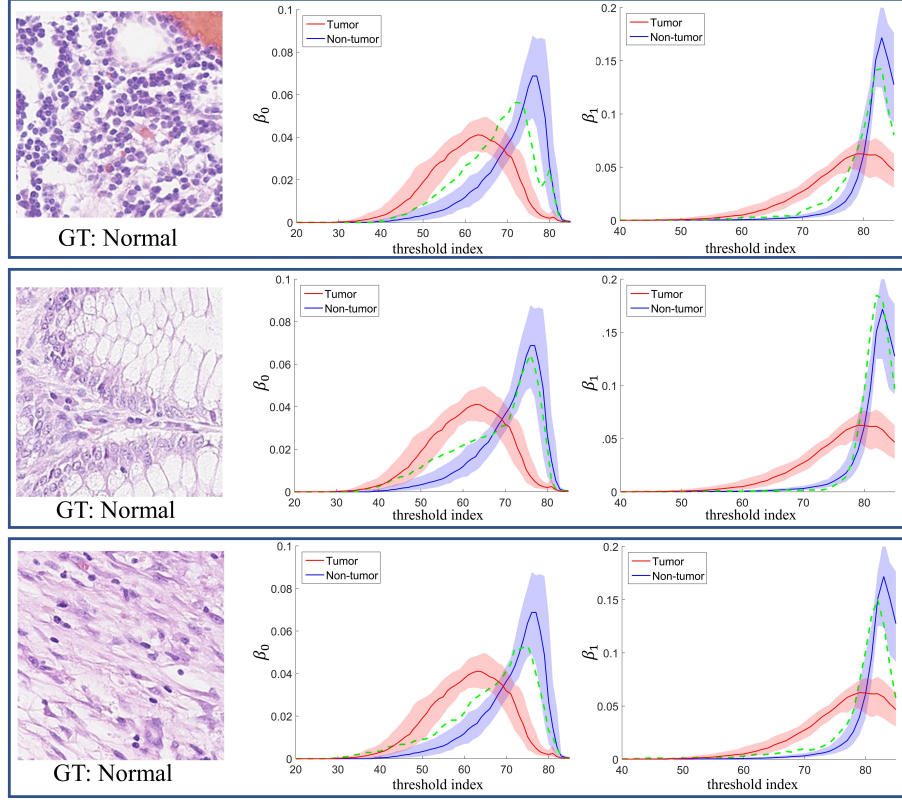
13

Figure 4: Another example of persistent homology profiles (PHP) for selected non-tumor patches (left): original images with ground truth (GT), (middle): PHP for β_0, (right) PHP for β_1 .

4.1.2. Selection of Exemplar Patches

To select exemplar patches for fast tumor segmentation, we first train a deep CNN model to predict whether a patch should be labeled as tumor or non-tumor. The CNN architecture is inspired by Krizhevsky et al. (2012) with some modifications, as shown in Figure 5. The objective here is to infer a set of representative patches from the entire training dataset for both tumor and non-tumor classes by exploring the learned activation maps from the last convolution layer. We then compute the PHPs of selected patches in order to measure the value of divergence from an input patch as described in the next section.

14

Let us consider a convolution layer with its corresponding activation maps $\alpha \in \mathbb{R}^{W \times H \times Z}$, where Z represents the depth of activation maps of spatial dimension $H \times W$. The activation maps from convolution layers emphasize different tissue parts for different layers Zeiler and Fergus (2014). The idea here is to collapse the 3D activation maps of the last convolution layer into a scalar value that can be later used as an indication of the significance of each patch with respect to the activation maps. We first define a mapping function closely related to Zagoruyko and Komodakis (2016) to flatten the 3D activation maps to 2D across the Z dimension as below,

$$F(w, h) = \sum_{z=1}^{Z} \left| \alpha_{(z)}^{w,h} \right|^2, \tag{2}$$

Here, the mapping function gives more weight to neurons with high activations. It is worth noting that normalizing the flatten 2D activation map is important for follow-up analysis. We compute the median value of $F(w, h)$ to find the central tendency of the 2D activation map, that later assist in exclusion of unimportant patches in the training dataset. Similarly, for the entire training dataset, we get an $M \times 1$ vector by computing the $median(F(w, h))$ for each patch, where M represents the number of patches.

The last step is to find a set of exemplar patches for tumor and non-tumor classes, separately. One way of choosing the exemplar patches is to find the highly activated patches from the $M \times 1$ matrix. The only pitfall of following such a method is that we may end up selecting patches representing a certain type of tumor or normal tissue. In order to avoid this scenario, we compute the interquartile range (IQR) Upton and Cook (1996) of $M \times 1$ matrix separately for the tumor and non-tumor classes. Furthermore, we split the IQR of each class into Q same size bins and select the value that lies closest to the median of corresponding bin, where Q represents the number of exemplar patches for each class. This method for selection of exemplar patches differs from Qaiser et al. (2016) and is capable of handling a much larger dataset as in Section 5.2.1.
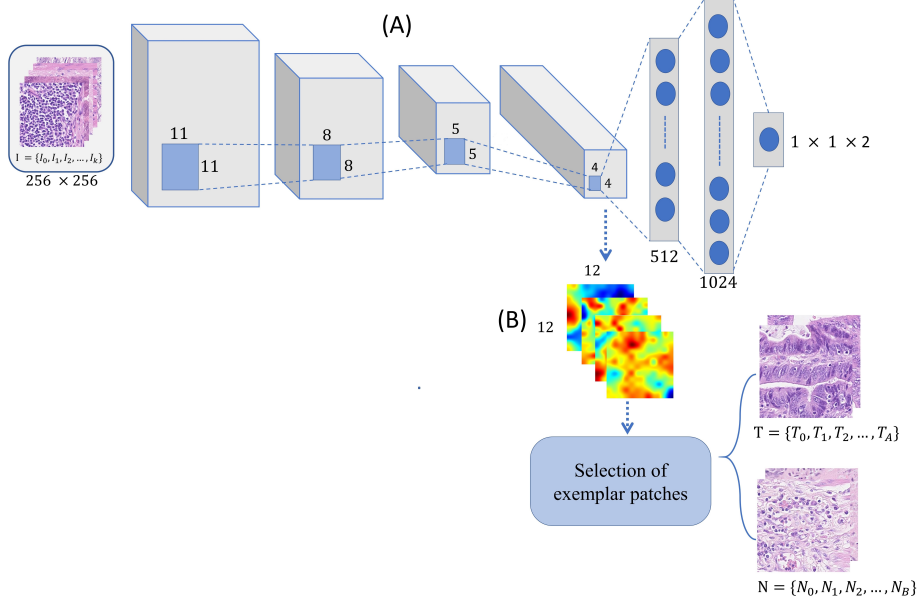
Figure 5: (A) A schematic illustration of our deep convolution neural network (B) An overview of the proposed exemplar selection algorithm based on activation maps of the last convolution layer.

4.1.3. Patch Classification

For patch classification, we first transform the PHPs into discrete probability distributions. We then compute the symmetric KLD to measure the distance between the PHPs of an input patch I and the PHPs of an exemplar patch E as defined below,

$$D_{KL}(I \parallel E) = \sum_i I(i) \log \frac{I(i)}{E(i)}, \tag{3}$$

where I represents the *true* representation of data and E represents an *approximation* of I. The symmetrised, non-negative KLD is defined as

$$d_{I,E} = D_{KL}(I \parallel E) + D_{KL}(E \parallel I). \tag{4}$$

We compute a vector of distance values to exemplar patches $D(I) = (d_{I,T_1}, d_{I,T_2}, ..., d_{I_A,T_A}, \ d_{I,N_1}, d_{I,N_2}, ..., d_{I,N_B})$, containing divergence values for input patch PHP profiles of I and all exemplar tumor $T = \{T_1, ..., T_A\}$ and non-tumor exemplar patches $N = \{N_1, ..., N_B\}$. We derive a similarity measure from the

KLD values computed in (4) and compare the total similarity scores to the k nearest tumor and non-tumor patches as follows:

$$\sum_{j=0}^{k_t} e^{-d_{I,T_j} c} > \sum_{j=0}^{k_n} e^{-d_{I,N_j} c}, \tag{5}$$

where c is a constant in the interval $(0, \max(d))$ and k_t, k_n denotes the number of nearest tumor and non-tumor patches according to (5), and $k = k_t + k_n$. In order to classify a given image patch as tumor, the total similarity scores of its k_t nearest tumor patches should be greater than that of the k_n nearest normal patches.

4.2. Accurate Tumor Segmentation

In this variant of our algorithm, we combine deep convolutional and persistent homology features. For this we trained a CNN as shown in Figure 5 to extract features from the last fully connected layer. We then fed the extracted features to a Random Forest (RF) regression model separately for the topological and deep convolutional features. Finally, we propose a multi-stage ensemble strategy to combine the two RF regression models. The main objective of this method is to combine *the best of both worlds*. The key contribution of the topological features is to capture the underlying connectivity whereas CNN tends to learn the data driven features.

4.2.1. Deep Convolutional Features

The CNN architecture contains four convolutional layers followed by an activation function and a max-pooling operation. Additionally, it contains two fully connected layers and the softmax classification layer to predict the label of each patch as a tumor or non-tumor. Instead of using a rectified linear unit ReLU as activation function, we use an exponential linear unit (ELU) Clevert et al. (2015), as it enables faster convergence and also reduces the vanishing gradient problem. A dropout layer at the end of the second fully connected layer is placed to overcome the overfitting problem. The CNN was trained to

minimize the overall cross entropy loss L, as given below

$$L(g, y) = -(y(x)log(g(x)) + (1 - y(x))log(1 - g(x)))\qquad(6)$$

315

where for input x, g represents the ground-truth label (0 for tumor and 1 for non-~~normal~~tumor) and y is the probability of the tumor predicted by the CNN. The fully connected layers contain non-linear combinations of learned features from the convolution layers. We extracted CNN features for the training dataset

320 after the last fully connected layer just before the softmax classification layer. For each patch of the training dataset, we obtained a feature vector of size $(1 \times 1 \times 1024)$.

4.2.2. Ensemble Strategy

After obtaining topological and convolutional features, we concatenate both

325 PHPs $(\beta 0, \beta 1)$ to form a combined topological feature vector. We then train the RF regression model separately for both types of feature. We optimize the RF model with an ensemble of 200 bagged trees, randomly selecting one third of the variables for each decision split and setting the minimum leaf size to 5.

We combine the probability of both regression models (O_1, O_2) as in (7),

330 where O_1 represents a regression model of topological features and similarly O_2 represents a regression model of convolutional features. The multi-stage ensemble strategy follows two alternative routes: a) averaging the outcome probabilities of O_1 and O_2 to predict the output label where both regression models agree b) for the remaining few patches ($\approx 1\%$ from the test data) where

335 the average probabilities lies in range $0.49 - 0.51$. We refer to these as *critical patches* and we assign the output label for those patches by rounding the probabilities from $O_1(x)$ as in our experiments based on two datasets we observed

18

the comparatively high F1-score with the setting.

$$\hat{O}(x) = \begin{cases} 0, & \text{if } \frac{(O_1+O_2)}{2} < 0.49. \\ 1, & \text{else if } \frac{(O_1+O_2)}{2} > 0.51. \\ \lfloor O_1(x) \rceil & \text{otherwise (rounding)} \end{cases} \qquad (7)$$

5. Experiments and Results

5.1. Dataset and Experimental Setup

The Warwick-UHCW Dataset. This dataset consists of 75 WSIs of H&E stained colorectal adenocarcinoma tissues. At the highest resolution, each WSI normally contains more than 10^{10} pixels. The WSIs were digitally scanned at a pixel resolution of 0.275µm/pixel (40×) using an Omnyx VL120 scanner. The ground-truth for tumor regions were handmarked by an expert pathologist. For each WSI, we randomly selected 1,500 patches including 750 from non-tumor and 750 patches from tumor regions. In total, we extracted, 75,000 patches for training from 50 WSIs and 37,500 patches for testing from 25 WSIs. The collected dataset for this study is roughly 20 times more than that in Qaiser et al. (2016) and at least 2 times more than in Qaiser et al. (2017). For generating the tumor probability map of a WSI, we first split the given WSI into patches and then applied our methods to each patch.

The Warwick-Osaka Dataset. This dataset contains 50 H&E stained histology WSIs of colorectal tissue. The WSIs were scanned at a pixel resolution of 0.23µm/pixel (40×) using a Hamamatsu NanoZoomer 2.0-HT scanner. The ground-truth for this dataset were handmarked by two expert pathologists and the cases were identified as belonging to 6 categories, including 11 cases of adenoma, 14 moderately differentiated, 6 poorly differentiated, 10 well-differentiated, 8 healthy and 1 signet case. The inclusion of normal cases in this dataset helps in evaluating the robustness of the proposed methods as discussed in Section 5.2.4. Similarly to the Warwick-UHCW dataset, we randomly selected 1,500 patches (750 tumor, 750 non-tumor) from each WSI.

19

Experimental Setting. For our proposed methodologies we split a WSI into manageable patches of 256×256 for training as well as testing. In order to counter overfitting, we performed data augmentation by rotating $(0°, 90°, 180°, 270°)$, flipping (horizontal or vertical axis), and perturbing the color distribution (hue variation) of both the training datasets. The weights for a CNN were initialized by using Xavier initialization as in (8).

$$l = \sqrt{\frac{3}{(N_{in} + N_{out})}} \qquad (8)$$

where N_{in} and N_{out} represents the number of input and output neurons, respectively and l represents the computed value for weights initialization. The
365 selected weight initialization approach tends to restrict the magnitude of gradients from excessive shrinking or growing during the training process. The network learns the weights by using the mini-batch gradient descent algorithm by selecting a batch size of 100. During the training phase, the initial learning rate was set to 0.0001 and an Adam optimizer was employed instead of conven-
370 tional gradient descent algorithm. In addition, a dropout layer (dropout rate 0.5) was placed between the two fully connected layers to overcome the interdependence among intermediate neurons and to increase the robustness of the trained network. For the fast tumor segmentation, we separately chose 128 exemplar patches for both tumor and non-tumor, $c = 0.2$ for similarity measures
375 and $k = 11$ for k-NN.

Evaluation. We compute the F1-score to evaluate the performance of different approaches as a harmonic mean of precision and recall as defined below,

$$F_1 = 2 \times \frac{P_r \times R_e}{(P_r + R_e)}; \qquad P_r = \frac{T_p}{T_p + F_p}; \qquad R_e = \frac{T_p}{T_p + F_n} \qquad (9)$$

where T_p, F_p, and F_n represents the number of true positive, false positive, and false negative patches. For a given test dataset, the correctly identified patches are classified as either true positives or true negative, misclassified predictions are categorized as false positives, and true negatives.
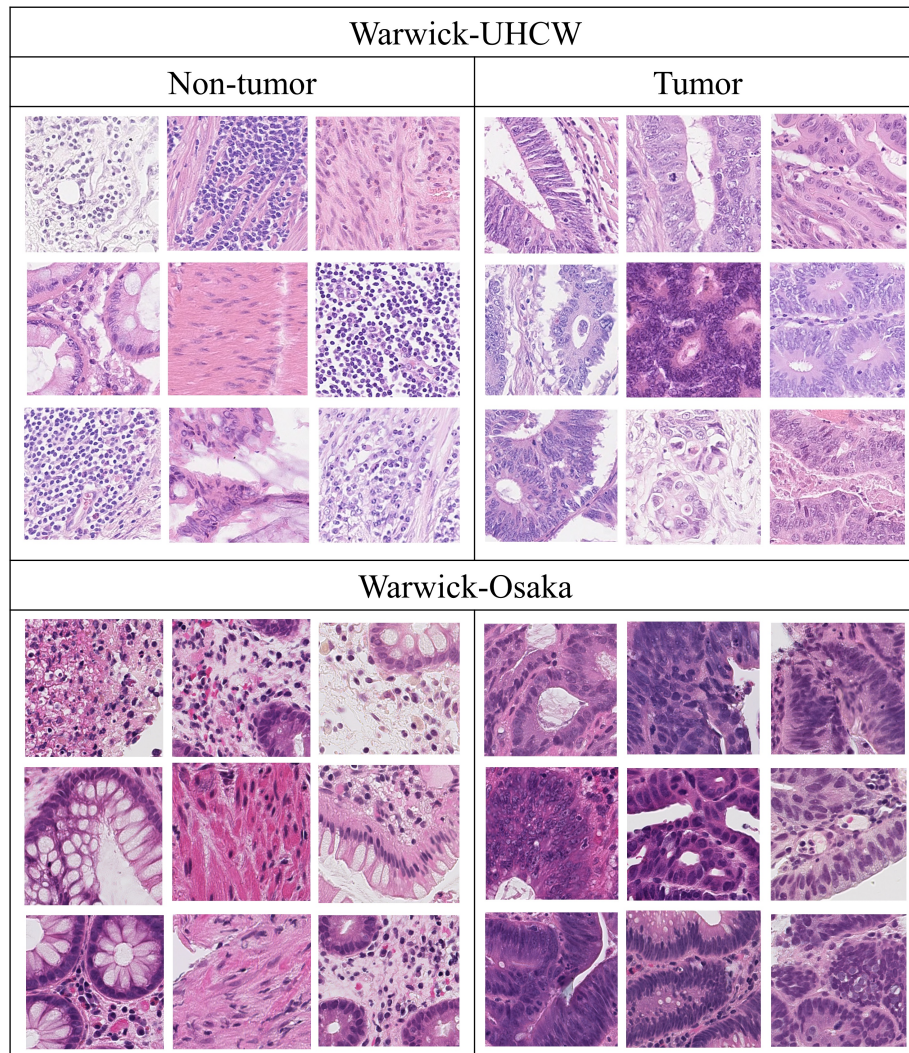
Figure 6: Representative tumor and non-tumor patches from the Warwick-UHCW and the Warwick-Osaka datasets.

Table 1: Comparison of various options for selecting the exemplar patches; reults on the Warwick-UHCW dataset.

Method	Precision	Recall	F1-score
CNN (Section 4.1.2)	0.9272	**0.8513**	**0.8999**
CNN (highly activated) Qaiser et al. (2016)	0.9112	0.8309	0.8692
k-means	**0.968**	0.7559	0.8489
Random selection	0.8812	0.8309	0.8553

380 *5.2. Comparative Analysis*

5.2.1. Selection of Exemplar Patches

One of the most critical parts in the fast tumor segmentation is the selection of exemplar patches. Hence the objective of this experiment is to investigate a few options for selection of exemplar patches. The experiment is conducted
385 on 75 colorectal adenocarcinoma WSIs from the Warwick-UHCW dataset. The following algorithms were selected for comparison and as an alternative to the proposed method of exemplar selection. We start with a random selection of exemplar patches from the training dataset for both tumor and non-tumor classes. We repeat this process 10 times before concluding the final results and the
390 reported results (Table 1) show the mean precision, recall and F1-score. In k-means clustering the training dataset was partitioned into k clusters with respect to their RGB intensities. Then we selected those patches that lie closest to the cluster centroids as exemplar patches, one patch per centroid. The third algorithm for comparison is to select highly activated patches as exem-
395 plars from a CNN as proposed in Qaiser et al. (2016). For a fair comparison, an equal number of exemplar patches were selected for each of the above mentioned algorithms.

Table 1 reports the patch based tumor segmentation results for different approaches. Overall the results are in favor of the proposed method. Figure
400 6 shows a sample of 9 representative patches for tumor and non-tumor classes

22

from both the datasets using the proposed method. The CNN (highly activated) Qaiser et al. (2016) seems a straightforward approach that selects only patches where its neurons produce high activations. However, for a relatively large dataset, the exemplars may be overemphasized by a particular atypical WSI, where we normally have thousands of patches from a single case. Thus, this kind of approach is more suitable for a small dataset containing a limited number of image patches from which to select exemplars. Additionally, this kind of approach may also be useful for tissue regions having relatively less tumor heterogeneity. Another downside is the inclusion of outliers as exemplars which can easily happen due to lack of precisely marked ground-truth. This argument remains valid for k-means and random selection approaches.

5.2.2. Tumor Segmentation on Adenocarcinoma Cases

In this experiment, we evaluate the performance on the UHCW-Warwick dataset of the proposed algorithms in comparison to some recently published algorithms for tumor segmentation. The experiment is conducted on all 75 adenocarcinoma WSIs where we used 75,000 randomly selected patches from 50 cases for training and the remaining 37,500 patches from 25 WSIs for testing. For a fair comparison, we retrained the selected algorithms on our dataset except for HyMap Khan et al. (2013), which is an unsupervised method. For comparative analysis we selected only algorithms that are closely related to CRC image analysis, provided that their source codes were released with proper implementation details and comments. The features of the selected algorithms are as follows:

- Multi-class texture analysis Kather et al. (2016) or MCTA: This method first computes a set of textural features on a given patch including lower-order and higher-order histogram statistics, local binary patterns, gray-level co-occurrence matrix, Gabor filter, preception-like features and then fed these features into radial-basis function (RBF) support vector machine (SVM).
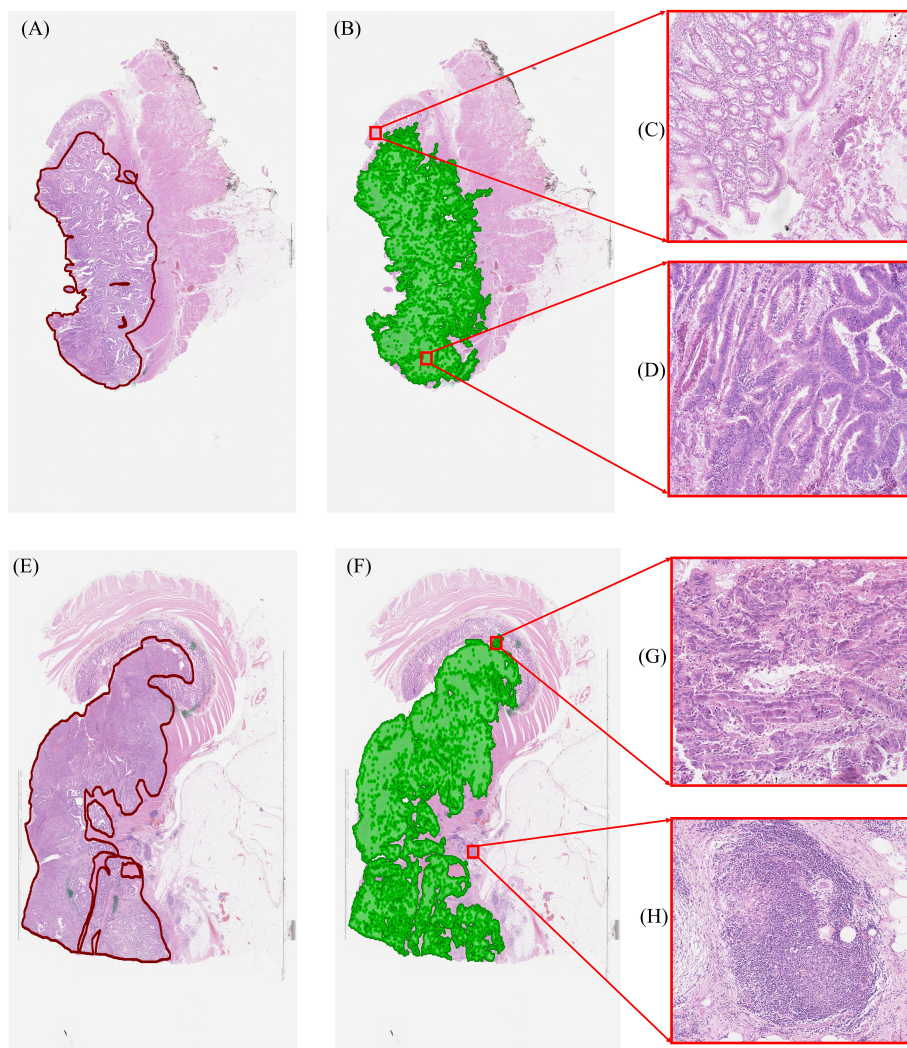
23

Figure 7: Results of accurate tumor segmentation on the whole-slide image (WSI) level. (A)&(E) input WSIs with annotated ground-truth, (B)&(F) tumor segmentation results, green region showing the predicted tumor areas, (C)&(H) zoom in regions containing true negatives (D)&(G) showing a sample of the true positive tumor segmented regions.

Table 2: Tumor segmentation results on the Warwick-UHCW dataset

Method	Precision	Recall	F1-score
Fast Tumor Segmentation (PHP)	**0.9272**	0.8513	0.8999
Accurate Tumor Segmentation (PHP+CNN)	0.9267	**0.922**	**0.9243**
HyMap Khan et al. (2013)	0.6851	0.8600	0.7626
ConvNet CNN $_3$ Cruz-Roa et al. (2017)	0.856	0.867	0.8615
MCTA Kather et al. (2016)	0.8701	0.8834	0.8767
TVIA Turkki et al. (2015)	0.8224	0.8641	0.8427

- HyMap Khan et al. (2013): This is an unsupervised algorithm for tumor segmentation that classifies each pixel as hypo or hyper cellular by computing Gabor filter, texture energy and phase-gradient features followed by ensembling the projections for each feature. Since this is an unsupervised method, we did not retrain it for our experiments.

- ConvNet CNN_3 Cruz-Roa et al. (2017): This method poses cancer detection as a two-class problem by assigning an 'invasive or non-invasive' label to each patch. The framework is a combination of convolutional layers followed by pooling operations and fully connected layers. The reported results suggest the CNN_3 outperformed the other counterparts on breast tissue so we used only CNN_3 for our experiment.

- Texture-based tumour viability image analysis Turkki et al. (2015) or TVIA: It classifies a given patch as viable or non-viable or as 'other tissue parts'. They computed local binary patterns and local contrast measures at the patch level and fed the extracted features into a SVM model. This algorithm is not directly related to tumor segmentation. We, therefore, evaluate it only as a two-class problem, that is, as a decision whether to classify a patch as tumor or non-tumor.

Table 2 reports comparative results from the experiment and Figure 7 shows qualitative results for the accurate tumor segmentation on the WSI level. Our

25

PHP+CNN based accurate tumor segmentation produces the best results in terms of recall and F1-score, whereas the PHP based fast tumor segmentation performs best for the precision metric. The PHP+CNN outperformed the other competing methods by a reasonable margin. It is encouraging that incorporating deep features along with topological features boosts the overall performance. Generally, CNN models struggle to capture the rotation or viewpoint invariance of the learned object. To overcome this deficiency, we need to perform flipping or several arbitrary rotations on our images in the form of augmentation to make our model more generalizeable. Data augmentation overcomes the rotational invariance to some extent but it is still a non-trivial task to produce all possible rotations while training a CNN classifier. In contrast, PHP captures the rotational invariance by emphasizing the merging and forming of homology classes so no matter how much a patch is rotated the PHP will remain persistent. The PHP also captures the biological phenomenon that the connectivity among tumor and non-tumor nuclei is significantly different. As compared to CNNs, the PHP based fast tumor segmentation algorithm only consults a number of exemplar patches from both classes in predicting the outcome of a patch. In our previous work Qaiser et al. (2016), we observed that a similar approach performs marginally better than CNN. This offers a trade-off between accuracy and efficiency with reliable outcomes.

5.2.3. Tumor Segmentation on Adenoma, Carcinoma and Healthy Cases

The goal of this experiment is to test the generalization of the proposed algorithms on another dataset that consists of different types of epithelial tumors and healthy cases. The experiment is conducted on 50 WSIs of the Warwick-Osaka dataset. We perform 2-fold cross-validation by selecting half of the dataset for training and the remaining half for testing. We then perform the same experiment by switching the training and test datasets. Similarly to the experiment with the Warwick-UHCW dataset we perform comparative analysis on the aforementioned selected algorithms and by retraining them on the Warwick-Osaka dataset.

Table 3: Tumor segmentation results on the Warwick-Osaka dataset

Method	Precision	Recall	F1-score
Fast Tumor Segmentation (PHP)	0.8259	0.8019	0.8137
Accurate Tumor Segmentation (PHP+CNN)	**0.8311**	**0.8235**	**0.8273**
HyMap Khan et al. (2013)	0.6469	0.7228	0.6827
ConvNet CNN_3 Cruz-Roa et al. (2017)	0.6927	0.8446	0.7612
MCTA Kather et al. (2016)	0.7050	0.7419	0.7229
TVIA Turkki et al. (2015)	0.6993	0.7240	0.7114

480   Table 3 shows the results from 2-fold cross-validation. It can be observed that the proposed methods for all 3 metrics perform well on an independent dataset. One of the arguments for ascribing relatively low performance by MCTA, HyMap, and TVIA is the selection of textural features, especially local binary patterns, local contrast measures, histogram statistics, and the gray-level

485 co-occurrence matrix, that are sensitive to stain variations and image blurring. An additional cause for the poor performance of these programs is that, due to varying staining protocols at different centers, the morphological appearance of lymphocytes and benign epithelial nuclei from one center can resemble that of malignant epithelial nuclei from another.

490 *5.2.4. Results for Healthy Cases*

This experiment compares the performance of different tumor segmentation algorithms for healthy cases in the Warwick-Osaka dataset. This is important for routine clinical practice as well as in cancer screening studies, which involve a large number of normal cases. The most challenging sections are healthy epithe-

495 lial and lymphocyte regions where nuclei are densely packed and pose difficulties in identifying those regions as non-tumor. We evaluate the performance of this experiment by the specificity metric (also known as the true negative rate), that measures the involvement of negative samples misclassified as positive. Figure 8 shows the summarized results for all the healthy cases involved in this study.
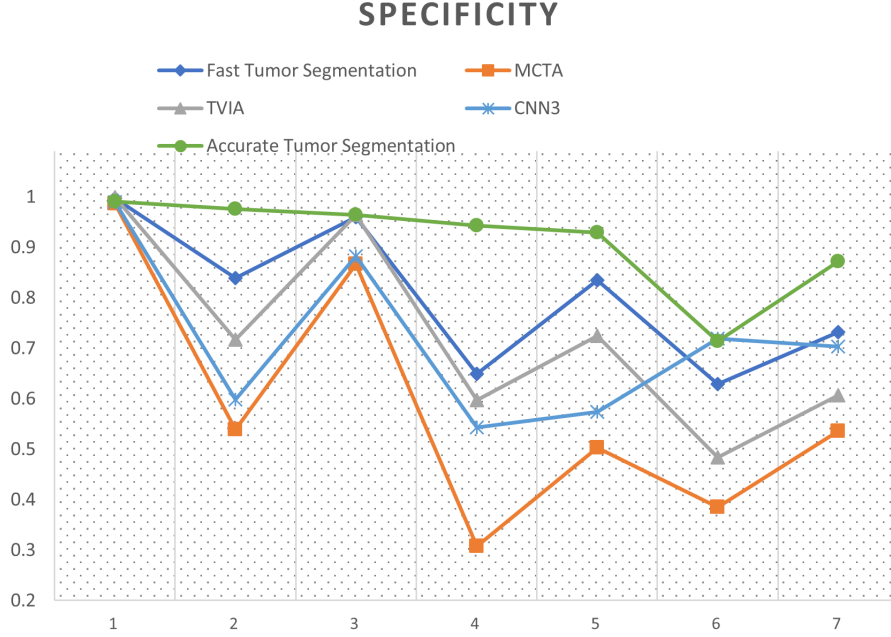
Figure 8: Specificity curve for tumor segmentation algorithms on healthy cases from the Warwick-Osaka dataset. The horizontal axis shows the number of healthy cases selected, and the vertical axis shows the specificity.

The specificity analysis demonstrates the effectiveness of the proposed ensemble strategy for accurate tumor segmentation, showing that a model relying on an agreement between topological and deep features outperforms all competing approaches.

5.2.5. Robustness Analysis

The aim of performing this experiment is to evaluate the robustness of the proposed tumor segmentation algorithms by training on the Warwick-UHCW dataset and testing on the Warwick-Osaka dataset which contains cancerous and healthy cases. One of the most challenging aspects of processing H&E WSIs is to overcome the non-standardized parameters involved in slide preparations like tissue sectioning, staining duration, dyes, and formalin concentration Veta et al. (2015). In order to become a part of the routine diagnosis, an automated tumor
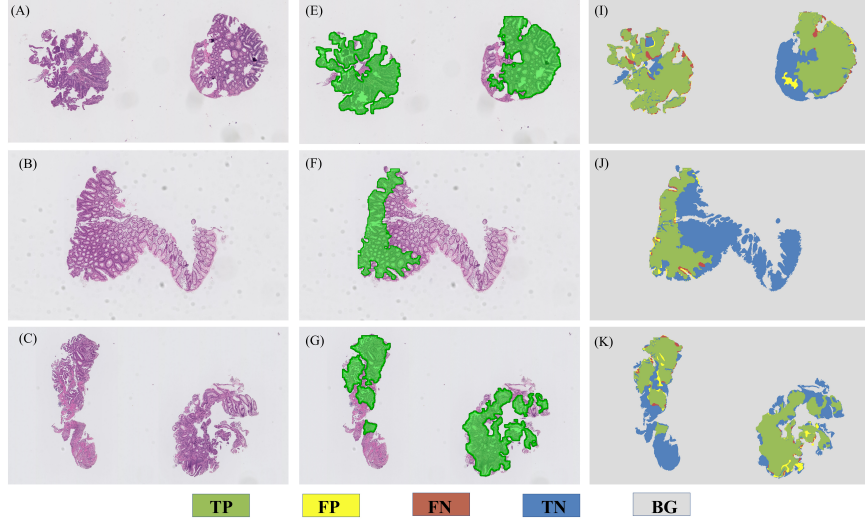
28

Figure 9: Robustness Analysis: results for the fast tumor segmentation on selected whole-slide images (WSI) from the Warwick-Osaka dataset. (A)-(C) input WSIs, (E)-(G) predicted tumor regions, (I)-(J) results for true positives (TP), false positive (FP), false negative (FN), and true negative (TN) regions.

segmentation algorithm should be resilient to these data variations. With that in mind, we perform this experiment by keeping the parameters of the algorithms fixed during the training and testing.

⁵¹⁵ Table 4 reports the results for robustness analysis and Figure 9 shows some qualitative results. The proposed methods attain the highest accuracy among different methods, demonstrating their robustness on strongly cross-validated data. The stain variability in both the datasets can be observed in Figure 6, which shows the selected exemplars from both the datasets. Regardless of stain ⁵²⁰ variations in both the datasets, the degree of connectivity among tumor and non-tumor class is notably distinguishable and that leads to better performance of the proposed methods. It is interesting to note that the PHP based fast tumor segmentation marginally outperformed the PHP+CNN based accurate tumor segmentation (0.3%). This bodes well for the potential generalizability ⁵²⁵ of fast tumor segmentation. Relatively smaller values of F1 measure for other

29

competing algorithms can also be noticed in Table 4. Excluding HyMap, an obvious indication for a considerable drop in F1-score is due to the precision measure. On the one hand, the competing algorithms perform well in predicting tumor patches but at the cost of a large number of false positives. In contrast, regardless of stain variations in the two datasets, the fast tumor segmentation algorithm maintains a balance between precision and recall. On the whole, there is a decrease in performance accuracy for all algorithms as compared to their performance when trained and tested on data from the same center, although the proposed algorithm still gives almost 8-10% higher F1-score than other competing algorithms. One potential strategy to increase the robustness of an underlying model is to train it on datasets from a number of different centers. Eventually, this will reduce the chances of overfitting and enables the learnable weights to optimize on a variety of datasets with varying contents and staining protocols.

Table 4: Results for robustness analysis of different tumor segmentation algorithms.

Method	Precision	Recall	F1-score
Fast Tumor Segmentation (PHP)	**0.772**	0.7890	**0.7804**
Accurate Tumor Segmentation (PHP+CNN)	0.7413	0.8172	0.7774
HyMap Khan et al. (2013)	0.6469	0.7228	0.6827
ConvNet CNN_3 Cruz-Roa et al. (2017)	0.6234	0.8167	0.7071
MCTA Kather et al. (2016)	0.5429	0.9586	0.6932
TVIA Turkki et al. (2015)	0.5334	**0.9747**	0.6895

5.2.6. Runtime Analysis

This section contains the runtime analysis of different tumor segmentation algorithms. Digitized WSIs are giga-pixel images, so fast tumor segmentation algorithms could play a crucial role in delivering 'real world' diagnostics. For all the algorithms, runtime analysis at the test stage is performed on a static machine having an 8-core processor with 3.1 GHz clock speed, 128 MB of memory

Table 5: Runtime analysis in milliseconds for different tumor segmentation algorithms

Method	Time
Fast Tumor Segmentation (PHP)	**28.8** ms
Accurate Tumor Segmentation (PHP+CNN)	151 ms
HyMap Khan et al. (2013)	931 ms
ConvNet CNN_3 Cruz-Roa et al. (2017)	97.1 ms
MCTA Kather et al. (2016)	228.12 ms
TVIA Turkki et al. (2015)	109 ms

and a GTX 1080 Ti graphical processing unit.

Table 5 reports the processing time for different algorithms on a patch of size $256 \times 256 \times 3$ at $20 \times$. The fast tumor segmentation is less computationally complex and an order of magnitude faster than competing algorithms, specifically ≈ 4.2 times faster than the CNN and ≈ 5.2 times faster than the accurate tumor segmentation. The algorithm for computing the PHPs is the foundation of fast tumor segmentation. The topological features (β_0, β_1) are computed by enumerating the connected components for a given filtered space. Consequently, the fast tumor segmentation algorithm is far less computationally complex than a multi-layer convolutional network.

6. Discussion and Conclusions

Visual examination of tissue slides under the microscope to analyze the morphological features is the 'gold standard' for cancer diagnosis Gurcan et al. (2009). This study aimed at improving the diagnostic workflow by introducing a novel automated tumor segmentation framework for colorectal cancer histology WSIs. Experimental results conducted on fairly challenging datasets collected from two independent pathology centers demonstrate the efficacy and generalizability of persistent homology in histopathological image analysis. In this work, we present novel topological signatures (PHPs) that, in some respects, resemble clinicians' approach for identification of tumor enrich areas. It is evident from

31

the performance that incorporating topological features and deep convolution features can enhance the overall performance of a CNN for the task of accurate tumor segmentation. The proposed framework was shown to work well on colorectal epithelial tumors of different histology grades.

Histological assessments are generally estimated visually and producing subjective measures for quantification of morphological features Webster et al. (2011). Computer-assisted image analysis requires precise annotations at high resolution to effectively train an underlying model. The inevitable fact is that the domain experts are generally not available for this laborious task of providing precise ground-truth at high resolutions. In such circumstances, the performance of a trained model may be affected by expert annotation that is carried out too rapidly or without close attention to detail, or by inexpert annotation. After careful consideration, we decided that this could be a relevant factor in both datasets, resulting in a difference in performance of the tumor segmentation algorithms (Table 2 & 3), regardless of stain and morphological variability. Nearly all the tumor segmentation algorithms experience difficulties in some of the benign epithelial and lymphocytic regions, especially where the intra-cellular region displays morphology that is similar to the cancer-distorted nucleoplasm.

WSI scanners are becoming more viable for routine analysis Snead et al. (2016), capable of producing hundreds of terabytes of data daily Cooper et al. (2012); Kamel (2011). The proposed framework offers a decent trade-off between speed and accuracy. With this change of paradigm, the fast tumor segmentation has enormous potential to overcome this ongoing challenge, reducing subjectivity and the pathologists' workload. It can be observed from the experimental results that careful selection of the exemplar patches can nearly obviate the need for retraining of our algorithm on a new dataset as shown in Table 4. One limitation of this work is that parameters like the number of exemplar patches and k (as in k-means) are empirically selected for this study and may need proper fine tuning, depending upon the data. The selection of a representative subset from a dataset is an active area of research and has several applications

32

in computer vision and natural language processing. The proposed method for selection of exemplar patches is presented as an application to deep convolution networks. In the literature, there exist dissimilarity based subset selection methods Elhamifar et al. (2016, 2012), but the selection of a dissimilarity measure is a non-trival task and computing a dissimilarity matrix for a large dataset is computationally expensive. Other approaches like loopy belief propagation Murphy et al. (1999) can handle a large dataset but does not offer any guarantee of convergence. In contrast, the proposed fast tumor segmentation method exploits the learned activation maps to deduce the representative patches and is less computationally expensive and more robust to outliers. It is also evident from the experimental results that the accurate tumor segmentation algorithm presents a simple yet meaningful way of combining our novel topological signatures with deep convolution features. An interesting direction could be to treat the homology profiles as temporal data and to explore such temporal information with recurrent networks. In this paper, we restrict ourselves to convolution networks which are better understood.

In conclusion, we presented an automated tumor segmentation framework for colorectal cancer histology WSIs based on persistent homology. The proposed framework is validated on two independent datasets, consisting of both malignant cases and healthy cases. Extensive comparative analysis demonstrated better than the state-of-the-art performance of the proposed algorithms. This study provides an insight into the topological persistence of an image and may constitute the first step towards interpretable incorporation of homology features in the domain of histopathology image analysis. The proposed homology profiles model the intrinsic phenomena of cell connectivity and may be applicable to other similar problems in the computational pathology.

References

Akbar S, Jordan L, Thompson AM, McKenna SJ. Tumor localization in tissue microarrays using rotation invariant superpixel pyramids. In: Biomedical

Imaging (ISBI), 2015 IEEE 12th International Symposium on. IEEE; 2015. p. 1292–5.

Altunbay D, Cigir C, Sokmensuer C, Gunduz-Demir C. Color graphs for automated cancer diagnosis and grading. IEEE Transactions on Biomedical Engineering 2010;57(3):665–74.

Baba AI, Câtoi C. Comparative oncology. Publishing House of the Romanian Academy Bucharest, 2007.

Bejnordi BE, Veta M, van Diest PJ, van Ginneken B, Karssemeijer N, Litjens G, van der Laak JA, Hermsen M, Manson QF, Balkenhol M, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. Jama 2017;318(22):2199–210.

Bendich P, Marron JS, Miller E, Pieloch A, Skwerer S. Persistent homology analysis of brain artery trees. The annals of applied statistics 2016;10(1):198.

Bhattacharya S, Ghrist R, Kumar V. Persistent homology for path planning in uncertain environments. IEEE Transactions on Robotics 2015;31(3):578–90.

Bianconi F, Álvarez-Larrán A, Fernández A. Discrimination between tumour epithelium and stroma via perception-based features. Neurocomputing 2015;154:119–26.

Cang Z, Wei GW. Integration of element specific persistent homology and machine learning for protein-ligand binding affinity prediction. International Journal for Numerical Methods in Biomedical Engineering 2017;.

Carlsson G. Topology and data. Bulletin of the American Mathematical Society 2009;46(2):255–308.

Chen H, Qi X, Yu L, Dou Q, Qin J, Heng PA. Dcan: Deep contour-aware networks for object instance segmentation from histology images. Medical image analysis 2017;36:135–46.

Clevert DA, Unterthiner T, Hochreiter S. Fast and accurate deep network learning by exponential linear units (elus). arXiv preprint arXiv:151107289 2015;.

Cooper LA, Carter AB, Farris AB, Wang F, Kong J, Gutman DA, Widener P, Pan TC, Cholleti SR, Sharma A, et al. Digital pathology: Data-intensive frontier in medical imaging. Proceedings of the IEEE 2012;100(4):991–1003.

Cruz-Roa A, Gilmore H, Basavanhally A, Feldman M, Ganesan S, Shih NN, Tomaszewski J, González FA, Madabhushi A. Accurate and reproducible invasive breast cancer detection in whole-slide images: A deep learning approach for quantifying tumor extent. Scientific Reports 2017;7:46450.

Curto C. What can topology tell us about the neural code? Bulletin of the American Mathematical Society 2017;54(1):63–78.

Dunaeva O, Edelsbrunner H, Lukyanov A, Machin M, Malkova D, Kuvaev R, Kashin S. The classification of endoscopy images with persistent homology. Pattern Recognition Letters 2016;83:13–22.

Dunne PD, McArt DG, Bradley CA, O'Reilly PG, Barrett HL, Cummins R, O'Grady T, Arthur K, Loughrey MB, Allen WL, et al. Challenging the cancer molecular stratification dogma: intratumoral heterogeneity undermines consensus molecular subtypes and potential diagnostic value in colorectal cancer. Clinical Cancer Research 2016;22(16):4095–104.

Edelsbrunner H, Harer J. Persistent homology-a survey. Contemporary mathematics 2008;453:257–82.

Elhamifar E, Sapiro G, Sastry SS. Dissimilarity-based sparse subset selection. IEEE transactions on pattern analysis and machine intelligence 2016;38(11):2182–97.

Elhamifar E, Sapiro G, Vidal R. Finding exemplars from pairwise dissimilarities via simultaneous sparse recovery. In: Advances in Neural Information Processing Systems. 2012. p. 19–27.

680 Ferlay J, Parkin D, Steliarova-Foucher E. Estimates of cancer incidence and mortality in europe in 2008. European journal of cancer 2010;46(4):765–81.

Ghrist R. Barcodes: the persistent topology of data. Bulletin of the American Mathematical Society 2008;45(1):61–75.

Gurcan MN, Boucheron LE, Can A, Madabhushi A, Rajpoot NM, Yener B.
685 Histopathological image analysis: A review. IEEE reviews in biomedical engineering 2009;2:147–71.

Kamel HM. Trends and challenges in pathology practice: choices and necessities. Sultan Qaboos University medical journal 2011;11(1):38.

Kather JN, Weis CA, Bianconi F, Melchers SM, Schad LR, Gaiser T, Marx
690 A, Zöllner FG. Multi-class texture analysis in colorectal cancer histology. Scientific reports 2016;6:27988.

Khan AM, El-Daly H, Simmons E, Rajpoot NM. Hymap: A hybrid magnitude-phase approach to unsupervised segmentation of tumor areas in breast cancer histology images. Journal of pathology informatics 2013;4(Suppl).

695 Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems. 2012. p. 1097–105.

Lamar-Leon J, Alonso-Baryolo R, Garcia-Reyes E, Gonzalez-Diaz R. Persistent homology-based gait recognition robust to upper body variations. In: Pattern
700 Recognition (ICPR), 2016 23rd International Conference on. IEEE; 2016. p. 1083–8.

Linder N, Konsti J, Turkki R, Rahtu E, Lundin M, Nordling S, Haglund C, Ahonen T, Pietikäinen M, Lundin J. Identification of tumor epithelium and stroma in tissue microarrays using texture analysis. Diagnostic pathology
705 2012;7(1):22.

Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, Hulsbergen-Van De Kaa C, Bult P, Van Ginneken B, Van Der Laak J. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Scientific reports 2016;6:26286.

Murphy KP, Weiss Y, Jordan MI. Loopy belief propagation for approximate inference: An empirical study. In: Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc.; 1999. p. 467–75.

Nakane K, Takiyama A, Mori S, Matsuura N. Homology-based method for detecting regions of interest in colonic digital images. Diagnostic pathology 2015;10(1):36.

Nakane K, Tsuchihashi Y, Matsuura N. A simple mathematical model utilizing topological invariants for automatic detection of tumor areas in digital tissue images. In: Diagnostic pathology. BioMed Central; volume 8; 2013. p. S27.

Otter N, Porter MA, Tillmann U, Grindrod P, Harrington HA. A roadmap for the computation of persistent homology. EPJ Data Science 2017;6(1):17.

Pokorny FT, Hawasly M, Ramamoorthy S. Topological trajectory classification with filtrations of simplicial complexes and persistent homology. The International Journal of Robotics Research 2016;35(1-3):204–23.

Qaiser T, Mukherjee A, Reddy PB C, Munugoti SD, Tallam V, Pitkaho T, Lehtimki T, Naughton T, Berseth M, Pedraza A, Mukundan R, Smith M, Bhalerao A, Rodner E, Simon M, Denzler J, Huang CH, Bueno G, Snead D, Ellis IO, Ilyas M, Rajpoot N. Her2 challenge contest: a detailed assessment of automated her2 scoring algorithms in whole slide images of breast cancer tissues. Histopathology 2018;72(2):227–38. URL: http://dx.doi.org/10.1111/his.13333. doi:10.1111/his.13333.

Qaiser T, Sirinukunwattana K, Nakane K, Tsang YW, Epstein D, Rajpoot

N. Persistent homology for fast tumor segmentation in whole slide histology images. Procedia Computer Science 2016;90:119–24.

Qaiser T, Tsang YW, Epstein D, Rajpoot N. Tumor segmentation in whole slide images using persistent homology and deep convolutional features. In: Annual Conference on Medical Image Understanding and Analysis. Springer; 2017. p. 320–9.

Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2015. p. 234–41.

Ruifrok AC, Johnston DA, et al. Quantification of histochemical staining by color deconvolution. Analytical and quantitative cytology and histology 2001;23(4):291–9.

Sabour S, Frosst N, Hinton GE. Dynamic routing between capsules. In: Advances in Neural Information Processing Systems. 2017. p. 3859–69.

Sirinukunwattana K, Pluim JP, Chen H, Qi X, Heng PA, Guo YB, Wang LY, Matuszewski BJ, Bruni E, Sanchez U, et al. Gland segmentation in colon histology images: The glas challenge contest. Medical image analysis 2017;35:489–502.

Smits AJ, Kummer JA, De Bruin PC, Bol M, Van Den Tweel JG, Seldenrijk KA, Willems SM, Offerhaus GJA, De Weger RA, Van Diest PJ, et al. The estimation of tumor cell percentage for molecular testing by pathologists is not accurate. Modern Pathology 2014;27(2):168–74.

Snead DR, Tsang YW, Meskiri A, Kimani PK, Crossman R, Rajpoot NM, Blessing E, Chen K, Gopalakrishnan K, Matthews P, et al. Validation of digital pathology imaging for primary histopathological diagnosis. Histopathology 2016;68(7):1063–72.

Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA: a cancer journal for clinicians 2015;65(2):87–108.

Turkki R, Linder N, Holopainen T, Wang Y, Grote A, Lundin M, Alitalo K, Lundin J. Assessment of tumour viability in human lung cancer xenografts with texture-based image analysis. Journal of clinical pathology 2015;:jclinpath–2015.

Turner K, Mukherjee S, Boyer DM. Persistent homology transform for modeling shapes and surfaces. Information and Inference: A Journal of the IMA 2014;3(4):310–44.

Upton G, Cook I. Understanding statistics. Oxford University Press, 1996.

Veta M, Van Diest PJ, Willems SM, Wang H, Madabhushi A, Cruz-Roa A, Gonzalez F, Larsen AB, Vestergaard JS, Dahl AB, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images. Medical image analysis 2015;20(1):237–48.

Viray H, Li K, Long TA, Vasalos P, Bridge JA, Jennings LJ, Halling KC, Hameed M, Rimm DL. A prospective, multi-institutional diagnostic trial to determine pathologist accuracy in estimation of percentage of malignant cells. Archives of Pathology and Laboratory Medicine 2013;137(11):1545–9.

Webster JD, Simpson ER, Michalowski AM, Hoover SB, Simpson RM. Quantifying histological features of cancer biospecimens for biobanking quality assurance using automated morphometric pattern recognition image analysis algorithms. Journal of biomolecular techniques: JBT 2011;22(3):108.

Wubie BA. Clustering Survival Data Using Random Forest and Persistent Homology. Ph.D. thesis; University of Alberta; 2016.

Xia K, Wei GW. Persistent homology analysis of protein structure, flexibility, and folding. International journal for numerical methods in biomedical engineering 2014;30(8):814–44.

Xu Y, Zhu JY, Chang E, Tu Z. Multiple clustered instance learning for histopathology cancer image classification, segmentation and clustering. In:

39

Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE; 2012. p. 964–71.

Xu Y, Zhu JY, Eric I, Chang C, Lai M, Tu Z. Weakly supervised histopathology cancer image segmentation and classification. Medical image analysis 2014;18(3):591–604.

Zagoruyko S, Komodakis N. Paying more attention to attention: Improving the performance of convolutional neural networks via attention transfer. arXiv preprint arXiv:161203928 2016;.

Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: European conference on computer vision. Springer; 2014. p. 818–33.

Zhu X. Persistent homology: An introduction and a new text representation for natural language processing. In: IJCAI. 2013. p. 1953–9.