

A Cascade-Learning Approach for Automated Segmentation of Tumor Epithelium in Colorectal Cancer

Mohammed M. Abdelsamea^{a,d}, Alain Pitiot^b, Ruta Barbora Grineviciute^{e,f},
Justinas Besusparis^{e,f}, Arvydas Laurinavicius^{e,f}, Mohammad Ilyas^c

^a*School of Computer Science, University of Nottingham, Nottingham, UK.*

^b*Laboratory of Image & Data Analysis, Ilixa Ltd., Nottingham, UK.*

^c*Nottingham Molecular Pathology Node and Academic Unit of Pathology, Division of Cancer and Stem Cells, University of Nottingham, Queen's Medical Centre, Nottingham, UK.*

^d*Mathematics Department, University of Assiut, Assiut, Egypt.*

^e*Faculty of Medicine, M.K.Ciurlionio 21, LT-03101 Vilnius University, Vilnius, Lithuania.*

^f*National Center of Pathology, affiliate of Vilnius University Hospital Santaros Clinics, P. Baublio 5, LT-08406 Vilnius, Lithuania.*

Abstract

Automated segmentation of tumor epithelial tissue from histological images is a fundamental aspiration of digital pathology to improve biomarker assessment and tissue diagnosis. Accurate tumor segmentation is an important step in many automated digital image analysis applications to be used in clinical practice. In particular, segmentation of tumor, non-tumor epithelium and stromal tissue compartments on immunohistochemistry images presents a challenge. Many artifacts, such as staining and/or illumination variations, can confound image analysis. In this paper, we propose a cascade-learning approach which can diminish the impact

Email addresses: m.abdelsamea@aun.edu.eg (Mohammed M. Abdelsamea),
alain.pitiot@ilixa.coms (Alain Pitiot), RutaBarbora.Grineviciute@vpc.lt
(Ruta Barbora Grineviciute), Justinas.Besusparis@vpc.lt (Justinas Besusparis),
Arvydas.Laurinavicius@vpc.lt (Arvydas Laurinavicius),
mohammad.ilyas@nottingham.ac.uk (Mohammad Ilyas)

of these artifacts. It consists of (a) a set of novel invariant features that encodes meaningful information about the appearance and shape of the region of interest and (b) a novel level set formulation where contour evolution is driven by a probabilistic model of the appearance of the region (based on fuzzy c-means). The merit of our approach is that it exploits both appearance and shape information and combines them in the tissue classification framework. We evaluate the performance of our approach on the segmentation of tumor epithelium in colorectal cancer. The experimental results show that our approach is robust to staining differences, additive noise, intensity inhomogeneities, and can cope with a limited number of training samples, when compared to the state-of-the-art tumor epithelial segmentation methods.

Keywords:

histology; colorectal cancer; tissue classification; machine learning; fuzzy pressure force; axis of least inertia; feature representation; neural network.

1. Introduction

Histopathology is the microscopic examination of thin sections of potentially diseased tissue, which have been fixed onto glass slides and stained to reveal particular structural or functional detail. With the availability of whole slide scanners, digitized images of those glass slides can be obtained, thereby making tissue histopathology amenable to the application of image analysis algorithms. By automating repetitive tasks, improving the accuracy and precision of feature measurements (such as cell count or gland area) and decreasing subjectivity, image analysis could help to improve tissue analysis for both biomarker assessment and tissue diagnosis. Furthermore, mining of the digital data could help identify new

diagnostic features for a variety of disease categories.

The demand for precise segmentation tool increases along with accumulating evidence on the diagnostic value that is produced by analyzing the spatial distribution of biomarkers in tumor tissue. Diagnostic tool Immunoscore® introduced solid evidence that quantity and spatial distribution of immune cells through the tumor is a reliable prognostic and predictive biomarker. Studies focusing on colorectal cancer note, that numerical expression characterizing CD3 + and CD8 + cells, calculated using Immunoscore® is superior to microsatellite stability status and add additional value to existing TNM system. Standardized way of evaluation is required in order to represent complex interaction of immune cells and tumor tissue. Conventional approach using light microscopy and manually estimating the number of positive stained immune cells in tumor stroma and epithelium is time consuming and lacks objectivity due to intra observer variability. Most cases precise percentage or ratio is required for clinical decision making. Automated segmentation of tissue into sections representing tumor epithelium followed by recognition of positive cells is a way to refine results and eliminate subjectivity of manual evaluation (Ogino et al., 2011; Mlecnik et al., 2016; Galon et al., 2012, 2014).

The automated segmentation of individual tissue compartments is a pivotal component of diagnostic histopathology. In cancer diagnosis, which is the focus of this manuscript, distinguishing epithelium from stroma and discriminating between normal and tumor epithelium are a necessary first step to the subsequent analysis of the features of each compartment (e.g. nuclear size or expression of biomarkers), which lead to the diagnosis (Gurcan et al., 2009). Tumor epithelium classification is especially challenging due to the very high morphological

and staining variability of the regions of interest (*ROI*). Most of the existing approaches were based on appearance features that have been designed in a way to simulate the visual perception of human pathologist in investigating tissue samples stained for a particular biomarker, previously proposed approaches are discussed in more details in Section 2. Indeed depending on the degree of differentiation of the tumor, it may look very similar to the normal tissue or it may be very different, it may stain positively or negatively with certain stains and biomarkers. Further sources of variation may lie in pre-analytic steps (such as tissue fixation), laboratory-dependent variations (such as thickness of tissue section and variation in tissue stains and scanner choice (Madabhushi and Lee, 2016)). Finally, scanned images are usually very large images (typically several giga-pixels), which makes it important to design computationally efficient approaches in order to produce results within an acceptable “diagnostic” time frame.

In this work, we aimed to develop an approach capable of dealing with the morphological and staining variability of epithelium in order to allow reliable segmentation of tumor epithelium, normal epithelium and stroma. We used colorectal cancer (*CRC*) as an exemplar. Fig. 1 shows some normal and tumor areas extracted from images of colorectal tissues that have undergone immunohistochemical staining for different biomarkers (see section 4.1 for details about the images). Note how the nuclei are regularly distributed in the normal epithelium (Fig. 1a) while they are randomly distributed in the tumor regions (Fig. 1b).

The proposed system has been designed to provide a generic framework for tumor epithelium segmentation in histology. It is based on learning the appearance and shape models of normal and tumor epithelial regions, provide histologically meaningful descriptors to deal with various images stained for different biomark-

ers. Additionally, our system provides a new level set formulation which allows extracting epithelial area from histology image in an unsupervised way by incorporating prior knowledge about the appearance model of the epithelium into the segmentation framework. Eventually, we used two self-organized maps for tissue classification to improve the robustness and efficiency of our system.

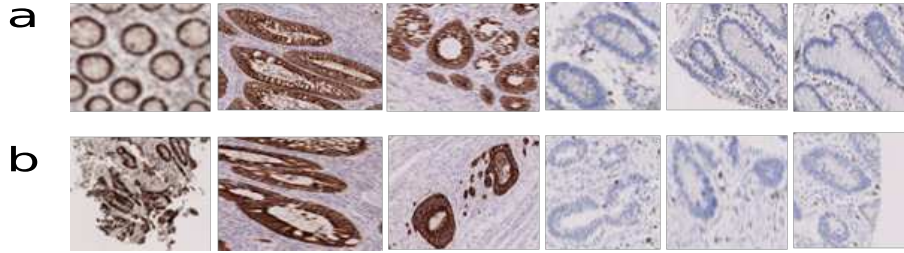


Figure 1: Positively-stained (first 3 columns) and negatively-stained (last 3 columns) normal (a) and tumorous (b) epithelium.

The rest of the paper is organized as follows. In Section 2, we briefly review recent approaches to tissue segmentation. We detail our processing pipeline in Section 3, before presenting experimental results on both real and synthetic examples in Section 4, and discussing advantages, limitations and future work in Section 5.

2. Related work

Recent histological tissue segmentation techniques can be split into two groups based on the nature of the features upon which they rely. In the first group we find methods based on feature engineering, with features linked either to pixel-/patch-level information (Linder et al., 2012; Kather et al., 2016; Mattfeldt et al., 2013; Doyle et al., 2012; Shu et al., 2016) or to histological primitives (Chang et al.,

2013; Bianconi et al., 2015; Chang et al., 2013; Sirinukunwattana et al., 2015). Because pixel-/patch- level features do not encode information specific to the biology of the *ROI*, the performance of those approaches is inherently dependent upon the characteristics (e.g. representatively, exhaustivity) of the learning set, unlike that of histological primitive-level features, which can provide more generic and robust solutions by encoding biologically meaningful information (e.g., size of nuclei, shape of glands). The second group of methods is based on feature learning (Zhou et al., 2014; Sirinukunwattana et al., 2016, 2015), a technique which exploits the availability of large, a priori annotated learning sets to construct discriminating features by itself. These methods often require extensive computational resources and, much like pixel-/patch- level techniques, are heavily dependent upon the characteristics of the learning set (Madabhushi and Lee, 2016; Jeong et al., 2010; Nayak et al., 2013).

A number of feature representations have been proposed in the literature. Intensity-based methods (Eramian et al., 2011; Mahmoud-Ghoneim, 2011; Ozdemir et al., 2012; Samsi et al., 2012; Kather et al., 2016; Bianconi et al., 2015) have been used to highlight the global structures/appearance of the stained tissues. Texture-based features (Datar et al., 2008; Wang et al., 2007; Caicedo et al., 2011; Cruz-Roa et al., 2011; Krishnan et al., 2011, 2012; Linder et al., 2012) are the most commonly used features due to their ability to describe the spatial arrangement of color/ intensities in the images, which reflect the local appearance of regions. Topology-based features (Chang et al., 2013; Tambasco et al., 2009; Ramesh et al., 2012; Wang et al., 2007) deal with the spatial arrangement of the regions (i.e., segmented regions) presented in the image. Lastly, feature learning-based methods (Zhong et al., 2017; Zhou et al., 2014; Sirinukunwattana et al.,

2016, 2015; Nayak et al., 2013) have been used to capture intrinsic morphometric patterns in stained tissues themselves.

For instance, Wang et al. (2007) used a patch-based hybrid method to segment the squamous epithelium in H&E-stained cervical tissue images. It relied on support vector machine to classify pixel patches based on texture information extracted at progressively higher levels of resolution. The results were overall good though the approach was computationally intensive and sometimes yielded over-segmentation. Datar et al. (2008) used Hierarchical Self-organizing Maps (*HSOM*) to segment immunohistochemistry images into four classes: glands, epithelium, stroma and nuclei. For each pixel a feature vector describing its texture and color was calculated and fed to the previously trained *HSOM*. Eramian et al. (2011) used colour histograms to segment the epithelium in H&E images, with an approach based on graph cuts. Linder et al. (2012) used a combination of texture features and contrast measure to differentiate, in an automated way, tumor epithelium from stroma using *SVM* in a large set of colorectal cancer. The designed features were significantly robust to illumination changes as compared to other common texture features (e.g., Haralick features and Gabor filters). Kather et al. (2016) compared the state-of-the-art texture features and classifiers on a collected dataset of all related types of tissues within colorectal cancer samples to discriminate between several tissue types in colorectal cancer. Results showed that global lower-order texture features performed well in segmenting tumor epithelium from different stromal compartments. Note that since those approaches often result in regions with discontinuities a post-processing phase is usually required to smooth the segmentation map. Chang et al. (2013) used a set of Cellular Morphometric Features (*CMFs*) as a way to encode biologically meaningful information to clas-

sify tumor tissue. A comprehensive study was also done in Zhong et al. (2017) to compare tissue architectures on two large series of tumor, i.e. Glioblastoma Multiforme (*GBM*) and Kidney Clear Cell Carcinoma (*KCCC*) and concluded that: (a) *CMF* and features based on feature learning (i.e., deep learning approaches) are outperforming other texture-/color-based features; (b) Cellular saliency (e.g., nuclei in tissue histology sections) incorporation weakens the performance for models built upon pixel-/patch-level features; and (c) Sparse feature encoder significantly improves the classification performance for models built upon *CMF*. Bianconi et al. (2015) used a few sets of appearance features based on visual perception to discriminate tumor epithelium from stroma in colorectal cancer. These features are related to human perception which can be interpreted by a pathologist in a meaningful way.

3. Proposed method

The proposed image analysis pipeline consists of 3 stages (see Fig. 2). In the first stage, we segment the epithelium using a novel approach based on fuzzy c-means and a signed pressure force level set: the Fuzzy Signed Pressure Force approach (*FSPF*), with the fuzzy c-means used to control the evolution of the contour over time. In the second stage, we calculate for each independent components from the produced epithelium map a set of novel morphometric features based on the Axis of Least Inertia (*ALI*), and a set of novel appearance descriptors based on Weighted Average Intensity profiles (*WAI*), which are invariant to staining. Finally, we distinguish between tumor and normal epithelium using two Self Organizing Maps (*SOMs*) of those two feature sets, previously trained on manually labelled images.

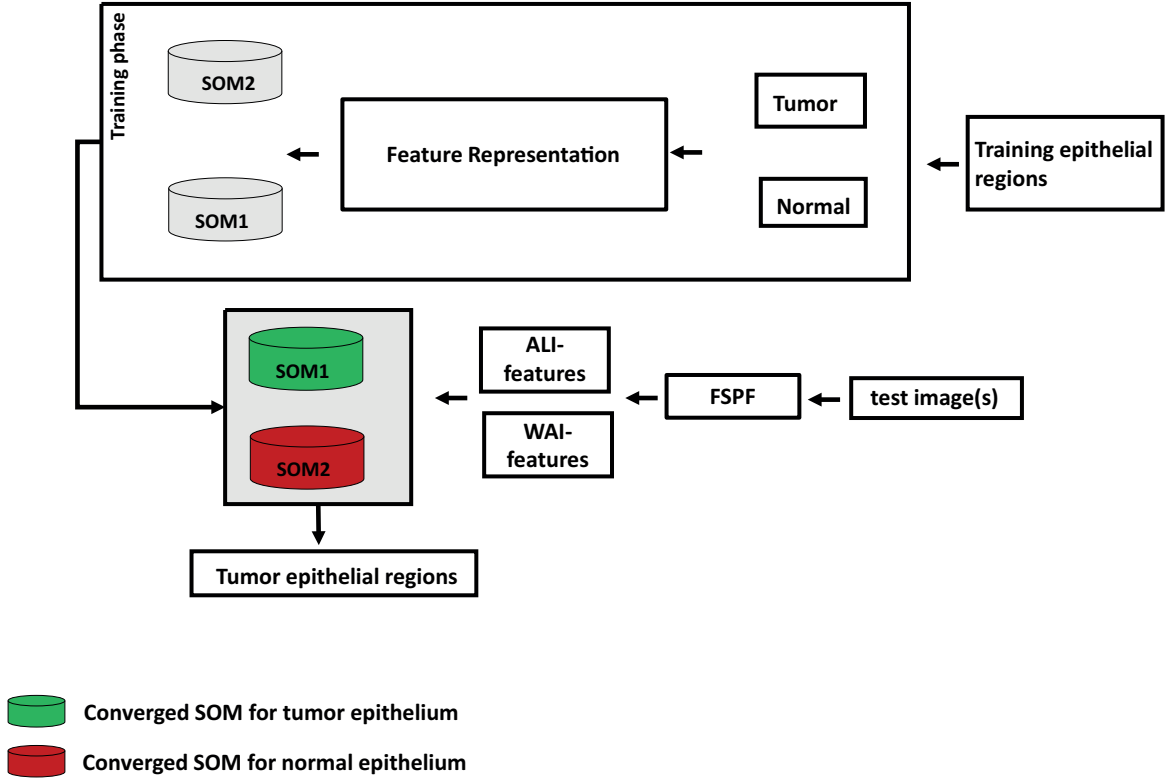


Figure 2: Graphical depiction of the proposed approach.

In the following, we discuss in detail each component of the proposed system, providing deep insights into the major contributions of the system.

3.1. Epithelium segmentation

We propose a novel machine-learning approach to epithelium segmentation capable of dealing with a wide variety of stains, and with both positively- and negatively- stained epithelium. It is based on a variant of the Signed Pressure Force method, and incorporates information about the epithelial area into the segmentation framework.

3.1.1. Signed Pressure Force (SPF) method

The *SPF* method (Zhang et al., 2010) combines the advantages of the Chan-Vese “active contour without edges” model ($C - V$), which is region-based (Chan and Vese, 2001), with those of the Geodesic Active Contour models (*GAC*), which is edge-based (Caselles et al., 1997). It statistically models the information inside and outside the contour to construct a region-based Signed Pressure Force (*SPF*) function. The *SPF* function is so-called because it tends to make the contour C shrink when it is outside the region of interest, and expand otherwise based on the signs of the forces inside and outside the contour. The evolution of the contour is controlled by the following *PDE*:

$$\frac{\partial \phi}{\partial t} = spf(I(x)) \cdot \alpha \|\nabla \phi\|, \quad (1)$$

where $I(x)$ is an input image, α is a balloon force parameter that modulates the propagation speed of the contour, ϕ is a level set function (Zhang et al., 2010), and the *SPF* function spf is defined as

$$spf(I(x)) := \frac{I(x) - \frac{c^+(C) + c^-(C)}{2}}{\max_{x \in \Omega} \left(\left| I(x) - \frac{c^+(C) + c^-(C)}{2} \right| \right)}, \quad (2)$$

where $c^+(C)$ and $c^-(C)$ are defined as the global average intensity inside and outside the contour, respectively.

SPF showed overall greater accuracy and lower computational time with respect to classical variational level set methods, essentially thanks to its ability to escape local minima and converge faster (Abdelsamea and Tsiftaris, 2013). However, as a global level set method which does not exploit prior knowledge *SPF* cannot successfully segment regions with overlapping intensity distributions. Moreover, *SPF* is not able to deal with the changes in the intensity distribution of the region to be segmented. Here, we propose to incorporate prior

knowledge in the SPF level set formulation to cope with the inherent complexity of histology images.

3.1.2. Fuzzy Signed Pressure Force method

The main idea of the $FSPF$ is to incorporate a probabilistic model of the epithelium into the SPF level sets formulation. Given a contour C , x the pixel location in the image $I(x)$, our proposed energy functional equation is defined as

$$E_{FSPF}(C, c^+(x), c^-(x)) := \int_{in(C)} \lambda^+ e^+(x) dx + \int_{out(C)} \lambda^- e^-(x) dx, \quad (3)$$

where

$$e^+(x) := |I(x) * P_{EPI}(x) - c^+(I * P_{EPI}(x))|^2, \quad (4)$$

$$e^-(x) := |I(x) * P_{EPI}(x) - c^-(I * P_{EPI}(x))|^2, \quad (5)$$

where the positive constants λ^- and λ^+ are the weight of the interior, e^- , and exterior, e^+ , terms, $c^-(x)$ and $c^+(x)$ are the local mean inside and outside the contour, and $P_{EPI}(x)$ is the local Epithelium probabilistic model that guides the movement of the contour. The latter is the optimal membership function results from the minimization of the following objective function:

$$E_{EPI}(x) := \sum_{j=1}^c P_j^2(x) \|I_{local}(x) - v_j(x)\|^2 \quad (6)$$

where c is the number of classes, $I_{local}(x)$ is the local intensity approximation function calculated at each pixel location x , $P_j^2(x)$ are the probabilistic model of

each class (e.g., $P_{EPI}(x) = P_3^2(x)$). $v_j(x)$ are the intensity centroids of the classes and are defined as follows:

$$I_{local}(x) := \frac{\int g_{\sigma''}(x-y) \cdot I(y) dy}{\int g_{\sigma''}(x-y) dy} \quad (7)$$

$$P_j(x) := \frac{1}{\sum_{k=1}^c \frac{\|I_{local}(x) - v_j(x)\|^2}{\|I_{local}(x) - v_k(x)\|^2}}, \quad (8)$$

$$v_j(x) := \frac{\sum_x P_j(x) \cdot I_{local}(x)}{\sum_x P_j(x)} \quad (9)$$

with $g_{\sigma''}$ a Gaussian smoothing kernel of standard deviation σ'' .

Following Zhang et al. (2010), we obtain:

$$fspf(I(x), EPI) = \alpha(P_{I(x), EPI}) \cdot spf(I(x), P_{EPI}), \quad (10)$$

where

$$\begin{cases} \alpha(P_{I(x), EPI}) = I_{min} + \frac{(I_{max} - I_{min})(P_{EPI_{min}} - P_{EPI}(x))}{P_{EPI_{min}} - P_{EPI_{max}}}, \\ spf(I(x), P_{EPI}) = \text{sign}(c^+ - c^-) \cdot \text{sign}(I * P_{EPI}(x) - \frac{c^{+2} - c^{-2}}{2c^+ - 2c^-}) \end{cases} \quad (11)$$

$\alpha(P_{I(x), EPI})$ is responsible for creating the external pressure forces for the evolution of the contour and $spf(I(x), P_{EPI})$ modulates the signs of the pressure forces.

Finally, the evolution of the contour can be described by the following *PDE*:

$$\frac{\partial \phi}{\partial t} = fspf(I(x), EPI) \cdot \alpha \|\nabla \phi\|, \quad (12)$$

Compare to *SPF* approach, *FSPF* is able to cope with i) intensity inhomogeneities; and ii) high visual variability of different tissues in the image. This is by integrating the probabilistic models of different tissues/classes (i.e., $P_j(x)$), which are calculated based on local information (i.e., $I_{local}(x)$), into the segmentation framework (Equation 12). For epithelium segmentation, the probabilistic model of epithelium (i.e., $P_{EPI}(x)$) is used to control the contour evolution for an effective segmentation. Moreover, one can benefit from this approach, as a multi-class object segmentation, by considering the relative probabilistic model to target a specific tissue (e.g., stroma only).

3.2. Histological image feature representation

We propose a novel set of invariant shape and appearance features, specifically designed for histology *ROIs*.

3.2.1. Axis of Least Inertia (*ALI*)-based features

The *ALI* (Yang et al., 2008) is the line that minimizes the integral of the square of the distances between each point in the region and the line. The *ALI* passes through the centroid of the region and is orientation independent.

For each connected component of the extracted epithelium we first estimate the *ALI* and then project each point, I , onto it. The Euclidean distance between the projected points furthest from each other is the length of the *ALI*. The *ALI* splits the ROI, $\{I(x)\}_x$, into left and right domains, Ω_{left} and Ω_{right} respectively (see Fig. 3).

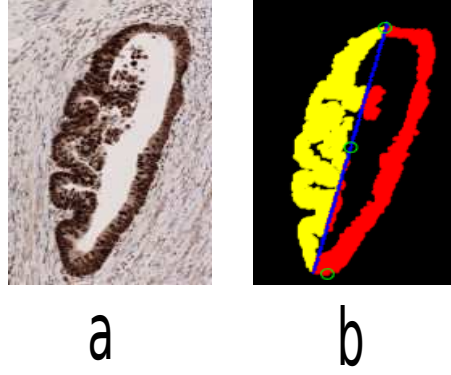


Figure 3: tumorous epithelial area (a) and its binary segmentation obtained by *FSPF* (b) with the estimated *ALI* in blue, and the left (yellow) and right (red) region domains.

We propose the following ALI-based descriptors:

$$ALI_{\Gamma} = \sqrt{\frac{\sum_n (\Gamma_{left} - \Gamma_{right})^2}{n}} \quad (13)$$

where, $\Gamma_{left}(\cdot) = \|I_{left}(\cdot) - P_{left}(\cdot)\|$, $\Gamma_{right}(\cdot) = \|I_{right}(\cdot) - P_{right}(\cdot)\|$, and n is the number of pixels in the region.

$$ALI_{\zeta} = \sqrt{\frac{\sum_n (\zeta_{left} - \zeta_{right})^2}{n}} \quad (14)$$

where, $\zeta_{left}(\cdot) = \|Centroid - P_{left}(\cdot)\|$ and $\zeta_{right} = \|Centroid - P_{right}(\cdot)\|$.

$$ALI_{\Psi} = \sqrt{\frac{\sum_n (\Psi_{left} - \Psi_{right})^2}{n}} \quad (15)$$

where, $\Psi_{left}(\cdot) = \|Centroid - I_{left}(\cdot)\|$ and $\Psi_{right} = \|Centroid - I_{right}(\cdot)\|$.

All three descriptors quantify the similarity of the left and right-hand domains: ALI_{Γ} by considering the distances to the *ALI*, ALI_{Ψ} by considering the distances

to the centroid of the region, and ALI_ζ by considering the distributions of the projected points.

To improve the robustness of the proposed segmentation system, we also proposed a set of appearance descriptors based on Weighted Average Intensity (WAI), which are sensitive to the size of the lumen region inside the glands (e.g. holes) and the distribution of the nuclei in the gland itself by considering the spatial dependencies between pixels in a neighbourhood.

3.2.2. Weighted Average Intensity (WAI)-based features

The weighted average intensity of a region at pixel x , $WAI(x)$, is defined as follows:

$$WAI(x) := \frac{\int_y g'_{\sigma'}(x-y)I(y)dy}{\int_y g'_{\sigma'}(x-y)dy} \quad (16)$$

where $g'_{\sigma'}$ is a Gaussian kernel function, $\int_{\mathbb{R}^2} g'_{\sigma(x)} dx = 1$.

We propose the following WAI -based descriptors:

$$WAI_1 = \frac{1}{n} \sum_x WAI(x) \quad (17)$$

$$WAI_2 = \begin{cases} WAI(x)_{\frac{n-1}{2}} & \text{when } n \text{ is odd} \\ \frac{1}{2}(WAI(x)_{\frac{n}{2}} + WAI(x)_{\frac{n}{2}+1}) & \text{when } n \text{ is even} \end{cases} \quad (18)$$

$$WAI_3 = \frac{1}{n} \sum_x (WAI(x) - WAI_1) \quad (19)$$

$$WAI_4 = \frac{\frac{1}{n} \sum_x (WAI(x) - WAI_1)^3}{(\sqrt{\frac{1}{n} \sum_x (WAI(x) - WAI_1)^2})^3} \quad (20)$$

$$WAI_5 = \frac{\sum_x (WAI(x) - WAI_1)^4}{\frac{1}{n}(\sum_x (WAI(x) - WAI_1)^2)^2} \quad (21)$$

WAI_1 encodes the symmetry of $WAI(\cdot)$; WAI_2 encodes the central tendency of $WAI(\cdot)$; WAI_3 encodes the dispersion/variability of $WAI(\cdot)$ distribution; WAI_4 encodes the frequency of extreme low and high values of $WAI(\cdot)$; and WAI_5 encodes the peakedness of $WAI(\cdot)$.

3.3. Tissue classification

Given a collection of training samples from normal and tumor cases, we set up two *SOMs* (Kohonen, 1990; Neagoe and Ropot, 2002) to learn the appearance and shape of the tumor and normal epithelium, independently. This is performed by topologically arranging the neurons of each *SOM* in the output layer on the basis of their weights using the following self-organized learning rule

$$w_n(t+1) := w_n(t) + \eta(t)h_{bn}(t)[x(t) - w_n(t)], \quad (22)$$

where $t = 0, 1, 2, 3, \dots$ is a time index, $w_n(t)$ is the weight of the neuron n at time t , $x(t)$ is the feature vector that contains both *ALI*-based and *WAI*-based features at time t , $\eta(t)$ is a learning rate, and $h_{bn}(t)$ is the neighbourhood kernel at time t of the neuron n around a specific neuron b , called best-matching unit (*BMU*). More precisely, in each *SOM* and at the time t , an input vector $x(t) \in \mathbf{R}^D$ is fed to the network, then the neurons in the map compete with each other to be the winner neuron b , which is then chosen as the one whose weight $w_b(t)$ is the closest to the input vector $x(t)$ in terms of a similarity measure, the Euclidean distance in our case. We get $\|x(t) - w_b(t)\|_2 := \min_n \|x(t) - w_n(t)\|_2$, where n varies in the set of neurons of the map.

Once the learning phase is completed, previously unseen epithelium will be assigned a class label using the minimum quantization error as per (Neagoe and Ropot, 2002).

Here the *SOMs* act as a dimensionality reduction approach. We could have used other classification methods such as Radial basis function *SVM* (*RBF – SVM*) or *k*-Nearest Neighbours (Bianconi et al., 2015)) but the topological preservation properties of *SOMs* make them more robust to noise and reduce the processing time.

4. Experimental results

4.1. Data and experiment

The image datasets used in this work were all from colorectal cancer cases and consisted of:

- 134 images containing both normal and tumor regions extracted from the high resolution scans of 15 Tissue Microarrays (*TMA*s) stained for the T-lymphocytes biomarker *CD3*, obtained using a Hamamatsu NanoZoomer with 40X magnification ($0.5\ \mu m$ resolution) and downsampled to approximately 400 x 400 pixels (2.5X magnification), for computational efficiency.
- 14 Whole Slide Images (*WSIs*) stained for four different biomarkers: MSH2 (Fig. 5 (b), Fig. 7 (c,d), and Fig. 7 (j)), AE1AE3 (Fig. 7 (g,h,i,m)), BerEP4 (Fig. 7 (f,k,l), and Fig. 7 (b)), and MLH1 (Fig. 5 (c), and Fig. 7 (a)), digitized in the same manner and downsampled to approximately 5000x5000 pixels (2.5X magnification).

All sections were obtained from previous clinical studies. All tissue sections underwent immunohistochemical staining using different protocols (optimized for each specific antibody) but all used 3'3-Diaminobenzidine as a chromogen and all were counterstained with Mayer's Haematoxylin. The staining pattern of the antibodies was as follows: CD3 – membranous staining of T-lymphocytes, BerEP4 – cytoplasmic staining of colonic epithelium (both tumor and non-tumor), AE1/3 – cytoplasmic staining of colonic epithelium (both tumor and non-tumor), MLH1 – strong staining of tumor nuclei and weak staining of nuclei of normal epithelium and stroma, MSH2 – strong staining of tumor nuclei and weak staining of nuclei of normal epithelium and stroma. The large array of antibodies with the varying staining patterns and the combination of *TMA/WSI* were chosen in order to ensure that the algorithms developed were not influenced by the immunohistochemical test being performed or the nature of the tissue section. Access to tissues and ethics approval were granted by Nottingham Health Sciences Biobank which has approval as an IRB from North West - Greater Manchester Central Research Ethics Committee REC reference: 15/NW/0685.

In order to compare the performance of our approach against that of competing methods from the literature, we implemented the following algorithms in Matlab on the basis of the information from the associated papers: the *CMF*-based algorithm of Chang et al. (2013), and the Lower-order Histogram (*LH*)-based and Higher-order Histogram (*HH*)-based algorithms of Kather et al. (2016). We also ran the publicly available code of the following methods: the Local binary patterns (*LBP*)-based algorithm of Linder et al. (2012), the Perception-based algorithm of Bianconi et al. (2015), the *SPF* method of Zhang et al. (2010), and the *RBF-SVM* method of Bianconi et al. (2015).

We adopted Precision, Recall, and Dice similarity coefficient metrics calculated at the pixel level to quantitatively assess the segmentation performance:

$$\text{Precision}(P) = \frac{TP}{TP + FP}, \quad (23)$$

$$\text{Recall}(R) = \frac{TP}{TP + FN}, \quad (24)$$

$$\text{Dice}(D) = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}, \quad (25)$$

where TP , FP , and FN represent the numbers of true positive, false positive, and false negative tumor epithelium or normal epithelium pixels. Dice coefficient measures the spatial overlap between machine vision result and ground truth and hence quantifies the overall accuracy. Precision and Recall are sensitive to the amount of over-segmentation and under-segmentation, respectively, in the sense that over-segmentation (of the epithelial area and/or tumor epithelium) is associated with a small Precision score, whereas under-segmentation leads to a small Recall score.

All experiments were run on a PC with the following configuration: 3.5 GHz Intel(R) Xeon(R), and 32.0 GB RAM in Matlab R2016b. Ground truth was obtained from the images by manual segmentation: a trained operator separated the epithelium from the stroma and discriminated between normal epithelium and tumor epithelium. The developed code and dataset will be publicly available at <http://aidpath.eu>.

4.2. Parameter settings and training

Our dataset has been divided into training and testing datasets (respectively 10% and 90% of the total dataset). For example, a set of 5 *WSIs* (Fig. 7) each contains both normal and tumor regions have been selected for training our approach in addition to a set of 8 randomly selected *TMA*s images (4 of which are

for normal cases). On the other hand, a set 8 *WSIs* and 126 *TMA* images used as a testing set for segmentation performance evaluation. We used the remaining images to train our model and fine-tune its parameters. Then the parameters were fixed for all testing images as follow: local Gaussian intensity of *FSPF* (σ'') was fixed to 0.1, contour smoothing parameter(σ) was fixed to 1.0, weight parameters (i.e., λ^+ , λ^-) were fixed to 1.0, local Gaussian intensity of *WAI* (σ') was fixed to 2, number of neurons in the output layer was fixed to 3×3 , and starting learning rate ($\eta(0)$) was fixed to 0.1.

The effect of all parameters on the behaviour of the proposed approach can be explained as follows:

- σ'' : local Gaussian intensity parameter of *FSPF* method. σ'' is used to control the amount of local information to deal with intensity inhomogeneity. If σ'' is zero, only global information will be used and under-segmentation problem may occur by our approach when intensity inhomogeneity is presented. On the other hand, if $\sigma'' > 0$, local information will be integrated into the segmentation framework. If σ'' is too large, edge leakage may occur, and computational time will increase dramatically.
- $\sigma > 0$: contour smoothing parameter. σ is used to control the smoothness degree of the contour during evolution. If σ is too small, our approach will be sensitive to noise, and hence the contour evolution will be unstable. Moreover, if σ is too large, over-segmentation problem may occur.
- $c > 1$: number of classes parameter. c is associated with the number of probabilistic maps of the given tissues to guide the evolution of the contour to extract each tissue separately (see Fig. 4), should be adjusted based on

the number of histology tissues in an image. In this paper, the local intensity distributions in each image can be categorised into three histology classes (i.e., stroma, epithelium, and background) where they can be effectively discriminated by the local probabilistic models of three classes (e.g., $P_j(x)$ of Equation 8), as illustrated by Fig. 4, therefore the number of classes (i.e., c) can be fixed to 3.

- σ' : local Gaussian intensity parameter of WAI distribution. Similar to σ'' , σ' is also used to control the amount of local information in the weighted average intensity distribution for the extraction of WAI -based descriptors. For example, if σ' is zero, then the weighted average intensity distribution can only describe the global situation in a region, and hence all WAI -based descriptors encode only global information about a region.
- $\lambda^+, \lambda^- \geq 0$: weights of the energy terms, respectively, inside and outside the contour. They used to assign relative importance to the inside and outside energy terms.
- Number of neurons in the output layers of the self-organizing neural map. This parameter is used to project the high dimensional feature space into a lower dimension (i.e. prototypes/weights of neurons in the output map). If the number of neurons is too large then an over-fitting problem may occur and the classification will be sensitive to noise. Likewise, an under-fitting problem may occur if the number of neurons is too small.
- $\eta_0 > 0$: starting learning rate of the neural network. η_0 is a constant that is used to initialize a learning rate function at the start of training, which is

then gradually decays over time so that during the last few iterations it is close to zero.

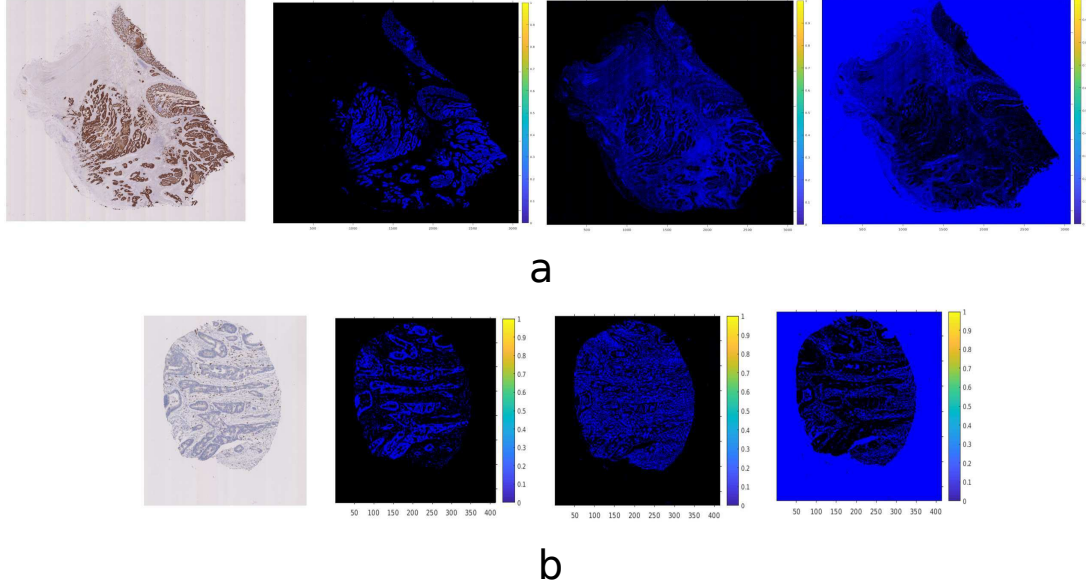


Figure 4: The probabilistic maps generated by our approach for *WSI* (a) and *TMA* (b). Arranged as columns are the original images, and the associated probabilistic maps of epithelium, stroma and background tissues, respectively.

4.3. Epithelium/stroma segmentation

Fig. 5 illustrates the ability of the proposed *FSPF* technique to segment both positively stained (a) or weakly stained (b,c) epithelium in *WSI*. Note how we managed to deal with artifacts (a), blurred boundaries (b) and intensity inhomogeneities (c).

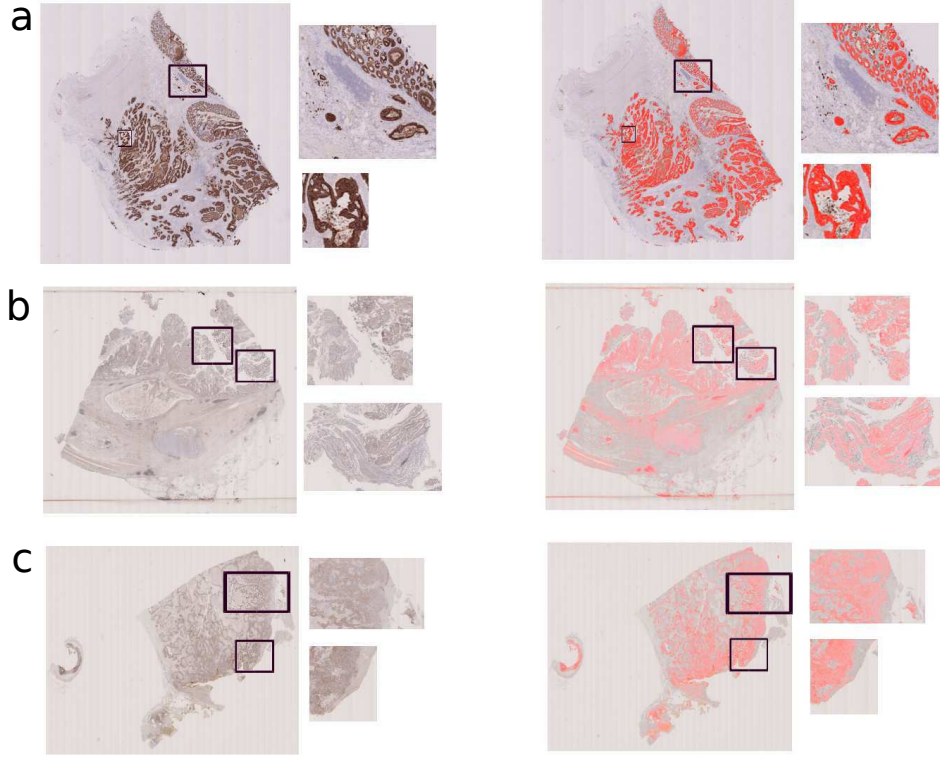


Figure 5: The epithelium segmentation results on some *WSIs* where the epithelium is positively stained (a) and weakly stained (b,c). Arranged as rows are the original images with different biomarkers, and the extracted epithelium highlighted in red.

Fig. 6 further illustrates the proposed *FSPF* approach's ability to deal with ill-defined boundaries (Fig. 6 b,c,f,g,j), and the presence of artifacts (Fig. 6 a,d,e,h,j) when segmenting the epithelium in *TMA* images and *WSIs*. Table 1 shows the superiority of the *FSPF* over a classical, unsupervised *SPF* in terms of precision, recall, and dice when dealing with *WSIs* of Fig. 5 and *TMA* cores of Fig. 6, except a *TMA* image of Fig. 6(h) where an under-segmentation problem has been occurred due to high similarity between the artifacts and epithelial regions.

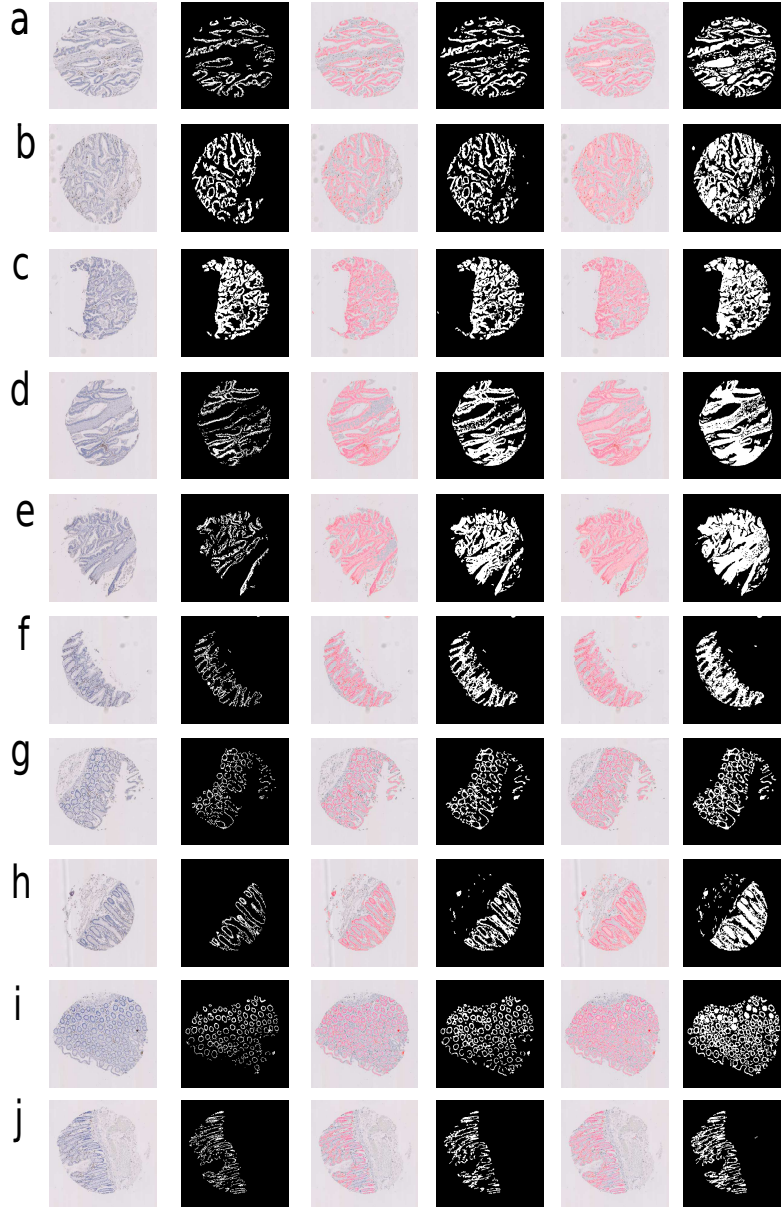


Figure 6: The epithelium segmentation results in some *TMA* cores with tumor epithelial areas (a,b,c,d,e) and normal epithelial areas (f,g,h,i,j). Arranged as columns are the original *TMA* cores, the ground truth images, the original *TMA* cores with the extracted epithelium highlighted in red obtained by our approach, segmentation masks by our approach, the original *TMA* cores with the extracted epithelium highlighted in red obtained by *SPF*, and the associated segmentation masks, respectively.

Table 1: The Precision, Recall, and Dice metrics for the *FSPF* and *SPF* approaches applied to *TMA* images of Fig. 6

Image	<i>FSPF</i>			Classical <i>SPF</i> method		
	<i>P</i> (%)	<i>R</i> (%)	<i>D</i> (%)	<i>P</i> (%)	<i>R</i> (%)	<i>D</i> (%)
Fig. 5 (a)	94.71	100	97.3	80.62	99.57	89.1
Fig. 5 (b)	90.71	100	95.33	43.95	99.75	61.02
Fig. 5 (c)	93.74	100	96.68	51.07	99.83	67.57
Fig. 6 (a)	100	79.08	88.31	71.06	85	77.41
Fig. 6 (b)	93.82	99.42	96.54	58.27	99.96	73.62
Fig. 6 (c)	97.16	99.66	98.4	82.2	99.98	90.22
Fig. 6 (d)	100	87.55	93.36	84.03	94.12	88.79
Fig. 6 (e)	96.19	91.33	93.7	28.02	99.98	43.77
Fig. 6 (f)	100	69.98	82.34	77.36	72.2	74.69
Fig. 6 (g)	99.98	79.17	88.36	83.26	81.35	82.29
Fig. 6 (h)	100	33.27	49.92	71.57	38.01	49.65
Fig. 6 (i)	100	79.8	88.76	65.12	84.41	73.52
Fig. 6 (j)	99.99	79.19	88.38	79.49	81.97	80.71

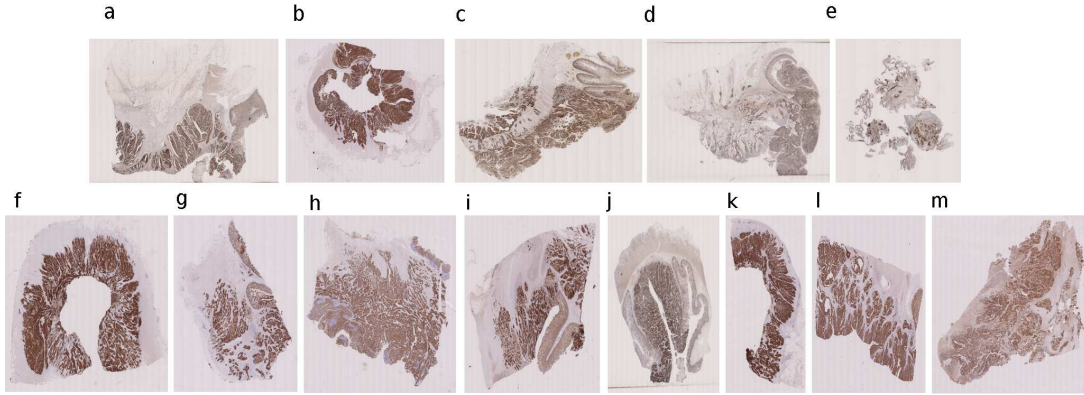


Figure 7: The *WSIs* used in the training (a-e) and testing (f-m) phases by our approach and other methods.

4.4. Tumor/normal segmentation

4.4.1. A few examples

Fig.8 illustrates the ability of our approach to segment tumor epithelium in *WSIs* that contain tumor epithelium only (a,d), and both normal and tumor epithelial regions (b,c). We used a single training image here, that of Fig. 7(a). By including more images in the training set as described in Subsection 4.2, the accuracy is slightly increased resulting in better overall performance, as shown in Table 2.

Table 3 compares the performance of all 6 approaches on the 8 *WSIs* of Fig. 7 (f-m). Note how our algorithm outperforms the others in terms of precision, recall, and dice metrics.

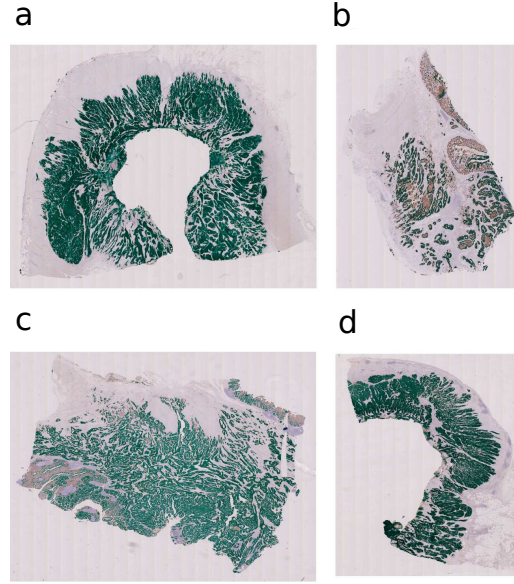


Figure 8: The tissue segmentation results on some *WSIs* where the tumor epithelium is highlighted in green.

Table 2: The Precision, Recall, and Dice metrics for our approach applied to *WSIs* of Fig. 8.

Image	Single training sample			Multiple training samples		
	P (%)	R (%)	D (%)	P (%)	R (%)	D (%)
Fig. 8 (a)	100	91.01	95.29	93.49	99.13	96.23
Fig. 8 (b)	94.88	94.46	94.67	93.53	97.79	95.57
Fig. 8 (c)	97.73	93	95.31	96.67	98.55	97.6
Fig. 8 (d)	100	92.45	96.08	100	93.28	96.53

Table 3: The Precision (%), Recall (%), and Dice (%) metrics for our approach, *LBP*-based, *LH*-based, *P*-based, *CMF*-based, and *HH*-based methods applied to *WSIs* of Fig. 7(f-m).

Ours			<i>LBP</i> -based			<i>LH</i> -based			<i>P</i> -based			<i>CMF</i> -based			<i>HH</i> -based		
<i>P</i>	<i>R</i>	<i>D</i>	<i>P</i>	<i>R</i>	<i>D</i>	<i>P</i>	<i>R</i>	<i>D</i>	<i>P</i>	<i>R</i>	<i>D</i>	<i>P</i>	<i>R</i>	<i>D</i>	<i>P</i>	<i>R</i>	<i>D</i>
93.49	99.13	96.23	100	55.04	71	100	91.61	95.62	100	31.36	47.75	100	85.35	92.09	100	64.42	78.36
93.53	97.7	95.57	90.48	83.46	86.83	89.04	95.69	92.24	85.11	52.8	65.17	97.31	83.09	89.64	93.73	70.06	80.18
96.67	98.55	97.6	96.48	86.3	91.11	95.46	98.74	97.07	95.94	51.5	67.03	80.07	83.42	81.9	99.14	51.09	67.43
97.24	97.26	97.25	96.66	89.09	92.72	93.42	90.01	91.68	94.81	29.68	45.21	96.61	86.15	91.08	95.68	83.02	88.9
99.44	74.53	85.2	94.3	28.07	43.27	91.41	59.14	71.82	90.77	19.5	32.11	99.94	85.01	91.88	97.9	76.88	86.12
100	93.28	96.53	100	88.4	93.84	100	64.88	78.7	100	43.82	60.94	100	89.99	94.73	100	64.69	78.56
97.73	99.31	98.52	99.05	97.42	98.23	97.45	99.74	98.59	98.88	88.23	93.25	81.13	78.43	79.87	99.53	73.39	84.48
98.94	95.13	97	100	88.81	94.07	100	94.7	97.28	100	50.99	67.54	100	73.56	84.77	100	67.65	80.7

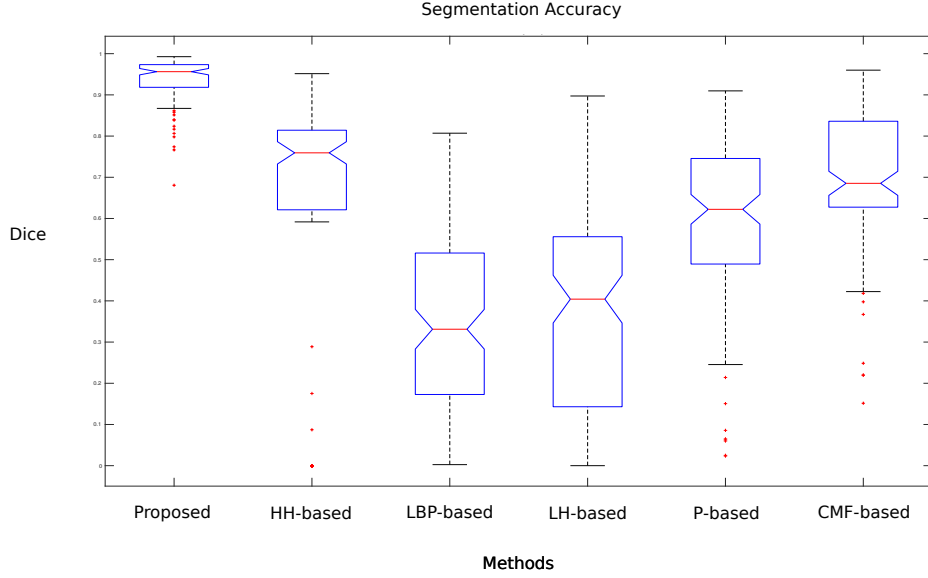


Figure 9: The segmentation performance of our approach and *HH*-based, *LBP*-based, *LH*-based, Perception-based, and *CMF*-based methods in 126 *TMA* images.

We also compared all 6 approaches on the *TMA* set. Our approach outperformed the others in terms of median, as shown in Fig. 9. A One-way, repeated measures ANOVA showed that the methods were statistically significantly different ($p < 0.001$) with planned comparisons indicating that our approach was statistically significantly better than the others (all Tukey corrected p -values < 0.01).

To illustrate the effect of *FSPF*, and our proposed *WAI*-based and *ALI*-based features, Fig.10 shows the segmentation result of a normal *TMA* image, where the normal region is highlighted in red while the misclassified region is highlighted in green, with different versions of the proposed approach to test each of its components independently. Note how the performance of our approach is significantly improved by combining both *ALI*-based and *WAI*-based features. The same conclusions can be confirmed when we compare the accuracy of these

Table 4: The Precision (%), Recall (%), and Dice (%) metrics for different versions of our proposed approach applied to a *TMA* image of Fig.10.

Features	Accuracies		
	P (%)	R (%)	D (%)
$ALI + WAI$	91.33	100.0	94.95
WAI	91.51	82.96	88.73
ALI	99.33	26.73	48.02

modified versions of the proposed approach, in terms of precision, recall, and dice metrics, as illustrated by Table 4.

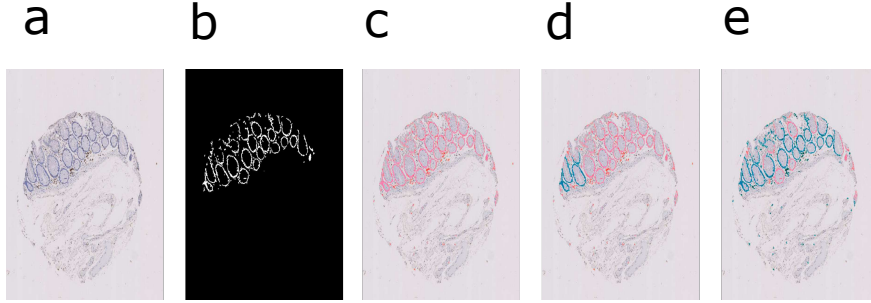


Figure 10: The segmentation results obtained by different versions of our proposed approach, where normal region is highlighted in red and incorrectly classified region is highlighted in green: (a) an original *TMA* image; (b) ground truth image; and (c-e) segmentation results obtained by a full version of our proposed approach, our approach without *ALI*-based features, and our approach without *WAI*-based features, respectively.

4.4.2. Robustness to noise

To illustrate the robustness of the *SOM*-based classifier to noise, we compared the behaviour of our approach to a similar one with a *REF-SVM* classifier

instead. This is by first corrupting a test *WSI* with different noise models, and then using the trained *RBF – SVM* and *SOM*-based classifiers to segment out tumor epithelium of a testing image. Fig. 11 illustrates the performance of our approach on a single *WSI* corrupted with Salt & Pepper noise model (e,j), and increasing levels of Gaussian noise: $SD = 5$ (a,f) to $SD = 35$ (d,i). Note that our approach is more robust to noise when using *SOMs* as a classifier than when using a *RBF – SVM* classifier, as confirmed in Table 5.

To illustrate the impact of noise at different features, we corrupted the individual features using white Gaussian noise, in both training and testing phases, and evaluated the classification performance of our proposed approach on a *WSI* of Fig. 11(a). Tables 6 and 7 show the effect of noise at each corrupted feature in terms of precision, recall, and dice metrics in the testing and training phases, respectively. Note how robust our proposed approach in dealing with attribute/feature noise. As demonstrated by Tables 6 and 7, we can see that attribute noise in the training phase has a higher impact on the classification performance than in the testing phase. Also ALI_{Γ} and WAI_5 are the most sensitive features to noise.

Despite the fact that *SOM*-based classifier has the ability to effectively preserve the topological structure of the training data using a few sets of prototypes and hence its robustness to noise. However, similar to other prototype-based classifiers, our approach is still sensitive to the very high-level of noise as shown in Tables 5 and 6. Interested readers may refer to (Zhu and Wu, 2004) for the different possible solutions to deal with noise, especially to detect and correct attribute noise.

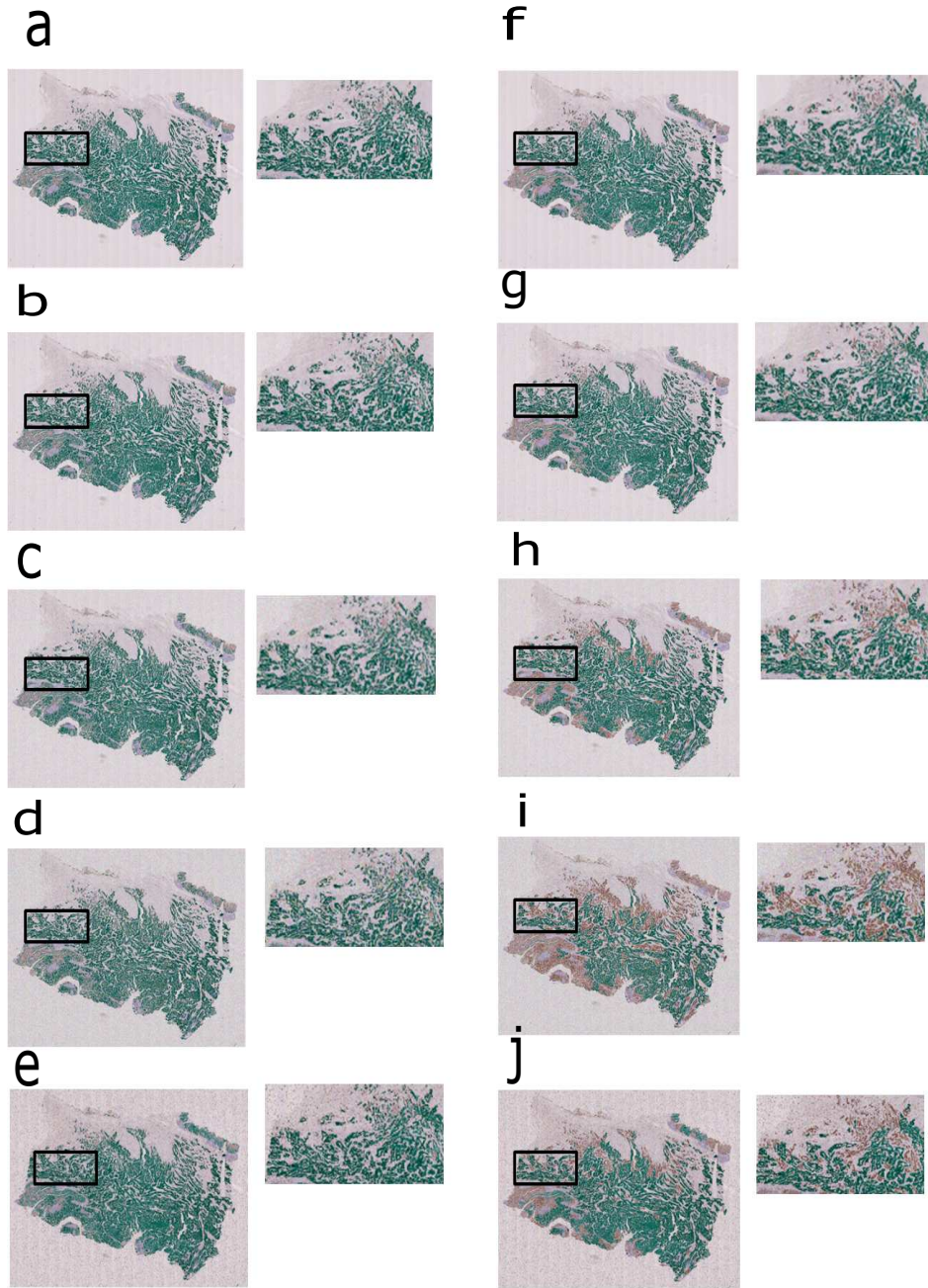


Figure 11: The robustness of our *SOMs* classifier (a-e) compare to *RBF-SVM* classifier (f-j) to additive noise (from Gaussian) and Salt & Pepper noise models on a *WSI* where the tumor area is highlighted in green.

Table 5: The robustness of the proposed approach to noise: the precision, recall, and dice metrics obtained when different Gaussian noise levels controlled by standard deviation (SD) and salt & pepper noise model.

Methods		SD				
		Salt & pepper	5	15	25	35
P(%)	SOMs	86.37	94.49	92.61	91.42	91.84
R (%)		97.71	98.91	96.72	93.3	86.17
D (%)		91.69	96.65	94.62	92.35	88.91
P(%)	RBFSVM	89.74	96.58	94.68	95.39	96.11
R (%)		84.08	95.43	93.53	82.41	58.89
D (%)		86.81	96	94.1	88.43	73.03

Table 6: The robustness of our approach to attribute noise (in the testing phase): the precision, recall, and dice metrics obtained when corrupting the individual features with white Gaussian noise model with different SNR values.

Accuracies		Features							
		ALI_{Γ}	ALI_{ζ}	ALI_{Ψ}	WAI_1	WAI_2	WAI_3	WAI_4	WAI_5
P(%)	$SNR = 30$	88.16	88.17	88.18	88.09	88.19	88.26	88.09	88.3
R (%)		96.51	96.47	96.5	96.35	96.44	96	96.21	96.15
D (%)		92.15	92.14	92.15	92.04	92.13	91.92	91.97	92.06
P(%)	$SNR = 10$	88.6	88.7	88.77	88.6	88.05	85.49	88.16	88.17
R (%)		89.88	94.64	94.03	95.44	96.09	99.93	96.51	96.48
D (%)		89.24	91.57	91.32	91.9	91.9	92.15	92.15	92.14

Table 7: The robustness of our approach to attribute noise (in the training phase): the precision, recall, and dice metrics obtained when corrupting the individual features with white Gaussian noise model with different SNR values.

Accuracies		Features							
		ALI_{Γ}	ALI_{ζ}	ALI_{Ψ}	WAI_1	WAI_2	WAI_3	WAI_4	WAI_5
P(%)	$SNR = 30$	88.20	88.16	88.14	88.07	88.2	88.24	87.85	87.97
R (%)		96.41	96.51	96.55	96.56	96.47	96.19	97.04	96.8
D (%)		92.12	92.15	92.15	92.12	92.15	92.04	92.22	92.17
P(%)	$SNR = 10$	88.46	88.17	88.24	87.95	88.24	88.33	89.41	90.05
R (%)		88.44	96.51	96.16	96.81	96.43	95.65	90.58	85.65
D (%)		88.45	92.15	92.03	92.17	92.16	91.84	89.99	87.8

4.5. Computational efficiency

The computational complexity of our approach to classify epithelial regions as normal or tumor is $O(NW + N^2)$, where N is the number of epithelial regions, and W is the number of units in the maps. The first term indicates the computations for searching the winning unit for each epithelial region while the second term indicates the computations for checking whether each pair of regions is projected on neighbouring prototypes in the output maps or not. Moreover, for epithelium region extraction, it can be observed from Equation that for every iteration, a convolution operation is performed. The convolution operation requires a high computational complexity $O(nP \times K)$, where nP and K are the size of the image (or number of pixels) and the Gaussian kernel, respectively. Compared to other approaches, several classifiers such as parameter-free 1- NN , Linear SVM , and Naïve Bayes have been used for tumor epithelium segmentation. However, RBF - SVM classifier provided the best results. The computational complexity of RBF kernel-based SVM is $O(nSV \times nP)$ where nSV is the number of support vectors. Moreover, the computational complexity of the Otsu method (as a segmentation method used in by CMF -based approach) is $O(L^4)$, where L is the number of gray levels in the processed image. In the feature extraction phase, our approach and CMF -based approach have the same computations of $O(N)$ while the computational complexity of both HH -based and P -based approach is $O(nP)$. On the other hand, LH -based features are required $O(nP \times K)$ for the convolution operation. Eventually, the LBP is calculated at each pixel location by estimating the binary differences of the values of a circular neighbourhood around the value of a central pixel, and hence LBP -based features requires a high computational complexity of $O(R \times P \times nP)$, where R is the radius of the

neighbourhood and P is the number of pixels in the neighbourhood.

For further supplement, we have calculated the CPU time (in seconds) used by our approach and other methods when dealing with the 126 *TMA* testing images. The average CPU time of our approach was 1.84 sec, which is close to the average of the *HH*-based method (1.82 sec), *P*-based method (1.95 sec), and *CMF*-based method (1.83 sec) and better than the *LBP*-based method (9.8 sec) and the *LH*-based method (3.15 sec).

5. Discussion and Conclusion

The segmentation of tumor epithelium in histopathology is a critical initial step for tumor quantification, biomarker assessment, and prognosis determination in colorectal cancer. In this paper we propose a novel cascade-learning approach to first distinguish epithelium from stroma and then discriminate between normal and tumor epithelium. Our method has been designed as a morphometry-based approach in a way to enable for a more informative high-level features that can encode biologically meaningful information.

The major contribution of our approach consists in four points. First, a novel unsupervised level set segmentation method is proposed for epithelium segmentation to deal with intensity inhomogeneity. This is by integrating local information about the appearance of the epithelium into the level set formulation as a prior knowledge. We demonstrated the superiority of the proposed level set method *FSPF*, compared to *SPF* as a reference method. Second, we developed and combined a new set of appearance (e.g. *WAI*-based) and shape (e.g. *ALI*-based) features. They can encode meaningful information to measure the regularity structure of epithelium in a way to be invariant to staining differences, and hence they

can deal with different biomarkers. We found that the segmentation performance obtained by our approach outperformed that of other methods (e.g. *HH*-based, *LBP*-based, *LH*-based, *P*-based, and *CMF*-based approaches) to confirm the ability of our solution in providing a generic solution to deal better with unseen patterns. Moreover, results indicated that combining both appearance and shape models led to significant performance improvement. Third, a robust classification framework based on self-organizing maps has been used, which is able to map the high dimensional feature space into a few numbers of prototypes (or weights) to improve the robustness of the classification to noise. Our classifier has demonstrated its robustness to different noise models when compared to *RBF-SVM*, as one of the most widely used classifier in histology tissue classification. Finally, our approach can provide an efficient solution in producing results within an acceptable diagnostic time on which clinical decisions can be efficiently made.

A potential limitation of our approach is that since it is a color-independent approach (in order to be able to deal with both positively and negatively stained tumor areas) it might be confused when other compartments such as lymphocytes are presented in the images.

In conclusion, this paper presents a novel approach to the segmentation of epithelium in colorectal cancer. Our method has been designed as a morphometry-based approach to enable for a more informative high-level features that can encode biologically meaningful information. This is by segmenting epithelium from a complicated background and then extracting a set of new shape and appearance descriptors from the segmented epithelial regions. In this work, two self-organizing maps have been used to reduce the dimensionality of the feature space of the training set and build a classifier to discriminate between normal and tumor

epithelium in an efficient and robust way. We focused on evaluating our approach on images stained for different biomarkers when a limited number of training images are provided. Results show that our method performs very well in segmenting epithelial regions and distinguishing between normal and tumor regions on both *TMA* and *WSI*. Moreover, our proposed framework has demonstrated its robustness to laboratory-dependent staining differences, noise, and scanner-dependent intensity inhomogeneities. Our approach could be adopted to deal with more complex cases of prostate and breast cancer histology (including different staining such as H&E) at different level of magnifications. As a future development, our system could be extended to deal with multi-class tissue segmentation by using additional *SOMs* to learn the appearance and shape model of the new tissues. This could help in overcoming the main limitation of our approach. Moreover, one could further improve the performance of our system by reducing the effect of noise and artifacts in the input images using stain normalization techniques (Ciompi et al., 2017). Also outlier removal methods and feature selection techniques (Zhu and Wu, 2004) could be integrated in the segmentation framework to improve the robustness of the proposed system.

Acknowledgment

The authors acknowledge financial support from the EC Marie Curie Actions, AIDPATH project (Contract No.612471).

References

Abdelsamea, M. M. and S. A. Tsaftaris (2013). Active contour model driven by globally signed region pressure force. In *Proceedings of the 18th IEEE*

International Conference on Digital Signal Processing (DSP), pp. 1–6.

Bianconi, F., A. Álvarez-Larrán, and A. Fernández (2015). Discrimination between tumour epithelium and stroma via perception-based features. *Neurocomputing* 154, 119–126.

Caicedo, J. C., F. A. González, and E. Romero (2011). Content-based histopathology image retrieval using a kernel-based semantic annotation framework. *Journal of biomedical informatics* 44(4), 519–528.

Caselles, V., R. Kimmel, and G. Sapiro (1997). Geodesic active contours. *International journal of computer vision* 22(1), 61–79.

Chan, T. F. and L. A. Vese (2001). Active contours without edges. *IEEE Transactions on image processing* 10(2), 266–277.

Chang, H., A. Borowsky, P. Spellman, and B. Parvin (2013). Classification of tumor histology via morphometric context. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2203–2210.

Chang, H., J. Han, A. Borowsky, L. Loss, J. W. Gray, P. T. Spellman, and B. Parvin (2013). Invariant delineation of nuclear architecture in glioblastoma multiforme for clinical and molecular association. *IEEE transactions on medical imaging* 32(4), 670–682.

Ciampi, F., O. Geessink, B. E. Bejnordi, G. S. de Souza, A. Baidoshvili, G. Litjens, B. van Ginneken, I. Nagtegaal, and J. van der Laak (2017). The importance of stain normalization in colorectal tissue classification with convolutional networks. In *Biomedical Imaging (ISBI 2017), 2017 IEEE 14th International Symposium on*, pp. 160–163. IEEE.

- Cruz-Roa, A., J. C. Caicedo, and F. A. González (2011). Visual pattern mining in histology image collections using bag of features. *Artificial intelligence in medicine* 52(2), 91–106.
- Datar, M., D. Padfield, and H. Cline (2008). Color and texture based segmentation of molecular pathology images using hsoms. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pp. 292–295. IEEE.
- Doyle, S., M. D. Feldman, N. Shih, J. E. Tomaszewski, and A. Madabhushi (2012, 2012). Cascaded discrimination of normal, abnormal, and confounder classes in histopathology: Gleason grading of prostate cancer. *BMC bioinformatics* 13, 282.
- Eramian, M., M. Daley, D. Neilson, and T. Daley (2011). Segmentation of epithelium in h&e stained odontogenic cysts. *Journal of microscopy* 244(3), 273–292.
- Galon, J., B. Mlecnik, G. Bindea, H. K. Angell, A. Berger, C. Lagorce, A. Lugli, I. Zlobec, A. Hartmann, C. Bifulco, et al. (2014). Towards the introduction of the ‘immunoscore’ in the classification of malignant tumours. *The Journal of pathology* 232(2), 199–209.
- Galon, J., F. Pagès, F. M. Marincola, H. K. Angell, M. Thurin, A. Lugli, I. Zlobec, A. Berger, C. Bifulco, G. Botti, et al. (2012). Cancer classification using the immunoscore: a worldwide task force. *Journal of translational medicine* 10(1), 205.
- Gurcan, M. N., L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and

- B. Yener (2009). Histopathological image analysis: A review. *IEEE Reviews in Biomedical Engineering* 2, 147–171.
- Jeong, W.-K., J. Schneider, S. G. Turney, B. E. Faulkner-Jones, D. Meyer, R. Westermann, R. C. Reid, J. Lichtman, and H. Pfister (2010). Interactive histology of large-scale biomedical image stacks. *IEEE Trans. Vis. Comput. Graph.* 16(6), 1386–1395.
- Kather, J. N., C.-A. Weis, F. Bianconi, S. M. Melchers, L. R. Schad, T. Gaiser, A. Marx, and F. G. Zöllner (2016). Multi-class texture analysis in colorectal cancer histology. *Scientific Reports* 6.
- Kohonen, T. (1990). The self-organizing map. *Proceedings of the IEEE* 78(9), 1464–1480.
- Krishnan, M. M. R., P. Shah, A. Choudhary, C. Chakraborty, R. R. Paul, and A. K. Ray (2011). Textural characterization of histopathological images for oral sub-mucous fibrosis detection. *Tissue and Cell* 43(5), 318–330.
- Krishnan, M. M. R., V. Venkatraghavan, U. R. Acharya, M. Pal, R. R. Paul, L. C. Min, A. K. Ray, J. Chatterjee, and C. Chakraborty (2012). Automated oral cancer identification using histopathological images: a hybrid feature extraction paradigm. *Micron* 43(2), 352–364.
- Linder, N., J. Konsti, R. Turkki, E. Rahtu, M. Lundin, S. Nordling, C. Haglund, T. Ahonen, M. Pietikäinen, and J. Lundin (2012). Identification of tumor epithelium and stroma in tissue microarrays using texture analysis. *Diagnostic pathology* 7(1), 22.

- Madabhushi, A. and G. Lee (2016). Image analysis and machine learning in digital pathology: Challenges and opportunities. *Medical Image Analysis* 33, 170 – 175.
- Mahmoud-Ghoneim, D. (2011). Optimizing automated characterization of liver fibrosis histological images by investigating color spaces at different resolutions. *Theoretical biology & medical modelling* 8, 25.
- Mattfeldt, T., P. Grahovac, and S. Lück (2013). Multiclass pattern recognition of the gleason score of prostatic carcinomas using methods of spatial statistics. *Image Analysis & Stereology* 32(3), 155–165.
- Mlecnik, B., G. Bindea, H. K. Angell, P. Maby, M. Angelova, D. Tougeron, S. E. Church, L. Lafontaine, M. Fischer, T. Fredriksen, et al. (2016). Integrative analyses of colorectal cancer show immunoscore is a stronger predictor of patient survival than microsatellite instability. *Immunity* 44(3), 698–711.
- Nayak, N., H. Chang, A. Borowsky, P. Spellman, and B. Parvin (2013). Classification of tumor histopathology via sparse feature learning. In *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*, pp. 410–413. IEEE.
- Neagoe, V.-E. and A.-D. Ropot (2002). Concurrent self-organizing maps for pattern classification. In *Cognitive Informatics, 2002. Proceedings. First IEEE International Conference on*, pp. 304–312. IEEE.
- Ogino, S., J. Galon, C. S. Fuchs, and G. Dranoff (2011). Cancer immunology—analysis of host and tumor factors for personalized medicine. *Nature reviews Clinical oncology* 8(12), 711.

- Ozdemir, E., C. Sokmensuer, and C. Gunduz-Demir (2012). A resampling-based markovian model for automated colon cancer diagnosis. *Biomedical Engineering, IEEE Transactions on* 59(1), 281–289.
- Ramesh, N., B. Dangott, M. E. Salama, T. Tasdizen, et al. (2012). Isolation and two-step classification of normal white blood cells in peripheral blood smears. *Journal of pathology informatics* 3(1), 13.
- Samsi, S., A. K. Krishnamurthy, and M. N. Gurcan (2012). An efficient computational framework for the analysis of whole slide images: Application to follicular lymphoma immunohistochemistry. *Journal of computational science* 3(5), 269–279.
- Shu, J., G. Dolman, J. Duan, G. Qiu, and M. Ilyas (2016). Statistical colour models: an automated digital image analysis method for quantification of histological biomarkers. *Biomedical engineering online* 15(1), 46.
- Sirinukunwattana, K., S. E. A. Raza, Y.-W. Tsang, D. Snead, I. Cree, and N. Rajpoot (2015). A spatially constrained deep learning framework for detection of epithelial tumor nuclei in cancer histology images. In *International Workshop on Patch-based Techniques in Medical Imaging*, pp. 154–162. Springer.
- Sirinukunwattana, K., S. E. A. Raza, Y.-W. Tsang, D. R. Snead, I. A. Cree, and N. M. Rajpoot (2016). Locality sensitive deep learning for detection and classification of nuclei in routine colon cancer histology images. *IEEE transactions on medical imaging* 35(5), 1196–1206.
- Sirinukunwattana, K., D. R. Snead, and N. M. Rajpoot (2015). A novel texture descriptor for detection of glandular structures in colon histology images. In

- SPIE Medical Imaging*, pp. 94200S–94200S. International Society for Optics and Photonics.
- Tambasco, M., B. M. Costello, A. Kouznetsov, A. Yau, and A. M. Magliocco (2009). Quantifying the architectural complexity of microscopic images of histology specimens. *Micron* 40(4), 486–494.
- Wang, Y., D. Crookes, J. Diamond, P. Hamilton, and R. Turner (2007). Segmentation of squamous epithelium from ultra-large cervical histological virtual slides. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pp. 775–778. IEEE.
- Yang, M., K. Kpalma, and J. Ronsin (2008). A survey of shape feature extraction techniques.
- Zhang, K., L. Zhang, H. Song, and W. Zhou (2010, April). Active contours with selective local or global segmentation: A new formulation and level set method. *Image and Vision Computing* 28(4), 668–676.
- Zhong, C., J. Han, A. Borowsky, B. Parvin, Y. Wang, and H. Chang (2017). When machine vision meets histology: A comparative evaluation of model architecture for classification of histology sections. *Medical Image Analysis* 35, 530 – 543.
- Zhou, Y., H. Chang, K. Barner, P. Spellman, and B. Parvin (2014). Classification of histology sections via multispectral convolutional sparse coding. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3081–3088.

Zhu, X. and X. Wu (2004). Class noise vs. attribute noise: A quantitative study.
Artificial intelligence review 22(3), 177–210.