A Stochastic Polygons Model for Glandular Structures in Colon Histology Images

Korsuk Sirinukunwattana, David R. J. Snead, and Nasir M. Rajpoot†, *Senior Member, IEEE*

***Abstract*—In this paper, we present a stochastic model for glandular structures in histology images of tissue slides stained with Hematoxylin and Eosin, choosing colon tissue as an example. The proposed Random Polygons Model (RPM) treats each glandular structure in an image as a polygon made of a random number of vertices, where the vertices represent approximate locations of epithelial nuclei. We formulate the RPM as a Bayesian inference problem by defining a prior for spatial connectivity and arrangement of neighboring epithelial nuclei and a likelihood for the presence of a glandular structure. The inference is made via a Reversible-Jump Markov chain Monte Carlo simulation. To the best of our knowledge, all existing published algorithms for gland segmentation are designed to mainly work on healthy samples, adenomas, and low grade adenocarcinomas. One of them has been demonstrated to work on intermediate grade adenocarcinomas at its best. Our experimental results show that the RPM yields favorable results, both quantitatively and qualitatively, for extraction of glandular structures in histology images of normal human colon tissues as well as benign and cancerous tissues, excluding undifferentiated carcinomas.**

***Index Terms*—Histology image analysis, Gland modeling, Random polygons, Bayesian inference, Reversible-Jump Markov chain Monte Carlo.**

I. Introduction

GLANDS in epithelial tissue normally have a tubular structure and consist of the lumen forming interior of the tubular structure surrounded by columnar epithelial cells with basally located nuclei, forming a radial epithelial nuclear boundary, as shown in Figure 1a. While there are inter-gland tissue constituents, including stromal nuclei and cytoplasm, arrangement of epithelial nuclei around the lumen can be used as a strong cue for the extraction of glandular structures. Glandular structures are important for diagnosis of several adenocarcinomas. Previous studies have suggested the usefulness of morphology of glands for grading of prostate, breast and colon adenocarcinomas [1]–[5]. One of the challenges of modern histopathology is achieving good intra-observer reproducibility in the grading of these cancers. The fusion of morphometric approach and automated histology image analysis[1] offers a means of doing so and possibly increasing the effectiveness of cancer grading as a result.

Glandular formation has been shown to reflect the degree of aggressiveness of colon tumors [9]. In adenoma and well differentiated (low grade) adenocarcinoma, the majority of tumor still exhibits glandular structures that have appearance similar to a normal gland (Figure 1d), whereas in moderately differentiated (intermediate grade) and poorly differentiated (high grade) adenocarcinomas, glandular structures become increasingly degenerated (Figures 1e and 1f).

We propose a Random Polygons Model (RPM) for modeling glandular structures in images of Hematoxylin and Eosin (H&E) stained histology slides, the most commonly used stain in the morphological assessment of cancers. We choose colon tissue including colonic carcinoma as an example. The proposed approach treats each glandular structure as a polygon made of a random number of vertices, where the vertices represent approximate locations of epithelial nuclei. Based on the biological fact that epithelial nuclei are present on the glandular boundary surrounding the lumen as shown in Figure 1b, edges connecting the vertices would result in a boundary of the glandular structure in the image being represented by a polygon, as shown in Figure 1c. We formulate the modeling of glandular structures using the RPM as a Bayesian inference problem, in which the prior distribution of a polygon is related to the spatial connectivity and arrangement of their vertices, while the likelihood term is related to the probability of a glandular structure represented by the polygon. We exploit the Reversible-Jump Markov Chain Monte Carlo (RJMCMC) method [10] using the Metropolis-Hastings algorithm [11] to infer all the *maximum a posteriori* polygons.

The RPM is devised to segment glandular structures on the whole spectrum of differentiation grades of histology samples. This ranges from glands in normal biopsies to those in poorly differentiated adenocarcinomas. To the best of our knowledge, all existing published algorithms for gland segmentation are mainly designed to work on healthy samples, adenomas, and well differentiated adenocarcinomas [12]–[16], except one of them which has been demonstrated to work on moderately differentiated adenocarcinomas [17]. However, it is worth noting that the work in [17] is concentrated on the segmentation of glandular structures on images of colon tissue slides stained with Hematoxylin and DAB. In practice, moderately

This paper has supplementary downloadable material available at http://ieeexplore.ieee.org, provided by the authors.

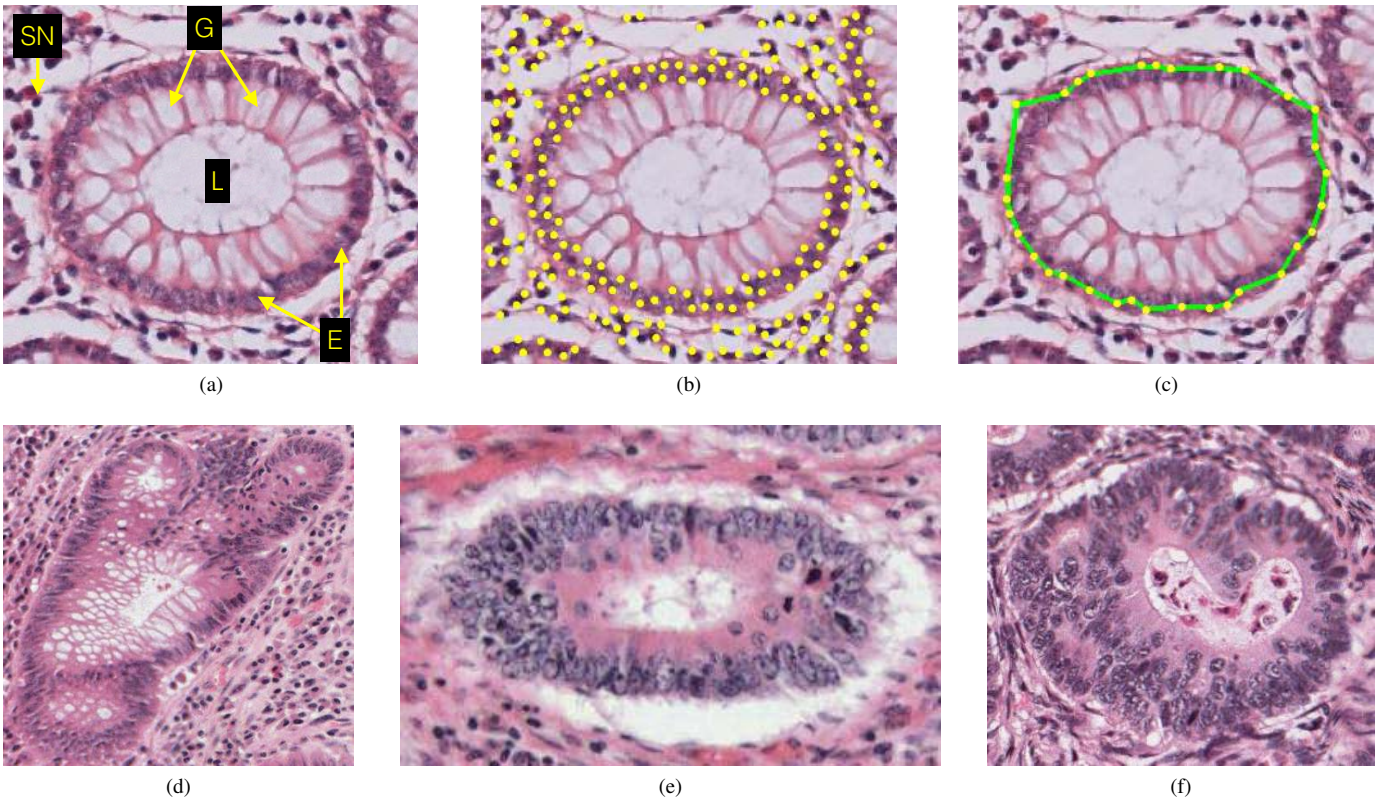
K. Sirinukunwattana is with the Department of Computer Science, University of Warwick, Coventry, UK. E-mail: k.sirinukunwattana@warwick.ac.uk.

D. R. J. Snead is a consultant histopathologist. He is with Department of Histopathology, University Hospitals Coventry and Warwickshire, NHS Trust, Coventry, UK, CV2 2DX. E-mail: david.snead@uhcw.nhs.uk. Contact number: (+44)7900954379.

N. M. Rajpoot is with the Department of Computer Science and Engineering, Qatar University, Doha, Qatar and Department of Computer Science, University of Warwick, Coventry, UK. E-mail: nasir.rajpoot@ieee.org. Contact number: (+44)2476980999.

† Corresponding Author.

[1]For broad reviews of automated histology image analysis, refer to [6]–[8].



(a) (b) (c)

(d) (e) (f)

Fig. 1. (a) A sample colon histology image showing various components (epithelial cell or E, stromal nucleus or SN, lumen or L, goblet cell or G). (b) Representative nuclei vertices are shown as yellow dots. (c) A *maximum a posteriori* polygon is shown as green boundary. Examples of glandular structures found in (d) adenoma, (e) moderately differentiated adenocarcinoma, and (f) poorly differentiated adenocarcinoma.

and poorly differentiated colorectal adenocarcinomas account for 70% and 20% of cases being diagnosed, respectively [9]. The main challenge in segmentation comes from the fact that glandular structures in moderately and poorly differentiated adenocarcinomas are deformed. Algorithms which assume the architectural regularity of glandular structures (Figure 1a), thus, are prone to fail. Note that in this work, we do not consider undifferentiated adenocarcinomas as in such cases, glandular morphology is totally lost and cannot offer any useful information for the inference of the glandular structures.

This paper is organized as follows. In Section II, we briefly review existing gland segmentation algorithms in the literature. The Bayesian formulation of the random polygons model and the framework for glandular structure segmentation are elaborated in Sections III and IV. In Section V, we present a comprehensive evaluation of the proposed gland segmentation framework on two different datasets consisting of healthy tissues as well as benign and malignant tumors.

II. Related Work

Existing methods for glandular structure segmentation can be categorized into two main classes: (a) texture based approaches and (b) structure based approaches. For textural methods, Farjam *et al*. [12] employed variance and Gaussian filters to extract texture features from glandular structures. Texture based methods do not employ any prior knowledge about the relationship between lumen and epithelial nuclei and, therefore, may result in poor segmentation, as shown in Section V-F.

In structural methods, Naik *et al*. [13] used a level set method to segment lumen areas in a gland. A Bayesian classifier is used to detect potential lumens and a level set curve is automatically initialized at the boundary of each detected luminal area. An evident limitation of the framework is that level sets often lead to erroneous segmentation in cases where lumen appears as a complex texture rather than a relatively smooth region such as healthy and adenomatous colon tissues. Nguyen *et al.* [14] employed the prior knowledge about glandular constituents in order to extract glandular regions. Their algorithm first jointly segments nuclei and cytoplasm to form a rough glandular boundary and then uses a region growing algorithm to expand the luminal area. Gunduz-Demir *et al*. [15] proposed object graphs for segmentation of glandular structures. Their algorithm first identifies nuclei and lumen using object-based information. A boundary of the gland is then constructed by connecting centroids of the nuclei objects. Fakhrzadeh *et al.* [16] employed low level color features to detect lumen and glandular boundaries. The outer borders of glands are delineated using geodesic distance transform. The major limitations of all the aforementioned algorithms are as follows. (a) They solely rely on pixel-level color information to detect different constituents of the tissues, which can be susceptible to stain variation. (b) They assume the architectural regularity of glandular structures, i.e., the

lumen is immediately surrounded by the cytoplasm and the cytoplasm is then surrounded by the epithelial boundary. This regularity may be degenerated in benign or cancerous colon tissues.

Fu *et al.* [17] devised a segmentation algorithm that makes use of the geometric property of the polar coordinate system. The algorithm first converts the image of a single gland embedded in the polar coordinate system into the Cartesian coordinate system, resulting in a close glandular boundary being transformed into a stretched curve. A contour representing the glandular boundary in the transform domain is then inferred through a conditional random field model. The algorithm has been mainly demonstrated on images of Hematoxylin and DAB stained slides from adenomas, and well to moderately differentiated adenocarcinomas. In contrast, for H&E stained slides, the algorithm has been demonstrated only on a limited amount of images belonging to healthy samples. Sabata *et al.* [18] tailored a gland segmentation method for PIN-4 stained slides of prostate needle biopsies. A graph-based segmentation method is run on the glandular probability map which is generated using a pixel classifier trained on Haar-like features. This method sometimes merges several glands together. The authors, thus, proposed to use *ad hoc* cues available via PIN-4 stain to separate them. Neither the information on tumor grade nor a comparative segmentation result is provided in the paper.

Previous image segmentation approaches that are close to the RPM in terms of methodology are based on Voronoi tessellation and point process [19]–[21]. These methods partition an image into a set of polygons represented by Voronoi cells, and each polygon corresponds to an area of homogeneous texture in the image. RJMCMC based on Metropolis-Hastings algorithms involving particular moves such as birth, death, position change and label change of generating points of Voronoi tessellation have been used to estimate the joint posterior distribution of the tessellation, labels of polygons, and the number of different textures. The RPM, however, is more well suited to the problem of gland segmentation due to the following facts: (a) It does not segment the glandular structures based on spatial homogeneity of texture since each glandular structure often comprises of an area with heterogeneous texture. (b) It separately generates a polygon for each individual glandular structure by sampling vertices from the epithelium boundary. It, therefore, can provide a better representative polygon for the glandular structure.

III. The Random Polygons Model

In this section, we will describe the inference of polygons for modeling glandular structures mathematically. To facilitate the reader's comprehension, frequently used mathematical notations are listed and defined in Table I. Given the evidence $y = (V, x(Z))$ available to us in the form of set of vertices representing approximate nuclei positions V and glandular probability map $x(Z)$, we infer the most likely polygon that encloses an individual glandular structure through Bayesian inference using the formulation below. For details on how V and $x(Z)$ are obtained, see Section IV.

TABLE I
Frequently used Notations in Section III

Notation	Definition
z	a vertex (pixel coordinates) on an image
Z	a set of all 2-dimensional pixel coordinates on an image
V	a set of vertices representing approximate nuclei locations, and $V \subset Z$
$x(\cdot)$	a function $x : Z \rightarrow [0, 1]$ such that $\forall z \in Z,\ x(z)$ is the probability of z belonging to a glandular structure
$x(Z)$	glandular probability map
y	an observation $(V, x(Z))$
S	a simple polygon
E_S	a set of all edges of polygon S
l_e	length of edge e of a polygon
$\overline{z_k z_{k+1}}$	an edge between vertices z_k and z_{k+1} of a polygon
θ_z	an angle that is less than 180 degrees at vertex z of a polygon
$\angle z_{k-1} z_k z_{k+1}$	a corner at vertex z_k of a polygon
Ω_S	a set of all pixel coordinates enclosed by polygon S
$P(S)$	prior probability for polygon S
$P(S\|y)$	posterior probability of polygon S given evidence y
$\mathcal{L}(y\|S)$	likelihood probability for evidence y given polygon S
$\mathcal{N}(\cdot\|\mu, \sigma)$	a probability density function of a normal distribution with mean μ and standard deviation σ
μ_l, σ_l	mean and standard deviation of a normal distribution for empha priori length of polygon edge
$\mu_\theta, \sigma_\theta$	mean and standard deviation of a normal distribution for *a priori* angle that is less than 180 degrees of polygon corner
σ_d	standard deviation of a normal distribution for the summation of the two edges to be built by adding a new vertex to a polygon
S_i, S_f	polygons at the current and proposed states of the Markov chain in the RJMCMC simulation, respectively
g	a proposal function, i.e., invertible deterministic function such that $(S_f, z') = g(S_i, z^*)$, where z' and z^* denote random variables required for dimension matching condition in the RJMCMC simulation
$Q(A\|B)$	proposal probability of state A given state B in the RJMCMC simulation
$P(A \rightarrow B)$	probability of choosing the type of move that allows the transition from state A to state B in the RJMCMC simulation

A. The Bayesian Formulation

Let $S = \{z_1, ..., z_n : z_j \in V\}$ be a simple polygon[2] consisting of n vertices, where the polygon edges are defined in the given order of vertices. According to the Bayes' rule, the posterior distribution of polygon S given observation y is written as follows,

$$P(S|y) \propto \mathcal{L}(y|S)P(S). \tag{1}$$

The prior probability $P(S)$ can be defined in terms of the length of edges of S and the angle at corners of S. Assume that both factors can be independently modeled by normal distributions. Then the prior probability $P(S)$ can be expressed as follows,

$$P(S) = \prod_{e \in E_S} \mathcal{N}\left(l_e \middle| \mu_l, \sigma_l\right) \prod_{z \in S} \mathcal{N}\left(\theta_z \middle| \mu_\theta, \sigma_\theta\right), \tag{2}$$

[2]A *simple polygon* is a polygon with non-intersecting edges.

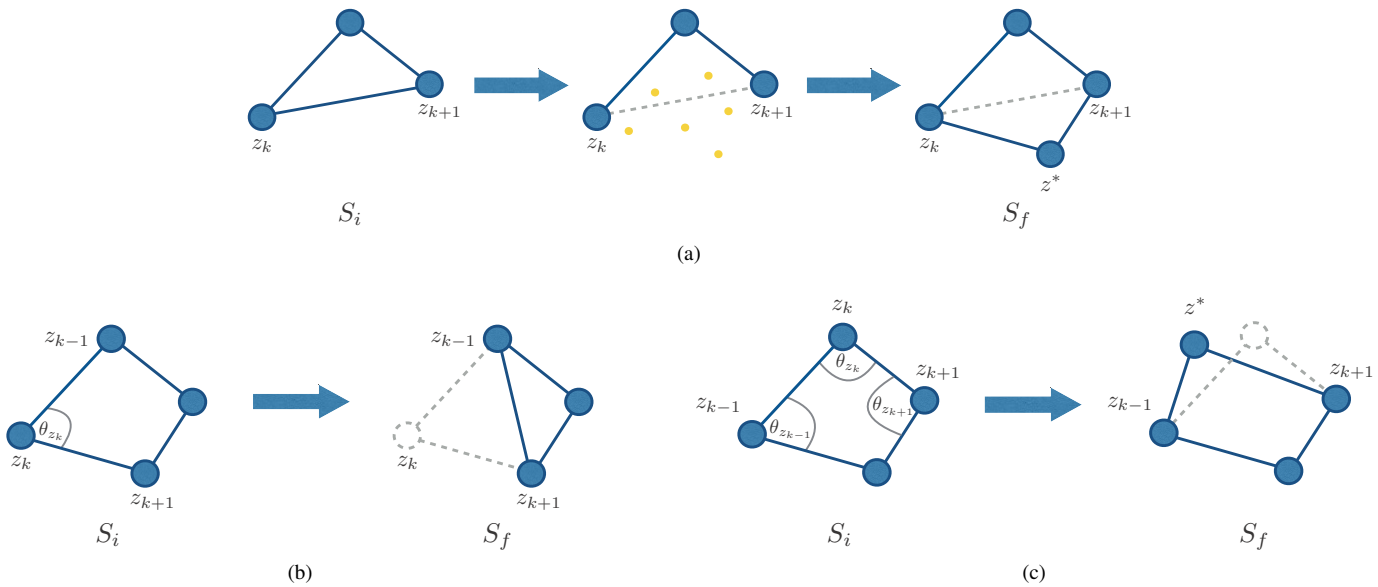


Fig. 2. Illustration of the proposed moves. (a) Addition of a vertex: edge $\overline{z_k z_{k+1}}$ of polygon S_i is sampled according to distribution $\mathcal{G}_{\text{edge}}$ (5). Then vertex z^* is uniformly drawn from set $U^m_{V\setminus S_i}$ (yellow dots) and added to S_i, resulting in polygon S_f. (b) Deletion of a vertex: vertex z_k of S_i is sampled according to distribution $\mathcal{G}_{\text{vertex}}$ (8) and removed from S_i, resulting in polygon S_f. (c) Change of a vertex: vertex z_k of polygon S_i is sampled according to distribution $\mathcal{P}$ (12) and is replaced by vertex z^* sampled according to distribution $\tilde{\mathcal{P}}$ (13), resulting in polygon S_f.

where the notation is as defined in Table I. The likelihood of observation y given polygon S, $\mathcal{L}(y|S)$, is defined in terms of how close the area inside S matches a glandular structure, while the area outside matches non-glandular structure. This information is available to us in the form of the glandular probability map $x(Z)$. The likelihood $\mathcal{L}(y|S)$ is defined as below,

$$\mathcal{L}(y|S) = \prod_{z\in\Omega_S} x(z) \prod_{z\notin\Omega_S} (1 - x(z)), \tag{3}$$

where the notation is as in Table I. Note that pixels $z \in Z$ such that $x(z) = 0$ or $x(z) = 1$ can rarely occur and cause (3) to be zero. It is preferable to transform such pixel probabilities or not include them in the calculation of (3).

B. The Reversible-Jump Markov Chain Monte Carlo

The posterior distribution $P(S|y)$ in (1) cannot be determined analytically. We, therefore, resort to Markov chain Monte Carlo (MCMC) simulation to approximate the posterior distribution. Reversible-Jump Markov chain Monte Carlo (RJMCMC) [10] using Metropolis-Hastings algorithm is employed due to its ability to accommodate changes in the dimensionality of the state space. This property is vital as it allows us to sample a polygon with varying number of vertices.

Suppose that the current state and the proposal state of the Markov chain are polygons S_i and S_f, respectively, and the transition from S_i to S_f is defined by a proposal function g such that $(S_f, z') = g(S_i, z^*)$, where z' and z^* are auxiliary random variables due to dimensional matching condition. The acceptance ratio for a transition from S_i to S_f according to the Metropolis-Hastings method is given by

$$\alpha(S_i, S_f) = \min\{1, R\}, \tag{4}$$

where

$$R = \frac{\mathcal{L}(y|S_f)}{\mathcal{L}(y|S_i)} \frac{P(S_f)}{P(S_i)} \frac{Q(S_i|S_f)}{Q(S_f|S_i)} \frac{P(S_f \to S_i)}{P(S_i \to S_f)} \left|\frac{\partial(S_f, z')}{\partial(S_i, z^*)}\right|,$$

$Q(S_f|S_i)$ $\left(Q(S_i|S_f)\right)$ denotes the proposal distribution of S_f given S_i (S_i given S_f), $P(S_i \to S_f)$ $\left(P(S_f \to S_i)\right)$ denotes the probability of selecting a type of move that allows the transition from S_i to S_f (S_f to S_i), and $|\partial(S_f, z')/\partial(S_i, z^*)|$ denotes a Jacobian arising from the change in dimensionality. In order to fulfill the reversibility condition of the Markov chain, we define three types of moves, including *addition*, *deletion*, and *change* of a vertex. Note that addition and deletion of a vertex are the reverse of each other.

Addition of a Vertex: In this move, we favor the addition of a new vertex to an edge of a polygon that is longer than the *a priori* mean length μ_l. The move is illustrated in Figure 2a. Assume that the current state of the Markov chain is a polygon $S_i = \{z_1, ..., z_{k-1}, z_k, z_{k+1}, z_{k+2}, ..., z_n\}$. The move proposes a new polygon S_f by first sampling an edge of polygon S_i according to a Gibbs distribution

$$\mathcal{G}_{\text{edge}}(e|S, \mu_l) = \frac{\exp(|l_e - \mu_l|/T_l)}{\sum_{e'\in E_S} \exp(|l_{e'} - \mu_l|/T_l)}, \tag{5}$$

where $T_l = \frac{1}{|E_S|} \sum_{e'\in E_S} |l_{e'} - \mu_l|$ is a temperature parameter for the Gibbs distribution, and $|E_S|$ denote the cardinality of set E_S. Here, the edge of polygon S with higher deviation from μ_l is more likely to be selected.

Suppose that edge $e^* = \overline{z_k z_{k+1}}$ of S_i is chosen. We define a score for each vertex $z \in V \setminus S_i$ as follows,

$$f_{e^*}(z) = \mathcal{N}\left(l_{\overline{z_k z}} + l_{\overline{z z_{k+1}}} \middle| l_{e^*}, \sigma_d\right) \times \mathcal{N}(\tilde{\theta}_{z_k}|\mu_\theta, \sigma_\theta)\mathcal{N}(\tilde{\theta}_z|\mu_\theta, \sigma_\theta)\mathcal{N}(\tilde{\theta}_{z_{k+1}}|\mu_\theta, \sigma_\theta), \tag{6}$$

where $\tilde{\theta}_{z_k}, \tilde{\theta}_z, \tilde{\theta}_{z_{k+1}}$ denote angles that are less than 180 degrees at corners $\angle z_{k-1} z_k z$, $\angle z_k z z_{k+1}$, $\angle z z_{k+1} z_{k+2}$, respec-

tively. The score $f_{e^*}(z)$ indicates (a) the proximity between vertex z and edge e^*, measured in terms of the deviation of the summation $l_{\overline{z_k z}} + l_{\overline{z z_{k+1}}}$ from l_{e^*}, and (b) the irregularity of shape of edges to be built if z is selected and added to the polygon, measured by the deviation of angles $\tilde{\theta}_{z_k}, \tilde{\theta}_z, \tilde{\theta}_{z_{k+1}}$ from the *a priori* mean angle μ_θ. The higher the value of $f_{e^*}(z)$, the closer the position of z to e^* and the values of $\tilde{\theta}_{z_k}, \tilde{\theta}_z, \tilde{\theta}_{z_{k+1}}$ to μ_θ.

Let $U^m_{V \setminus S}$ be a set of m vertices $z \in V \setminus S$ with the highest values of f_{e^*} score in the descending order. A new vertex z^* is then uniformly drawn from $U^m_{V \setminus S_i}$ and added into S_i between vertices z_k and z_{k+1}. Hence, $S_f = g(S_i, z^*) = \{z_1, ..., z_k, z^*, z_{k+1}, ..., z_n\}$, $|\partial S_f / \partial (S_i, z^*)| = 1$, and

$$Q(S_f|S_i) = \mathcal{G}_{\text{edge}}(e^*|S_i, \mu_l) \frac{1}{|U^m_{V \setminus S_i}|}. \quad (7)$$

We assume that this move occurs with probability $P(S_i \rightarrow S_f) = r_{\text{add}}$. Note that even though (5) also favors the selection of an edge shorter than μ_l, the addition of a vertex to the shorter edge is less likely to be accepted due to the small acceptance ratio (4).

Deletion of a Vertex: In this move, we favor the deletion of a narrow angle from a polygon. The move is illustrated in Figure 2b. Assume that polygon $S_i = \{z_1, ..., z_{k-1}, z_k, z_{k+1}, ..., z_n\}$ is the current state of the Markov chain. We sample vertex $z \in S_i$ and remove it from S_i according to a Gibbs distribution

$$\mathcal{G}_{\text{vertex}}(z|S, \mu_\theta) = \frac{\exp(|\theta_z - \mu_\theta|/T_\theta)}{\sum_{z' \in S} \exp(|\theta_{z'} - \mu_\theta|/T_\theta)}, \quad (8)$$

where $T_\theta = \frac{1}{|S|} \sum_{z' \in S} |\theta_{z'} - \mu_\theta|$. From (8), the likelihood of vertex $z \in S$, whose angle θ_z is more deviated from the *a priori* mean angle μ_θ, to be selected is higher than those with smaller deviation. Given that in general, μ_θ is an obtuse angle close to 180 degrees, a narrow angle is more likely to be selected.

Without loss of generality, suppose that z_k is being sampled. It then follows that $(S_f, z_k) = g(S_i)$ where $S_f = \{z_1, ..., z_{k-1}, z_{k+1}, ..., z_n\}$, $|\partial (S_f, z_k)/\partial S_i| = 1$, and

$$Q(S_f|S_i) = \mathcal{G}_{\text{vertex}}(z_k|S_i, \mu_\theta). \quad (9)$$

Here, we assume that the deletion occurs with probability $P(S_i \rightarrow S_f) = r_{\text{delete}}$.

Change of a Vertex: Figure 2c shows an illustration of the move. Suppose that the current state of the Markov chain is given by polygon $S_i = \{z_1, ...z_{k-1}, z_k, z_{k+1}, ..., z_n\}$. This move proposes a new state of the chain, S_f, by replacing a vertex in S_i that makes it less conform to the prior knowledge of polygons (2) with a new vertex. Let us consider vertex $z_k \in S_i$. The variables involved in the calculation of (2) that depend on z_k are adjacent edges $\overline{z_{k-1} z_k}$ and $\overline{z_k z_{k+1}}$ and angles $\theta_{z_{k-1}}, \theta_{z_k}$, and $\theta_{z_{k+1}}$. Thus, the probability of selecting z_k is defined based on these terms as follows,

$$\mathcal{P}(z_k|S_i, \mu_l, \mu_\theta) = \prod_{e \in \{\overline{z_{k-1} z_k}, \overline{z_k z_{k+1}}\}} \mathcal{G}_{\text{edge}}(e|S_i, \mu_l) \times \prod_{z \in \{z_{k-1}, z_k, z_{k+1}\}} \mathcal{G}_{\text{vertex}}(z|S_i, \mu_\theta). \quad (10)$$

This probability takes into account the deviation of $\overline{z_{k-1} z_k}$ and $\overline{z_k z_{k+1}}$ from the *a priori* mean length μ_l, and the deviation of $\theta_{z_{k-1}}, \theta_z$, and $\theta_{z_{k+1}}$ from the *a priori* mean angle μ_θ. Next, we draw a vertex $z^* \in V \setminus S_i$ to replace z_k according the probability

$$\widetilde{\mathcal{P}}(z^*|S_f, \mu_l, \sigma_l, \mu_\theta, \sigma_\theta) = \prod_{e \in \{\overline{z_{k-1} z^*}, \overline{z^* z_{k+1}}\}} \mathcal{N}(e|\mu_l, \sigma_l) \times \prod_{z \in \{z_{k-1}, z^*, z_{k+1}\}} \mathcal{N}(z|\mu_\theta, \sigma_\theta). \quad (11)$$

Consequently, we have that $(S_f, z_k) = g(S_i, z^*)$ where $S_f = \{z_1, ..., z_{k-1}, z^*, z_{k+1}, ..., z_n\}$, $|\partial (S_f, z_k)/\partial (S_i, z^*)| = 1$, and

$$Q(S_f|S_i) = \mathcal{P}(z_k|S_i, \mu_l, \mu_\theta) \widetilde{\mathcal{P}}(z^*|S_f, \mu_l, \sigma_l, \mu_\theta, \sigma_\theta) \quad (12)$$

We assume that this move occurs with a probability $P(S_i \rightarrow S_f) = r_{\text{change}}$.

For the reverse transition from S_f to S_i, we simply select z^* according to the distribution $\mathcal{P}(z|S_f, \mu_l, \mu_\theta)$ and replace it with vertex z_k drawn from the distribution $\widetilde{\mathcal{P}}(z|S_i, \mu_l, \sigma_l, \mu_\theta, \sigma_\theta)$. The proposal distribution of S_i given S_f, therefore, can be expressed by

$$Q(S_i|S_f) = \mathcal{P}(z^*|S_f, \mu_l, \theta_l) \widetilde{\mathcal{P}}(z_k|S_i, \mu_l, \sigma_l, \mu_\theta, \sigma_\theta). \quad (13)$$

We also assume that this transition occurs with the probability $P(S_f \rightarrow S_i) = r_{\text{change}}$.

IV. The Glandular Modeling Framework

An illustration of the operation of the proposed framework is shown in Figure 3. Given an H&E image (Figure 3a), first we generate the glandular probability map (Figure 3b) and nuclei vertices (Figure 3c), two evidences necessary to perform an inference for random polygons. Seed areas (Figure 3d) to initialize the random polygons used in the RJMCMC simulation are derived from the glandular probability map. *Maximum a posteriori* polygons are then estimated using the RJMCMC simulation (Figure 3e). Finally, as a post-processing operation for the RPM, false positive polygons are removed, and a relatively smoother contour of the glandular structures is generated (Figure 3f).

A. The Glandular Probability Map

Following [22], [23], the generation of glandular probability map $x(Z)$ for a given image is a 3-step process, as described below.

Step 1: Superpixel Segmentation. We segment the image into superpixels using the Simple Linear Iterative Clustering (SLIC) algorithm [24].

Step 2: Feature Extraction. For each superpixel i, let $\mathcal{B}_i$ denote a union of intermediate neighboring superpixels of i. A feature vector for each superpixel i,

$$\mathbf{f}_i = \left[\mathbf{f}_i^{\text{color}\top}, \mathbf{f}_i^{\text{texture}\top} \right]^\top, \quad (14)$$

comprises two cues, namely color and texture, from the superpixel i itself and its neighboring superpixels. The color feature is calculated as $\mathbf{f}_i^{\text{color}} = [h(i), h(\mathcal{B}_i)]^\top$, where $h(i)$ and $h(\mathcal{B}_i)$ are normalized histograms of quantized L*a*b* color

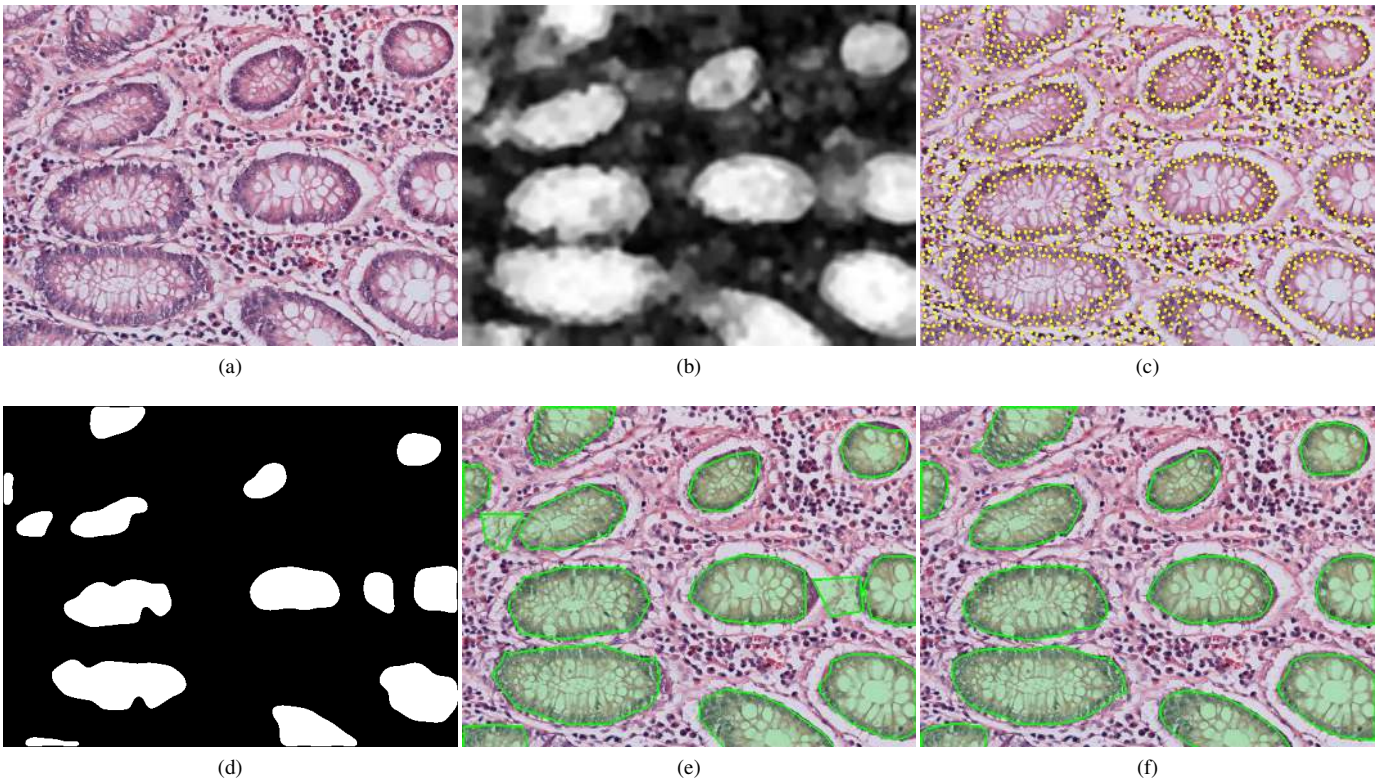


Fig. 3. The Random Polygons Model framework. (a) A sample Hematoxylin and Eosin stained colon histology image. (b) A glandular probability map. The brighter the intensity, the higher the probability of the area belonging to glandular structures. (c) Representative vertices of nuclei are shown as yellow dots. (d) Seed areas for initializing random polygons in the Reversible-Jump Markov chain Monte Carlo simulation. (e) *Maximum a posteriori* polygons are shown in green. (f) Post-processed polygons.

intensities calculated over superpixels i and $\mathcal{B}_i$, respectively. The texture feature is defined as $\mathbf{f}_i^{\text{texture}} = [\bar{S}p(i), \bar{S}p(\mathcal{B}_i)]^\top$, where $\bar{S}p(i)$ and $\bar{S}p(\mathcal{B}_i)$ denote scattering symmetric positive definite descriptors [22] based on gray-scale intensities of pixels in superpixels i and $\mathcal{B}_i$, respectively.

Step 3: Glandular Probability Assignment. A glandular probability for each superpixel is obtained from a random forest classifier trained on the extracted features. This results in a map indicating the likelihood for each pixel in the image belonging to the glandular region. The map is then smoothed by a spatial average filter, e.g., Gaussian filter, yielding the glandular probability map $x(Z)$.

B. The Nuclei Vertices

We employ a color-deconvolution method [25] to extract the Hematoxylin channel from the image. By thresholding the Hematoxylin channel using Otsu's threshold [26], we obtain a binary image corresponding to the approximate locations of nuclei in the image. To construct a set of vertices V representing nuclei locations, we sample vertices from a set of coordinates of boundary pixels of the binary mask. The sampling without replacement is done in a greedy fashion where, at each step, a single vertex z^* is uniformly randomly drawn. If there exists $z \in V$ such that $\|z^* - z\| \leq d$ where d is a desired minimum distance between any two vertices, then z^* is rejected, otherwise it is put in V.

C. The Seed Areas

These provide an initial guess about the number of glands and their locations in the image. We threshold the probability map to produce a binary image that coarsely represents the location of glandular structures in the image. The value of the glandular probability in the region close to the boundary of the glandular structures is often less reliable than that of the inner region. We, therefore, process each connected component in the binary image by a morphological operation which consists of boundary erosion and boundary smoothing to get seed areas which can provide more reliable information regarding the location of individual glandular structure. Each seed area is used to initialize an individual polygon in the RJMCMC simulation.

D. Inference for Random Polygons

Let C_i denote a set of all coordinates in the ith seed area. To infer a random polygon corresponding to the glandular structure that is represented by the ith seed area, we first restrict the analysis to the neighborhood of C_i. This neighborhood is formed by expanding all sides of the minimum bounding box of C_i by m_{expand} pixels. We define W_{C_i} to be a set of all coordinates in the expanded bounding box, i.e.,

$$W_{C_i} = \{(a,b) \in Z : a_{\min} - m_{\text{expand}} \leq a \leq a_{\max} + m_{\text{expand}} \text{ and } b_{\min} - m_{\text{expand}} \leq b \leq b_{\max} - m_{\text{expand}}\}, \quad (15)$$

Algorithm 1 Metropolis-Hastings and RJMCMC Simulation

Input: Evidence $y = (V|_W, x(Z)|_W)$, initial polygon S_0, total number of iterations t_{total}, probabilities of moves $r_{\text{add}}, r_{\text{delete}}, r_{\text{change}}$ (not all inputs are listed here)

Output: Set of sample polygons $\{S_t\}_{t=1}^{t_{\text{total}}}$, posterior distribution $P(S|y)$

for $t = 1$ to t_{total} **do**
  Draw a type of move based on probabilities $\{r_{\text{add}}, r_{\text{delete}}, r_{\text{change}}\}$
  if addition of a vertex **then**
    $\tilde{S} \leftarrow$ addition of a vertex (S_{t-1}) (Section III-B)
    Calculate $\alpha(S_{t-1}, \tilde{S})$ in (4)
  else if deletion of a vertex **then**
    $\tilde{S} \leftarrow$ deletion of a vertex (S_{t-1}) (Section III-B)
    Calculate $\alpha(S_{t-1}, \tilde{S})$ in (4)
  else
    $\tilde{S} \leftarrow$ change of a vertex (S_{t-1}) (Section III-B)
    Calculate $\alpha(S_{t-1}, \tilde{S})$ in (4)
  end if
  if $\tilde{S}$ is a simple polygon **then**
    $r \leftarrow$ a random number uniformly drawn from $[0, 1]$
    if $r > \alpha(S_{t-1}, \tilde{S})$ **then**
      $\tilde{S} \leftarrow S_{t-1}$
    end if
  else
    $\tilde{S} \leftarrow S_{t-1}$
  end if
  $S_t \leftarrow \tilde{S}$
  Calculate $P(S_t|y)$ in (1)
end for

where $a_{\min} = \min\{a|(a,b) \in C_i\}, a_{\max} = \max\{a|(a,b) \in C_i\}, b_{\min} = \min\{b|(a,b) \in C_i\}$, and $b_{\max} = \max\{b|(a,b) \in C_i\}$. Further, we define $x(Z)|_{W_{C_i}}$ to be the probability map restricted to W_{C_i}, and $V|_{W_{C_i}}$ to be the set of all vertices restricted to W_{C_i} that are closest to the ith seed area than the others, i.e.,

$$V|_{W_{C_i}} = \{z \in V : z \in W_{C_i} \text{ and for all } j \neq i, \min_{z' \in C_i} \|z - z'\| < \min_{z' \in C_j} \|z - z'\|\} \quad (16)$$

Next, we form an initial polygon S_0 from vertices $z \in V|_{W_{C_i}}$ that are close to the boundary of the ith seed.

Then we use the evidence $y = (V|_{W_{C_i}}, x(Z)|_{W_{C_i}})$ and the initial polygon S_0 as inputs to the RJMCMC simulation (Algorithm 1). The cardinality of $V|_{W_{C_i}}$ reflects the size of the sample space for random polygons. Thus, we define the total number of iterations t_{total} for the RJMCMC to be proportional to the cardinality of $V|_{W_{C_i}}$, i.e., $t_{\text{total}} = \beta \cdot \left|V|_{W_{C_i}}\right|$, where β is a constant factor. Assume that a sequence of polygons $\{S_t\}_{t=1}^{t_{\text{total}}}$ is sampled from the posterior distribution $P(S|y)$ by the RJMCMC. We use the *maximum a posteriori* polygon

$$S^{\text{MAP}} = \arg\max_{S \in \{S_t\}_{t=1}^{t_{\text{total}}}} P(S|y), \quad (17)$$

as an estimator of the glandular structure. Note that it is straightforward to adapt the above procedure to utilize parallel computing since random polygon modeling is performed separately on each seed area.

E. Post-Processing

It is possible that false positive polygons are generated due to imperfections in the glandular probability map. We remove these polygons using a heuristic criterion related to the number of vertices and the size of the polygons as will be explained in more detailed in Section V-D. As can be seen in Figure 3e, false positive polygons may occupy some vertices on glandular structures, preventing the true positive polygons to model the whole glandular structure that they correspond to. After false positive removal, we therefore treat the remaining polygons as seed areas and rerun the procedure in Section IV-D for a small number of iterations. Finally, we smooth the boundary of the obtained polygons by cubic spline interpolation.

V. Experimental Setup, Experimental Results and Discussion

A. Datasets

1) The Bilkent Dataset: This dataset used in [15] consists of 72 microscopic images of H&E stained healthy colon tissues. The images were captured at 20× magnification using Nikon Coolscope Digital Microscope with resolution 480×640. The expert annotation of a total of 1,009 glandular structures are available for this dataset.

2) The Warwick-QU Dataset: This dataset consists of 165 images extracted from H&E stained colon histology slides. The slides were digitally scanned at 20× magnification by Zeiss MIRAX MIDI Slide Scanner. A total of 52 visual fields were extracted, and an expert pathologist graded these fields into five categories: healthy, adenomatous, moderately differentiated, moderately-to-poorly differentiated, and poorly differentiated (well-differentiated adenocarcinoma is not observed). Handmarked ground truth for glandular structures in each visual field is also provided. This results in a total of 165 non-overlapping images taken from the graded visual fields, where the grade of the image is assigned according to the original field. This gives 42 healthy, 32 adenomatous, 47 moderately differentiated, 20 moderately-to-poorly differentiated, and 24 poorly differentiated images, respectively containing 670, 298, 287, 135, and 195 glandular structures present in them. In terms of dimensions, 151 of the images have the resolution of 520×775, and 14 have the resolution of 430×575. At the time of publication, this dataset is being released as part of the Gland Segmentation (GlaS) challenge contest to be held in conjunction with MICCAI 2015.

B. Comparative Algorithms

To compare the performance of the RPM framework against others gland segmentation approaches proposed in the literature, we implemented the methods by Farjam *et al.* [12], Naik *et al.*, Nguyen *et al.* [14], as well as the RPM in Matlab. We did not implement the approach by Gunduz-Demir *et al.* [15], but carefully followed the protocol given by the authors in order to make a fair comparison for the

Bilkent dataset. The glandular probability map was also used for segmentation, whereby pixels are labeled as belonging to glands if their probability is greater than or equal to 0.5. This threshold was empirically found to be the best for majority of our experiments (Table S1). We will denote this method by Thresholded Glandular Probability Map (TGPM), henceforth.

C. Data Pre-processing

To reduce stain variation across different images which may affect the performance of the algorithms that use color cues, we normalize all images in both datasets to the same target stain distribution using a recently proposed algorithm [27]. Moreover, white background outside tissue regions is cropped and not considered in the performance evaluation of all algorithms.

D. Implementation Details and Parameters

The values of the following parameters are determined on the training part of the Bilkent dataset and fixed in all the experiments. In glandular probability map generation, we construct superpixels on a regular grid of 15 pixels apart and set the compactness parameter in SLIC [24] to 10. Superpixels constructed in this way allow us to capture a thin epithelial boundary in healthy glands. As suggested in [22], for color histogram, we quantize pixels based on their $L^*a^*b^*$ intensity into 20 bins. In the calculation of scattering symmetric positive definite descriptor, we use a filter bank consisting of Gaussian filters at 6 scales and Gabor filters at 21 scales and 8 orientations. In the probability assignment, the number of trees in the random forest classifier is set to 50.

For constructing the set V of representative nuclei locations, we sample vertices which are at least $d = 8$ pixels apart from each other. For the seed area generation, after thresholding the glandular probability map at 0.5, we shrink the binary mask to 60% of the original area to get seed areas which provide more reliable information regarding the location of each individual glandular structure in the image. Note that the RPM is not too sensitive to the size of the shrunk seed areas as long as they can provide the approximate location of glandular structures (Table S2). The expanded bounding box for each seed area is constructed using $m_{\text{expand}} = 100$.

For the prior distribution of polygons given in (2), we calculate all the required statistics from the manually constructed polygons on the training part of the Bilkent dataset. This construction is simply done by connecting vertices from V lying along the boundary of a glandular structure to form a polygon surrounding each glandular structure. The following values of parameters are found: $\mu_l = 12.7, \sigma_l = 4.2, \mu_\theta = 157.2, \sigma_\theta = 23.3, \sigma_d = 10$. In the proposal moves of the RJMCMC, we empirically set $m = 10, r_{\text{add}} = r_{\text{delete}} = 0.45, r_{\text{change}} = 0.1$. The RPM is not sensitive to these parameter values (Table S3). We found the scale factor $\beta = 20$ to be a good compromise between the quality of the estimated random polygons and computational complexity of the RJMCMC simulation.

For false positive removal, we consider the following polygons as extreme and remove them: (a) polygons with the number of vertices less than or equal to 4, and (b) polygons with square root of their area less than or equal to the 10th percentile of the square root of the area of glands in training images. Both values were empirically found for the training part of the Bilkent dataset.

E. Evaluation Indices

We calculate pixel-wise segmentation accuracy in order to evaluate performance of different algorithms. For each image, the calculation is done at the pixel level and at the object level to examine the segmentation quality of the whole image as well as individual glands.

For the pixel-level segmentation accuracy, we use two evaluation indices, i.e. Jaccard [28] and Dice [29] indices. Given G a set of pixels marked as ground truth and O a set of pixels segmented as glandular structures, both indices measure the similarity between G and O. The Jaccard index is formulated as

$$\text{Jaccard}(G, O) = \frac{|G \cap O|}{|G \cup O|}, \tag{18}$$

and the Dice index is formulated as

$$\text{Dice}(G, O) = \frac{2|G \cap O|}{|G| + |O|}, \tag{19}$$

where $|\cdot|$ denotes the set cardinality. Both indices produce scores ranging between 0 and 1, where 1 indicates perfect segmentation.

For the object-level segmentation accuracy, we are concerned with: (A1) how well each segmented object overlaps with the ground truth objects, and (A2) how well each ground truth object overlaps the segmented objects. These two terms are not equivalent. Let O_i denote the set of pixels of the ith segmented object in O and G_i denote the set of pixels of ground truth objects in G that intersect O_i. Further, let $\tilde{G}_i$ denote the set of pixels of the ith ground truth object in G and $\tilde{O}_i$ denote the set of pixels of segmented objects in O that intersect $\tilde{G}_i$. We defined the object-level Jaccard index as

$$\text{Jaccard}_{\text{object}}(G, O) = \frac{1}{2}\Big[\sum_{i=1}^{n_O} \omega_i \text{Jaccard}(G_i, O_i) + \sum_{i=1}^{n_G} \tilde{\omega}_i \text{Jaccard}(\tilde{G}_i, \tilde{O}_i)\Big], \tag{20}$$

and the object-level Dice index as

$$\text{Dice}_{\text{object}}(G, O) = \frac{1}{2}\Big[\sum_{i=1}^{n_O} \omega_i \text{Dice}(G_i, O_i) + \sum_{i=1}^{n_G} \tilde{\omega}_i \text{Dice}(\tilde{G}_i, \tilde{O}_i)\Big], \tag{21}$$

where $\omega_i = |O_i| / \sum_{j=1}^{n_O} |O_j|$, $\tilde{\omega}_i = |\tilde{G}_i| / \sum_{j=1}^{n_G} |\tilde{G}_j|$, and n_O and n_G are the total number of segmented objects and the total number of ground truth objects, respectively. On the right-hand side of (20) and (21), the first term corresponds to (A1) and the second term corresponds to (A2). Both terms are weighted by the relative area of the objects. As shown in Figure S1, the object-level accuracy is a more appropriate measure of performance than the pixel-level accuracy for the task of gland segmentation.

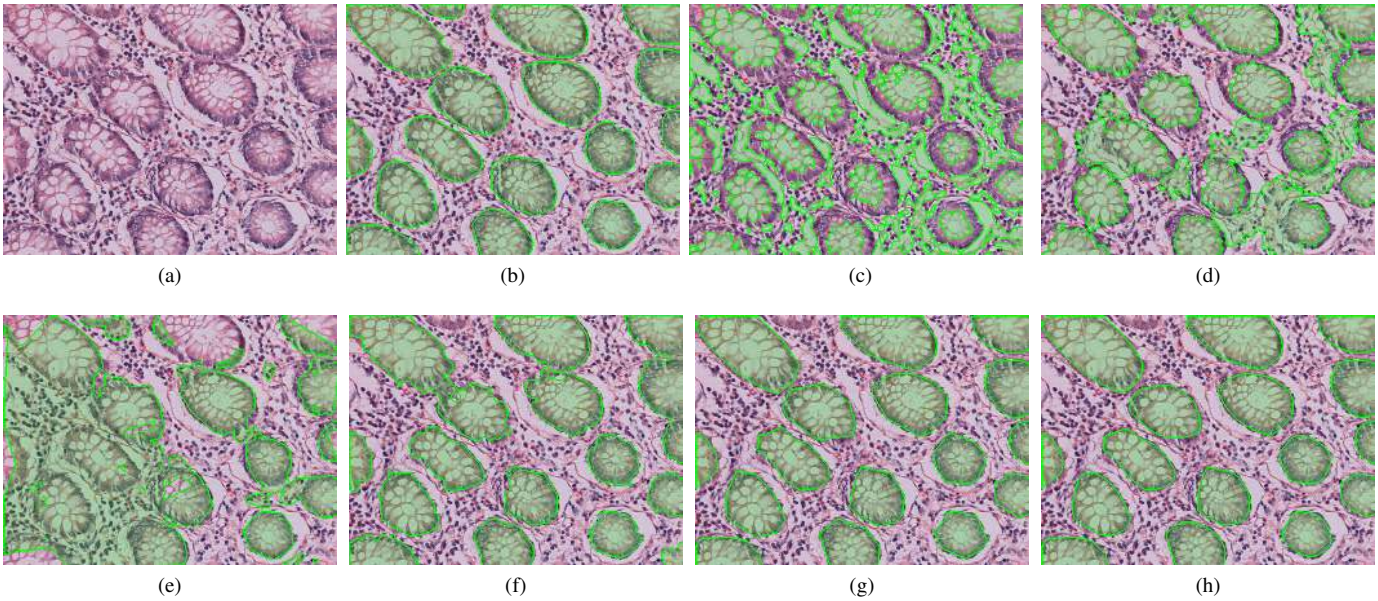


Fig. 4. Results for different gland segmentation approaches on an example healthy case in the Bilkent dataset: (a) Original image; (b) Ground truth; (c) Farjam *et al.* [12]; (d) Naik *et al.* [13]; (e) Nguyen *et al.* [14]; (f) TGPM; (g) RPM; (h) RPM (post-processed). Glandular structures are shown in green in (b)-(h).

F. Comparative Results

In this section, we present a variety of experiments to demonstrate the effectiveness and degree of generalization of the RPM. Due to stochastic nature of the RJMCMC simulation employed by the RPM, for each image, we run the algorithm for 5 repetitions and aggregate the mean segmentation accuracy result. All the competing algorithms are described in Section V-B. We did not perform the false positive removal routine used for the RPM on these algorithms due to the fact that they do not significantly benefit from the routine (Table S4). Other false positive removal methods may provide better quantitative results. All the experiments are conducted on a 2.8 GHz Intel Core i7 machine with 16GB RAM.

TABLE II
SEGMENTATION ACCURACIES ON THE BILKENT DATASET (EXPERIMENT 1)

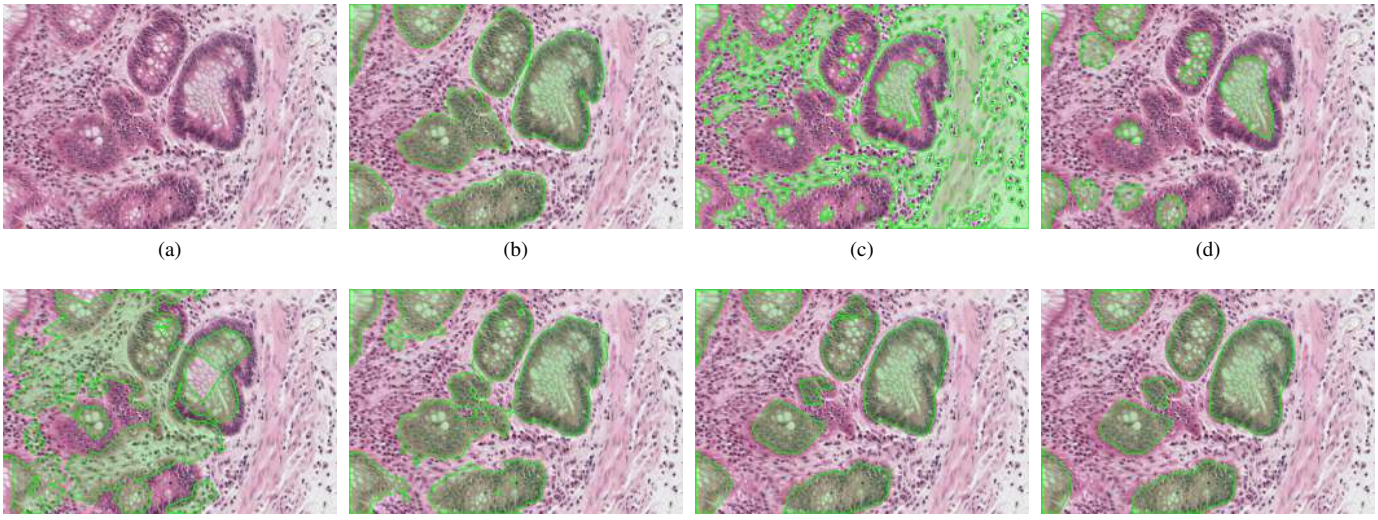
Approach	Pixel-Level		Object-Level	
	Jaccard	Dice	Jaccard	Dice
Gunduz-Demir *et al.* [15][a]	NA	0.89 ± 0.05	NA	NA
Farjam *et al.* [12]	0.43 ± 0.12	0.59 ± 0.12	0.39 ± 0.14	0.52 ± 0.15
Naik *et al.* [13]	0.57 ± 0.14	0.72 ± 0.12	0.43 ± 0.14	0.56 ± 0.14
Nguyen *et al.* [14]	0.52 ± 0.12	0.67 ± 0.12	0.35 ± 0.09	0.46 ± 0.09
TGPM	0.78 ± 0.07	0.87 ± 0.05	0.68 ± 0.12	0.77 ± 0.11
RPM	**0.82** ± 0.07	**0.90** ± 0.05	**0.78** ± 0.10	**0.85** ± 0.08

[a] The result is excerpted from [15].
NA means not available. The reported figures are the average ± standard deviation across images. Bold figures indicate the best results.

1) Experiment 1: Healthy Samples: In this experiment, we evaluate the segmentation accuracy of different algorithms on the Bilkent dataset [15]. All 72 images in the dataset are assessed as healthy by the experts. We follow the protocol given in [15], where 24 images are used for training and the rest for testing. Table II shows comparative results from the experiment. At the pixel level, the RPM produces a comparable result to that of Gunduz-Demir *et al.* [15], which is a benchmark algorithm on this dataset. As compared with all the other algorithms, the RPM yields the best results both at the pixel and the object levels. An important observation here is that the accuracy of the TGPM at the pixel level is one of the factors attributing to the success of the RPM. However, the main benefit of the RPM over the TGPM is that it tends not to fuse neighboring glands together (Figures 4f,4h). This owes to the fact that the RPM models each individual glandular object separately by utilizing nuclei vertices surrounding the glandular structure. Thus, the RPM achieves statistically significant segmentation accuracy at both pixel and object levels as compared to TGPM (Table S5).

Figure 4 shows an example of the segmentation results. Farjam *et al.* [12] (Figure 4c) labels a region as gland or non-gland based on statistics computed from its pixel-level texture features. This information alone, however, is not sufficient to assign an accurate label because some areas in the stromal region can have similar texture statistics as those of the luminal areas. Naik *et al.* [13] and Nguyen *et al.* [14] only utilize pixel-level color cue to label pixels as part of different constituents of a tissue. Consequently, areas with similar color characteristic but belonging to different tissue components will be assigned the same label. As shown in Figures 4d and 4e, some white areas in stroma are being mislabeled as lumen. The glandular probability map (Figure 4f), on the other hand, is generated using superpixel-based color and texture information. Thus, it provides a more reliable observation for the RPM.

2) Experiment 2: Healthy and Adenomatous Samples: The goal of this experiment is to compare the effectiveness of different segmentation algorithms on segmenting glandular structures in healthy and adenomatous samples. The experiment is conducted on 74 images (42 healthy, 32 adenomatous) from



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 5. Results for different gland segmentation approaches on an example adenomatous case in the Warwick-QU dataset: (a) Original image; (b) Ground truth; (c) Farjam *et al.* [12]; (d) Naik *et al.* [13]; (e) Nguyen *et al.* [14]; (f) TGPM; (g) RPM; (h) RPM (post-processed). Glandular structures are shown in green in (b)-(h).

TABLE III
SEGMENTATION ACCURACIES ON THE WARWICK-QU DATASET: HEALTHY AND ADENOMATOUS CASES (EXPERIMENT 2)

Approach	Pixel-Level		Object-Level	
	Jaccard	Dice	Jaccard	Dice
Farjam *et al.* [12]	0.33 ± 0.12	0.48 ± 0.13	0.30 ± 0.13	0.42 ± 0.15
Naik *et al.* [13]	0.46 ± 0.17	0.62 ± 0.17	0.46 ± 0.14	0.58 ± 0.14
Nguyen *et al.* [14]	0.45 ± 0.15	0.60 ± 0.16	0.29 ± 0.11	0.40 ± 0.12
TGPM	**0.77** ± 0.12	**0.87** ± 0.08	0.71 ± 0.13	0.80 ± 0.11
RPM	**0.77** ± 0.11	**0.87** ± 0.08	**0.74** ± 0.11	**0.82** ± 0.09

The reported figures are the average ± standard deviation across images. Bold figures indicate the best results.

the Warwick-QU dataset. In adenomas, damaged glands often are associated with morphological changes such as branching of a new gland due to regeneration, or enlargement of gland with serrated boundary due to inflammation, etc. Nonetheless, the glandular architecture still resembles closely to that of normal glands, and an individual gland is distinguishable.

Table III reports the results from 2-fold cross-validation, and Figure 5 shows qualitative comparison of the results. The RPM overall performs better than the others in this experiment. Despite producing comparable results to the TGPM at the pixel level, there is a statistically significance difference between the results of the RPM and the TGPM at the object level (Table S5). This is in line with the argument given in the Experiment 1. It is also worth noting that the object-level performance is more important than the pixel-level performance since the main goal here is to segment objects rather than pixels.

3) Experiment 3: Generalization for Healthy and Adenomatous Samples: We demonstrate the generalization of the RPM framework on a strong cross-validation experiment, where 72 images classified as healthy from the Bilkent dataset are used for training and 74 images classified as healthy and adenomatous from the Warwick-QU dataset are treated as the unseen data. The results are reported in Table IV. The highest segmentation accuracy among different methods on the strong cross-validation indicates a high degree of generalization of the RPM framework. Note that there is a statistically significance difference between the performance of the RPM and the TGPM at both pixel and object levels (Table S5).

TABLE IV
SEGMENATION ACCURACIES WHEN SEGMENTATION APPROACHES ARE TRAINED AND TESTED ON DIFFERENT DATASETS (EXPERIMENT 3)

Approach	Pixel-Level		Object-Level	
	Jaccard	Dice	Jaccard	Dice
Farjam *et al.* [12]	0.33 ± 0.12	0.48 ± 0.13	0.30 ± 0.13	0.42 ± 0.15
Naik *et al.* [13]	0.52 ± 0.13	0.67 ± 0.12	0.33 ± 0.12	0.45 ± 0.12
Nguyen *et al.* [14]	0.45 ± 0.15	0.60 ± 0.16	0.29 ± 0.11	0.40 ± 0.12
TGPM	0.69 ± 0.11	0.81 ± 0.09	0.64 ± 0.12	0.75 ± 0.10
RPM	**0.73** ± 0.12	**0.84** ± 0.09	**0.69** ± 0.13	**0.78** ± 0.12

The reported figures are the average ± standard deviation across images. Bold figures indicate the best results.

4) Experiment 4: Moderately Differentiated Samples: This experiment evaluates the viability of the gland segmentation algorithms on segmenting glandular structures in moderately differentiated adenocarcinomas. Disarray of glandular architecture renders the segmentation problem very challenging. Absence of goblet cells leaves the glandular lumen empty. Sometimes it is filled with necrotic debris or red blood cells, making the luminal texture close to that of stroma. Moreover, some glandular structures are purely consisting of proliferate nuclei.

The experiment is conducted on 47 moderately differentiated images from the Warwick-QU dataset. In the experiment, we employ 2-fold cross-validation, where one fold consists of 23 images and the other consists of 24 images. The summarized results are shown in Table V, and the results of an example case are shown in Figure 6. One of the potential reasons ascribing for the relatively low segmentation

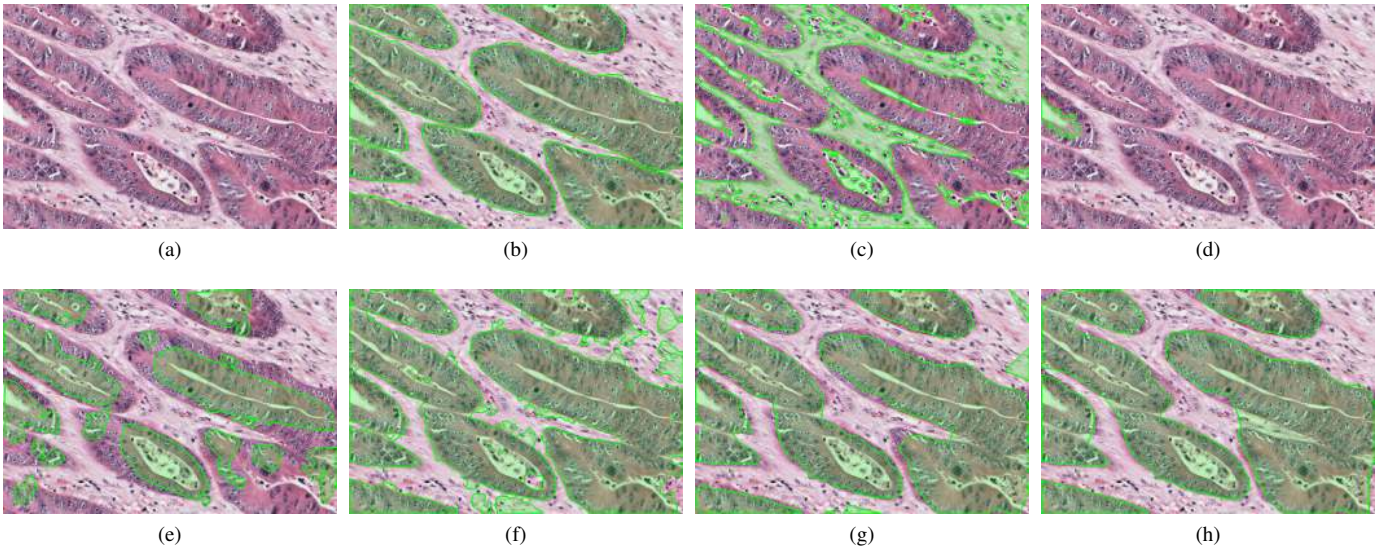


Fig. 6. Results for different gland segmentation approaches on an example moderately differentiated case in the Warwick-QU dataset: (a) Original image; (b) Ground truth; (c) Farjam *et al.* [12]; (d) Naik *et al.* [13]; (e) Nguyen *et al.* [14]; (f) TGPM; (g) RPM; (h) RPM (post-processed). Glandular structures are shown in green in (b)-(h).

accuracies attained by Naik *et al.* [13] and Nguyen *et al.* [14] is the simplistic structural assumption. They assume that, in glandular structure, luminal area is immediately followed by cytoplasmic area, and the cytoplasmic area is enclosed by a chain of epithelial boundary. This assumption does not hold true, in general, in intermediate and high grade cancer. On the contrary, the RPM does not make such an assumption regarding the arrangement of glandular components, but it instead infers the presence of glandular structures as a whole, based on texture and color cues, and the spatial connectivity of epithelial cells. Consequently, the RPM yields better performance among the considered algorithms in this experiment, and there is a statistically significant difference at the object-level results between the TGPM and the RPM (Table S5).

TABLE V
SEGMENTATION ACCURACIES ON THE WARWICK-QU DATASET: MODERATELY DIFFERENTIATED CASES (EXPERIMENT 4)

Approach	Pixel-Level		Object-Level	
	Jaccard	Dice	Jaccard	Dice
Farjam *et al.* [12]	0.18 ± 0.10	0.30 ± 0.14	0.15 ± 0.11	0.23 ± 0.14
Naik *et al.* [13]	0.16 ± 0.15	0.24 ± 0.21	0.16 ± 0.13	0.23 ± 0.18
Nguyen *et al.* [14]	0.39 ± 0.16	0.54 ± 0.18	0.27 ± 0.11	0.39 ± 0.13
TGPM	**0.72** ± 0.11	**0.83** ± 0.08	0.64 ± 0.15	0.74 ± 0.13
RPM	**0.72** ± 0.14	**0.83** ± 0.11	**0.67** ± 0.17	**0.76** ± 0.14

The reported figures are the average ± standard deviation across images. Bold figures indicate the best results.

5) Experiment 5: Moderately-to-Poorly and Poorly Differentiated Samples: In this experiment, we assess the practicability of the gland segmentation approaches on high grade adenocarcinomas. As compared to the gland segmentation in healthy to moderately differentiated cases, the nature of this problem is more convoluted, as high proliferation rates cause glandular structures highly indistinguishable. We conduct the experiment on 20 moderately-to-poorly differentiated and 24 poorly differentiated images, using 2-fold cross-validation. The quantitative results together with qualitative results on an example case are shown in Table VI and Figure 7, respectively. The RPM achieves significantly better results among all the algorithms considered in the comparison (Table S5).

TABLE VI
SEGMENTATION ACCURACIES ON THE WARWICK-QU DATASET: MODERATELY-TO-POORLY AND POORLY DIFFERENTIATED CASES (EXPERIMENT 5)

Approach	Pixel-Level		Object-Level	
	Jaccard	Dice	Jaccard	Dice
Farjam *et al.* [12]	0.18 ± 0.09	0.29 ± 0.12	0.13 ± 0.10	0.20 ± 0.13
Naik *et al.* [13]	0.22 ± 0.14	0.34 ± 0.19	0.18 ± 0.12	0.27 ± 0.15
Nguyen *et al.* [14]	0.46 ± 0.12	0.62 ± 0.12	0.31 ± 0.12	0.44 ± 0.13
TGPM	0.76 ± 0.08	0.86 ± 0.05	0.65 ± 0.13	0.76 ± 0.11
RPM	**0.79** ± 0.09	**0.88** ± 0.06	**0.70** ± 0.13	**0.79** ± 0.11

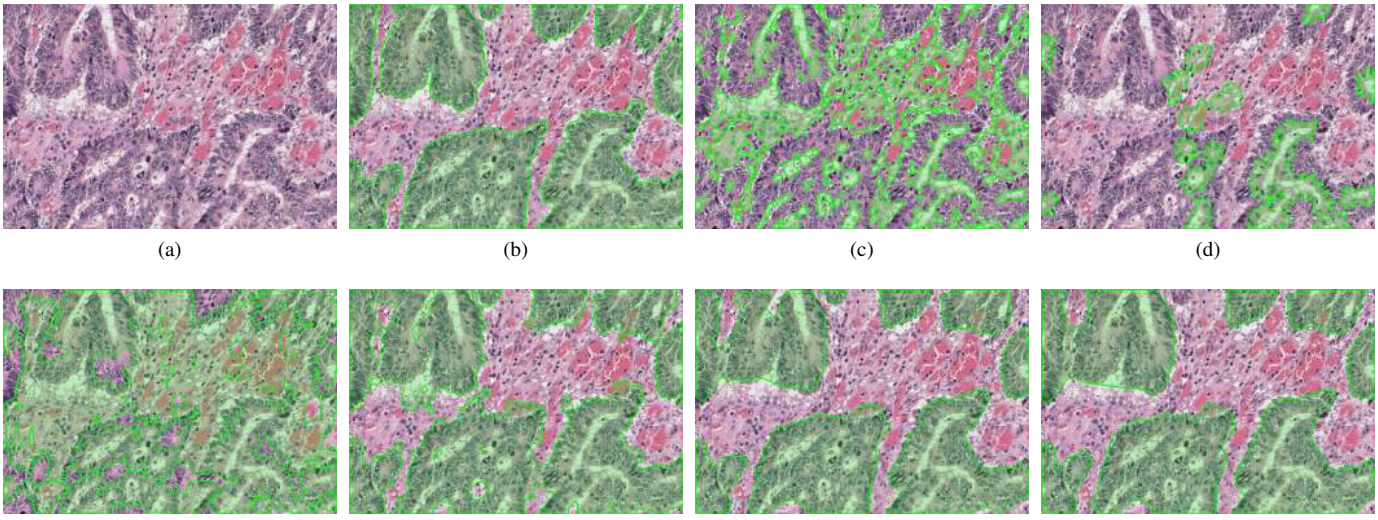
The reported figures are the average ± standard deviation across images. Bold figures indicate the best results.

G. Execution Times

In this section, we provide an analysis on computational efficiency of different segmentation algorithms in terms of the execution time. We only consider the execution time taken to segment glandular structures in test images, as in practice a training process is done offline. Table VII shows the execution time for each algorithm for the Experiment 2 (strong cross-validation). Similar to other stochastic modeling process, the RPM has higher computational complexity than deterministic algorithms, which results in the algorithm taking longer time than its competing algorithms to complete the segmentation.

VI. CONCLUSIONS

In this work, we have presented a novel random polygons model (RPM) for modeling glandular structures in Hematoxylin and Eosin stained histology images of human



(a) (b) (c) (d)

(e) (f) (g) (h)

Fig. 7. Results for different gland segmentation approaches on an example poorly differentiated case in the Warwick-QU dataset: (a) Original image; (b) Ground truth; (c) Farjam *et al.* [12]; (d) Naik *et al.* [13]; (e) Nguyen *et al.* [14]; (f) TGPM; (g) RPM; (h) RPM (post-processed). Glandular structures are shown in green in (b)-(h).

TABLE VII
EXECUTION TIMES AT TEST FROM EXPERIMENT 2 IN SECONDS

Approach	Execution Time
Farjam *et al.* [12]	3.0 ± 0.5
Naik *et al.* [13]	108.2 ± 74.4
Nguyen *et al.* [14]	6.1 ± 2.3
TGPM	36.6 ± 11.8
RPM	206.4 ± 332.5

The reported figures are the average ± standard deviation across images.

colon tissue. The model employs Bayesian inference and the Reversible-Jump Markov chain Monte Carlo (RJMCMC) simulation to estimate the glandular structures as polygons whose vertices are sampled from epithelial nuclei. Without a strict assumption about architectural regularity of glands, the RPM is shown to be broadly applicable for modeling glandular structures in all grades of colon cancer, excluding undifferentiated carcinomas. Results from comprehensive evaluation show the effectiveness and the degree of generalization of the RPM over other state-of-the-art gland segmentation algorithms. Due to its stochastic modeling to infer polygons, the RPM takes relatively higher amount of time to segment glandular structure. However, the polygon inference procedure can be easily adapted to take advantage of parallel computation.

ACKNOWLEDGMENTS

This paper was made possible by NPRP grant number NPRP5-1345-1-228 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors. Korsuk Sirinukunwattana acknowledges the partial financial support provided by the Department of Computer Science, University of Warwick, UK. We are grateful to the Warwick Impact Fund for funding an earlier version of this work covered by the US patent application number 61452293. The authors would like to thank Dr. Cigdem Gunduz-Demir for kindly providing us the dataset used in [15]. We are also grateful to Dr. Adnan Mujahid Khan and Dr. Khalid Masood for their inputs during the earlier stages of this work, to Prof. Ian Cree for useful discussions on the topic of gland segmentation, and to Dr. Richard Savage and Dr. Mudassar Iqbal for their comments on an earlier draft of this paper. Last but not the least, we are grateful to the anonymous reviewers for their constructive criticism, insightful comments, and positive suggestions which have undoubtedly resulted in improved quality of the manuscript.

REFERENCES

[1] C. W. Elston and I. O. Ellis, "Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: experience from a large study with long-term follow-up," *Histopathology*, vol. 19, no. 5, pp. 403–410, 1991.

[2] D. F. Gleason and G. T. Mellinger, "Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging," *The Journal of Urology*, vol. 167, no. 2, pp. 953–958, 2002.

[3] P. A. Humphrey, "Gleason grading and prognostic factors in carcinoma of the prostate," *Modern Pathology*, vol. 17, no. 3, pp. 292–306, 2004.

[4] S. R. Hamilton, L. A. Aaltonen, International Agency for Research on Cancer, World Health Organization, and others, *Pathology and genetics of tumours of the digestive system*. France: IARC Press, 2000.

[5] K. Derwinger, K. Kodeda, E. Bexe-Lindskog, and H. Taflin, "Tumour differentiation grade is associated with TNM staging and the risk of node metastasis in colorectal cancer," *Acta Oncologica*, vol. 49, no. 1, pp. 57–62, 2009.

[6] M. McCann, J. Ozolek, C. Castro, B. Parvin, and J. Kovacevic, "Automated histology analysis: Opportunities for signal processing," *IEEE Signal Processing Magazine*, vol. 32, no. 1, pp. 78–87, 2015.

[7] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Reviews in Biomedical Engineering*, vol. 2, pp. 147–171, 2009.

[8] T. J. Fuchs and J. M. Buhmann, "Computational pathology: Challenges and promises for tissue analysis," *Computerized Medical Imaging and Graphics*, vol. 35, no. 7, pp. 515–530, 2011.

[9] M. Fleming, S. Ravula, S. F. Tatishchev, and H. L. Wang, "Colorectal carcinoma: pathologic aspects," *Journal of Gastrointestinal Oncology*, vol. 3, no. 3, pp. 153–173, 2012.

[10] P. Green, "Reversible jump Markov chain Monte Carlo computation and Bayesian model determination," *Biometrika*, vol. 82, no. 4, pp. 711–732, 1995.
[11] W. K. Hastings, "Monte Carlo sampling methods using Markov chains and their applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, 1970.
[12] R. Farjam, H. Soltanian-Zadeh, K. Jafari-Khouzani, and R. Zoroofi, "An image analysis approach for automatic malignancy determination of prostate pathological images," *Cytometry Part B: Clinical Cytometry*, vol. 72, no. 4, pp. 227–240, 2007.
[13] S. Naik, S. Doyle, A. Madabhushi, J. Tomaszewski, and M. Feldman, "Gland segmentation and Gleason grading of prostate histology by integrating low-, high-level and domain specific information," in *Workshop on Microscopic Image Analysis with Applications in Biology*, 2007.
[14] K. Nguyen, A. Jain, and R. Allen, "Automated gland segmentation and classification for Gleason grading of prostate tissue images," in *IEEE International Conference on Pattern Recognition (ICPR)*, 2010, pp. 1497–1500.
[15] C. Gunduz-Demir, M. Kandemir, A. Tosun, and C. Sokmensuer, "Automatic segmentation of colon glands using object-graphs," *Medical Image Analysis*, vol. 14, no. 1, pp. 1–12, 2010.
[16] A. Fakhrzadeh, E. Spörndly-Nees, L. Holm, and C. L. Luengo Hendriks, "Analyzing tubular tissue in histopathological thin sections," in *IEEE International Conference on Digital Image Computing Techniques and Applications*, 2012, pp. 1–6.
[17] H. Fu, G. Qiu, J. Shu, and M. Ilyas, "A novel polar space random field model for the detection of glandular structures," *IEEE Transaction On Medical Imaging*, pp. 764–776, 2014.
[18] B. Sabata, B. Babenko, R. Monroe, and C. Srinivas, "Automated analysis of PIN-4 stained prostate needle biopsies," in *Prostate Cancer Imaging: Computer-Aided Diagnosis, Prognosis, and Intervention*, 2010, pp. 89–100.
[19] I. L. Dryden, M. R. Scarr, and C. C. Taylor, "Bayesian texture segmentation of weed and crop images using reversible jump Markov chain Monte Carlo methods," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 52, no. 1, pp. 31–50, 2003.
[20] I. L. Dryden, R. Farnoosh, and C. C. Taylor, "Image segmentation using Voronoi polygons and MCMC, with application to muscle fibre images," *Journal of Applied Statistics*, vol. 33, no. 6, pp. 609–622, 2006.
[21] Y. Li, J. Li, and M. A. Chapman, "Segmentation of SAR intensity imagery with a Voronoi tessellation, Bayesian inference, and reversible jump MCMC algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 4, pp. 1872–1881, 2010.
[22] K. Sirinukunwattana, D. Snead, and N. Rajpoot, "A novel texture descriptor for detection of glandular structures in colon histology images," in *SPIE Medical Imaging*. International Society for Optics and Photonics, 2015, pp. 94 200S–94 200S.
[23] A. Lucchi, K. Smith, R. Achanta, V. Lepetit, and P. Fua, "A fully automated approach to segmentation of irregularly shaped cellular structures in EM images," in *Medical Image Computing and Computer-Assisted Intervention MICCAI*, 2010, pp. 463–471.
[24] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.
[25] M. Macenko, M. Niethammer, J. Marron, D. Borland, J. T. Woosley, X. Guan, C. Schmitt, and N. E. Thomas, "A method for normalizing histology slides for quantitative analysis," in *International Symposium on Biomedical Imaging*, vol. 9, 2009, pp. 1107–1110.
[26] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
[27] A. Khan, N. Rajpoot, D. Treanor, and D. Magee, "A nonlinear mapping approach to stain normalization in digital histopathology images using image-specific color deconvolution," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 6, pp. 1729–1738, 2014.
[28] P. Jaccard, "Étude comparative de la distribution florale dans une portion des Alpes et des Jura," *Bulletin de la Société Vaudoise des Sciences Naturelles*, vol. 37, pp. 547–579, 1901.
[29] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.