Computerized Medical Imaging and Graphics 93 (2021) 101974

Contents lists available at ScienceDirect

Computerized Medical Imaging and Graphics

journal homepage: www.elsevier.com/locate/compmedimag

An unsupervised method for histological image segmentation based on tissue cluster level graph cut

Hongming Xu [a], Lina Liu [b], Xiujuan Lei [c], Mrinal Mandal [b], Cheng Lu [d,*]

[a] School of Biomedical Engineering at Faculty of Electronic Information and Electrical Engineering, Dalian University of Technology, Dalian, Liaoning 116024, China
[b] Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB T6G 1H9, Canada
[c] College of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi 710119, China
[d] Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH 44106, USA

ARTICLE INFO

Keywords:
Histological image analysis
Unsupervised segmentation
Objects clustering
Graph cut

ABSTRACT

While deep learning models have demonstrated outstanding performance in medical image segmentation tasks, histological annotations for training deep learning models are usually challenging to obtain, due to the effort and experience required to carefully delineate tissue structures. In this study, we propose an unsupervised method, termed as tissue cluster level graph cut (TisCut), for segmenting histological images into meaningful compartments (e.g., tumor or non-tumor regions), which aims at assisting histological annotations for downstream supervised models. The TisCut consists of three modules. First, histological tissue objects are clustered based on their spatial proximity and morphological features. The Voronoi diagram is then constructed based on tissue object clustering. In the last module, morphological features computed from the Voronoi diagram are integrated into a region adjacency graph. Image partition is then performed to divide the image into meaningful compartments by using the graph cut algorithm. The TisCut has been evaluated on three histological image sets for necrosis and melanoma detections. Experiments show that the TisCut could provide a comparative performance with U-Net models, which achieves about 70% and 85% Jaccard index coefficients in partitioning brain and skin histological images, respectively. In addition, it shows the potential to be used for generating histological annotations when training masks are difficult to collect for supervised segmentation models.

1. Introduction

Histological examination of tissue pathology slides is the gold standard for diagnosing and grading various cancers such as skin melanoma (Xu et al., 2017). Traditionally, pathologists visually examine cytological features and tissue distributions under a microscope to determine whether a glass slide contains any malignant region and if so, the malignancy level. However, visual examinations performed by pathologists are typically subjective and prone to inter- and intra-observer variations. To achieve an objective judgement, quantitative image analysis using computerized methods is crucial. Histological image segmentation is usually the first step in quantitative analysis of tumor pathology slides that include heterogenous tissue components. For example, identifying melanoma regions from Ki-67 stained lymph node histological slides is the prerequisite in quantifying tumor proliferation index (Alheejawi et al., 2019; Xing et al., 2013). Distinguishing necrotic regions from lymphocytic infiltrations helps in constructing histological features for tumor subtype identification and patients' survival prognosis (Saltz et al., 2018). Therefore, segmenting histological image into meaningful compartments plays a significant role in computer-aided diagnosis and grading system of cancers.

In the literature, one category of histology image segmentations was based on pixel- or grid-level analysis. These studies first characterized each pixel or grid (i.e., a small image patch) of an image using low-level color/texture descriptors. The classifier was then applied to assign pixels or grids into different labels. The connected pixels or grids with the same label were considered to be the same tissue. Doyle et al. (2012) proposed to compute texture features around a neighborhood of every pixel, and then used Boosted Bayesian classifiers to detect cancerous regions in prostate biopsy images. Wang et al. (2009) proposed a method for squamous epithelium segmentation, which first divides the image into a number of blocks and then performs segmentation by using the SVM classifier with extracted texture features within small blocks. Manivannan et al. (2016) proposed a method to segment glands in colon

* Corresponding author.
E-mail address: cxl884@case.edu (C. Lu).

https://doi.org/10.1016/j.compmedimag.2021.101974
Received 27 April 2021; Received in revised form 11 July 2021; Accepted 17 August 2021
Available online 21 August 2021

histopathology images, which first extracts grid-level features and then performs classification by using a structure learning approach. Since this category of techniques relies on pixel- or grid-level texture analysis, their performances are likely be affected by color variation and image noise.

Another category of studies performed object-level analysis instead of pixel-level analysis for histological image segmentation. These studies first defined objects to represent tissue components, and then extracted object-level features for image partitioning algorithm. Tosun et al. (2009) and Gunduz-Demir et al. (2010) proposed to represent colon histological image by a set of primitive circular objects. Region growing methods integrated with object-level features were then utilized to segment the image. The segmentation performances shown in these two studies depend on initially selected seed points used with the region growing method. Simsek et al. (2012) proposed a multilevel segmentation method that first constructs a graph based on randomly selected tissue objects and then cluster graph nodes based on object co-occurrence features, such that colon histological image is divided into different regions. Because graph nodes were selected by random sampling, this method did not guarantee to provide a global optimum segmentation.

Recently deep learning techniques have been explored for histological image segmentation and provided superior performances over existing methods. Xu et al. (2015) applied transfer learning on deep convolutional neural network to extract features that were used by SVM classifier for brain histology image segmentation. Bejnordi et al. (2017) reported a number of deep learning methods for detection of lymph node metastases in H&E stained breast cancer pathology slides. Ronneberger et al. (2015) designed a U-Net deep learning model which has been used for different biomedical image segmentations. Chan et al. (2019) designed a HistoSegNet that performs histological image segmentation based on convolutional neural networks. Although deep learning methods are powerful to learn image features, they generally require a large number of training images and associated annotations, which are difficult to collect due to the effort and experience required to carefully delineate tissue structures, especially for histological images (Xing et al., 2017; Amgad et al., 2019).

In this paper, we propose an unsupervised method, termed as tissue cluster level graph cut (TisCut), which automatically partitions the histology image into meaningful compartments. The proposed method is based on the observation that different tissue compartments in pathology slides usually vary in terms of cellular structure, density and type. Overall there are three main contributions for this study. First, we design a novel unsupervised image segmentation method that performs image partition based on tissue cluster level features computed from the Voronoi diagram. Second, the proposed method has been evaluated on three histological image sets and has shown a promising performance in necrosis and melanoma detections, which would be able to assist in computer-aided disease diagnosis and grading by quantitative histological image analysis. Finally, we show that the TisCut technique has the potential to be used for annotating histological images to train the U-Net model, which would be able to overcome the difficulties in collecting histological annotation masks for supervised learning-based segmentations.

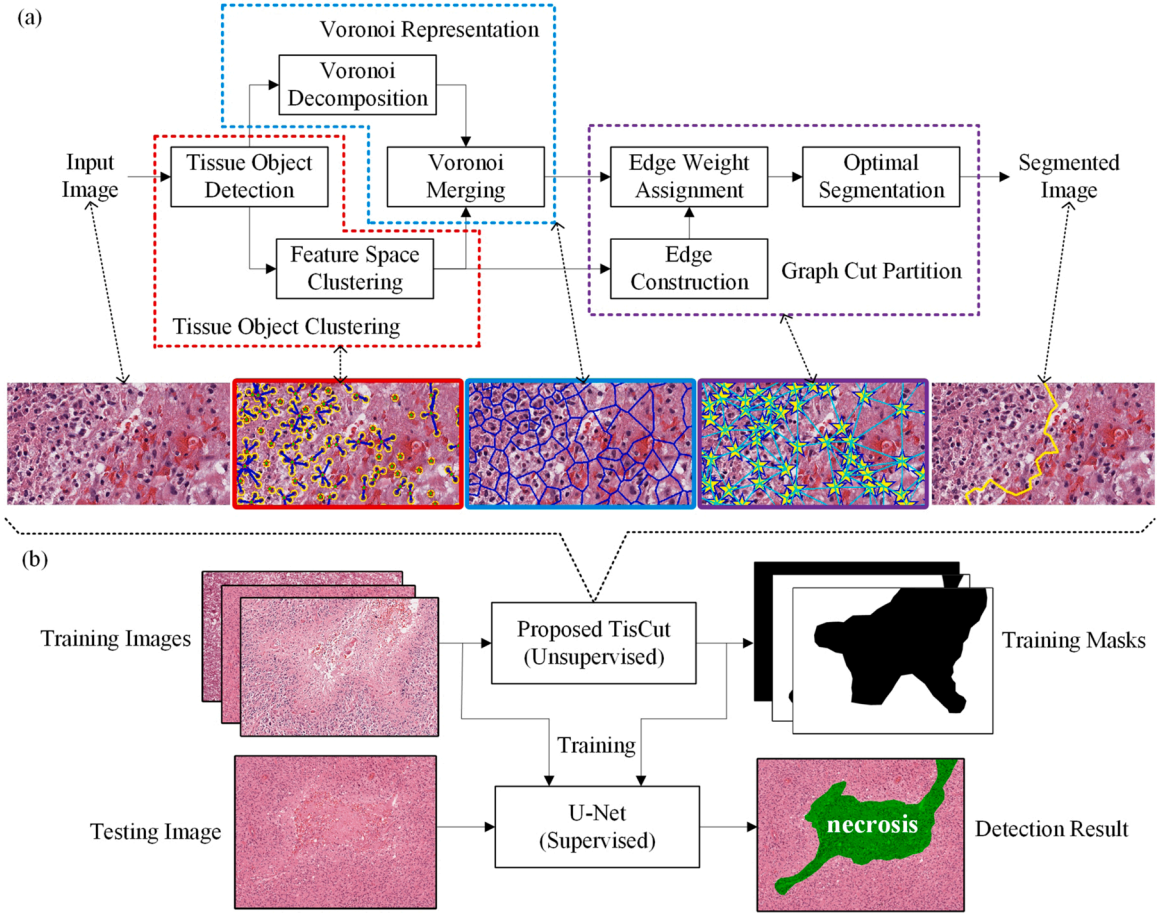

Fig. 1. Pipeline of the TisCut technique and its applications. (a) Schematic of the TisCut technique. First cell nuclei with spatially close distance and similar size are clustered together. The image is then represented by the Vonoroi diagram based on nuclei clustering. Finally a region adjacency graph is constructed such that the image is partitioned by using the graph cut algorithm. (b) Potentially extended applications of the TisCut technique to train supervised models for histological tissue detection.

The organization of this paper is as follows. Section 2 describes the TisCut, followed by performance evaluations in Section 3. The conclusion is presented in Section 4.

2. Tissue cluster level graph cut (TisCut)

The schematic of the proposed TisCut technique is shown in Fig. 1(a). It is observed from Fig. 1(a) that the TisCut consists of three main modules. The first module is the tissue object clustering, where a set of objects of interest (e.g., cell nuclei) are detected and grouped into different clusters based on extracted object features. The second module is the Voronoi representation, which divides the image into a number of regional components by merging Voronoi polygons based on tissue object clustering. The third module is the graph cut partition, where the graph is built based on tissue object clustering and then cut off by utilizing histological features computed in the Voronoi diagram such that the image is divided into different meaningful compartments. Fig. 1(b) illustrates the potentially extended applications of the TisCut technique in training supervised models for histological tissue segmentation. Table 1 lists symbols and notations employed in this paper. In the following, the TisCut technique is explained in the context of brain tumor pathology image partition, which aims at segmenting cell necrosis regions.

2.1. Tissue object clustering

The proposed TisCut technique considers the optimal image partition as a tissue cluster level graph cut problem, where the features of local tissue components are utilized. To this end, this module first extracts objects of interest (i.e., cell nuclei) in the histological image, and then groups them into local object clusters. The details of this module are as follows.

2.1.1. Tissue object detection

In histology image analysis, objects of interest are usually cell nuclei or glands, which may differ in different tissue slides and applications. In this work, objects of interest in brain and skin histological images are cell nuclei (see Fig. 2(a)). To efficiently segment cell nuclei, a local optimal thresholding method (Lu et al., 2012) that segments cell nuclei based on area and shape analysis is applied. This method adaptively determines the optimal threshold by minimizing a cost function to ensure that segmented cell nuclei have elliptical shape and acceptable size. Fig. 2(a) shows a brain tumor histological image, and Fig. 2(b) shows nuclei segmentation in the image where nuclei boundaries are highlighted by yellow contours. Let us denote the set of segmented cell nuclei as $\mathcal{O} = \{o_i\}_{i=1,\dots,N}$, where N is the total number of cell nuclei.

2.1.2. Feature space clustering

After segmenting cell nuclei, spatial and morphological features of cell nuclei are computed for feature space analysis. Since each nucleus corresponds to a feature vector $x_i \in \boldsymbol{R}^d$, where d indicates the feature dimension, all segmented cell nuclei correspond to a set of features $\mathcal{X} = \{x_i\}_{i=1,\dots,N}$. To group cell nuclei into clusters, mean-shift algorithm (Comaniciu and Meer, 2002) is applied on feature set $\mathcal{X}$. Mean-shift is selected mainly because it does not require a prior specification for the number of clusters. To segment necrotic regions in brain histological images, we explore a 3D feature space which includes 2D spatial coordinates (i.e., centroid location) of cell nuclei in the image and nuclei size. These features are chosen based on the observation that the same type cell nuclei usually locate closely and have similar size. The bandwidth h for the mean-shift algorithm is set as 60 in brain histological image segmentation (at 40× magnification). This indicates that for each cell nucleus in 3D feature space, its neighbors within a sphere with a radius of 60 are used to estimate the density for clustering. Let us denote $\mathcal{T} = \{t_j\}_{j=1,\dots,K}$ as the set of tissue object clusters, and $\mathcal{C} = \{c_j\}_{j=1,\dots,K}$ as the set of 2D spatial components of cluster centers, where K is the number of clusters. Fig. 2(c) and (d) illustrate nuclei clustering results with binary and original color images, where green stars indicate 2D spatial locations of cluster centers (i.e., point set $\mathcal{C}$). As observed in Fig. 2 (c) and (d), nuclei with close distances and similar sizes have been clustered together.

Table 1
List of symbols employed in this paper.

Symbols	Descriptions
$\mathcal{O} = \{o_i\}_{i=1,\dots,N}$	Set of tissue objects
$\mathcal{X} = \{x_i\}_{i=1,\dots,N}$	Set of tissue object features
$\mathcal{T} = \{t_j\}_{j=1,\dots,K}$	Set of tissue object clusters
$\mathcal{C} = \{c_j\}_{j=1,\dots,K}$	Set of 2D spatial components of clusters centers
$\mathcal{V} = \{v_i\}_{i=1,\dots,N}$	Set of Voronoi polygons for tissue objects
$\mathcal{U} = \{u_j\}_{j=1,\dots,K}$	Set of Voronoi regions for tissue objects clusters
$\mathcal{G} = (\mathcal{R}, \mathcal{E}, \mathcal{W})$	Graph representation
S, T	Two terminal vertices for the graph structure
$\mathcal{R} = \mathcal{C} \cup \{S, T\}$	Set of RAG vertices
$\mathcal{E}$	Set of edges connecting graph nodes
$\mathcal{W}$	Set of cost weights on graph edges

2.2. Voronoi representation

Voronoi diagram has been applied for extracting spatial features of tissue objects in histology image analysis (Gurcan et al., 2009; Xu et al., 2018). Unlike existing studies such as (Xu et al., 2018) that constructs Voronoi diagram by using nuclear seeds directly, in this module, we propose to decompose the whole image into a collection of Voronoi regions based on tissue object clustering (i.e., nuclei clusters). The details of this module are explained below.

2.2.1. Voronoi decomposition

Let us consider tissue object centers (i.e., nuclear seeds) as the 2D points in the plane. The whole histological image is then decomposed into a set of Voronoi polygons, which are denoted as $\mathcal{V} = \{v_i\}_{i=1,\dots,N}$. Such decomposition has the property that an arbitrary pixel p in the image is assigned to the polygon v_i such that it has the closest distance to the spatial center of object o_i than any other object centers. Fig. 2(e) shows the Voronoi diagram constructed by using nuclear seeds in the image. It is observed that each Voronoi polygon encompasses a nucleus and its surrounding region.

2.2.2. Voronoi merging

Based on pre-computed tissue object clusters $\mathcal{T}$, in this step, the Voronoi polygons $\mathcal{V}$ are merged into larger Voronoi regions which are denoted as $\mathcal{U} = \{u_j\}_{j=1,\dots,K}$. Specifically, a Voronoi region u_j consists of several Voronoi polygons v_i such that,

$$u_j = \{v_i | o_i \in t_j\}, \quad i \in 1, 2, \dots, N \tag{1}$$

In other words, u_j is composed of several Voronoi polygons v_i, where cell nuclei within polygons v_i belong to the same cluster t_j. The reasons behind representing an image as a set of larger Voronoi regions are two folds. First there is a large number of cell nuclei in histology image, and hence clusters of cell nuclei could lead to efficient partition by subsequent graph cut algorithm. Second clusters of cell nuclei help in capturing local region properties and reducing feature variants and noise that commonly exist in histology images. Fig. 2(f) shows the example of Voronoi regions constructed based on tissue object clusters. As observed in Fig. 2(f), the Voronoi polygons in Fig. 2(e) have been merged together based on nuclei clustering. The Voronoi regions in Fig. 2(f) is a higher level representation of Voronoi polygons shown in Fig. 2(e). These larger Voronoi regions are treated as basic elements for subsequent image segmentation.

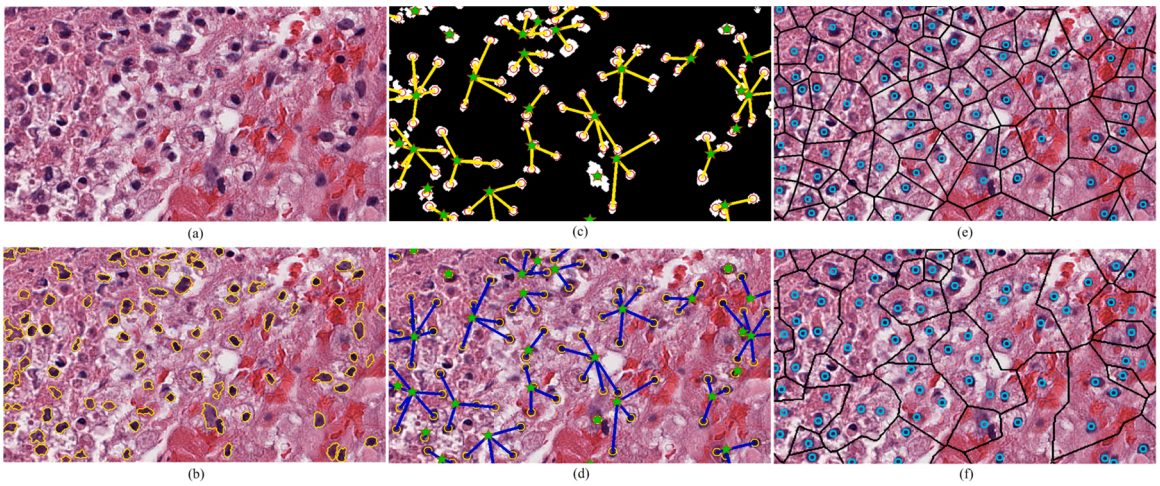

Fig. 2. Illustration of tissue object clustering and Voronoi representation. (a) A H&E stained brain histological image. (b) Nuclei segmentation result. (c) Nuclei clustering shown with a binary image. (d) Nuclei clustering shown with the color image. (e) Voronoi decomposition. (f) Voronoi merging. Note that green stars in (c) and (d) indicate 2D spatial components of cluster centers. Cyan circles in (e) and (f) indicate the locations of cell nuclei.

2.3. Graph cut partition

A graph is a structure that consists of vertices, edges and weights on the edges, which is denoted by $\mathcal{G} = (\mathcal{R}, \mathcal{E}, \mathcal{W})$. Unlike existing techniques that use pixels or object centers as vertices to construct a simple grid graph for an image, we propose to incorporate tissue object clusters into the region adjacency graph (RAG) (Trémeau and Colantoni, 2000) to describe a histology image. Details of this module are as follows.

2.3.1. Edge construction

Let us consider 2D spatial components of cluster centers (i.e., point set $\mathcal{C}$) as graph nodes. We assign edges and weights between these nodes to construct RAG. For graph node connection, Delaunay triangulation is applied to build interconnected edges. The Delaunay triangulation for the set $\mathcal{C}$ with K points in a plane is a triangulation DT($\mathcal{C}$) such that no point in $\mathcal{C}$ is inside the circumcircle of any triangle in DT($\mathcal{C}$). Fig. 3(a) shows an example of generated graph edges by Delaunay triangulation. Note that in Fig. 3(a) the interconnected edges between graph nodes (e.g., yellow stars) are called *n-links* which encode the neighborhood/proximity information. Since Delaunay triangulation is used to construct neighboring system, there are three interconnected *n-links* for each node in the RAG. Let $\mathcal{N}$ denote the set of all pairs of $\{m, n\}$ neighboring interconnected edges, where $1 \leq m, n \leq K$. Besides graph nodes $\mathcal{C}$, there are two specially designed terminal nodes S(*source*) and T(*sink*) that represent "foreground" and "background" labels. The edges connecting graph nodes to terminals are called *t-links*. Each node c_j has two *t-links* which are denoted by $\{\{c_j, S\}$ and $\{c_j, T\}\}$. Fig. 3(b) illustrates an example of constructing graph neighboring system including both *n-* and *t-links*.

2.3.2. Edge weight assignment

Let $A = [A_1, A_2, \ldots, A_K]$ be a binary vector whose components A_j, $1 \leq j \leq K$, specify the class label (e.g., necrosis or non-necrosis) of object clusters $t_j \in \mathcal{T}$. The vector A describes a segmentation for tissue object clusters in the image. Since the vector A has K components and each component has a binary value, there exist 2^K possible segmentations. To obtain an optimal segmentation, an energy function that encodes the regional and boundary properties (Boykov and Funka-Lea, 2006) of tissue object clusters is defined as follows:

$$E(A) = \lambda R(A) + B(A) \tag{2}$$

where $R(A)$ and $B(A)$ are functions encoding regional and boundary terms, respectively. λ is a coefficient that balances the importance of regional and boundary terms. In this work, λ is empirically set as 0.25. Note that a smaller λ value indicates a relatively less importance for regional term, which helps in producing large connected regions during image segmentation. The regional term corresponds to the assigned weights/costs on *t-links*, while the boundary term corresponds to the assigned weights/costs on *n-links*.

The regional term $R(A)$ in Eq. (2) is computed as:

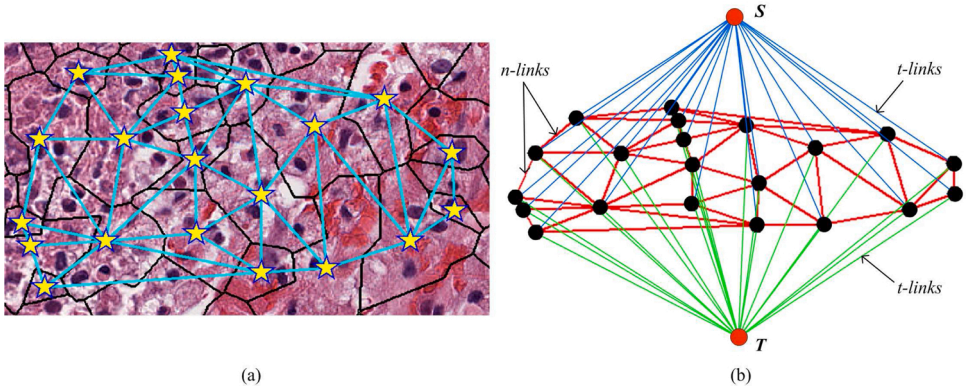

Fig. 3. Illustration of RAG construction. (a) Delaunay triangulation on tissue object clusters. (b) Example of constructing graph neighboring system. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$R(A) = \sum_j R_j(A_j) \tag{3}$$

where each regional term $R_j(A_j)$ computes the penalty for assigning object cluster t_j to "foreground" or "background" labels. Motivated by the MAP-MRF model (Boykov et al., 2001), a negative log-likelihood function is used to compute the penalty, which is defined as follows:

$$R_j(\mathcal{L}) = -lnP(F_j|\mathcal{L}) \tag{4}$$

where $P(F_j|\mathcal{L})$ is a Gaussian probability, that is,

$$P(F_j|\mathcal{L}) = \frac{1}{\sqrt{(2\pi)^q|\Sigma|}} exp\left(-\frac{1}{2}(F_j - \mu_{\mathcal{L}})^T \Sigma^{-1}(F_j - \mu_{\mathcal{L}}) \right) \tag{5}$$

where F_j is the feature/property for the jth object cluster t_j, and $\mu_{\mathcal{L}} \in \mathcal{R}^q$ is the mean feature vector (with dimension q) for one class. For binary case where $\mathcal{L} = \{1, 0\}$ or {*positive*, *negative*}, $\mu_{\mathcal{L}} = \mu_1$ or $\mu_{\mathcal{L}} = \mu_0$. $\Sigma \in \mathcal{R}^{q\times q}$ is the covariance matrix, and $|\Sigma|$ is the determinant. In this work, the parameters $\mu_{\mathcal{L}}$ and $|\Sigma|$ are adaptively estimated by fitting features F_j, $1 \le j \le K$, with two Gaussian mixture models using the Expectation-Maximization (EM) algorithm (Lu et al., 2013). The feature F_j can be adaptively computed based on different applications. For necrosis segmentation evaluated in this study, a simple but effective feature that estimates the ratio between each Voronoi area and its enclosed nuclei area, is proposed as follows:

$$F_j = \frac{\text{Area}(u_j)}{\text{Area}(\text{nuclei} \in u_j)}, \quad F_j \in \mathcal{R}^1 \tag{6}$$

This feature is motivated based on the observation that cell nuclei in non-necrosis tumor regions tend to have larger sizes and higher densities than those of necrosis regions. Thus a larger value of F_j in Eq. (6) is expected for necrosis regions.

The boundary term in Eq. (2) is computed as:

$$B(A) = \sum_{\{m,n\}\in\mathcal{N}} B_{m,n} \cdot \mathbf{1}(A_m \neq A_n) \tag{7}$$

where $\mathbf{1}(\cdot)$ is an indicator function, i.e.,

$$\mathbf{1}(\text{expression}) = \begin{cases} 1 & \text{if expression is true} \\ 0 & \text{if expression is false} \end{cases} \tag{8}$$

The boundary term $B_{m,n}$ encodes boundary penalties for a potential segmentation. Since we are trying to find the minimum cut from the whole RAG, a large penalty of $B_{m,n}$ value should indicate high similarity between object clusters t_m and t_n. Meanwhile, a small penalty of $B_{m,n}$ value should be used when object clusters t_m and t_n have significantly different properties. The boundary term $B_{m,n}$ can be computed based on object cluster features. For two class segmentation problem in this study, we define the boundary penalty function as follows:

$$B_{m,n} = exp\left(-\left| \frac{(F_m - F_n)^2}{2(\mu_1 - \mu_0)^2} \right| \right) \tag{9}$$

Note that the denominator, inside the exponential function, encodes the feature mean difference between two classes. The numerator, inside the exponential function, computes the distance between two object clusters t_m and t_n in feature space. The largest penalty occurs as a value of 1, when features of two object clusters t_m and t_n are identical, i.e., $F_m = F_n$.

2.3.3. Optimal segmentation

In this section, we seek for the minimum cost/energy cut on the RAG to achieve an optimal segmentation. It is noted that in the RAG the edge set $\mathcal{E}$ consists of *n-links* and *t-links*, i.e.,

$$\mathcal{E} = \mathcal{N} \bigcup_{c_j \in \mathcal{C}} \{\{c_j, S\}, \{c_j, T\}\} \tag{10}$$

Table 2 lists RAG edge weights. The RAG is partitioned by using the graph cut algorithm (Boykov and Funka-Lea, 2006), which provides a global minimum cost cut on the graph. The global minimum cost cut is obtained in polynomial time by the max-flow algorithm, given that all edge weights are non-negative. For more details, readers are referred to Boykov and Funka-Lea (2006). After graph cut segmentation, the small holes that might exist in foreground or background regions are filled by morphological operations. Fig. 4 shows a cropped example for tissue cluster level graph cut partition. In Fig. 4, the yellow contour indicates the segmentation boundary, while letters A, B indicate two classes of Voronoi regions obtained by the proposed segmentation algorithm. The regions overlapped with letters B belong to necrotic tumors.

3. Evaluation results

The proposed TisCut is a general framework for histological image segmentation. It can be adaptively adjusted and applied for different segmentation problems according to image characteristics. In this section, we provide evaluations and comparisons on three different test cohorts, which consist of brain, skin and lung histological images, respectively. We aim at detecting necrosis and melanoma regions from brain and skin testing sets, respectively. The lung data set is independently used for verifying whether the proposed technique could help in generating annotations for supervised deep learning models.

3.1. Datasets description

Dataset-I: The first dataset consists of 35 H&E stained brain histological images diagnosed with glioblastoma multiforme. These images have different sizes, ranging from 500 to 5000 pixels in each dimension. The pixel resolution for this dataset is about 0.50 μm/pixel. The brain histological images, along with ground truth annotations (i.e., necrosis vs non-necrosis), were obtained by following Xu et al. (2015). This dataset was used to design necrosis detection algorithms, which would help in brain tumor diagnosis and prognosis.

Dataset-II: The second dataset consists of 30 Ki-67 stained skin histological images with the average size about 1000 × 1000 pixels per image. These images were collected from the Cross Cancer Institute at University of Alberta, for skin melanoma diagnosis and grading study. The images including melanoma and non-melanoma regions were cropped from the whole slide images at 40× magnification (Alheejawi et al., 2019; Xu et al., 2017), where the pixel resolution is about 0.25 μm/pixel. The ground truth melanoma regions in this dataset were manually labeled by an experienced pathologist. These cropped images are regions of interest which are used for evaluating tumor proliferation index (Xing et al., 2013). For skin histological image partition, we adaptively compute another feature from Voronoi regions of nuclei clusters, which is provided as follows:

$$F_j = \frac{1}{M} \sum_{m=1}^{M} R(x_m), \quad x_m \in u_j \tag{11}$$

where $R(x_m)$ represents the gray intensity (in Red channel) of the mth pixel in Voronoi region u_j, and M is the total number of pixels belonging

Table 2
Edge weights in the RAG.

Edges	Weights (costs)
$\{c_m, c_n\}$	$B_{m,n} = exp\left(-\frac{(F_m - F_n)^2}{2(\mu_1 - \mu_0)^2} \right)$
$\{c_j, S\}$	$R_j(negative) = -\ln P(F_j\|negative)$
$\{c_j, T\}$	$R_j(positive) = -\ln P(F_j\|positive)$

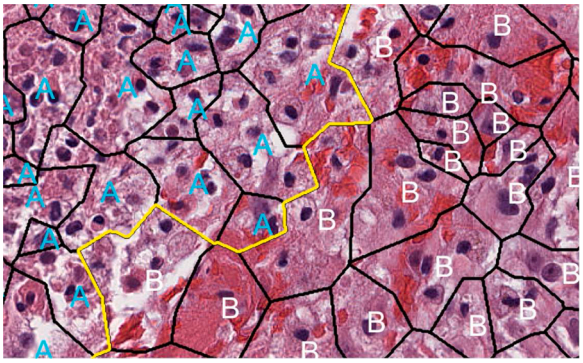

Fig. 4. Illustration of graph cut partition. Note that superimposed letters A (cyan) and B (white) in the image indicate segmented two classes, and the yellow contour indicates the segmentation boundary. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to u_j. The feature computed in Eq. (11) replaces that of Eq. (6), and is integrated into the RAG to perform skin image partition.

Dataset-III: The third dataset consists of 80 H&E stained lung adenocarcinoma histological images, where each image has the size of 500×500 pixels and the pixel resolution is about 1.0 μm/pixel. These lung histological images, along with corresponding ground truth necrosis regions (i.e., necrosis vs non-necrosis), were obtained from Saltz et al. (2018). These images were used to build deep learning models for necrosis detection, such that necrosis could be distinguished from regions of tumor infiltrating lymphocytes.

3.2. Quantitative comparisons and evaluations

As existing works such as Simsek et al. (2012), segmentation sensitivity (*SEN*) and precision (*PRE*) are computed to evaluate the performance. In addition, the accuracy of automatic segmentations is evaluated by using the Dice similarity coefficient (*DSC*) and Jaccard index coefficient (*JIC*), respectively, which are defined as follows:

$$\mathrm{DSC} = \frac{2|\mathrm{GT} \cap \mathrm{AS}|}{|\mathrm{GT}| + |\mathrm{AS}|} \times 100\% \tag{12}$$

$$\mathrm{JIC} = \frac{|\mathrm{GT} \cap \mathrm{AS}|}{|\mathrm{GT} \cup \mathrm{AS}|} \times 100\% \tag{13}$$

where *GT* and *AS* are ground-truth and automatic segmentations, respectively, for any histological image. Note that if the image contains necrotic (or melanoma) regions, the foregrounds in *GT* and *AS* separately correspond to necrosis (or melanoma) regions obtained by manual labeling and automatic segmentation. Otherwise, the foregrounds in *GT* and *AS* correspond to non-necrosis (or non-melanoma) regions accordingly. The *DSC* is the ratio equaling to twice the number of pixels common to both ground truth and automatic segmentation sets divided by the sum of the number of pixels in each set. The *JIC* is the ratio between the number of pixels which are correctly segmented by the method and the number of pixels which are included in either the automatic segmentation or ground truth masks.

The proposed TisCut has been compared with three baseline supervised techniques, which were implemented as follows:

(1) Block based Textual Analysis with SVM classifier (BTA-SVM) (Wang et al., 2009): In our implementation, each histological image in brain or skin dataset was divided into a set of 120×120 non-overlapping image blocks. 278 textural features were then computed from each image block. These features included 7 histogram features, 88 Haralick features and 44 gray level run length matrix related features from red and blue channels (Xu et al., 2018), respectively. The whole image was partitioned by predicting image blocks into different categories such as necrosis or non-necrosis with the trained SVM classifier.

(2) Transfer Learning on InceptionV3 model (TL-InceptionV3) (Kermany et al., 2018): Likewise the BTA-SVM, each histological image was divided into a set of 120×120 non-overlapping image blocks. Training was then performed by fine-tuning parameters of the last fully connected layer of the InceptionV3 model. Image augmentation such as random flipping and rotation were applied along with the training. Stochastic gradient descent with a batch of 100 image blocks per step and a learning rate of 0.01 were applied for training. The whole image was partitioned by predicting image blocks into different categories with the TL-InceptionV3 model.

(3) U-Net model (Ronneberger et al., 2015): In our implementation, the depth of each convolutional layer in the original U-Net model was reduced by 4 times to avoid over-fitting (i.e., the deepest convolutional layer had 256 kernels instead of 1024). In addition, to ensure more contextual information captured by convolutional kernels, the whole image was scaled to a smaller size (i.e., 192×192 pixels). The kernel size of convolutional layers was increased from 3×3 to 5×5. Due to the small dataset, we performed aggressive image augmentations to train the U-Net model. First, every training image was augmented to 18 images by cropping from the original image using different size settings. Second, random rotation, flipping, elastic distortion and zooming were applied during the training process. The U-Net model was trained at most 40 epoches (with early stopping if validation performance did not improve) with the batch size of 16 images per step. During testing, an empirical threshold of 0.5 was applied on the obtained probability map to segment the image into foreground and background.

Table 3 lists quantitative evaluations of different techniques on Dataset-I and II. Note that fivefold cross validation was separately applied on two datasets for three supervised techniques. To ensure the fair comparison, the divisions for training and testing sets during fivefold cross validation were the same across three supervised techniques. Both BTA-SVM and TL-InceptionV3 techniques first make predictions on 120×120 non-overlapping image blocks, and then predictions were stitched together to get whole image segmentation result. The U-Net model provided extremely poor performance when the high resolution original images were used, mainly because it was hard for convolutional operations of the U-Net model to capture the contextual information during tissue region segmentation. Based on trial and error tests, the high resolution images were scaled down to the size of 192×192 pixels for U-Net model training and testing such that a satisfied segmentation performance could be achieved. As observed in Table 3, the BTA-SVM provides the poorest performance among four techniques, with *JIC*

Table 3
Quantitative evaluations and comparisons on Dataset-I and II.

Techniques	Dataset-I/II (Brain/skin histological images)			
	SEN (%)	*PRE* (%)	*DSC* (%)	*JIC* (%)
BTA-SVM (Wang et al., 2009)	75.93/84.94	70.23/91.48	68.79/87.60	55.87/78.48
TL-InceptionV3 (Kermany et al., 2018)	80.92/88.85	78.05/91.47	77.05/89.65	65.05/81.77
U-Net (Ronneberger et al., 2015)	81.22/92.62	80.62/92.74	78.58/91.38	67.22/85.76
Proposed TisCut	79.91/91.28	88.15/93.38	80.64/91.82	70.24/85.44

values of 55.87% and 78.48% on Dataset-I and II, respectively. The BTA-SVM performs image partition by using hand-crafted texture features, and hence its performance is relatively sensitive to the training dataset. The TL-InceptionV3 provides better performances than the BTA-SVM, with *JIC* values of 65.05% and 81.77% on two datasets, respectively. The U-Net and proposed TisCut provide comparative results, both of which are superior to other two techniques. Specifically, the U-Net and TisCut separately provide 67.22% and 70.24% *JIC* values on brain dataset, and 85.76% and 85.44% on skin dataset. The TisCut integrates cytological features into the RAG model and performs image partition directly on histological images at high resolution. By contrast, the U-Net model tends to provide a good performance on low resolution images for histological tissue segmentation, although it is likely to provide poor histological segmentation results close to image borders where contextual information is not enough.

3.3. Qualitative and sensitive evaluations

Fig. 5 shows two visual examples of brain and skin histological image partitions. In Fig. 5, the first row compares brain necrosis and non-necrosis partitions by four techniques, where cyan contours indicate automatically segmented necrosis regions and yellow contours indicate ground truth segmentations. The second row compares skin melanoma and non-melanoma partitions, where cyan contours indicate automatically segmented non-melanoma regions and yellow contours indicate ground truth segmentations. As shown in Fig. 5(a) and (b), the BTA-SVM and TL-InceptionV3 provide many false segmentations. Note that the jagged segmentation boundaries in Fig. 5(a) and (b) are generated because the whole image is divided into 120×120 blocks for supervised classification. The U-Net model and proposed TisCut, as shown in Fig. 5 (c)and (d), provide more accurate segmentations than the other two techniques. However, since regions close to image borders lack of contextual information, the U-Net model may make incorrect predictions for those image border regions.

In the TisCut technique, there are two adjustable parameters which are mean-shift bandwidth h (see Section 2.1.2) and weighing coefficient λ (see Eq. (2)). In our evaluations, they are empirically set as: $h = 60$ and $\lambda = 0.25$. To explore sensitivities of these two parameters, we did experiments by fixing one parameter and progressively adjusting another parameter. Specifically, when $h = 60$, λ was tested with [0.15, 0.20, 0.25, 0.30, 0.35]. When $\lambda = 0.25$, h was tested with [40, 50, 60, 70, 80]. Fig. 6(a) and (b) shows sensitivity evaluations by tuning parameters λ and h on Datasets I and II, respectively. As observed in Fig. 6(a), the *JIC* value on Dataset-I reduces greatly to about 66.5% when λ value is reduced to 0.15. This indicates that an appropriate λ value balancing regional and boundary terms in Eq. (2) is important for achieving a good performance. The *JIC* value on Dataset-I reduces about 1.5% when parameter h increases to 80. The slight performance drop with the increase of parameter h is mainly because larger nuclei clusters are likely to result in less accurate segmentation boundaries. However, it should be noted that the larger h value for mean-shift clustering produces less nuclei clusters (or Voronoi regions) in the image, which helps in reducing the computational complexity for RAG segmentation. Compared with Dataset-I, Dataset-II (see Fig. 6(b)) is relatively insensitive to parameters h and λ, where *JIC* values only fluctuate within about 3% when h or λ value was tuned.

3.4. Annotation generation for deep learning model

The TisCut is a fully unsupervised histological image segmentation method, which could assist in histological annotations for training deep learning models. To verify the annotation capability, we compare U-Net segmentations based on training with manual annotations and automatic annotations, respectively. Since both Dataset-I and III include histological images containing necrosis and non-necrosis regions, they were used for this evaluation. First, we trained the U-Net model by using Dataset-I along with its manual annotations (i.e., U-Net-MA) and tested it on Dataset-III. Second, we trained the U-Net model by using Dataset-I along with its automatic segmentations (i.e., U-Net-AS) generated by the TisCut technique. We then tested the U-Net-AS on Dataset-III. Fig. 7 lists an example of histological image segmentation in Dataset-III, where the overlapped yellow contours and cyan contours indicate manual annotations and automatic segmentations, respectively. Compared with the TisCut, it is found from Fig. 7 that U-Net models tend to generate false segmentations around image borders due to the lack of contextual information.

To verify the data robustness in model training, we also evaluated training on Dataset-III based on manual annotations and automatic segmentations, respectively, and then testing on Dataset-I. Table 4 lists quantitative evaluations on Dataset-III and I by using the proposed TisCut and U-Net models trained with manual annotations and automatic segmentations, respectively. As observed in Table 4, the U-Net models trained with manual annotations and automatic segmentations provide a comparative performance on Dataset-III, with the *JIC* values of 58.14% and 57.25%, respectively. Fig. 8 shows the boxplot of *JIC* values for all images in Dataset-III in terms of three methods, which indicates that there are no significant performance differences (i.e., $p > 0.05$) among three methods. By contrast, the U-Net-AS provides the *JIC* value of 61.58% on Dataset-I, which is about 5% lower than that of the U-Net-MA. This is mainly because the Dataset-III is more heterogenous and challenging for automatic segmentation by the TisCut, and hence automatic annotations include more noisy regions during training the U-Net-AS. This also indicates that training data with well annotations is crucial for the U-Net model to achieve a good performance. Nevertheless, it could be concluded from these evaluations that the proposed TisCut method can provide a comparative necrosis segmentation performance compared with U-Net models. In addition, the TisCut method has the potential to be used in generating histological annotations for

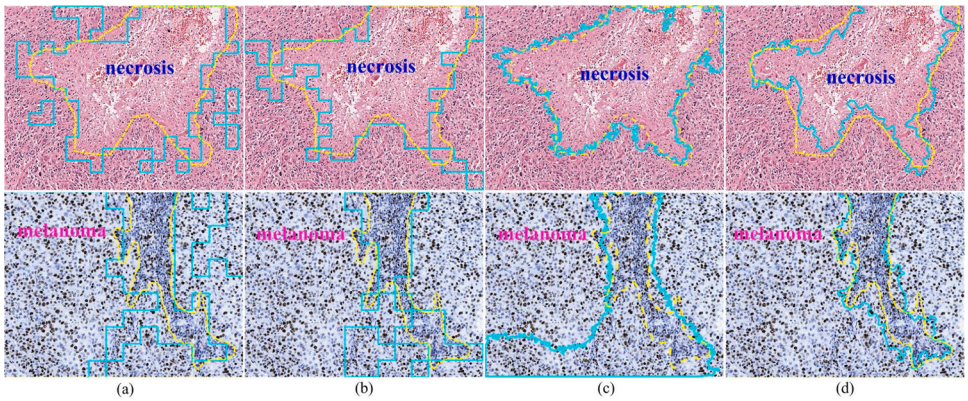

Fig. 5. Examples of histological image partitions by using (a) BTA-SVM. (b) TL-InceptionV3. (c) U-Net model. (d) Proposed TisCut. Note that the first row shows H&E stained brain histological image partition, where yellow contours indicate ground truth necrosis regions. The second row shows Ki-67 stained skin histological image partition, where yellow contours indicate non-melanoma regions. In all images, automatically segmentations are labeled by cyan contours. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

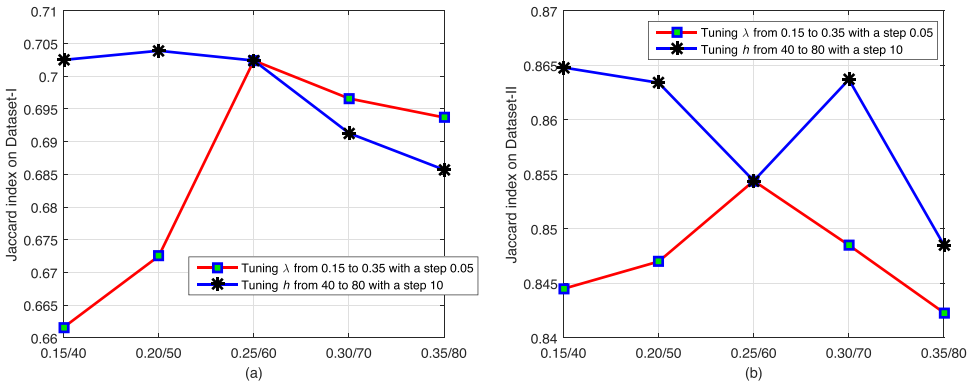

Fig. 6. Sensitivity evaluations by tuning parameters λ and h. (a) *JIC* values on Dataset-I. (b) *JIC* values on Dataset-II. In our evaluations one parameter value was fixed during progressively adjusting another parameter value.

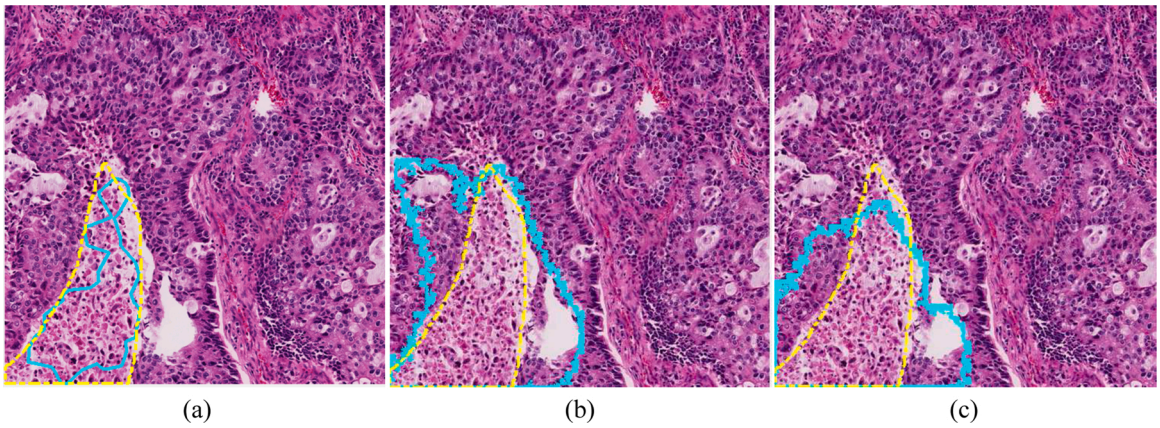

Fig. 7. Histological image partition on an example of Dasetset-III. (a) TisCut. (b) U-Net-MA. (c) U-Net-AS. In images, the manually labeled necrosis regions are indicated by yellow dashed contours, while the automatically segmented necrosis regions are indicated by cyan contours. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Quantitative evaluation of necrosis segmentations on Dataset-III/I (lung/brain histological images). Note that the performance of Dataset-III is obtained by using the U-Net model trained on Dataset-I with either manual annotations or automatic segmentations, and vice versa.

Techniques	*SEN* (%)	*PRE* (%)	*DSC* (%)	*JIC* (%)
TisCut	82.99/79.91	66.20/88.15	64.42/80.64	54.00/70.24
U-Net-MA	82.11/79.87	69.89/81.45	68.99/77.01	58.14/66.70
U-Net-AS	84.82/78.65	67.62/77.30	67.96/73.15	57.25/61.58

training supervised segmentation methods, when manual histological annotations are time-consuming and expensive to collect.

The proposed TisCut method was implemented and tested on a 2.90 GHz Intel Core i7-10700 CPU with 32.0 GB RAM using Matlab2020a. The total time consumed for unsupervised histological segmentations on Dataset-III (including 80 histological images) was about 151.03 s, which indicates that each image on average takes about 1.88 s for histological segmentation by the TisCut method. For comparison, we implemented U-Net models for histological image segmentation by using the Tensorflow toolbox with Python3.6. The U-Net model trained with either manual annotations or automatic segmentations on Dataset-I takes around 20 minutes on our deep learning server which is equipped with Intel Xeon 8353H CPU, 128.0 GB RAM and NVIDIA RTX 3090 GPU card. The trained U-Net model then on average takes about half of second to perform histological image segmentation on Dataset-III. Compared with the supervised segmentation models such as the U-Net model, the TisCut method mainly has the advantage that it is fully

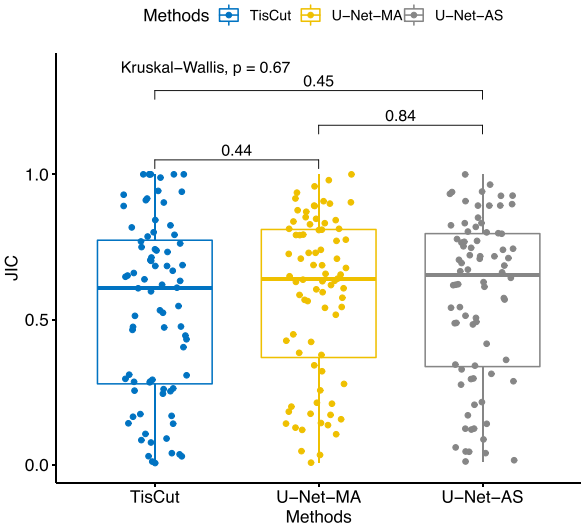

Fig. 8. The boxplot of *JIC* values for three methods on Dataset-III. Note that there are no significant performance differences among three methods, which indicates that the proposed TisCut technique could provide a comparative performance with U-Net models and also has the potential to be used in generating histological image annotations for training supervised segmentation models.

unsupervised, and hence it does not take any time for training.

4. Conclusion

This paper presents a novel unsupervised technique, termed as Tissue Cluster Level Graph Cut (TisCut), for histological image partition. The TisCut method first groups nuclei into different clusters based on nuclear similarities and spatial distance. The image is then represented by Voronoi diagrams, where histological features are computed from constructed Voronoi diagrams. Finally the computed features are integrated into a region adjacency graph. Image partition is performed by using a graph cut algorithm. Experimental results evaluated on three different histology image sets with different color staining show that the TisCut method can effectively partition the testing histological image into meaningful compartments. In particular, it provides a comparative segmentation performance with U-Net models in terms of necrosis and melanoma detections. The TisCut method is a general framework for image partition, which has the potential to be extended for solving various image segmentation problems. In addition, because of its unsupervised property, the TisCut method could help in generating image annotations for training supervised segmentation models such as the U-Net when histological annotations are difficult to collect.

Authors' contribution

Hongming Xu: methodology, visualization, experiment, writing, reviewing and editing. Lina Liu: visualization, validation, experiment, writing. Xiujuan Lei: conceptualization software, reviewing and editing. Mrinal Mandal: data curation, reviewing, supervision. Cheng Lu: conceptualization, data curation, methodology, supervision, writing, reviewing and editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

Dr. Xu is partially supported by the Fundamental Research Funds for the Central Universities (DUT21RC(3)038). Dr. Lei is partially supported by National Natural Science Foundation of China (No.61972451) the Fundamental Research Funds for the Central Universities (GK201901010), and the Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2019JM-179). Dr. Lu is partially supported by the DoD Breast Cancer Research Program Breakthrough Level 1 Award, W81XWH-19-1-0668 and NIH-NCI R21 CA253108-01.

Conflict of interest: There is no conflict of interest for all the authors.

References

Alheejawi, S., Xu, H., Berendt, R., Jha, N., Mandal, M., 2019. Novel lymph node segmentation and proliferation index measurement for skin melanoma biopsy images. Comput. Med. Imaging Graph. 73, 19–29.

Amgad, M., Elfandy, H., Hussein, H., Atteya, L.A., Elsebaie, M.A., Abo Elnasr, L.S., Sakr, R.A., Salem, H.S., Ismail, A.F., Saad, A.M., et al., 2019. Structured crowdsourcing enables convolutional segmentation of histology images. Bioinformatics 35 (18), 3461–3467.

Bejnordi, B.E., Veta, M., et al., 2017. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. JAMA 318 (22), 2199–2210.

Boykov, Y., Funka-Lea, G., 2006. Graph cuts and efficient nd image segmentation. Int. J. Comput. Vision 70 (2), 109–131.

Boykov, Y., Veksler, O., Zabih, R., 2001. Fast approximate energy minimization via graph cuts. IEEE Trans. Pattern Anal. Mach. Intell. 23 (11), 1222–1239.

Chan, L., Hosseini, M.S., Rowsell, C., Plataniotis, K.N., Damaskinos, S., 2019. Histosegnet: semantic segmentation of histological tissue type in whole slide images. Proceedings of the IEEE International Conference on Computer Vision (ICCV) 10662–10671.

Comaniciu, D., Meer, P., 2002. Mean shift: A robust approach toward feature space analysis. IEEE Trans. Pattern Anal. Mach. Intell. 24 (5), 603–619.

Doyle, S., Feldman, M., Tomaszewski, J., Madabhushi, A., 2012. A boosted Bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies. IEEE Trans. Biomed. Eng. 59 (5), 1205–1218.

Gunduz-Demir, C., Kandemir, M., Tosun, A.B., Sokmensuer, C., 2010. Automatic segmentation of colon glands using object-graphs. Med. Image Anal. 14 (1), 1–12.

Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., Yener, B., 2009. Histopathological image analysis: a review. IEEE J. Rev. Biomed. Eng. 2, 147–171.

Kermany, D.S., Goldbaum, M., et al., 2018. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 172 (5), 1122–1131.

Lu, C., Mahmood, M., Jha, N., Mandal, M., 2012. A robust automatic nuclei segmentation technique for quantitative histopathological image analysis. Anal. Quant. Cytol. Histol. 34, 296–308.

Lu, C., Mahmood, M., Jha, N., Mandal, M., 2013. Automated segmentation of the melanocytes in skin histopathological images. IEEE J. Biomed. Health Informatics 17 (2), 284–296.

Manivannan, S., Li, W., Akbar, S., Zhang, J., Trucco, E., McKenna, S.J., 2016. Local structure prediction for gland segmentation. IEEE 13th International Symposium on Biomedical Imaging (ISBI) 799–802.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. International Conference on Medical Image Computing and Computer-Assisted Intervention 234–241.

Saltz, J., Gupta, R., Hou, L., Kurc, T., Singh, P., Nguyen, V., Samaras, D., Shroyer, K.R., Zhao, T., Batiste, R., et al., 2018. Spatial organization and molecular correlation of tumor-infiltrating lymphocytes using deep learning on pathology images. Cell Rep. 23 (1), 181–193.

Simsek, A.C., Tosun, A.B., Aykanat, C., Sokmensuer, C., Gunduz-Demir, C., 2012. Multilevel segmentation of histopathological images using cooccurrence of tissue objects. IEEE Trans. Biomed. Eng. 59 (6), 1681–1690.

Tosun, A.B., Kandemir, M., Sokmensuer, C., Gunduz-Demir, C., 2009. Object-oriented texture analysis for the unsupervised segmentation of biopsy images for cancer detection. Pattern Recogn. 42 (6), 1104–1112.

Trémeau, A., Colantoni, P., 2000. Regions adjacency graph applied to color image segmentation. IEEE Trans. Image Process. 9 (4), 735–744.

Wang, Y., Crookes, D., Eldin, O.S., Wang, S., Hamilton, P., Diamond, J., 2009. Assisted diagnosis of cervical intraepithelial neoplasia (cin). IEEE J. Sel. Top. Signal Process. 3 (1), 112–121.

Xing, F., Su, H., Neltner, J., Yang, L., 2013. Automatic ki-67 counting using robust cell detection and online dictionary learning. IEEE Trans. Biomed. Eng. 61 (3), 859–870.

Xing, F., Xie, Y., Su, H., Liu, F., Yang, L., 2017. Deep learning in microscopy image analysis: a survey. IEEE Trans. Neural Networks and Learning Systems 29 (10), 4550–4568.

Xu, Y., Jia, Z., Ai, Y., Zhang, F., Lai, M., Eric, I., Chang, C., 2015. Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 947–951.

Xu, H., Berendt, R., Jha, N., Mandal, M., 2017. Automatic measurement of melanoma depth of invasion in skin histopathological images. Micron 97, 56–67.

Xu, H., Lu, C., Berendt, R., Jha, N., Mandal, M., 2018. Automated analysis and classification of melanocytic tumor on skin whole slide images. Comput. Med. Imaging Graph. 66, 124–134.