Invariant Delineation of Nuclear Architecture in Glioblastoma Multiforme for Clinical and Molecular Association

Hang Chang*, *Member, IEEE*, Ju Han, *Member, IEEE*, Alexander Borowsky, Leandro Loss, *Member, IEEE*, Joe W. Gray, Paul T. Spellman, and Bahram Parvin*, *Senior Member, IEEE*

***Abstract*—Automated analysis of whole mount tissue sections can provide insights into tumor subtypes and the underlying molecular basis of neoplasm. However, since tumor sections are collected from different laboratories, inherent technical and biological variations impede analysis for very large datasets such as The Cancer Genome Atlas (TCGA). Our objective is to characterize tumor histopathology, through the delineation of the nuclear regions, from hematoxylin and eosin $(H\&E)$ stained tissue sections. Such a representation can then be mined for intrinsic subtypes across a large dataset for prediction and molecular association. Furthermore, nuclear segmentation is formulated within a multi-reference graph framework with geodesic constraints, which enables computation of multidimensional representations, on a cell-by-cell basis, for functional enrichment and bioinformatics analysis. Here, we present a novel method, multi-reference graph cut (MRGC), for nuclear segmentation that overcomes technical variations associated with sample preparation by incorporating prior knowledge from manually annotated reference images and local image features. The proposed approach has been validated on manually annotated samples and then applied to a dataset of 377 Glioblastoma Multiforme (GBM) whole slide images from 146 patients. For the GBM cohort, multidimensional representation of the nuclear features and their organization have identified 1) statistically significant subtypes based on several morphometric indexes, 2) whether each subtype can be predictive or not, and 3) that the molecular correlates of predictive subtypes are consistent with the literature.**

Data and intermediaries for a number of tumor types (GBM, low grade glial, and kidney renal clear carcinoma) are available at: http://tcga.lbl.gov for correlation with TCGA molecular data. The website also provides an interface for panning and zooming of whole mount tissue sections with/without overlaid segmentation results for quality control.

***Index Terms*—Molecular pathology, nuclear segmentation, subtyping, tumor histopathology.**

Manuscript received August 21, 2012; revised October 22, 2012; accepted November 03, 2012. Date of publication December 04, 2012; date of current version March 29, 2013. This work was supported by NIH U24 CA1437991 carried out at Lawrence Berkeley National Laboratory under Contract DE-AC02-05CH11231. *Asterisk indicates corresponding author.*

*H. Chang and *B. Parvin are with the Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (e-mail: hchang@lbl.gov; b_parvin@lbl.gov).

J. Han and L. Loss are with the Life Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA (e-mail: jhan@lbl.gov; laloss@lbl.gov).

A. Borowsky is with the Center for Comparative Medicine, University of California, Davis, CA 95616 USA (e-mail: adborowsky@ucdavis.edu).

J. W. Gray and P. T. Spellman are with the Center for Spatial Systems Biomedicine, Oregon Health Sciences University, Portland, OR 97239 USA (e-mail: grayjo@ohsu.edu; spellmap@ohsu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMI.2012.2231420

I. Introduction

OUR main *motivation* for quantifying morphometric composition from histology sections is to gain insight into cellular morphology, organization, and sample tumor heterogeneity in a large cohort. In tumor sections, robust representation and classification can identify mitotic cells, cellular aneuploidy, and autoimmune responses. More importantly, if tissue morphology and architecture can be quantified on a very large scale dataset, then it will pave the way for constructing databases that are prognostic, the same way that genome-wide array technologies have identified molecular subtypes and predictive markers. Genome-wide molecular characterization (e.g., transcriptome analysis) has the advantage of standardized techniques for data analysis and pathway enrichment, which can enable hypothesis generation for the underlying mechanisms. However, array-based analysis 1) can only provide an average measurement of the tissue biopsy, 2) can be expensive, 3) can hide occurrences of rare events, and 4) lacks the clarity for translating molecular signature into a phenotypic signature. Though nuclear morphology and context are difficult to compute as a result of intrinsic cellular characteristic and technical variations, histology sections can offer insights into tumor architecture and heterogeneity (e.g., mixed populations), in addition to, rare events. Moreover, in the presence of a very large dataset, phenotypic signatures can be used to identify intrinsic subtypes within a specific tumor bank through unsupervised clustering. This facet is orthogonal to histological grading, where tumor sections are classified against known grades. The tissue sections are often visualized with hematoxylin and eosin stains, which label DNA content (e.g., nuclei) and protein contents, respectively, in various shades of color. Even though there are inter- and intra-observer variations [1], a trained pathologist can characterize the rich content, such as the various cell types, cellular organization, cell state and health, and cellular secretion. If hematoxylin and eosin $(H\&E)$ stained tissue sections can be quantified in terms of cell type (e.g., epithelial, stromal), tumor subtype, and histopathological descriptors (e.g., necrotic rate, nuclear size and shape), then a richer description can be linked with genomic information for improved diagnosis and therapy. This is the main benefit of histological imaging since it can capture tumor architecture.

Ultimately, our goal is to mine a large cohort of tumor data in order to identify morphometric indexes (e.g., nuclear size) that have prognostic and/or predictive subtypes. The Cancer Genome Atlas (TCGA) offers such a collection; however, the

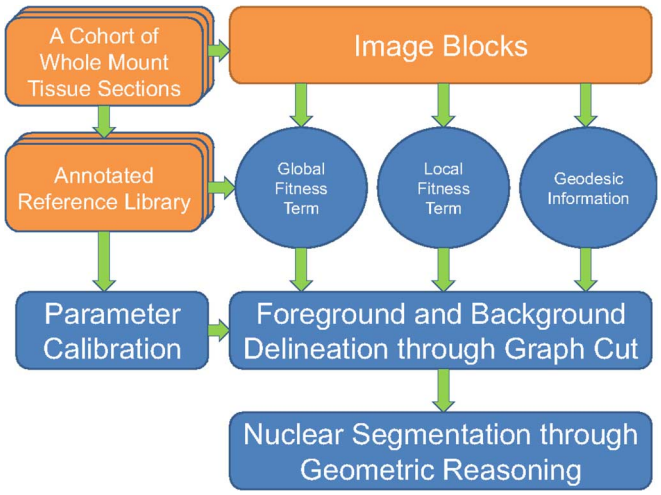


Fig. 1. Work flow in nuclear segmentation for a cohort of whole mount tissue sections.

main *issue* with processing a large cohort, is the inherent variations as a result of 1) the sample preparation protocols (e.g., fixation, staining), practiced by different laboratories, and 2) the intrinsic tumor architecture (e.g., cell type, cell state). For example, with respect to heterogeneity in the tumor architecture, the nuclear color in the RGB space found in one tissue section may be similar to the cytoplasmic color in another tissue section. Simultaneously, the nuclear color intensity (e.g., chromatin content) can vary within a whole slide image. Therefore, image analysis should be tolerant and robust, with respect to variations in sample preparation and tumor architecture, within the entire slide image and across the tumor cohort.

Stained whole mount tissue sections are scanned at either at 20X or 40X, which results in larger images in the order of 40 k$\times$40 k pixels or higher. Each image is partitioned into blocks of 1 k $\times$1 k pixels for processing, and cells at the borders of each block are excluded during the processing. The details of the computational pipeline can be found in our earlier paper [2]. Our approach evolved from our observation that simple color decomposition and thresholding misses or overestimates some of the nuclei in the image, i.e., nuclei with low chromatin contents are excluded. Further complications ensue as a result of diversity in nuclear size and shape (e.g., the classic scale problem).

The general approach is shown in Fig. 1, where the primary *novelty* is in the image-based modeling of inherent ambiguities that are associated with technical variations and biological heterogeneity. Image-based modeling captures prior knowledge from a diverse set of annotated images (e.g., a dictionary) needed in order to model the foreground and background representations. Each annotated image is independent of other images and signifies one facet (e.g., color space, nuclear shape and size) of the diversity within the cohort. Moreover, each image is represented in the feature-space as the Gaussian Mixture Model (GMM) of the Laplacian of Gaussian (LoG) and RGB responses. Collectively, the reference dictionary of annotated images provides the means for color normalization and for capturing global statistics for segmenting test images. The computed global statistics can then be coupled, through a graph cut formulation, with the intrinsic local image statistics and spatial continuity for binarization. Having segmented an input test image, each segmented foreground region is subsequently validated for nuclear shape. If needed, it is decomposed through geometric reasoning. A secondary novelty is in the details of the computational pipeline. For example, we introduce the concept of 1) "color map normalization" for registering a test image against each of the images in the reference library, and 2) "blue ratio image" for mapping RGB images into the gray space; thus, LoG responses can be computed efficiently in one channel. All important free parameters are selected through cross-validation. Thus far, close to 1000 whole slide images have been processed, and the data has been made publicly available through our website at http://tcga.lbl.gov. In addition, segmentation results, from the whole mount tissue sections, are available for quality control through a web-based zoomable interface.

Essentially, nuclear segmentation provides the basis for morphometric representation on a cell-by-cell basis. As a result, tumor histology can be represented as a meaningful data matrix, where well-known bioinformatics and statistical tools can be readily applied for hypotheses generation. For example, a large cohort facilitates tumor subtyping based on computed morphometric features. Each subtype can then be 1) tested for its prognostic value, and 2) utilized for identifying molecular basis of each subtype for hypothesis generation. In the case of GBM, prognostic and/or predictive subtypes have also been posted on our website.

Organization of this paper is as follows. Section II reviews previous research with a focus on quantitative representation of the $H\&E$ sections for translational medicine. Sections III and IV describes the details of the image-based modeling for nuclear segmentation and experimental validation, respectively. Section V examines one application of nuclear segmentation of morphometric subtyping and molecular association for hypothesis generation. Lastly, Section VI concludes the paper.

II. Review of Previous Work

Several excellent reviews for the analysis of histology sections can be found in [3] and [4]. From our perspective, four distinct works have defined the trends in tissue histology analysis. 1) One group of researchers proposed nuclear segmentation and organization for tumor grading and/or prediction of tumor recurrence [5]–[8]. 2) A second group of researchers focused on patch level analysis (e.g., small regions) [9]–[11], using color and texture features, for tumor representation. 3) A third group focused on block-level analysis to distinguish different states of tissue development using cell-graph representation [12], [13]. 4) Finally, a fourth group has suggested detection and representation of the auto-immune response as a prognostic tool in cancer [14]. In contrast to previous research, our strategy is based on processing a large cohort of tumors, to compute morphometric subtypes, and to examine whether computed subtypes are predictive of outcome. Since tumor histology is characterized in terms of nuclear and cellular features, a more detailed review of nuclear segmentation strategies follows.

The main barriers in nuclear segmentation are technical variations (e.g., fixation) and biological heterogeneity (e.g., cell type). These factors are visible in TCGA dataset. Present

techniques have focused on adaptive thresholding followed by morphological operators [15], [16]; fuzzy clustering [17], [18]; level set method using gradient information [14], [19]; color separation followed by optimum thresholding and learning [20], [21]; hybrid color and texture analysis followed by learning and unsupervised clustering [6]; and representation of nuclei organization in tissues [22], [23] that is computed from either interactive segmentation or a combination of feature detector. Some applications combine the above techniques; Several examples are given below. In [24], iterative radial voting [25] was used to estimate seeds for partitioning perceptual boundaries between neighboring nuclei. Subsequently, seeds were used to segment each nucleus through the application of multiphase level sets [26], [27]. In [28], the input image was initially binarized into foreground and background regions with a graph cut framework, the seeds were then selected from a binarized image using a constrained multi-scale LoG filter, with the combined results being refined using a second iteration of the graph cut. Similarly, in [29], the input image was first normalized through histogram equalization, and then binarized based on color-texture extracted from the most discriminant color space. This was followed by an iterative operation to split touching nuclei based on concave-points and radial-symmetry. In their experiment, they had 21 images where five of them were annotated. Nuclei, in all images, had similar size with high chromaticity. Recently, a spatially constrained expectation maximization algorithm [30] was demonstrated to be robust to "color nonstandardness" in histological sections with color being represented in the HSV space. However, our analysis of the GBM cohort indicates that strict incorporation of color and spatial information will not be sufficient as demonstrated in Section IV-B (MRGC versus MRGC-CF). A more related work, described in [31], was based on a voting system that uses multiple classifiers built from different reference images; we will refer to this method as MCV, for short, in the rest of the paper. Compared to the previous approaches, MCV provides a better way to handle the variation among different batches. However, due to the lack of smoothness constraints and local statistical information, the classification results can be noisy and erroneous, as demonstrated in Fig. 8. Some of these concepts have also been utilized in our earlier paper [2], but the results posted on our website are for the current implementation outlined in this paper.

In summary, the main limitations of the above techniques are that they are often applied to a small dataset that originate from a single laboratory, ignore technical variations that are manifested in both nuclear and background signals, and are insensitive to cellular heterogeneity (e.g., variation in chromatin contents). Our goal is to address these issues by processing whole mount tissue sections, from multiple laboratories, to construct a large database of morphometric features, and to enable subtyping and genomic association.

III. Approach

Details of the proposed approach are shown in Fig. 2, which leverages several key observations for segmenting nuclear regions: 1) global variations across a large cohort of tissue sections can be captured by a representative set of reference images, 2) local variations within an image can be captured by local foreground(nuclei)/background samples detected by LoG filter, and 3) color normalization, against a reference image, reduces variations in image statistics and batch effects between a test and a reference image. These concepts are integrated within a graph cut framework to delineate nuclei or clumps of nuclei from the background. Having performed foreground and background segmentation, we then partitioned potential clumps of nuclei through geometric reasoning. In the rest of this section, we summarize (a) the representation of prior models from a diverse set of reference images, (b) the methodology for color normalization, (c) an effective approach for color transformation for dimensionality reduction, (d) the details of feature extraction from each test image, (e) the multi-reference graph cut formalism for nuclei/background separation, and (f) the partitioning of a clump of nuclei into individual nucleus.

A. Construction and Representation of Priors

The purpose of this step is to capture the global variations for an entire cohort from a reference library. For bioinformatics analysis, the target dataset consists of 377 individual tissue sections, and a representative of N $(N = 20)$ reference images of 1 k $\times$1 k pixels at 20X have been selected. Each reference image is selected to be an exemplar of tumor phenotypes based on staining and morphometric properties. Therefore, it is reasonable to suggest that each reference image has its own unique feature space, in terms of RGB and LoG responses, which leads to $2N$ feature spaces for all reference images

$$\left\{\mathbb{F}^1_{RGB_1}, \mathbb{F}^2_{RGB_2}, \cdots, \mathbb{F}^N_{RGB_N}, \mathbb{F}^{N+1}_{\mathrm{LoG}_1}, \mathbb{F}^{N+2}_{\mathrm{LoG}_2}, \cdots, \mathbb{F}^{2N}_{\mathrm{LoG}_N}\right\} \tag{1}$$

where $\mathbb{F}^i_{RGB_i}$ and $\mathbb{F}^{N+i}_{\mathrm{LoG}_i}$ are RGB feature space and LoG feature space for the ith reference image, $1 \leq i \leq N$. Subsequently, each reference image is hand segmented and processed with a LoG filter (please refer to Section III-C for the details on our LoG integration), at a single scale, followed by the collection of foreground (nuclei) and background statistics in both the RGB space and LoG response. Our experience indicates that even within a single reference image, there could be distinct modes in terms of RGB color and nuclear size. One way to capture these heterogeneities is to represent foreground and background distributions with GMM. Hence, the conditional probability for pixel p, with feature $f^k(p)$ in the kth $(k \in [1, 2N])$ feature space, belonging to $\mathrm{nuclei}(l = 1)/\mathrm{background}(l = 0)$ can be expressed as a mixture with D component densities

$$\mathrm{GMM}^k_l(p) = \sum_{j=1}^{D} \tilde{p}\left(f^k(p)|j\right) P(j) \tag{2}$$

where a mixing parameter $P(j)$ corresponds to the weight of component j and $\sum_{j=1}^{D} P(j) = 1$. Each mixture component is a Gaussian with mean μ and covariance matrix Σ in the corresponding feature space (e.g., 3×3 and 1×1 matrices in RGB and single scale LoG spaces, respectively)

$$\tilde{p}\left(f^k(p)|j\right) = \frac{1}{(2\pi)^{\frac{3}{2}}|\Sigma|_j^{\frac{1}{2}}} \cdot \exp\left(-\frac{1}{2}\left(f^k(p) - \mu_j\right)^T \Sigma_j^{-1}\left(f^k(p) - \mu_j\right)\right) \tag{3}$$

$P(j)$ and (μ_j, Σ_j) for $\tilde{p}(C_p|j)$ were estimated by expectation maximization (EM) algorithm [32].

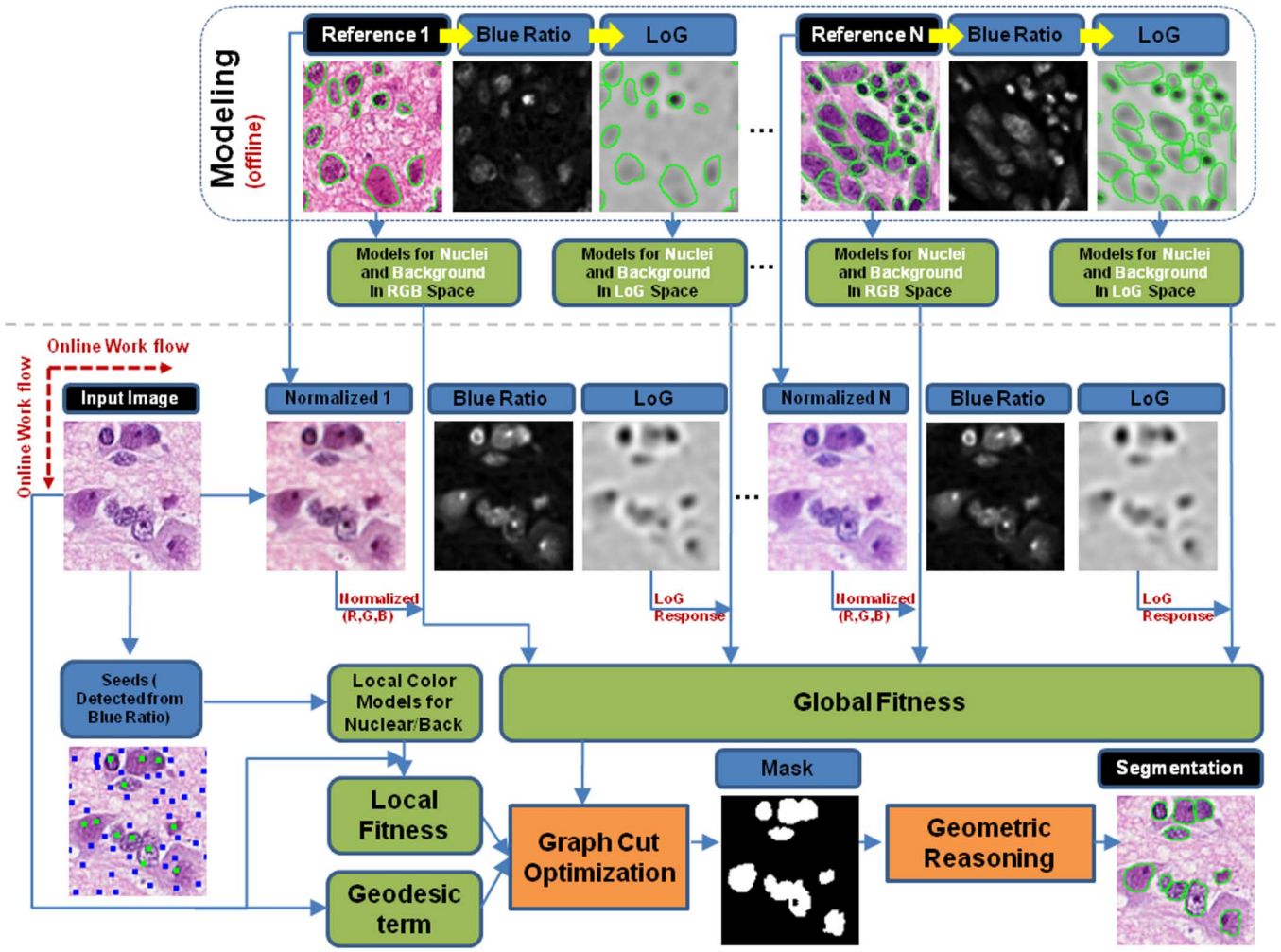


Fig. 2. Steps in nuclear segmentation.

B. Color Normalization

The purpose of color normalization is to close the gap, in color space, between an input test image and a reference image. As a result, the prior models, constructed from each reference image, can be better utilized. We evaluated a number of color normalization methods and chose the color map normalization described in [31] for its effectiveness in handling histological data.

- Let input image I and reference image Q have K_I and K_Q unique color triplets in terms of (R, G, B), respectively.
- Let $\mathbb{R}_C^{I/Q}$ be a monotonic function, which maps the color channel intensity, $C \in \{R, G, B\}$, from Image I/Q to a rank that is in the range $[0, K_I)/[0, K_Q)$.
- Let (r_p, g_p, b_p) be the color of pixel p, in image I, and $(\mathbb{R}_R^I(r_p), \mathbb{R}_G^I(g_p), \mathbb{R}_B^I(b_p))$ be the ranks for each color channel intensity.
- Let the color channel intensity values r_{ref}, g_{ref}, and b_{ref}, from image Q, have ranks

$$\mathbb{R}_R^Q(r_{\text{ref}}) = \left\lfloor \frac{\mathbb{R}_R^I(r_p)}{K_I} \times K_Q + \frac{1}{2} \right\rfloor$$
$$\mathbb{R}_G^Q(g_{\text{ref}}) = \left\lfloor \frac{\mathbb{R}_G^I(g_p)}{K_I} \times K_Q + \frac{1}{2} \right\rfloor$$
$$\mathbb{R}_B^Q(b_{\text{ref}}) = \left\lfloor \frac{\mathbb{R}_B^I(b_p)}{K_I} \times K_Q + \frac{1}{2} \right\rfloor.$$

As a result of color map normalization, the color for pixel p: (r_p, g_p, b_p), will be normalized as $(r_{\text{ref}}, g_{\text{ref}}, b_{\text{ref}})$. In contrast to standard quantile normalization, which utilizes all pixels in the image, color map normalization is based on the unique color in the image, thereby, excluding the frequency of any color. Our experience suggests that this method is quite powerful for normalizing histology sections, since the color frequencies vary widely as a result of technical variations and tumor heterogeneity. Examples of color map normalization can be found in Fig. 2.

C. Color Transformation

In order to reduce the computational complexities for integrating the LoG responses, the RGB space is transformed into a gray level image to accentuate the nuclear dye. While several techniques for color decomposition have been proposed [34], [33], they are either too time-consuming or do not yield favorable outcomes. The color transformation policy needs to enhance the nuclear stain while attenuating the background stain. One way to realize such a transformation is by: $BR(x, y) = ((100{*}B(x, y))/(1{+}R(x, y){+}G(x, y))){\times}(256/(1{+}B(x, y){+}R(x, y) + G(x, y)))$, where $B(x, y)$, $R(x, y)$, and $G(x, y)$ are the blue, red, and green intensities at position (x, y). We refer to this transformation as the blue ratio image in the rest of this

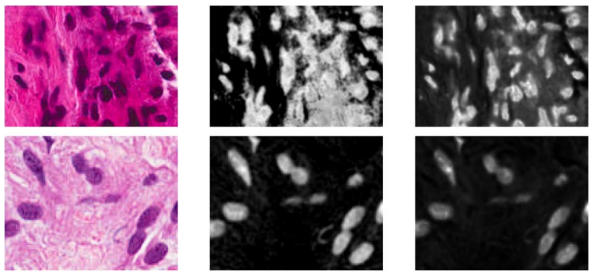


Fig. 3. (a) Two diverse pinhole of tumor signatures. (b) Decompositions by [33]. (c) Blue ratio images.

manuscript. In this formulation, the first and second terms accentuate and attenuate nuclear and background signals, respectively. Subsequently, the LoG responses are always computed at a single scale from the blue ratio image. Fig. 3 demonstrates that the blue ratio image method has an improved performance compared to an alternative method [33].

D. Feature Extraction

Our approach integrates both color and scale information, where the scale is encoded by the LoG response.

1) Normalization of the input test image against every reference image, as described in Section III-B.
2) Conversion of each normalized image into the blue ratio image, as described in Section III-C.
3) Application of a LoG filter on each of the blue ratio images, at a single scale.
4) Representation of each pixel, from the test image, by its RGB color in each of the normalized images and LoG response from each of the blue ratio images.

As a result, each pixel in the test input image is represented by $2N$ features, where the first N features are RGB colors from the normalized images, and the last N features are LoG responses computed from the blue ratio of the normalized images. All $2N$ features are assumed to be independent per selection of images in Section III-A. The rational for integrating both color and scale information is that: 1) in some cases, color information is insufficient to differentiate nuclear regions from background; 2) the scales (e.g., LoG responses) of the background structure and nuclear region are typically different; and 3) the nuclear region responds well to blob detectors, such as a LoG filter [28].

E. Multi-Reference Graph Cut Model

In this section, we first present the background material on graph cut formalism, and then proceed to the details of the image-based modeling for incorporating intrinsic and extrinsic variations.

Within the graph cut formulation, an image is represented as a graph $G = \langle \bar{V}, \bar{E} \rangle$, where $\bar{V}$ is the set of all nodes, and $\bar{E}$ is the set of all arcs connecting adjacent nodes. Usually, the nodes and edges correspond to pixels ($\mathcal{P}$) and their adjacency relationship, respectively. Additionally, there are special nodes known as terminals, which correspond to the set of labels that can be assigned to pixels. In the case of a graph with two terminals, the terminals are referred to as the source (S) and the sink (T), which correspond to specific labels. The labeling problem is to assign a unique label x_p (0 for background, and 1 for foreground) for each node $p \in \bar{V}$, and the image cutout is performed by minimizing the Gibbs energy E [35]

$$E = \sum_{p \in \bar{V}} E_{\text{fitness}}(x_p) + \beta \sum_{(p,q) \in \bar{E}} E_{\text{smoothness}}(x_p, x_q) \quad (4)$$

where $E_{\text{fitness}}(x_p)$ is the likelihood energy, encoding the data fitness cost for assigning x_p to p, and $E_{\text{smoothness}}(x_p, x_q)$ is the prior energy, denoting the cost when the labels of adjacent nodes, p and q, are x_p and x_q, respectively; β is the weight for $E_{\text{smoothness}}$.

The optimization algorithms could be classified into two groups: Goldberg–Tarjan "push-relabel" methods [36], and Ford–Fulkerson "augmenting paths" [37]. The details of the two methods can be found in [38].

We recognize that the training data set cannot fully capture the intrinsic variations of the nuclear signature. Therefore, the data fitness term is expressed as a combination of the intrinsic local probability map and learned global property map. The local probability map has the advantage of capturing local intrinsic image property in the absence of colormap normalization, thus, diversifying the data fitness term. Equation (1) is rewritten as

$$E = \sum_{p \in \bar{V}} (E_{gf}(x_p) + E_{lf}(x_p)) + \beta \sum_{(p,q) \in \bar{E}} E_{\text{smoothness}}(x_p, x_q) \quad (5)$$

where E_{gf} is the global data fitness term encoding the fitness cost for assigning x_p to p, E_{lf} is the local data fitness term encoding the fitness cost for assigning x_p to p. Each term together with the optimization process is discussed below.

1) Global Fitness Term: The global fitness is established based on manually annotated reference images. Let us assume N reference images: Q_i, $i \in [1, N]$, and for each reference image, GMMs are used to represent the nuclei and background in both RGB space and LoG response space, respectively: $\text{GMM}^k_{\text{Nuclei}}$, $\text{GMM}^k_{\text{Background}}$, in which $k \in [1, 2N]$, and the first N GMMs are for RGB space, and the last N GMMs are for LoG response space. Details can be found in Section III-A.

An input test image I is first normalized as U_i with respect to every reference image, Q_i. Subsequently, RGB color and LoG responses of U_i are collected to construct $2N$ features per pixels, where the first N features are from the normalized color (RGB) space, and the second N features are from LoG response.

- Let p be a node corresponding to a pixel.
- Let $f^k(p)$ be kth feature of p.
- Let α be the weight of LoG response.
- Let $\mathbf{p}^k_i$ be the probability function of f^k being $\text{Nuclei}(l = 1)/\text{Background}(l = 0)$

$$\mathbf{p}^k_l(p) = \frac{\text{GMM}^k_l(p)}{\sum_{j=0}^{1} \text{GMM}^k_j(p).}$$

- Let λ_i be the weight for Q_i

$$\lambda_i = \frac{1}{3} \sum_{C}^{C \in \{R,G,B\}} \lambda^C_i$$

$$\lambda^C_i = H^C(Q_i) \cdot H^C(U_i) / \left(\left\| H^C(Q_i) \right\| \cdot \left\| H^C(U_i) \right\| \right)$$

where $\|.\|$ is L_2 norm, $H^C(\cdot)$ is the histogram function on a single color channel $C \in \{R, G, B\}$ of an image. Intuitively, λ measures similarity between two histograms derived from Q_i and U_i, which are represented with 20 bins. Based on our experiments, the λs become stable when the number of bins reaches 20; conversely, histograms with less than 20 bins are considered to have insufficient resolution. The similarity parameter weighs the fitness of the prior model, constructed from Q_i, to the features extracted from the normalized image U_i.

The global fitness term is now defined as

$$E_{gf}(x_p = i) = -\sum_{k=1}^{N} \lambda_k log\left(\mathbf{p}_i^k\left(f^k(p)\right)\right) - \alpha \cdot \sum_{k=N+1}^{2N} \lambda_{k-N} log\left(\mathbf{p}_i^k\left(f^k(p)\right)\right) \quad (6)$$

where the first and second terms integrate normalized color features and LoG responses, respectively.

2) Local Fitness Term: While the global fitness term utilizes both color and LoG information in the normalized space, it does not utilize information in the original color space of the input image. As a result, local variation may be lost for a number of reasons, i.e., nonuniformity in the tissue sections, local lesions, etc. The local data fitness of a pixel, p, is computed from foreground and background seeds in a local neighborhood around p that corresponds to peaks detected by a LoG filter on the blue ratio image, where positive and negative peaks often, but not always, correspond to the background and foreground (nuclei), respectively. The accuracy can be improved by a cascade of filters as follows.

1) Seeds detection: This step aims to collect local foreground and background seeds by incorporating local and global image statistics. Typical positive and negative peak responses, associated with the LoG filter, are shown in Fig. 4(a). Most of the time, the LoG filter detects foreground and background locations correctly, but there is a potential for errors. The protocol consists of three steps.
   a) Create a blue ratio image (Section III-C): In this transformed space, the peak of the intensity histogram always corresponds to the preferred frequency of the background intensity as shown in Fig. 4(b).
   b) Construct distributions of the foreground and background: Apply the LoG filter on the blue ratio image, detect peaks, and construct a distribution of the blue ratio intensity at the peaks corresponding to the negative and positive LoG responses. A small subset of seeds can be mislabeled, but most can be corrected in the following step.
   c) Constrain the seed selection: Seeds (e.g., peaks of the LoG response) are constrained by three criteria: 1) the LoG responses must be above a minimum conservative threshold for removing strictly noisy artifacts; 2) the intensity associated with the peak of the negative LoG responses (e.g., foreground peaks) must concur with the background peak, specified in step (a); and 3) within a small neighborhood of $w_1 \times w_1$, the minimum blue ratio intensity, at the location of negative seeds, is set as the threshold for background peaks, as shown in Fig. 5.
2) Local foreground/background color modeling: For each pixel, p, foreground and background statistics within a local neighborhood, $w_2 \times w_2$, is represented by two GMMs in the original color space. These GMMs correspond to the nuclei and background models (e.g., $\mathrm{GMM}^{\mathrm{Local}}_{\mathrm{Nuclei}}$ and $\mathrm{GMM}^{\mathrm{Local}}_{\mathrm{Background}}$), respectively.

The local fitness term is defined as

$$E_{lf}(x_p = i) = -\gamma log\left(\mathbf{p}_l\left(f(p)\right)\right) \quad (7)$$

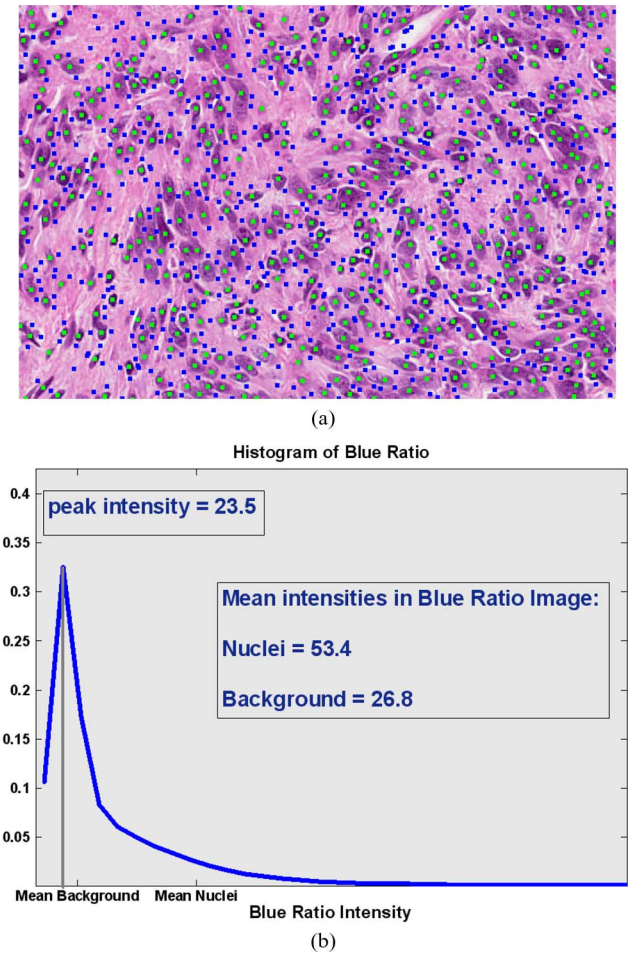


Fig. 4. (a) An example of the LoG response for detection of foreground (green dot) and background (blue dot) signals indicates an excellent performance on the initial estimate. (b) Histogram of the blue ratio intensity derived from image (a) indicates that the peak of the distribution corresponds to the occurrence frequency of the background pixels.

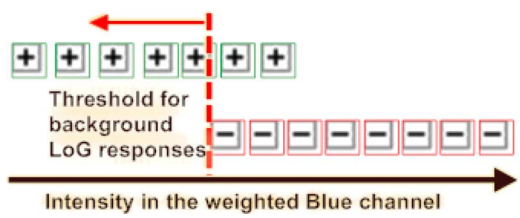
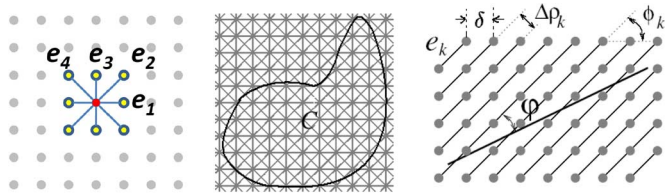


Fig. 5. LoG responses can be either positive (e.g., potential background) or negative (e.g., foreground or part of foreground) in the transformed blue ratio image. In the blue ratio image with the most negative LoG response, the threshold is set at the minimum intensity.



(a) (b) (c)

Fig. 6. (a) Eight-neighborhood system: $n_G = 8$. (b) Contour on eight-neighborhood 2-D grid. (c) One family of lines formed by edges of the graph.

where $f(p)$ refers to the RGB feature of node p in the original color space, γ is the weight for local fitness, $\mathbf{p}_l$ is the probability function of f being $\text{Nuclei}(l=1)/\text{Background}(l=0)$

$$\mathbf{p}_l(p) = \frac{\text{GMM}_l^{\text{Local}}(p)}{\sum_{j=0}^{1} \text{GMM}_j^{\text{Local}}(p)}.$$

3) Smoothness Term: While both local and global data fitness terms are encoded by t-links (links between node and terminals) in the graph, the smoothness term, which ensures the smoothness of labeling between adjacent nodes, is represented by n-links (links between adjacent nodes). Here, we adopt the setup from [39] for n-links, which approximates a continuous Riemannian metric by a discrete weighted graph so that the max-flow/min-cut solution for the graph corresponds to a local geodesic or minimal surface in the continuous case. Consider a weighted graph constructed in Section III-E: $G = \langle \bar{V}, \bar{E} \rangle$, where $\bar{V}$ is the set of image pixels, and $\bar{E}$ is the set of all edges connecting adjacent pixels.

- Let $\{e_k | 1 \le k \le n_G\}$ be a set of vectors for the neighborhood system, where n_G is the neighborhood order, and the vectors are ordered by their corresponding angle ϕ_k with respect to the $+x$ axis, such that $0 \le \phi_1 < \phi_2 \cdots < \phi_{n_G} < \pi$. For example, when $n_G = 8$, we have $e_1 = (1,0)$, $e_2 = (1,1)$, $e_3 = (0,1)$, $e_4 = (-1,1)$, as shown in Fig. 6(a).
- Let w_k be the weight for the edge between pixels: p and q, where p and q belong to the same neighborhood system, and $\vec{pq} = \pm e_k$.
- Let L be a line formed by the edges in the graph, as shown in Fig. 6(c).
- Let C be a contour in the same 2-D space where the graph G is embedded, as shown in Fig. 6(b).
- Let $|C|_G$ be the cut metric of C

$$|C|_G = \sum_{e \in \bar{E}_C} w_e$$

  where $\bar{E}_C$ is the set of edges intersecting contour C.
- Let $|C|_R$ be the Riemannian length of contour C.
- Let $D(p)$ be the metric (tensor), which continuously varies over points p in the 2-D Riemannian space.

Based on Integral Geometry [40], the Crofton-style formula for Riemannian length $|C|_R$ of contour C can be written as

$$\int \frac{detD(p)}{2\left(u_L^T \cdot D(p) \cdot u_L\right)^{\frac{3}{2}}} n_C dL = 2|C|_R$$

where u_L is the unit vector in the direction of the line L, and n_C is a function that specifies how many times line L intersects contour C. Following the approach in [39], the local geodesic can be approximated by the max-flow/min-cut solution ($|C|_G \to |C|_R$) with the following edge weight setting:

$$w_k(p) = \frac{\delta^2 \cdot |e_k|^2 \cdot \Delta\phi_k \cdot detD(p)}{2 \cdot \left(e_k^T \cdot D(p) \cdot e_k\right)^{\frac{3}{2}}} \tag{8}$$

where δ is the cell-size of the grid, $\triangle\phi_k$ is the angular difference between the kth and $(k+1)$th edge lines, $\triangle\phi_k = \phi_{k+1} - \phi_k$, and

$$D(p) = g(|\nabla I|) \cdot \mathbf{I} + (1 - g(|\nabla I|)) \cdot \mathbf{u} \cdot \mathbf{u}^T \tag{9}$$

where $\mathbf{u} = \nabla I / |\nabla I|$ is a unit vector in the direction of image gradient at point p, $\mathbf{I}$ is the identity matrix, and $g(x) = exp(-(x^2/2\sigma^2))$

4) Optimization: The construction of the graph, with two terminals, source S and sink T, is defined in Table I. This graph is partitioned via the max-flow/min-cut algorithm proposed in [41] to label the input image into foreground and background. The optimization method belongs to a class of algorithms based on augmenting paths, and the details can be found in [41].

TABLE I
EDGE WEIGHTS FOR THE GRAPH CONSTRUCTION, WHERE $\mathbb{N}$ IS THE NEIGHBORHOOD SYSTEM, AND β IS THE WEIGHT FOR SMOOTHNESS

Edge	Weight	For
$p \to S$	$E_{gf}(x_p = 1) + E_{lf}(x_p = 1)$	$p \in \mathcal{P}$
$p \to T$	$E_{gf}(x_p = 0) + E_{lf}(x_p = 0)$	$p \in \mathcal{P}$
$w_e(p,q)$	$\beta \cdot w_k(p)$	$\{p,q\} \in \mathbb{N}$, $\phi_{\vec{pq}} \in \{\phi_k, \pi + \phi_k\}$

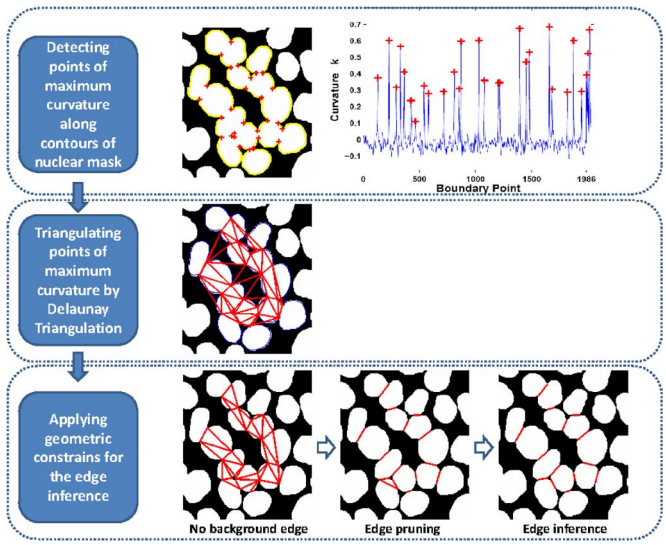


Fig. 7. Steps in the delineation of overlapping nuclei: (Top row) identifying points of maximum curvature where potential folds are formed, (middle row) formation of partitioning hypotheses through triangulation, (bottom row) stepwise application of geometric constraints for deleting and pruning edges.

F. Nuclear Mask Partitioning

A key observation we made is that the nuclear shape is typically convex. Therefore, ambiguities associated with the delineation of overlapping nuclei could be resolved by detecting concavities and partitioning them through geometric reasoning. The process, shown in Fig. 7, consists of the following steps.

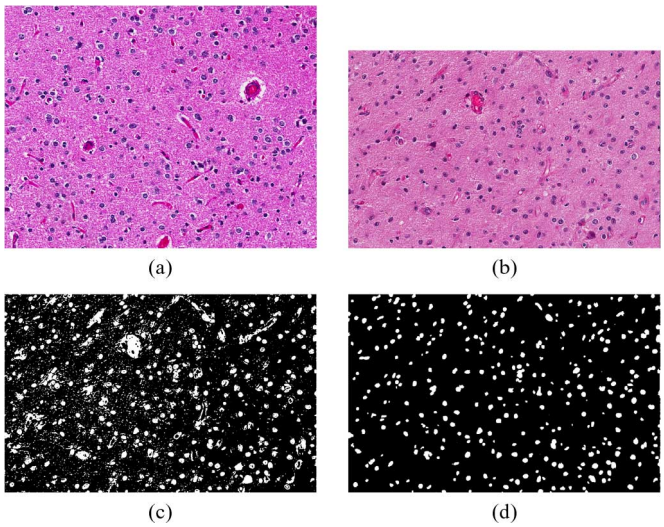


Fig. 8. Comparison between MCV and MRGC (as shown in (c) and (d), respectively) based on the same reference image, as shown in (a). Even though the test image and the reference image are slightly different in color space, compared with MCV, MRGC still produces 1) more accurate classification, due to the encoding of statistics from test image's color space via local probability map; 2) less noisy classification due to the smoothness constrain. (a) Reference image; (b) test image; (c) results via MCV; (d) results via MRGC.

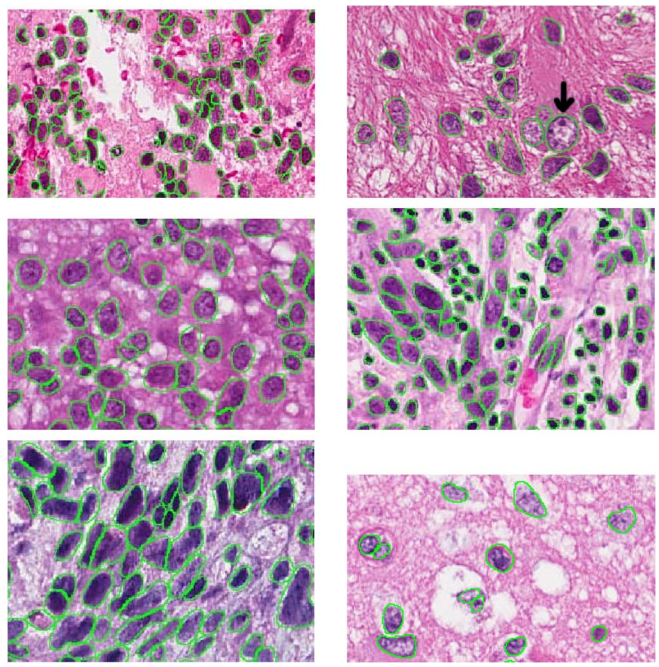


Fig. 9. A subset of reference image ROI, with manual annotation overlaid as green contours, indicating significant amounts of technical variation. Nuclei with white hollow regions inside are pointed out by arrows.

1) Detection of Points of Maximum Curvature: The contours of the nuclear mask were extracted, and the curvature along the contour was computed by using $k = (x'y'' - y'x'')/(x'^2 + y'^2)^{3/2}$, where x and y are coordinates of the boundary points. The derivatives were then computed by convoluting the boundary with derivatives of Gaussian. An example of detected points of maximum curvature is shown in Fig. 7.
2) Delaunay Triangulation (DT) of Points of Maximum Curvature for Hypothesis Generation and Edge Removal: DT was applied to all points of maximum curvature to hypothesize all possible groupings. The main advantage of DT is that the edges are nonintersecting, and the Euclidean minimum spanning tree is a sub-graph of DT. This hypothesis space was further refined by removing edges based on certain rules, e.g., no background intersection.
3) Geometric reasoning: Properties of both the hypothesis graph (e.g., degree of vertex), and the shape of the object (e.g., convexity) were integrated for edge inference.

This method is similar to the one proposed in our previous work [42]; however, a significant performance improvement has been made through triangulation and subsequent geometric reasoning. Please refer to [43] for details.

IV. Experimental Results and Discussion

In this section, we 1) discuss parameter setting, and 2) evaluate performance of the system against previous methods.

A. Experimental Design and Parameter Setting

In order to capture the technical variation, we manually selected and annotated 20 reference images of the size of 1 k×1 k pixels at 20X, and a subset is shown in Fig. 9. Nuclear segmentation was also performed at 20X, and only the top $M = 10$ reference images with the highest weight of λ were used. Essentially, this was a trade-off between performance and computational time cost (see in Fig. 13). The number of components for GMM was selected to be $D = 20$, while the parameters for GMM were estimated via EM algorithm. Other parameter settings were: $\alpha = 0.1$, $\beta = 10.0$, $\gamma = 0.1$, $w_1 = 100$, $w_2 = 100$, and $\sigma = 4.0$ (the scale for both seeds detection and LoG feature extraction), in which σ was determined based on the preferred nuclear size at 20X, w_1 was selected to minimize the seeds detection error on the annotated reference images, and all other parameters were selected to minimize the cross validation error from the following discretization: $D \in \{5, 10, 15, 20, 25, 30\}$, $\alpha \in \{0.05, 0.10, \ldots, 0.95, 1.00\}$, $\beta \in \{5, 10, \ldots, 95, 100\}$, $\gamma \in \{0.05, 0.10, \ldots, 0.95, 1.00\}$, $w_2 \in \{50, 60, \ldots, 190, 200\}$. The optimal γ value is relatively small, which can be attributed to the fact that the global statistics from the well-constructed reference images, cover most of the heterogeneity in our dataset, and the role of local statistics is simply to assist the global statistics with improved discriminating powers.

B. Evaluation

Two-fold cross validation, with optimized parameter settings, was applied to the reference images, and a comparison of average classification performance was made between our approach, random forest [44], and the most related work (Here, we refer it to MCV: multi-classifier voting, for short) in [31], as shown in Table II. Our experiment indicates the following.

1) By incorporating both global and local statistics (MRGC versus MRGC-GF), our system better characterizes the variation in the data.
2) By incorporating the LoG response as a feature (MRGC versus MRGC-CF), we can encode the prior scale information into the system.

TABLE II
COMPARISON OF AVERAGE CLASSIFICATION PERFORMANCE AMONG OUR APPROACH(MRGC), OUR PREVIOUS APPROACH [2], MCV APPROACH IN [31], AND RANDOM FOREST. FOR MCV, ONLY COLOR IN RGB SPACE IS USED, WHICH IS IDENTICAL TO [31]. FOR RANDOM FOREST, THE SAME FEATURES ARE USED: $\{R, G, B, LoG\}$, AND THE PARAMETER SETTINGS ARE $ntree = 100$, $mtry = 2$, $node = 1$

Approach	Precision	Recall	F-Measure
MRGC-MS (Multi-Scale LoG)	0.77	0.82	0.794
MRGC	0.79	0.78	0.785
MRGC-CF (Color Feature Only)	0.72	0.83	0.771
MRGC-GF (Global Fitness Only)	0.80	0.71	0.752
Our Previous approach	0.78	0.65	0.709
MCV	0.69	0.75	0.719
Random Forest	0.59	0.76	0.664

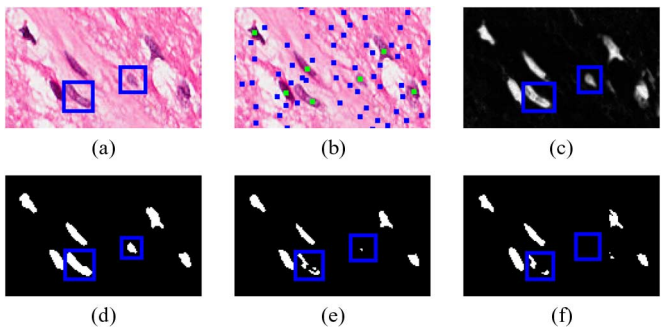


Fig. 10. A comparison among our approach, MCV, and random forest. (a) Original image patch. (b) Detected seeds, green: nuclei region; blue: background. (c) Local nuclei probability established based on seeds. (d) Classification by our approach. (e) Classification by MCV. (f) Classification by random forest.

3) As a result, ambiguous background structures are excluded, which leads to an increase of precision. However, there is also a decrease in the recall when compared to MRGC-CF, which is due to the fact that the tiny fragments inside the nuclei, as indicated by Fig. 9, can also be eliminated.
4) MRGC with multi-scale LoG features (MRGC-MS) has the best performance. We evaluated LoG responses at three scales, $\sigma \in \{2, 4, 6\}$, to compensate for a wide variation in the nuclear size. Improvement in segmentation is marginal, and it comes with a significant increase in the computational cost of about 40%. The LoG filter is simply used for seed detection to represent the underlying image statistics, and as long as a single scale can provide sufficient statistics, multiscale LoG is redundant. Besides, in processing whole slide images, computational throughput is an important factor.

We also provide an intuitive example, shown in Fig. 10, demonstrating the effectiveness of the local probability map. It is clear that the local probability map [Fig. 10(c)] helps to characterize nuclei with the low chromatin content, as shown in the blue bounding boxes. Another example, shown in Fig. 11, further demonstrates the effectiveness of our approach on the segmentation of low chromatin nuclei.

Finally, a comparison of the segmentation performance between our current approach and our previous approach [2] is indicated in Table III, where the correct nuclear segmentation is defined as follows.

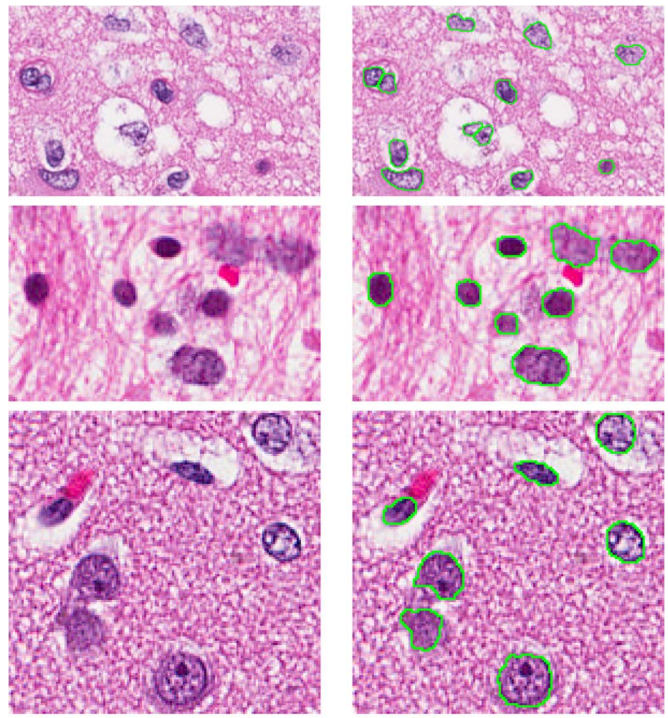


Fig. 11. Segmentation on low chromatin nuclei. (a) Original image patch. (b) Segmentation by our approach.

TABLE III
COMPARISON OF AVERAGE SEGMENTATION PERFORMANCE BETWEEN OUR CURRENT APPROACH(MRGC), AND OUR PREVIOUS APPROACH [2], IN WHICH $precision = \#correctly\_segmented\_nuclei/\#segmented\_nuclei$, AND $recall = \#correctly\_segmented\_nuclei/\#manually\_segmented\_nuclei$

Approach	Precision	Recall	F-Measure
MRGC	0.75	0.85	0.797
Our previous approach	0.63	0.75	0.685

- Let $MaxSize(a, b)$ be the maximum nuclear size of nuclei a and b.
- Let $Overlap(a, b)$ be the amount of overlap between nuclei a and b.

Subsequently, for any nucleus, n_G, from ground truth, if there is one and only one nucleus, n_S, in the segmentation result, that satisfies $\text{overlap}(n_G, n_S)/\text{maxsize}(n_G, n_S) > T$, then n_S is considered to be a correct segmentation of n_G. The threshold was set to be $T = 0.8$.

The reader may question the classification performance since both precision and recall are not very high. The major reason is that the ground truth (annotation) for the reference images is created at the object (nucleus) level, which means the hollow regions (loss of chromatin content for various reasons) inside the nuclei will be marked as the nuclear region rather than the background, as indicated by Fig. 9.

V. ANALYSIS OF TCGA GBM COHORT

Having evaluated the performance of the system, we applied our method to a cohort of 377 GBM whole slide images, from 146 patients, for bioinformatics analysis. Fig. 12 shows a few

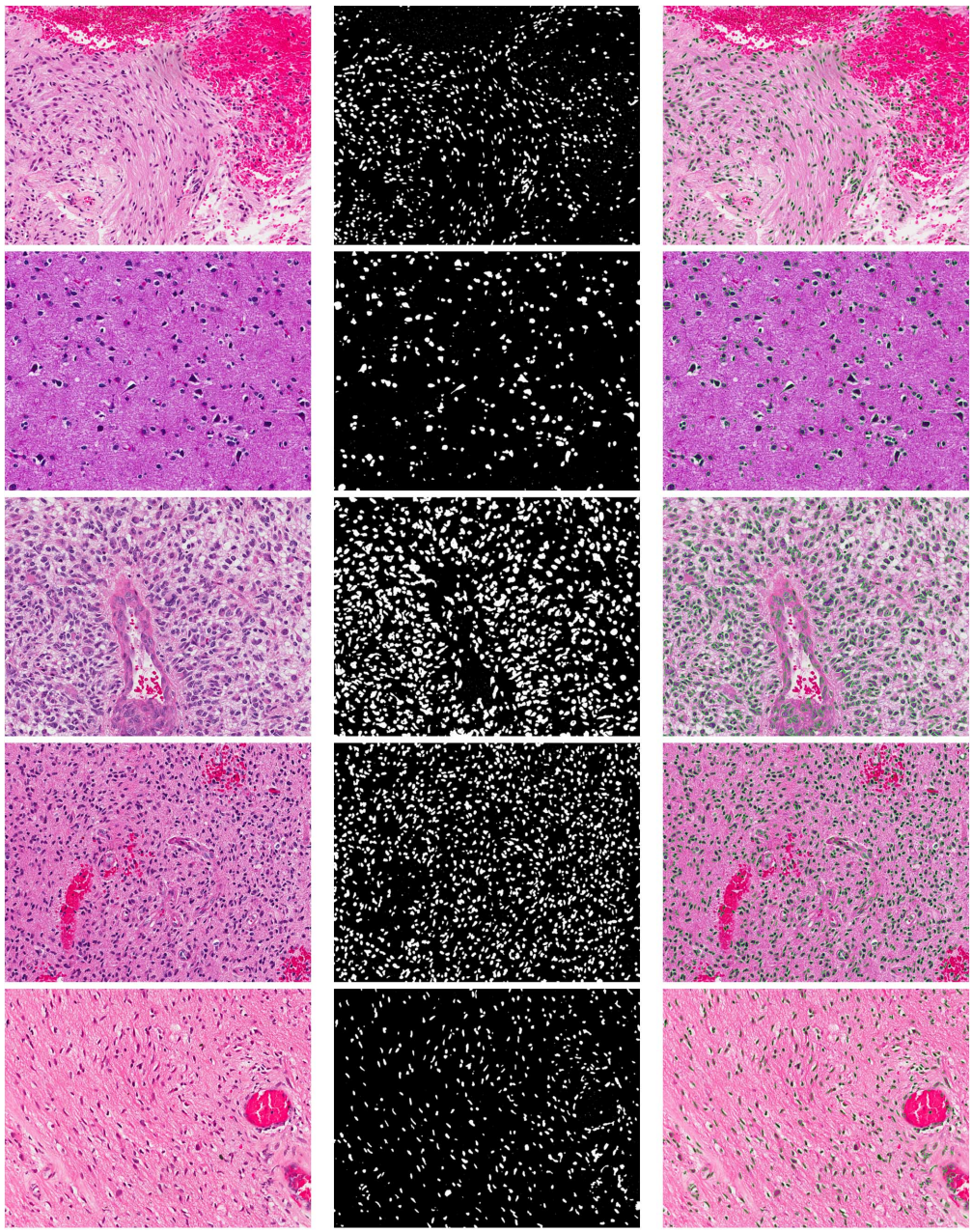


Fig. 12. Classification and segmentation results indicates tolerance to intrinsic variations. (a) Original images. (b) Nuclear/Background classification results via our approach(MRGC). (c) Nuclear partition results via geometric reasoning.

snapshots of our classification and segmentation results; Complete results for all the GBM tissue sections (and a few other tumor types) are available through the NIH web site at http://tcga-data.nci.nih.gov/tcga/. Following segmentation, each nucleus is represented by a multidimensional feature vector, which includes over 52 morphometric indexes such as nuclear size, cellularity, cytoplasmic features, etc., [2]. The density distribution of each index is then computed per histology section and aggregated per patient.

A particular aspect of bioinformatics analysis relies on subtyping based on a subset of computed morphometric indexes (e.g., cellular density), where subtyping is performed through consensus clustering [45], [46]. In our experiment, we evaluated all morphometric indexes and discovered that subtyping based on 1) nuclear size and cellularity, and 2) nuclear intensity and gradient, are statistically stable, where four and two subtypes were inferred, respectively. Fig. 14 shows the computed subtypes based on nuclear size and cellularity, where one of the subtypes is predictive of the outcome based on the clinical data. In addition, the computed subtypes from nuclear intensity and gradient were also predictive of the outcome. The patients in the GBM cohort received one of the two types of therapies 1) an intensive therapy with either concurrent radiation and chemotherapy, or four or more cycles of chemotherapy only, or 2) a less intensive therapy of either nonconcurrent radiation and chemotherapy or less than four cycles of chemotherapy only [47]. Although the sample size for the patient receiving the less intensive therapy is small, survival analyses [48] for

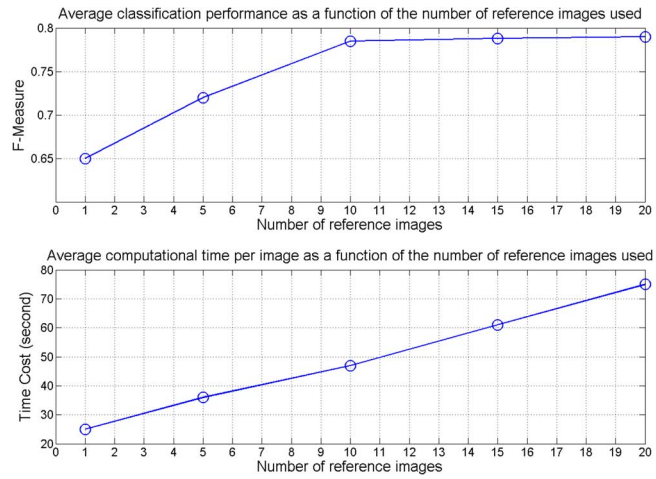


Fig. 13. Top and bottom rows show average classification performance and computational time as a function of number of reference images used. It is clear that the top $M = 10$ reference images with highest λ is a reasonable trade-off between performance and computational time.

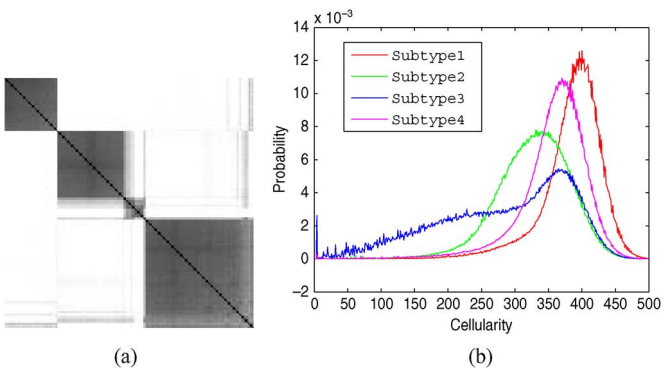


Fig. 14. Morphometric subtyping reveals four subtypes based on cellularity index and nuclear area: (a) visualization of consensus clustering with four clusters; and (b) distribution of cellularity index per subtype.

one of the subtypes in each of the clustering experiments points to a trend in an improved survival for patients receiving the more intensive therapy, as shown in Fig. 15. In addition, several computed subtypes, based on other morphometric indices, have also been found to be predictive of the outcome. We also examined molecular correlates of the predictive subtypes. With respect to predictive subtype computed from nuclear size and cellularity indexes, we used moderated t-test [49] and identified a set of differentially regulated transcripts for subtype 2 (e.g., predictive subtype) as shown in Fig. 16. A total of 10 differentially regulated transcripts were then subject to further bioinformatics analysis for subnetwork enrichment analysis using *Pathway Logic*, which computes and ranks hubs according to their p-values, as shown in Table IV (e.g., IL1, IL6), which impacts tumor proliferation and migration in both normal and malignant cells [50], [51] and the recruitment of the immune response. The relationships between these hubs and the genes associated with them are shown in Fig. 17. Among the common regulators, MAPK1 and FN1, which are involved in the proliferation, are highly ranked transcripts in TCGA's gene tracker for GBM. Furthermore, FN1 is 1) implicated in the invasion and

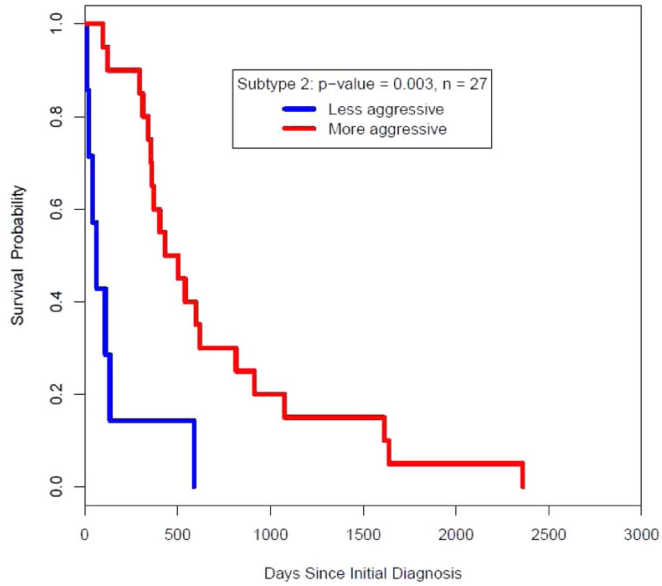


Fig. 15. Computed subtypes with cellularity and nuclear size is predictive as a result of more aggressive therapy.

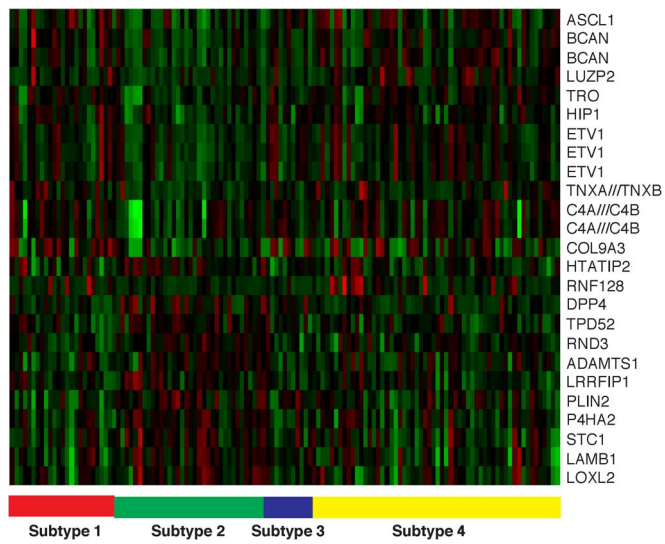


Fig. 16. Heat map representing a subset of differentially regulated transcripts for Subtype 2.

TABLE IV
KEY HUBS IDENTIFIED THROUGH PATHWAY ENRICHMENT ANALYSIS

Hub name	p-value
IL1A	0.0003
MAPK1	0.0005
FN1	0.0005
TNF	0.003
TGBF1	0.009
IL6	0.03

angiogenesis, and 2) validated as differentially expressed transcripts in GBM versus benign tumors [52]. Finally, TGFB1 is well known to be involved in tumor maintenance and progression through suppression of the immune response and is abundantly produced by GBM [53]. These molecular associations reflect that morphometric subtyping can hypothesize relevant

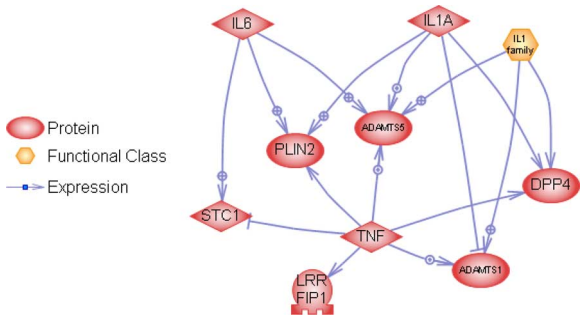


Fig. 17. Subnetwork enrichment analysis, for the predictive subtype in Fig. 15(a), reveals inflammatory hubs that promote tumor differentiation and invasiveness in GBM.

transcripts that are potential targets of therapy, which is consistent with current literature. An example being, FN1, and its role in the induction of angiogenesis. With respect to the predictive subtype computed from nuclear intensity and gradient indexes, subnetwork enrichment analysis revealed a large number of hubs from a set of differentially regulated transcripts. In this case, VEGF was discovered to be at the intersection of all pathways curated through enrichment analysis. VEGF is well known to be the hallmark of glioblastoma for the induction of microvasculature formation [54] and has been suggested as a therapuetic target in GBM [55].

VI. Conclusion

We have shown that morphometric representation of cellular architecture from a large cohort of histology sections can provide new opportunities for hypothesis generation. The main barriers are the batch effect and tumor heterogeneity which hinders nuclear segmentation. However, through image-based modeling, technical and tumor variations can be captured for robust nuclear segmentation from whole slide images. Subsequently, segmented nuclei and corresponding computed morphometric representation enables characterization of tumor histopathology. Our approach for nuclear segmentation addresses technical and biological variations by 1) utilizing global information from a diverse set of annotated reference images, 2) normalizing the test image against the reference images in the color space, and 3) incorporating local variations in the test image. Segmentation is formulated within a graph cut framework with geodesic constraint for improved accuracy of the nuclear boundaries. The method has been validated against annotated data and applied to a large dataset of GBM tumor cohort to identify subtypes as a function of cellularity and nuclear size. One of these subtypes is shown to have an increase in survival as a result of a more aggressive therapy with an underlying molecular signature that is consistent with invasiveness and proliferation.

References

[1] L. Dalton, S. Pinder, C. Elston, I. Ellis, D. Page, W. Dupont, and R. Blamey, "Histolgical gradings of breast cancer: Linkage of patient outcome with level of pathologist agreements," *Modern Pathol.*, vol. 13, no. 7, pp. 730–735, 2000.

[2] H. Chang, G. Fontenay, J. Han, G. Cong, F. Baehner, J. Gray, P. Spellman, and B. Parvin, "Morphometric analysis of TCGA gliobastoma multiforme," *BMC Bioinformat.*, vol. 12, no. 1, 2011.

[3] C. Demir and B. Yener, Automated cancer diagnosis based on histopathological images: A systematic survey Dept. Comput. Sci., Rensselaer Polytech. Inst., Tech. Rep., 2009.

[4] M. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, and Y. Bulent, "Histopathological image analysis: A review," *IEEE Rev. Biomed. Eng.*, vol. 2, pp. 147–171, 2009.

[5] D. Axelrod, N. Miller, H. Lickley, J. Qian, W. Christens-Barry, Y. Yuan, Y. Fu, and J. Chapman, "Effect of quantitative nuclear features on recurrence of ductal carcinoma in situ (DCIS) of breast," *Cancer Informat.*, vol. 4, pp. 99–109, 2008.

[6] M. Datar, D. Padfield, and H. Cline, "Color and texture based segmentation of molecular pathology images using HSOMs," in *Proc. Int. Symp. Biomed. Imag.*, 2008, pp. 292–295.

[7] A. Basavanhally, J. Xu, A. Madabhushu, and S. Ganesan, "Computer-aided prognosis of ER+ breast cancer histopathology and correlating survival outcome with oncotype DX assay," in *Proc. Int. Symp. Biomed. Imag.*, 2009, pp. 851–854.

[8] S. Doyle, M. Feldman, J. Tomaszewski, N. Shih, and A. Madabhushu, "Cascaded multi-class pairwise classifier (CASCAMPA) for normal, cancerous, and cancer confounder classes in prostate histology," in *Proc. Int. Symp. Biomed. Imag.*, 2011, pp. 715–718.

[9] R. Bhagavatula, M. Fickus, W. Kelly, C. Guo, J. Ozolek, C. Castro, and J. Kovacevic, "Automatic identification and delineation of germ layer components in $h\&e$ stained images of teratomas derived from human and nonhuman primate embryonic stem cells," in *Proc. Int. Symp. Biomed. Imag.*, 2010, pp. 1041–1044.

[10] J. Kong, L. Cooper, A. Sharma, T. Kurk, D. Brat, and J. Saltz, "Texture based image recognition in microscopy images of diffuse gliomas with multi-class gentle boosting mechanism," in *Proc. Int. Conf. Acoust., Speech, Signal Process.*, 2010, pp. 457–460.

[11] J. Han, H. Chang, L. Loss, K. Zhang, F. Baehner, J. Gray, P. Spellman, and B. Parvin, "Comparison of sparse coding and kernel methods for histopathological classification of glioblastoma multiforme," in *Proc. Int. Symp. Biomed. Imag.*, 2011, pp. 711–714.

[12] E. Acar, G. Plopper, and B. Yener, "Coupled analysis of in vitro and histology samples to quantify structure-function relationships," *PLoS One*, vol. 7, no. 3, p. e32227, 2012.

[13] C. Bilgin, S. Ray, B. Baydil, W. Daley, M. Larsen, and B. Yener, "Multiscale feature analysis of salivary gland branching morphogenesis," *PLoS One*, vol. 7, no. 3, p. e32906, 2012.

[14] H. Fatakdawala, J. Xu, A. Basavanhally, G. Bhanot, S. Ganesan, F. Feldman, J. Tomaszewski, and A. Madabhushi, "Expectation-maximization-driven geodesic active contours with overlap resolution (EMaGACOR): Application to lymphocyte segmentation on breast cancer histopathology," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 7, pp. 1676–1690, Jul. 2010.

[15] P. Phukpattaranont and P. Boonyaphiphat, "Color based segmentation of nuclear stained breast cancer cell images," *ECTI Trans. Elect. Eng., Commun.*, vol. 5, no. 2, pp. 158–164, 2007.

[16] B. Ballaro, A. Florena, V. Franco, D. Tegolo, C. Tripodo, and C. Valenti, "An automated image analysis methodology for classifying megakaryocytes in chronic myeloproliferative disorders," *Med. Image Anal.*, vol. 12, pp. 703–712, 2008.

[17] L. Latson, N. Sebek, and K. Powell, "Automated cell nuclear segmentation in color images of hematoxylin and eosin-stained breast biopsy," *Anal. Quant. Cytol. Histol.*, vol. 26, no. 6, pp. 321–331, 2003.

[18] W. Land, D. McKee, T. Zhukov, D. Song, and W. Qian, "A kernelized fuzzy support verctor machine CAD system for the diagnostic of lung cancer from tissue," *Int. J. Funct. Inf. Personal. Med.*, vol. 1, no. 1, pp. 26–52, 2008.

[19] D. Glotsos, P. Spyridonos, D. Cavouras, P. Ravazoula, P. Dadioti, and G. Nikiforidis, "Automated segmentation of routinely hematoxyli-eosin stained microscopic images by combining support vector machine, clustering, and active contour models," *Anal. Quant. Cytol. Histol.*, vol. 26, no. 6, pp. 331–340, 2004.

[20] H. Chang, R. Defilippis, T. Tlsty, and B. Parvin, "Graphical methods for quantifying macromolecules through bright field imaging," *Bioinformatics*, vol. 25, no. 8, pp. 1070–1075, 2009.

[21] E. Cosatto, M. Miller, H. Graf, and J. Meyer, "Grading nuclear plemorphism on histological micrographs," in *Int. Conf. Pattern Recognit.*, 2008, pp. 1–4.

[22] S. Petushi, F. Garcia, M. Haber, C. Katsinis, and A. Tozeren, "Large-scale computations on histology images reveal grade-differentiation parameters for breast cancer," *BMC Med. Imag.*, vol. 6, no. 14, pp. 1070–1075, 2006.

[23] S. Doyle, S. Agner, A. Madabhushi, and M. F. Tomaszewski, "Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features," in *Proc. Int. Symp. Biomed. Imag.*, 2008, pp. 496–499.

[24] F. Bunyak, A. Hafiane, and K. Palanippan, "Histopathology tissue segmentation by combining fuzzy clustering with multiphase vector level set," *Adv. Exp. Med. Biol.*, vol. 696, pp. 413–424, 2011.

[25] B. Parvin, Q. Yang, J. Han, H. Chang, and B. R. Barcellos-Hoff, "Iterative voting for inference of structural saliency and characterization of subcellular events," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 615–623, Mar. 2007.

[26] S. Nath, K. Palaniappan, and F. Bunyak, "Cell segmentation using coupled level sets and graph-vertex," in *Proc. Med. Image Comput. Comput.-Assist. Intervent.*, 2006, pp. 101–108.

[27] H. Chang and B. Parvin, "Multiphase level set for automated delineation of membrane-bound macromolecules," in *Proc. Int. Symp. Biomed. Imag.*, 2010, pp. 165–168.

[28] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 4, pp. 841–852, Apr. 2010.

[29] H. Kong, M. Gurcan, and K. Belkacem-Boussaid, "Partitioning histopathological images: An integrated framework for supervised color-texture segmentation and cell splitting," *IEEE Trans. Med. Imag.*, vol. 30, no. 9, pp. 1661–1677, Sep. 2011.

[30] J. Monaco, J. Hipp, D. Lucas, S. Smith, U. Balis, and A. Madabhushi, "Image segmentation with implicit color standardization using spatially constrained expectation maximization: Detection of nuclei," in *Proc. Med. Image Comput. Comput.-Assist. Intervent.*, 2012, pp. 365–372.

[31] S. Kothari, J. H. Phan, R. A. Moffitt, T. H. Stokes, S. E. Hassberger, Q. Chaudry, A. N. Young, and M. D. Wang, "Automatic batch-invariant color segmentation of histological cancer images," in *Proc. Int. Symp. Biomed. Imag.*, 2011, pp. 657–660.

[32] C. Tomasi, Estimating Gaussian mixture densities with EM—A tutorial [Online]. Available: http://www.cs.duke.edu/courses/spring04/cps196.1/handouts/EM/tomasiEM.pdf 2004

[33] A. Ruifork and D. Johnston, "Quantification of histochemical staining by color decomposition," *Anal. Quant. Cytol. Histol.*, vol. 23, no. 4, pp. 291–299, 2001.

[34] A. Rabinovich, S. Agarwal, C. Laris, J. H. Price, and S. Belongie, "Unsupervised color decomposition of histologically stained tissue samples," in *Conf. Neural Inf. Process. Syst.*, 2003, pp. 667–674.

[35] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 6, no. 6, pp. 721–741, Nov. 1984.

[36] A. V. Goldberg and R. E. Tarjan, "A new approach to maximum-flow problem," *J. Assoc. Comput. Mach.*, vol. 35, no. 4, pp. 921–940, 1988.

[37] L. Ford and D. Fullkerson, *Flows in Networks*. Princeton, NJ: Princeton Univ. Press, 1962.

[38] W. J. Cook, W. H. Cunningham, W. R. Pulleyblank, and A. Schrijver, *Combinatorial Optimization*. New York: Wiley, 1998.

[39] Y. Boykov and V. Kolmogorov, "Computing geodesics and minimal surfaces via graph cuts," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2003, vol. 1, pp. 26–33.

[40] L. A. Santalo, *Integral Geometry and Geometric Probability*. Reading, MA: Addison-Wesley, 1979.

[41] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 9, pp. 1124–1137, Sep. 2004.

[42] S. Raman, C. Maxwell, M. Barcellos-Hoff, and B. Parvin, "Geometric approach segmentation and protein localization in cell cultured assays," *J. Microscopy*, pp. 427–436, 2007.

[43] Q. Wen, H. Chang, and B. Parvin, "A Delaunay triangulation approach for segmenting clumps of nuclei," in *Proc. Int. Symp. Biomed. Imag.*, 2009, pp. 9–12.

[44] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[45] S. Monti, P. Tamayo, J. Mesirov, and T. Golub, "Consensus clustering—A resampling-based method for class discovery and visualization of gene expression microarray data," in *Mach. Learn.*, 2003, pp. 91–118.

[46] J. Han, H. Chang, O. Giricz, G. Lee, F. Baehner, J. Gray, M. Bissell, P. Kenny, and B. Parvin, "Molecular predictors of 3D morphogenesis by breast cancer cells in 3D culture," *PLoS Computat. Biol.*, vol. 6, no. 2, p. e1000684, 2010.

[47] R. G. Verhaak, K. A. Hoadley, E. Purdom, V. Wang, Y. Qi, M. D. Wilkerson, C. R. Miller, L. Ding, T. Golub, J. P. Mesirov, G. Alexe, M. Lawrence, M. O'Kelly, P. Tamayo, B. A. Weir, S. Gabrie, W. Winckler, S. Gupta, L. Jakkula, H. S. Feiler, J. G. Hodgson, C. D. James, J. N. Sarkaria, C. Brennan, A. Kahn, P. T. Spellman, R. K. Wilson, T. P. Speed, J. W. Gray, M. Meyerson, G. Getz, C. M. Perou, D. N. Hayes, and T. C. G. A. R. Network, "Integrated genomic analysis identifies clinically relevant subtypes of glioblastoma characterized by abnormalities in PDGFRA, IDH1, EGFR, and NF1," *Cancer Cell*, vol. 17, no. 1, pp. 98–110, 2010.

[48] P. Meier and E. Kaplan, "Nonparametric estimation from incomplete observations," *J. Am. Stat. Assoc.*, vol. 53, pp. 457–481, 1958.

[49] G. Smyth, "Linear models and empirical Bayes methods for assessing differential expression in microarray experiments," *Stat. Appl. Genet. Mole. Biol.*, vol. 3, no. 3, 2004.

[50] B. Paugh, L. Bryan, S. Paugh, K. Wilczynska, S. Alvarez, S. Singh, D. Kapitonov, H. Rokita, S. Wright, I. Griswold-Prenner, S. Milstien, S. Spiegel, and T. Kordula, "Interleukin-1 regulates the expression of shpingosone kinase 1 in glioblastoma cells," *J. Biol. Chem.*, vol. 284, no. 6, pp. 3408–3417, 2009.

[51] Q. Liu, R. Li, J. Shen, Q. He, L. Deng, C. Zhang, and J. Zhang, "Il-6 promotion of glioblastoma cell invasion and angiogenesis in u251 and t98 cell lines," *J. Neurooncol.*, vol. 100, no. 2, pp. 165–176, 2010.

[52] C. Colin, N. Baeza, C. Bartoli, F. Fina, N. Eudes, I. Nanni, P. Martin, L. Ouafik, and D. Figarella-Branger, "Identification of genes differentially expressed in glioblastoma versus pilocytic astrocytoma using suppression subtractive hybridization," *Oncogenomics*, vol. 25, pp. 2818–2826, 2006.

[53] M. Barcellos-Hoff, E. Newcomb, D. Zagzag, and A. Narayana, "Therapeutic targets in malignant glioblastoma microenvironment," *Seminal Radiat. Oncol.*, vol. 19, pp. 163–170, 2009.

[54] R. Jain, T. Di, D. Duda, J. Loeffler, A. Sorensen, and T. Batchelor, "Angiogenesis in brain tumors," *Nature Rev. Neurosci.*, vol. 8, no. 8, pp. 610–622, 2007.

[55] A. Hormigo, B. Ding, and S. Rafii, "A target for antiangiogenic therapy: Vascular enothelium derived from glioblastoma," *Proc. Nat. Acad. Sci.*, vol. 108, no. 11, pp. 4271–4272, 2011.