Leveraging unsupervised training sets for multi-scale compartmentalization in renal pathology

Brendon Lutnick, John Tomaszewski, Pinaki Sarder

Brendon Lutnick, John E. Tomaszewski, Pinaki Sarder, "Leveraging unsupervised training sets for multi-scale compartmentalization in renal pathology," Proc. SPIE 10140, Medical Imaging 2017: Digital Pathology, 101400I (1 March 2017); doi: 10.1117/12.2254750

SPIE.

Event: SPIE Medical Imaging, 2017, Orlando, Florida, United States

Leveraging Unsupervised Training Sets for Multi-Scale Compartmentalization in Renal Pathology

Brendon Lutnick[a], John E. Tomaszewski[a,b], Pinaki Sarder[a,c,d,*]

Departments of [a]Pathology and Anatomical Sciences, [b]Biomedical Informatics, [c]Biomedical Engineering, and [d]Biostatistics

University at Buffalo – The State University of New York,
207 Farber Hall, 3435 Main Street
Buffalo, NY, 14214, USA

[*]Address all correspondence to: Pinaki Sarder
Tel: 716-829-2265; E-mail: pinakisa@buffalo.edu

ABSTRACT

Clinical pathology relies on manual compartmentalization and quantification of biological structures, which is time consuming and often error-prone. Application of computer vision segmentation algorithms to histopathological image analysis, in contrast, can offer fast, reproducible, and accurate quantitative analysis to aid pathologists. Algorithms tunable to different biologically relevant structures can allow accurate, precise, and reproducible estimates of disease states. In this direction, we have developed a fast, unsupervised computational method for simultaneously separating all biologically relevant structures from histopathological images in multi-scale. Segmentation is achieved by solving an energy optimization problem. Representing the image as a graph, nodes (pixels) are grouped by minimizing a Potts model Hamiltonian, adopted from theoretical physics, modeling interacting electron spins. Pixel relationships (modeled as edges) are used to update the energy of the partitioned graph. By iteratively improving the clustering, the optimal number of segments is revealed. To reduce computational time, the graph is simplified using a Cantor pairing function to intelligently reduce the number of included nodes. The classified nodes are then used to train a multiclass support vector machine to apply the segmentation over the full image. Accurate segmentations of images with as many as 10^6 pixels can be completed only in 5 sec, allowing for attainable multi-scale visualization. To establish clinical potential, we employed our method in renal biopsies to quantitatively visualize for the first time scale variant compartments of heterogeneous intra- and extra-glomerular structures simultaneously. Implications of the utility of our method extend to fields such as oncology, genomics, and non-biological problems.

Keywords: Segmentation, Unsupervised, Hamiltonian, Potts Model, Graph Theory, Glomerulus

1. INTRODUCTION

Segmentation is crucial in any large-scale data analysis problem. The ability to cluster data into significant classes is important in many domains of health science, providing quantitative insight into the fundamental data structure. Unsupervised learning techniques are often preferable, as data groups can be learned from data relationships (determined from features) with no prior knowledge, meaning that segmentation is determined from the unique data structure without any bias from predetermined class labels. This is particularly useful in image segmentation tasks as large training sets are not required for segmentation, which can reveal unique communities not immediately apparent. When viewing tissue slides, the pathologist relies on the minds inherent separation of the image into distinct regions. The expert's ability to identify and separate hidden biologically relevant compartments and resolutions is why manual segmentation and diagnosis is still today's golden standard. However, generation of expert segmentations specific enough for quantitative analysis is extremely time consuming, and a waste of talent. For this reason, glomerular diagnosis has thus far been limited to qualitative analysis of glomerular features. Capitalizing on the multi-scale abilities of unsupervised Potts Model Hamiltonian[1] based segmentation methods[2, 3], we intend to develop a quantitative computational toolbox to aid expert diagnosis. Our algorithm mimics experts' ability to cluster regions in multiple scales, but examines images on a pixel by pixel basis, allowing the discovery of multi-scale biologically relevant clusters, with extractable quantitative features. Unsupervised segmentation using a Potts model Hamiltonian is known in the literature to be precise and accurate[4]. Previously this method produced good results, but was computationally unobtainable for large images[3, 5]. Previous applications took ~72 hr to complete segmentation of *RGB* images with 256×256 pixels. Our algorithm combines this

Medical Imaging 2017: Digital Pathology, edited by Metin N. Gurcan, John E. Tomaszewski, Proc. of SPIE Vol. 10140, 101400I · © 2017 SPIE · CCC code: 1605-7422/17/$18 · doi: 10.1117/12.2254750

unsupervised method with well-established supervised machine learning techniques, in an image analysis pipeline, requiring only one optional scale parameter for image segmentation. The multi-scale segmentations achieved by our method have never before been feasible on the timescale that we can achieve (*seconds*). Due to the computational speed, and advantage of obtaining multi-scale segmentation without parameter tuning, our method provides the opportunity to segment diverse scale structures in clinical histopathology images.

2. METHOD

At the heart of our method is the Potts Model Hamiltonian[1]. Representing an image using graph theoretic approaches (pixels are nodes, and pixel relationships edges), unsupervised segmentation can be achieved by minimizing the resulting Hamiltonian, a cost function adopted from theoretical physics used to group electrons by modeling the interactions of their spins. Pixel relationships (modeled as edges) are used to update the energy of the partitioned graph, which is iteratively improved until convergence. Traditional graph based image segmentation methods include all image pixels as nodes, and large images become quickly unmanageable as the number of edges increases with ~$pixels^2$. To increase computational efficiency, traditional methods may reduce the number of edges included, or create graphs based on region adjacency[6]; leading to inaccurate region dependent segmentation. By only considering unique pixels, our method achieves segmentation accurate on the pixel scale with unprecedented speed. By intelligently reducing the included nodes, efficiency is optimized while preserving segmentation accuracy.

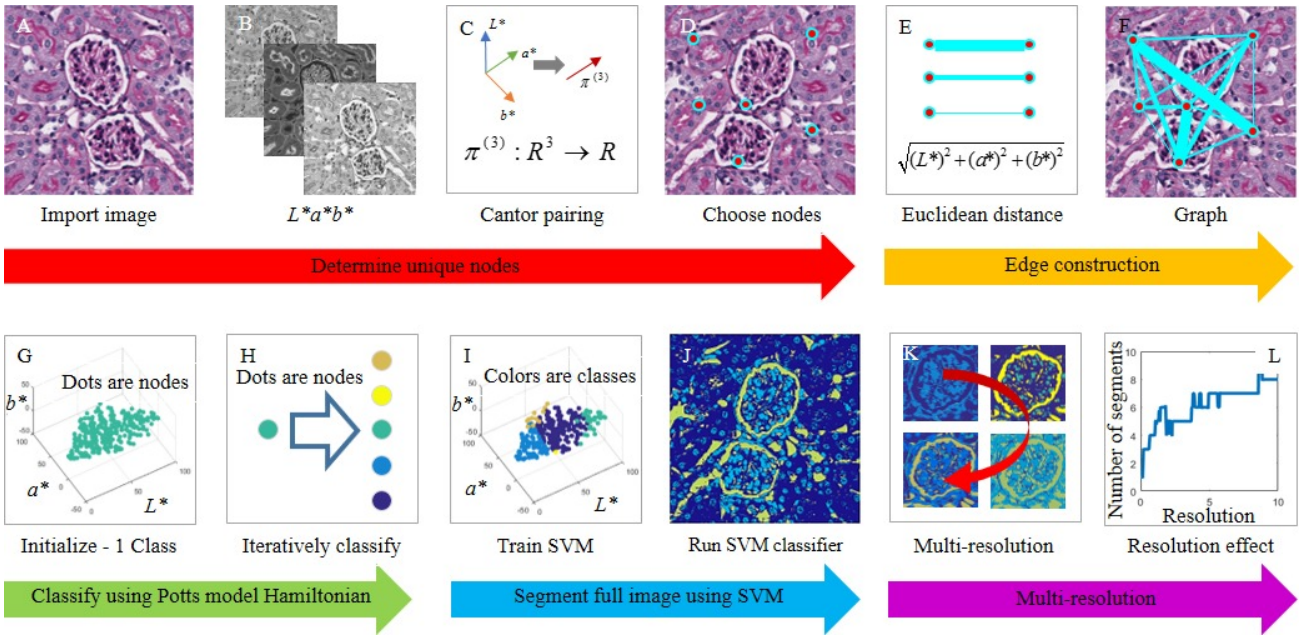

Figure 1. Demonstration of the multi-scale segmentation pipeline. (A) RGB image of a renal histopathology image. (B) *L*a*b** color space image[7, 8]. (C) Reduction of three unique image dimensions to one using Cantor pairing[9]. (D) Unique nodes chosen from the RGB image shown in (A). (E) Calculation of the edges using Euclidean distance measure. (F) Sparse graph representation of the image using unique nodes in (D). (G) Nodes initially given for classification to the Potts model Hamiltonian based multi-scale segmentation algorithm[2, 3]. (H) Optimal classes obtained based on Potts model based segmentation. (I) Classification of nodes using Potts model, which are used to train a support vector machine (SVM)[10]. (J) Clustered image using the resulting SVM. (K) Segmentation output in multi-scale (resolution). (L) Number of classes plotted against scale (resolution).

Segmentation is achieved by our method in two parts. Firstly, an unsupervised method is used to obtain an initial segmentation of a selected set of pixels (Figs. 1A-1H). This segmentation is then used to train a support vector machine (SVM)[10], which then applies this segmentation to the full image (Figs. 1I-1J). The initial segmentation is achieved by solving an energy optimization problem. This segmentation method uses only one optional scale parameter to tune to important structures in varying scale[2, 3]. This parameter effectively scales the algorithm's focus, zooming in at higher scales to segment finer structures. Representing the image as a graph, nodes (pixels) are grouped by minimizing a Potts model Hamiltonian[11], adopted from theoretical physics, modeling interacting electron spins. Pixel relationships (modeled as edges) are used to update the energy of the partitioned graph. The iteration step starts with all nodes in the same class, and using an iterative strategy, the optimal number of segments are determined, offering lowest Potts model energy of the

partitioned graph. Due to the large number of nodes and edges in the graph of a typical tissue histology image, a sparse graph is used to improve the computational speed.

2.1. Dataset definition

The input for our method is a vectorized three-channel RGB image $p(R,G,B)$ (Fig. 1A) of size $M \times K$, given by,

$$p(R,G,B) \in \mathbb{N} \mid R,G,B \in \{0,1,2,\ldots,255\}, \tag{1}$$

where $N = MK$ is the number of pixels. Our method first converts the image color-space $p(R,G,B) \rightarrow p(L,a,b)$ (Fig. 1B). The $L^*a^*b^*$ color space[7, 8] is designed to be perceptually uniform, mimicking the nonlinear response of the human eye. Therefore, perceptual color differences can be estimated using Euclidean distance[12].

2.2. Preliminary reduction of data using rounding

We round the data to the nearest 10 first, and obtain $p(L^{'},a^{'},b^{'})$, where $L^{'}$, $a^{'}$, and $b^{'}$ are given by,

$$L^{'} = \frac{\lfloor 10L \rfloor}{10}, \tag{2}$$

$$a^{'} = \frac{\lfloor 10a \rfloor}{10}, \tag{3}$$

$$b^{'} = \frac{\lfloor 10b \rfloor}{10}, \tag{4}$$

where $\lfloor \cdot \rfloor$ denotes the flooring operation. Using this method, the number of nodes in the resultant graph is reduced by ~1000 times, while $p(L,a,b)$ and $p(L^{'},a^{'},b^{'})$ have structural similarities of ~0.95, giving accurate segmentation, using the sparse graph created from Cantor pairing[9] of $p(L^{'},a^{'},b^{'})$, as discussed in Section 2.5.

2.3. Graph definition

To perform clustering, the vectorized input data must be represented as a graph (Fig. 1F), where Φ are nodes, and E are edges. These metrics are given by,

$$\Phi = \{1,2,...,n\} \mid n > 1, \tag{5}$$

$$E = \{1,2,...,e\}, \tag{6}$$

$$e = \binom{n}{2}, \tag{7}$$

where n is the number of nodes and e the number of edges. Using an efficient node selection process, the clustering speed is optimized.

2.4. Pixel selection

In a 3 channel 8-bit image, there are 256^3 unique colors, the full image color space therefore has a maximum of 16777216 unique colors. In a common histological slide, the color space is much reduced, including many repeated colors. By representing only unique pixels as nodes, the sample size can be reduced from the original image pixel dimensions. For the image with N pixels as defined in Section 2.1 and using the reduced $p(L^{'},a^{'},b^{'})$ in Section 2.2, we have:

$$p(L^{'},a^{'},b^{'}) = \{1,2,...,N\}. \tag{8}$$

Pixels with unique colors are assigned as nodes Φ, as defined in Eq. (5), where:

$$\Phi \subseteq \{1,2,...,N\}. \tag{9}$$

To find unique pixels, we use a Cantor Pairing function[9] (Fig. 1C) to transform the three-dimensional color-space (*L*a*b**) into a one-dimensional vector, where unique colors are given arbitrary values, given by:

$$\pi(k_1, k_2) = \frac{1}{2}(k_1 + k_2)(k_1 + k_2 + 1) + k_2, \tag{10}$$

where k_1 and k_2 are intensity values in two image channels. The above equation is repeated twice to transform each *RGB* color into a unique value $\pi(L, a, b)$, where:

$$\Phi \equiv p(\pi(L, a, b)). \tag{11}$$

Nodes are selected by choosing pixels with unique Cantor values $\pi(L, a, b)$ (Fig. 1D).

2.5. Potts model definition

Using the unique image super-pixels, a sparse graph can be constructed and clustered, assuming that similar pixels will be clustered into the same segment. The image super-pixels found by Cantor pairing are considered as nodes Φ. Edges between the i^{th} and j^{th} nodes are given by,

$$E_{ij} = \sqrt{(\Delta L)^2 + (\Delta a)^2 + (\Delta b)^2}, \tag{12}$$

where ΔL, Δa, and Δb are the Euclidean distances between the respective *L*a*b** components of the nodes (Fig. 1E). Graph model formed using this way contains very low number of unique nodes, and therefore for the Potts model based optimization, a full set of edges can be included, increasing the resulting classification accuracy.

The graph is clustered by minimizing a Potts Model Hamiltonian[11],

$$\hbar = \frac{1}{2}\sum_{i \neq j}(E_{ij} - \bar{E})\left[\Theta(\bar{E} - E_{ij}) + \gamma\Theta(E_{ij} - \bar{E})\right]\delta(\Phi_i, \Phi_j), \tag{13}$$

where $\bar{E}$ denotes the background (average) edge value, and Φ_i and Φ_j denote the i^{th} and j^{th} nodes, respectively. The Heaviside function is given by,

$$\Theta(E_{ij} - \bar{E}) = \begin{cases} 1, & E_{ij} > \bar{E} \\ 0, & \text{otherwise} \end{cases}, \tag{14}$$

and γ is a scale parameter. Decreasing γ results in segments with lower intra-community density, revealing larger communities. The Kronecker delta is given by,

$$\delta(\Phi_i, \Phi_j) = \begin{cases} 1, & \Phi_i = \Phi_j \\ 0, & \text{otherwise} \end{cases}, \tag{15}$$

where $\delta(\Phi_i, \Phi_j)$ ensures that each node (spin) Φ_i interacts only with nodes in its own segment. We define $\Phi_i \in \{1, 2, \ldots, S\}$ as the segment identity for the i^{th} node, with S the maximum number of segments, optimized by minimizing the Hamiltonian energy $\hbar$. Here minimization of the Hamiltonian $\hbar$ partitions the graph space into connected segments. Unknowns involved are the segment identities Φ_i and S.

2.6. Potts model optimization

We optimize the Potts model Hamiltonian to obtain segments $C_{\Phi_1}, C_{\Phi_2}, \ldots, C_{\Phi_S}$, where

$$C_\Phi = \bigcup_i C_{\Phi_i}, \tag{16}$$

and,

$$C_\Phi \subseteq \Phi. \tag{17}$$

Initially the nodes are assigned the same starting class C_Φ (Fig. 1G). The algorithm then moves pixel by pixel, and determine whether a pixel remains in its own segment, moves to a different segment, or forms a new segment, by computing the cost defined in Eq. (13) for each of such cases (Fig. 1H). Namely, suppose the segments are $\{C_{\Phi_1}, C_{\Phi_2}, \ldots, C_{\Phi_s}\}$ in the t^{th} iteration with number of segments to be $s \leq S$. The optimization process scans each node/pixel in the i^{th} segment C_{Φ_i} and investigates whether this pixel will remain in C_{Φ_i} or in C_{Φ_j} with $i \neq j$ or can form a new segment $C_{\Phi_{s+1}}$ to deliver a smaller Hamiltonian score than the current score. The process continues until no energy lowering move happens in a complete cycle of scanning all the pixels. Practically, we found that optimal histopathological image segmentation is obtained within 5 such iterations.

2.7. Full image segmentation

The sparse graph clustering is used to intelligently divide the color space of the image into relevant segments. To obtain a segmented image, the Potts model Hamiltonian based clustering is applied to the full image using multiclass SVMs[10] trained on the Potts model output (Fig. 1I). The resulting multi-class SVM is used to determine the optimal segments of the input image pixels based on their *L*a*b** values (Fig. 1J). Using this method, the initial unsupervised segmentation can be reused for images of a similar color space, and thus extends the initial segmentation to adjacent tissue regions.

2.8. Multi scale segmentation

To achieve multi-scale segmentation the above process is repeated using different scale parameters γ (Fig. 1K). Here γ effects the sensitivity of the algorithm to small structure by tuning the algorithms sensitivity to image background. Segmentation using higher γ values results in smaller communities and a greater number of segments (Fig. 1L), given by,

$$\gamma \nearrow \rightarrow S \nearrow . \tag{18}$$

It is useful to change γ exponentially, as to maintain a constant proportionality between scales. This helps to prevent generation of identical segmentations. Therefore, to select *R* scales:

$$\gamma_i = \exp(i) \,|\, i \in \{0, 1, 2, \ldots, R\}. \tag{19}$$

3. RESULTS

Intact kidney tissues from normal healthy untreated mice were collected (generously obtained by Dr. Rabi Yacoub, Department of Medicine – Nephrology, University at Buffalo) and sacrificed under their institutionally approved animal protocol. Tissues were formalin fixed, processed in a standard fashion, embedded in paraffin blocks, microtome sectioned at 2μm and stained with Periodic Acid-Schiff (PAS). Imaging was conducted using a whole slide bright field microscope (Aperio, Leica, Buffalo Grove, IL) using a 40X objective with 0.75NA. The computational method proposed here is independent of this imaging configuration, and no parameter tuning is required for obtaining the segmentation result presented in Fig. 2.

To validate our method, we use histopathological glomerular images of healthy mice. The biological relevance of glomerular segmentation has been validated by board certified renal pathologist (co-author J.E.T.). We present an example of our multi-scale segmentation in Fig. 2. Fig. 2A shows the original RGB image of a renal tissue histopathology image, depicting renal glomeruli and tubules. Application of our computational segmentation method reveals reproducible structures by tuning the scale parameter γ. By tuning sensitivity to image background the scale parameter can be used to reveal unique information, with structures segmented at higher scales. Structures revealed are Bowman's space (γ = 0.3); intra-glomerular nuclei (γ = 0.53); glomerular basement membrane, Bowman's capsule, and mesangial matrix (γ = 0.75); mesangial matrix (γ = 2.0); intra- and extra-glomerular nuclei (γ = 8.3), and inner borders of tubular lumen (γ = 9.75). The segmented structures can be easily quantified to classify disease states; particularly, to determine digital bio-markers to classify early disease states for developing therapeutic intervention strategies to slow down disease progression.

To analyze the performance of our method, we compared the performance against expert annotation. Renal tissue slices from 3 mice were stained using PAS and imaged. Staining was automated to reduce variability in the data. Ten glomeruli were randomly selected per mouse to be annotated by expert pathologist (J.E.T), to generate a gold standard. The best Bowman's, lumen, and mesangial segmentations generated by our method were compared to this gold standard annotation,

yielding 98.9% precision and 93.7% accuracy. Our method accomplished this result with an average of 85× faster speed than the expert annotation per image.

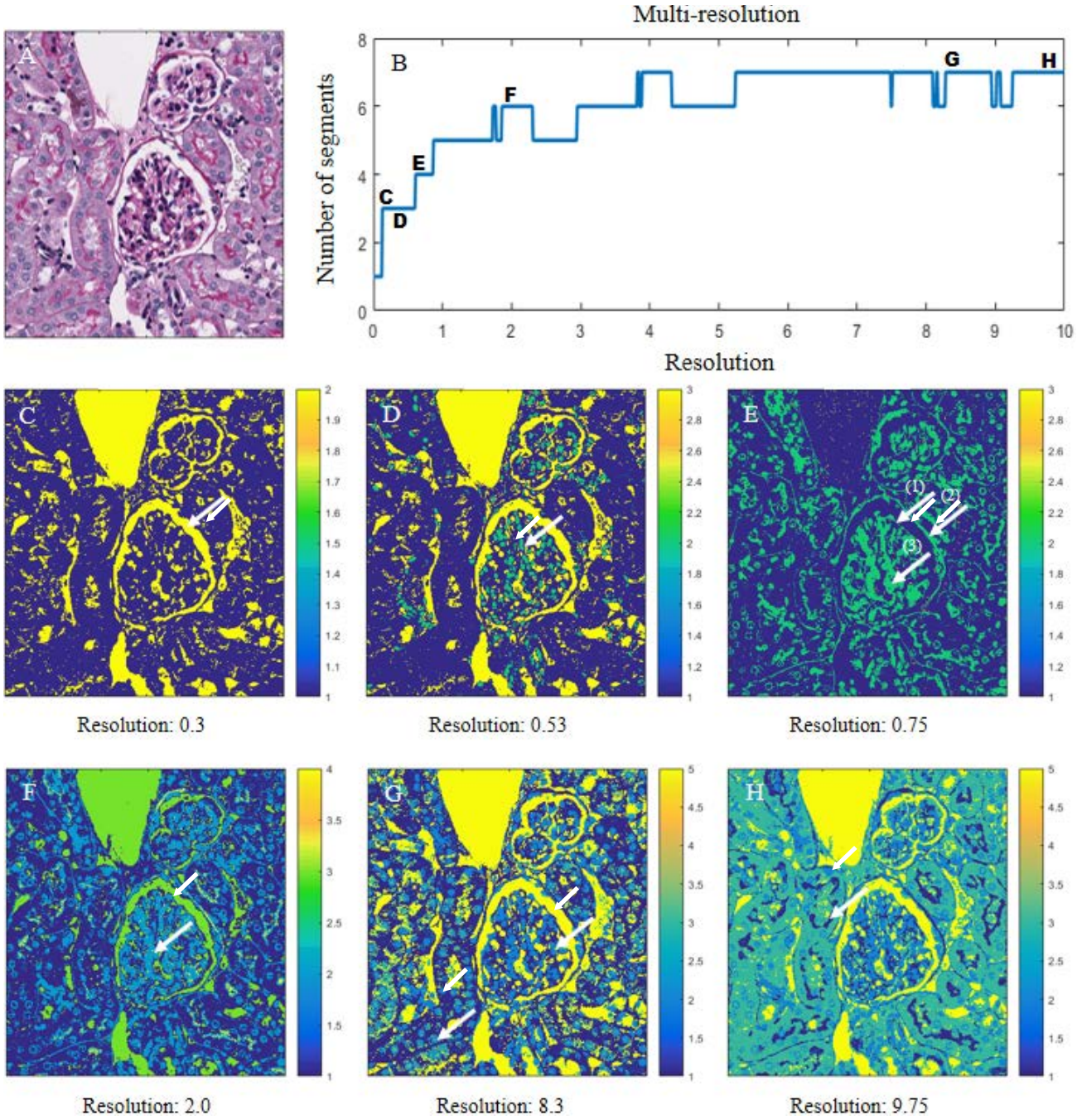

Figure 2. Segmentation of structures with a varying scale parameter. The scale parameter is tuned to reveal pertinent structures/scales in the image. Processing time is 5-15sec for a given scale. (A) RGB histopathology image of a renal glomeruli and extended tubules, containing ~10^6 pixels. (B) Number of segments estimated by varying the scale parameter between 0 and 10. (C-H) The segmented structures with increasing scale. Higher scale reveal increasing number of structures. Relevant structures are marked with white arrows. Bowman's capsule and space, intra- and extra-glomerular nuclei, mesangial matrix, basement membrane, and inner border of the tubule are revealed at varying scale. NOTE: Segment color is used to delineate the compartments. It is not consistent across scales.

4. CONCLUSION

We have developed a robust method for extracting image features from histopathological images with computational efficiency unparalleled by similar unsupervised methods. Our computational method is a step towards simultaneous and rapid quantification of pertinent structures in multi-scale for computational renal pathology. The strength of our method is its lack of parameters, segmentations achieved by our method can be easily selected and classified by experts simplifying quantification tasks in renal pathology, with potential to be a basis of powerful tools for diagnostic aid. The potential of unsupervised methods for segmentation tasks in pathology, and all of medicine remains relatively unexplored, and with proper development, has potential to greatly improve the quality of care through more concrete (quantified) evaluation of disease states.

Algorithmically, we have exponentially improved the performance of Potts Model based image segmentation. Our statistical assumptions reduce the complexity of the graphical image representation, allowing complex segmentation tasks which were previously unobtainable in a reasonable timescale. We hope this discovery will not only advance the use Potts Model based segmentation, but apply to other similar unsupervised methods. To better optimize our method, the effects of scale on segmentation output needs to be studied, and a method to choose applicable scales for obtaining biologically relevant structures developed. We believe that multi-scale segmentation is what sets Potts Model segmentation apart, and a concrete understanding of the scale parameter therefore needs to be established. Further work needs to be done to validate the accuracy of our method against established approaches such as spectral clustering[13]. We hope to apply our method for data segmentation in fields such as oncology, genomics, as well as non-biological problems, and aim to call on experts in their respective fields to evaluate such applications.

ACKNOWLEDGMENTS

This project was supported by the faculty start-up funds from the Pathology & Anatomical Sciences Department, Jacobs School of Medicine and Biomedical Sciences, University at Buffalo. We thank the histology core facility (Pathology & Anatomical Sciences, University at Buffalo) for performing histopathological staining of tissues.

REFERENCES

[1] Yu, Y., Cao, Z. and Feng, J., [Continuous Potts Model Based SAR Image Segmentation by Using Dictionary-Based Mixture Model], Springer International Publishing, Cham (2014).

[2] Hu, D., Sarder, P., Ronhovde, P. *et al.*, "Community detection for fluorescent lifetime microscopy image segmentation," Proc. of SPIE, 8949, 89491K: 1-13 (2014).

[3] Hu, D., Sarder, P., Ronhovde, P. *et al.*, "Automatic segmentation of fluorescence lifetime microscopy images of cells using multiresolution community detection-a first study," Journal of microscopy, 253(1), 54-64 (2014).

[4] Hu, D., Ronhovde, P. and Nussinov, Z., [A Replica Inference Approach to Unsupervised Multi-Scale Image Segmentation], Washington University in St. Louis, St. Louis (2011).

[5] Hu, D., Ronhovde, P. and Nussinov, Z., "Replica inference approach to unsupervised multiscale image segmentation," Physical Review E, 85(1), 016101 (2012).

[6] Peng, B., Zhang, L. and Zhang, D., "A survey of graph theoretical approaches to image segmentation," Pattern Recognition, 46(3), 1020-1038 (2013).

[7] Hunter, R. S., " Photoelectric color-difference meter," JOSA, 38(7), 661 (1948).

[8] Hunter, R. S., "Accuracy, precision, and stability of new photo-electric color-difference meter," JOSA, 38(12), 1094 (1948).

[9] Hopcroft, J. E. and Ullman, J. D., [Introduction to Automata Theory, Languages, and Computation.], Addison Wesley, Reading, MA (1979).

[10] Hastie, T., Tibshirani, R., Friedman, J.H., [The Elements of Statistical Learning: Data Mining, Inference, and Prediction], Springer, 768 (2009).

[11] Wu, F. Y., "The Potts model," Reviews of Modern Physics, 54(1), 235-268 (1982).

[12] Velde, K. V., "Multi-scale color image enhancement," Proceedings 1999 International Conference on Image Processing, 584-587 (1999).

[13] Ng, A., Jordan, M. and Weiss, Y., [On spectral clustering: analysis and an algorithm], MIT Press, Cambridge (2002).