2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Semi-Supervised Segmentation of Renal Pathology: An Alternative to Manual Segmentation and Input to Deep Learning Training

Adrienne Kline[1,2], Hyun Jae Chung[2,3], Waleed Rahmani[2,3], Justin Chun[2,3]

***Abstract*— Kidney biopsy interpretation is the gold standard for the diagnosis and prognosis for kidney disease. Pathognomonic diagnosis hinges on the correct assessment of different structures within a biopsy that is manually visualized and interpreted by a renal pathologist. This laborious undertaking has spurred attempts to automate the process, offloading the consumption of temporal resources. Segmentation of kidney structures, specifically, the glomeruli, tubules, and interstitium, is a precursory step for disease classification problems. Translating renal disease decision making into a deep learning model for diagnostic and prognostic classification also relies on adequate segmentation of structures within the kidney biopsy. This study showcases a semi-automated segmentation technique where the user defines starting points for glomeruli in kidney biopsy images of both healthy normal and diabetic kidney disease stained with Nile Red that are subsequently partitioned into four areas: background, glomeruli, tubules and interstitium. Five of 30 biopsies that were segmented using the semi-automated method were randomly selected and the regions of interest were compared to the manual segmentation of the same images. Dice Similarity Coefficients (DSC) between the methods showed excellent agreement; Healthy (glomeruli: 0.92, tubules: 0.86, intersititium: 0.78) and diabetic nephropathy: (glomeruli: 0.94, tubules: 0.80, intersititium: 0.80). To our knowledge this is the first semi-automated segmentation algorithm performed with human renal biopsies stained with Nile Red. Utility of this methodology includes further image processing within structures across disease states based on biological morphological structures. It can also be used as input into a deep learning network to train semantic segmentation and input into a deep learning algorithm for classification of disease states.**

***Keywords* - segmentation, kidney biopsy, Nile Red, diabetic kidney disease, deep learning**

I. INTRODUCTION

A. Kidney Biopsy with Nile Red Staining

Renal biopsy interpretation remains the gold standard for the diagnosis and staging severity of many kidney diseases [1]. Visual morphological assessment of the renal parenchyma provides useful information for disease categorization by pathologists. Current kidney biopsy tissue preparation for light microscopy, immunohistochemistry and electron microscopy are complex, time consuming and technically challenging. Nile Red is a fluorescent lipophilic dye that can be used as a fast staining alternative for structural investigation [2],[3]. Comparing morphology, composition and distribution of structures in various kidney compartments (glomerulus, tubule, and interstitium) across different disease states using Nile Red may provide mechanistic insight into their prognostic value of evaluating patients with kidney disease. However, such an analysis requires reproducible identification and segmentation of the different kidney compartments, a time-consuming process limited by poor intra- and inter-reader manual assessment of renal biopsy specimens [4],[5]. Figure 1 shows a normal human kidney tissue specimen processed with Nile Red stain. This example shows two glomeruli - denoted with 'G', multiple tubules are also visible, some labeled 'T', and interstitium 'I' which is non-glomeruli or tubule morphology.

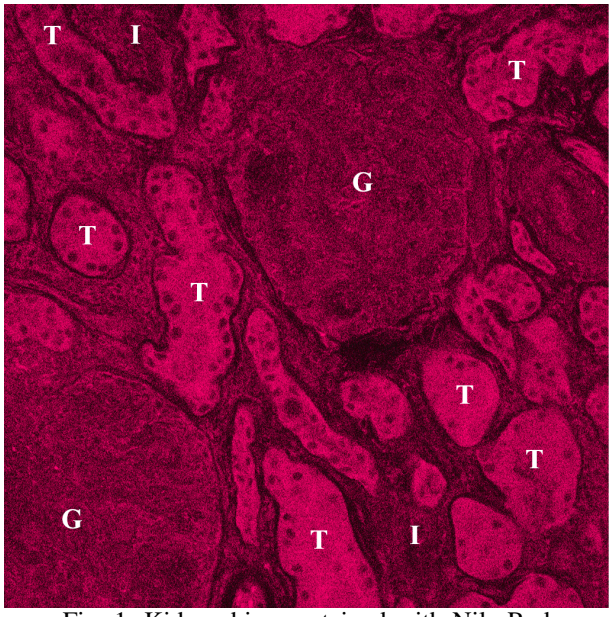

Fig. 1: Kidney biopsy stained with Nile Red

B. Segmentation

Partitioning a digital image has been applied in very diverse fields of study such as area detection in satellite images, traffic control, autonomous cars, medical imaging, face/iris recognition and object detection/classification. To our knowledge, it has not been used, in segmenting Nile Red stained renal biopsy specimens. There are currently several options for segmentation including region-based, edge detection, clustering, mask R-CNN (regional convolution neural network), and fast marching [6]. Within these options several techniques are subsumed. Region-based utilizes objects based on threshold value(s). This method works well when high contrast exists between the foreground and background, and has the advantage of efficient calculations. However, region based is limited when the foreground and

Corresponding author: Adrienne Kline, askline1@gmail.com
[1]Faculty of Biomedical Engineering, University of Calgary, Canada
[2]Department of Medicine, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
[3]Calvin, Phoebe and Joan Snyder Institute for Chronic Diseases, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

background are homogeneous. Edge detection (eg. water shedding) uses discontinuous local change within an image to define boundaries. However it is not suitable when there is an overabundance of edges in the image even after applying filters. Clustering (eg. K-mean) divides the image into a user defined 'n' number of homogeneous areas based on distances and works well with small data sets. Drawbacks to this approach are computation time and that distance algorithms are not suitable for non-convex clusters. While R-CNN models entail high training time require many samples, they are most generalizable therafter and have been and have been used in kidney pathologic segmentation [7],[8]. A fast-marching method (eg. Chan-Vese) finds a segmentation that optimizes an energy functional. A limitation is that it is highly susceptible to starting position. However, it can be used for regional or global segmentation. Our images contain an abundance of edges, poor contrast between distinct morphology, and too few samples on which to train a neural network adequately. Hence, we selected a combination of Chan-Vese and a regional based semi-supervised approach thus benefiting from Chan-Vese generalizability while retaining the benefit of user input. The latter alleviates the issue of starting position and thus allows for multiple glomeruli location(s) and fluctations in image scale (size of glomeruli).

II. Methods

A. Image Acquisition

Images were acquired from human healthy normal nephrectomy samples and patients with diabetic renal pathology (society classification class 3) acquired from the Biobank for the Molecular Classification of Kidney Disease (PMID: 28747168). Usage of human samples were conducted in accordance with guidelines set forth by the Research Ethics Board at the University of Calgary and Alberta Health Services. Samples were fixed with 4% paraformaldehyde for 10 minutes followed by staining with Nile Red for 10 minutes, using a Zeiss LSM 880 confocal microscope with a 20X objective, NA 0.8 equipped with a 24-channel detector (485 nm to 691 nm) for spectroscopy. Specifically channel 691 was used to capture structural information for purposes of this study. Diabetic disease states as well as healthy normal controls where acquired and included in the image set.

B. Validation Protocol

To establish the ground truth, a nephrologist manually segmented regions of interest (ROIs) for a sub-sample (n=5) from the acquired set of 30 images used in development. The sample of images for manual segmentation was selected using a random non-repeating number generator $\{n \in 30 | n = 5\}$. This was done for both normal and diabetic samples. A freehand segmentation in Matlab was used to define borders and generate masks of glomeruli and tubules with the remaining non-background being allocated at interstitial morphology. These manually-generated ROIs were used to verify the results of the same ROIs using the semi-supervised segmentation procedure. The Sørensen–Dice coefficient colloquially known as dice similarity coefficient (DSC) was used to calculate the degree of overlap between the ground truth mask segmentation (X) and the semi-supervised segmentation (Y), as shown in Eq. 1. A DSC value of greater than 0.70 is considered excellent [9]. DSC values were calculated for the structures of glomerli, tubules, interstitium and overall.

$$DSC = \frac{2 * |X \cap Y|}{|X| + |Y|} \tag{1}$$

C. Image Pre-processing

Images were in '.tif' format and were processed in Matlab 2019a. Kidney biopsy images were loaded into the Matlab environment as Red Green Blue (RGB). Prior to the algorithm running, the user completed three tasks: 1) selecting the center of all glomeruli in the image by double clicking, 2) drawing a ROI around a portion of a tubule (inclusive), and 3) defining the size of the glomeruli relative to the image in decimal format to 3 decimal places. RGB images were subsequently coverted to greyscale using a luminosity conversion. A 2-D Guassian filter (Eq. 2) was applied to the images with with an adaptive kernel based on the size of the glomeruli relative to the image size. Lastly, background was isolated using a threshold based on the Guassian processed images.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+Y^2}{2\sigma^2}} \tag{2}$$

D. Segmentation of Glomeruli

Both the location and size of glomeruli established in the preprocessing phase provide a preliminary starting mask of the glomeruli from which a Chan-Vese optimization algorithm (Eq. 3) finds a local minimum of entropy over a set number of iterations, where $H(\phi)$ is the Heaviside equation, $u_0(x, y)$ is the input image, and c_1, c_2, ϕ are updated recursively. The number of times the user selects center(s) of glomeruli also serves as a counter for the number of times the algorithm will run a glomerulus segmentation centered at that point. Following this, a series of image morphological processing occurs including closing, dilation and area filtering steps. This is then followed by background subtraction.

$$\begin{aligned} F(c_1, c_2, \phi) = &\int_\Omega (u_0(x, y) - c_1)^2 H(\phi)\, dxdy \\ &+ \int_\Omega (u_0(x, y) - c_1)^2 (1 - H(\phi))\, dxdy \\ &+ v \int_\Omega |\nabla H(\phi)| \end{aligned} \tag{3}$$

E. Segmentation of Tubules and Interstitium

After the glomeruli are segmented, the mask of their inverse can be used to further isolate the tubules and intersititium. Tubules are selected for utilizing the mask of the user-defined tubule sub-sample from the RGB image. This mask is imposed on the Guassian filtered grayscale image, to generate an average threshold of the 'tubule' class. To ensure inclusiveness, any value above 1.5 standard deviations

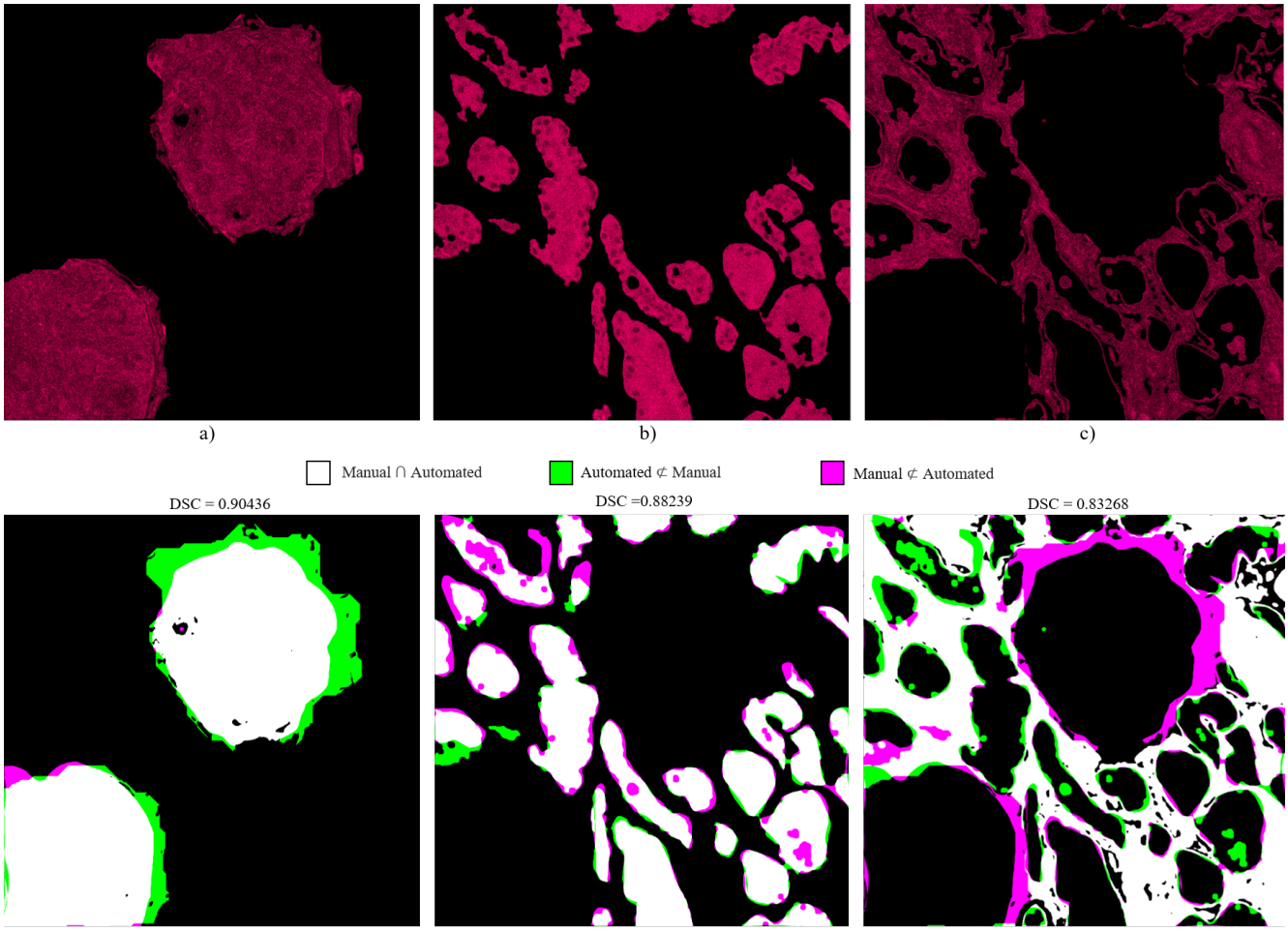

Fig. 2: Semi-Automated Segmentation of RGB Nile Red image (a-c), Semi-Automated-Manual Agreement (d-f)

below the average value within the segment was classified as a 'tubule'. After the glomeruli, tubules, and background were established, the inverse of the superposition of these segment masks were the only structure left - intersititum. An overview of the algorithm is shown in Algorithm 1.

Algorithm 1: Segmentation Code Overview

Result: Glomeruli, Tubulese, Interstitium
gnum, gsize, $\text{tubule}_{thresh} \leftarrow$ User defined
for $j = 1 : numberofimages$ **do**
  $j \leftarrow$ `rgb2gray` $(j), 0.2989R + 0.587G + 0.114B$
  $j \leftarrow$ `guassfilt` $(j), G(x,y) \otimes j(x,y)$
  for $i = 1 : len(gnum)$ **do**
    glom-mask$\leftarrow$(glom-mask=[])+`Chan-Vese` $(j(i))$
  end for
  j $\leftarrow$ j - glom-mask
  if $j \geq tubule_{thresh} - 0.5 * tubulesample_{\sigma}$ **then**
    tubule-mask $\leftarrow j$
  else
    interstitium-mask $\leftarrow j$
  end if
end for

III. Results

A sample of one the images (Fig. 1) that was run through the pipeline described is shown in Figure 2. Fig. 2a) are glomeruli, 2b) are tubules and 2c) is the interstitium. Figure 2d)-2f) shows the results of DSC calculations comparing the agreement between the semi-supervised segmentation and full manual segmentation of the three morphological structures. The white portion of the mask denotes areas of overlap between the two modalities, $X \cap Y$; green where $X \nsubseteq Y$ and magenta where $Y \nsubseteq X$. Figure 2 d), e) and f) denote the glomeruli, tubules and interstitium, respectively. DSC values for the 5 randomly selected samples of disease and healthy states are reported in Table 1. An independent samples t-test showed that there was no difference in the DSC values between healthy control and diabetic kidney images t(28)=0.32, p=0.75. The two groups included 5 images with the 3 compartments (glomeruli, tubule, and interstitium) making (n=15) healthy control samples and (n=15) diabetic nephropathy instances.

IV. Discussion

We showcase a semi-automated segmentation tool that allows for rapid and reproducible segmentation of glomeruli, tubules and interstitium in Nile Red stained human kidney

TABLE I: Dice Similarity Coefficient

Structure	Control DCS ± SE	Diabetes DCS ± SE
Glomeruli	0.92 ± 0.01	0.94 ± 0.01
Tubules	0.86± 0.01	0.80 ± 0.02
Interstitium	0.78± 0.02	0.80 ± 0.03
Overall	0.85	0.85

biopsy specimens. Nile Red is a fast and inexpensive tool to investigate structural differences making it ideally suited for investigating any changes in glomeruli, tubules and interstitium of healthy versus disease states. Although it was possible to stain compartments for specific markers we selected a more difficult problem so the process would more readily generalize. Continued research into this important area relies on the correct segmentation and subsequently classification of different morphological renal structures/compartments that are expected to vary with disease state such as diabetes.

The results of the DSC analysis showed 85% agreement between automated and manual segmentation techniques for both healthy and diabetic images. No discernible difference in the agreement rate was found when comparing healthy control and diabetic disease state images. These results demonstrate that this is very promising alternative method to the full manual segmentation approach. While we obtained excellent agreement, there were a couple of notable areas of discrepancy. First, in the glomerulus, segmentations were predominantly located around the glomerular tuft. This is an area that is sometimes included in the glomerulus but other times not [10]. With this knowledge, however, the user could change the starting size of the semi-supervised segmentation so as reduce the expansion of the Chan-Vese algorithm over the curve set. Additionally, depending on the amount of transection of tubules, a doughnut or 'blob' shape may emerge if the transection is in yz plane versus xy or xz. Such lumens were excluded by the program instead of attributing the area to the intersititum. Limitations with this method include that segmentation of the tubules is based on applying a threshold and therefore is tailored to this particular type of image capture. Chan-Vese segmentation is also prone to errors in starting position so it is crucial the user defines the centers for the glomeruli at the outset. Additionally this method was not compared with other open source options ie. ilastik.

Despite the shortcomings, this study is a prime example of using a human-in-the-loop automation. All deep learning based segmentations require training, this method works by offloading the segmentation work of pathologists without losing their expertise in structure identification [11]. Secondly, should users suffer from too small a sample size for deep learning this represents an alternative to investigating morphological differences in kidney disease states. The rapid rise of image analysis as part of the artificial intelligence suite of methodologies has barely been exploited in medicine [12]. However, "incorporation of new tools into traditional histopathological evaluation of renal tissue is needed to improve diagnostic precision and predictive value of renal biopsy in kidney disease" [13]. It also offers an alternative to direct deep learning segmentation, which is both computationally and temporally expensive to train, greatly limiting its potential for use. Deep learning is an approach that can be trained to recognize patterns in images through a network of artificial neurons. The outcomes of the processes described in the current study can be used as input into a deep learning classifier, substantially reducing the training burden. The script and supplementary notes can be found at `https://github.com/adriennekline/renal-pathology-segmentation`

V. CONCLUSION

This showcases a semi-supervised automated segmentation method as applied to renal biopsies using Nile Red. To our knowledge this is the first of its kind that requires minimal user input with robust segmentation results, reducing the temporal burden on human experts. This methodology can be used in other image processing contexts to discern structural differences based on morphological characteristics. In addition, resulting images can be used as input into a deep learning classifier for training segmentation or identifying regions of interest for classification of disease states.

ACKNOWLEDGMENT

Dr. Chun is supported by CIHR and the Kidney Research Scientist Core Education and National Training Program.

REFERENCES

[1] S. Amoueian et al., "Renal Biopsy Interpretation," *in Renal Biopsy and Pathology*, M. Mubarak, Ed., Rijeka, Croatia: InTech, 2012, pp. 45-64

[2] P. Greenspan et. al, "Nile Red: A selective fluorescent stain for intracellular lipid droplets", *J. Cell Biol.*, vol.100, no.3, pp.965–73, 1985

[3] P. Walker, et. al, "Practice guidelines for the renal biopsy", *Mod Pathol.*, vol. 17, no.12, pp.1555-1563, 2004

[4] L. Barisoni, et. al, "Digital pathology imaging as a novel platform for standardization and globalization of quantitative nephropathology", *Clinical Kidney Journal*, vol. 10, no. 2, pp.176–187, 2017

[5] L. Oni, et al, "Inter-observer variability of the histological classification of lupus glomerulonephritis in children". *Lupus*, vol. 26, no. 11, pp.1205-11, 2017

[6] W. Kang, et. al, "The Comparative Research on Image Segmentation Algorithms," *First International Workshop on Education Technology and Computer Science*, Wuhan, Hubei, 2009, pp.703-707

[7] C. Jayapandian, et. al, "Development and evaluation of deep learning–based segmentation of histologic structures in the kidney cortex with multiple histologic stains", *Kidney International*, vol. 99 no. 1, pp.86-101, 2021

[8] M. Hermsen et. al, "Deep Learning–Based Histopathologic Assessment of Kidney Tissue", *J. Am. Soc. Nephrol.*, vol. 30, no.10, pp.1968–1979, 2019

[9] D. Muruve, "The biobank for the molecular classification of kidney disease: research translation and precision medicine in nephrology", *BMC Nephrol.*, vol. 18 no.1, pp.252, 2017

[10] K. Zou et. al, "Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index", *Acad. Radiol*, vol 11, no. 2, pp.178-189, 2004

[11] SM Sheehan et al. "Automatic glomerular identification and quantification of histological phenotypes using image analysis and machine learning", *Am. J. Physiol. Renal Physiol.* vol 315, pp.1644-1651, 2018

[12] A Holzinger, "Interactive machine learning for health informatics: when do we need the human-in-the-loop?" *Brain Inform.*, vol.3, pp.119-131, 2016

[13] J. Topol et. al, "High performance medicine: the convergence of human and artificial intelligence Nature Medicine", vol.25, pp.44-56, 2019

[14] SM Bagnacio, "Beyond the microscope: interpreting renal biopsy findings in the era of precision medicine". *Am. J. Physiol. Renal Physiol.* vol. 315, pp. 1652-1655, 2018