Analyzing Tubular Tissue in Histopathological Thin Sections

Azadeh Fakhrzadeh*, Ellinor Spörndly-Nees†, Lena Holm†and Cris L. Luengo Hendriks*
*Center for Image Analysis
†Department of Anatomy, Physiology and Biochemistry
Swedish University of Agricultural Sciences, Uppsala, Sweden
{*azadeh.fakhrzadeh, cris*}*@cb.uu.se*, {*Ellinor.Sporndly-Nees, Lena.Holm*}*@slu.se*

***Abstract*—We propose a method for automatic segmentation of tubules in the stained thin sections of various tissue types. Tubules consist of one or more layers of cells surrounding a cavity. The segmented tubules can be used to study the morphology of the tissue. Some research has been done to automatically estimate the density of tubules. To the best of our knowledge, no one has been able to, fully automatically, segment the whole tubule. Usually the border between tubules is subtle and appears broken in a straight-forward segmentation. Here we suggest delineating these borders using the geodesic distance transform. We apply this method on images of Periodic Acid Shiffs (PAS) stained thin sections of testicular tissue, delineating 89% of the tubules correctly.**

I. Introduction

Many organs in vertebrates, such as the kidneys, the prostate and the testis, contain tubular tissue. Such tissue is composed in large part of tubules, which are an epithelium surrounding a lumen, or, in other words, one layer or more of cells surrounding a cavity. Epithelial cells secrete some substance into the lumen, through which this substance is transported. These tissues are frequently studied because of their propensity for cancer, but also, for example, to study other diseases, or damage due to environmental pollutants. Commonly, tubular tissue is fixated, sectioned and stained for observation in a microscope. A tubular structure can be seen in Figure 1. This is a histological image of thin section of testicle. These thin sections show lumens surrounded by epithelial cells, as well as other cells and connective tissue. The lumens appear as the brightest areas in the image, and are therefore often easy to detect automatically. However, distinguishing the different cell types is a much larger challenge, as is determining which cells belong to which tubule. A proper staining is necessary for this task, but no stain can make the task trivial.

Detecting different structures such as nuclei and glands in histopathology image is a recurring theme in the literature [1], [2], [3], [4], [5], [6], [7], [8], [9]. There have been some publications detailing the identification and characterization of tubules. Hafiane et al. [10] used a combination of fuzzy c-means clustering with spatial constraints to delineate glandular structures in prostate cancer histopathology. Basavanhally et al. [11] used domain information to detect the glands. Naik et al. [12] used a Bayesian classifier to detect candidate gland regions by using low-level image features to find the lumen, epithelial cell cytoplasm, and epithelial nuclei of the tissue. In probabilistic methods, however, a larger amounts of training data is needed, to be able to model the prior distribution. Petushi et al. [13] first segment the cell nuclei on the images of histology slides and classifies them into three categories. The spatial positions of cell nuclei are then used to detect higher order tissue structures, such as tubular cross sections and the boundaries of high nuclei density regions.

In all of these methods, the high pixel intensity regions are first detected, and then the cells around each area are used to decide if it is a lumen of a tubule or not. None of these methods, however, try to detect the outer border of the tubules. This means that, with these methods, we are limited in the type of morphological measurement that can be performed. For example, to measure size of tubules in the tissue or the height of the epithelium we need to delineate the outer border of the tubule. By measuring the height of the epithelium in testicular tissue an indirect value of possible sperm output can be estimated [14], [15], [16]. The height of the epithelium is the distance from the lumen of tubule to its outer border as shown in Figure 1. In this paper, we suggest a method to find the outer border of tubules, and thereby fully segment individual tubule cross-sections. This method does not rely on identifying the individual cells forming the tubule, instead it uses the stain color to detect as much of the interstitial tissue (inter-tubular space) as possible, then infers boundaries for individual tubules using lumens and interstitium as hints.

II. Algorithm

A. Delineating lumens

To delineate the outer border of tubules, we will start by identifying the lumens. The lumen will serve as an anchor point to pick out suitable tubule cross-sections. As it is shown in Figure 1, the brightest areas of the image correspond to lumens. Here, we will use an active contour, which can be seeded inside the lumen and then be stopped at the edges of the lumen, to distinguish lumens from other types of holes in the tissue.

The active contour is an evolving interface whose motion is guided by some internal and external forces. We used the level set method to segment lumens. In this segmentation method, we define a level set function $\phi(x, y, t)$ and the evolving curves are zero iso-contours of this function, i.e.

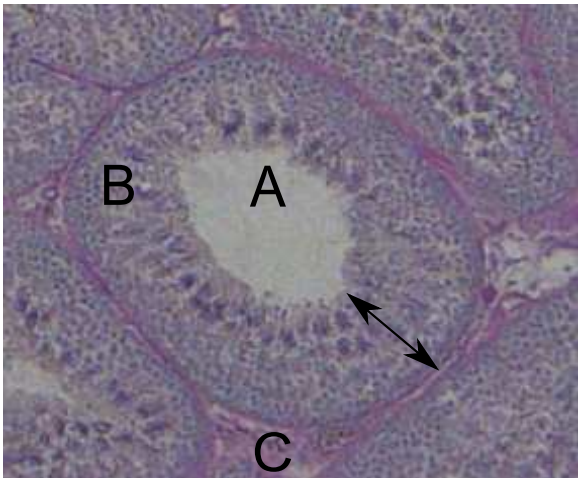


Fig. 1. Histological image of seminiferous tubule in testis; A: lumen B: epithelium C: interstitial tissue. The arrow shows the epithelial height.

$C = \{(x,y)|\phi(x,y,t) = 0\}$. The equation for the motion of the interface is

$$\frac{\partial \phi}{\partial t} + F|\nabla \phi| = 0 \ . \tag{1}$$

where F is a function that models the desired velocity on the interface and ∇ is the gradient operator. The interior and exterior of ϕ in region Ω are defined as $\{(x,y) \in \Omega|\phi(x,y,t) < 0\}$ and $\{(x,y) \in \Omega|\phi(x,y,t) > 0\}$, respectively. The most difficult task is to establish a level set function that can make a meaningful segmentation of the image. Li et al. [17] suggested an energy function for the level set which includes both the edge and region information. If image I belongs to domain Ω, the edge indicator can be defined as

$$g \triangleq \frac{1}{1 + |\nabla G_\sigma * I|^2} \ , \tag{2}$$

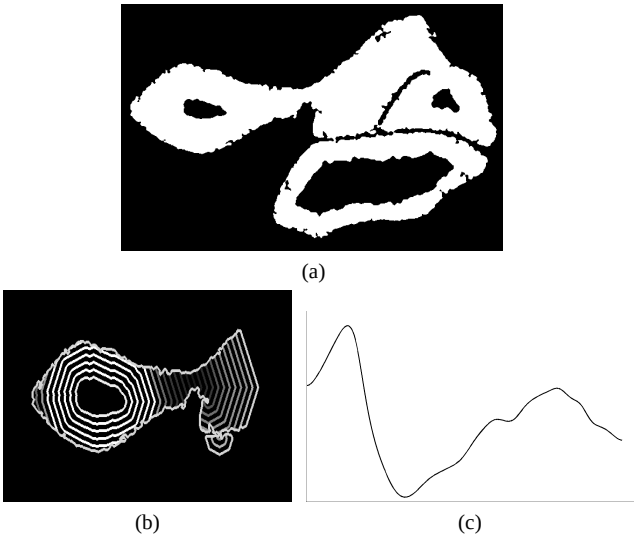
where G_σ is a zero-mean Gaussian distribution with standard deviation of σ. The function g usually has smaller values at object boundaries than at other locations. For a level set function $\phi : \Omega \longrightarrow R$ Li et al. suggested energy functions as

$$\begin{aligned} \varepsilon(\phi) &= \mu \int_\Omega p(|\nabla \phi|)dx + \lambda \int_\Omega g\delta(\phi)|\nabla \phi|dx \\ &+ \alpha \int_\Omega gH(-\phi)dx \ . \end{aligned} \tag{3}$$

The first part of this equation is the level set regularization term, p is the energy density function, μ, λ and α are the energy coefficients, and δ and H are the Dirac delta and the Heaviside functions, respectively. This energy function is minimized when the contours reach the boundary of the object. We used the level set algorithm implemented by Li et al. [17]. For the initialization of the level set, we use k-means clustering on the RGB image. The potential lumens are large high pixel intensity regions. As result of the clustering we have all the high pixel intensity regions as one class. We use morphological filtering to remove small areas, the result is used for initialization of the level set algorithm. The coefficient λ is used to smooth the object boundary. Some of these white areas that we obtained here are not lumen. Most of these regions will be excluded at a later step.

B. Delineating the outer border of tubules

Using a straight-forward segmentation of the interstitium (e.g. using k-means clustering), the boundary between tubules is usually detected only partially, or sometimes not at all. This means that tubules segmented this way are clustered and usually not seen as individual tubules. If one assumes each tubule has a lumen, then one can use the watershed method and let the lumens compete with each other. However, due to the complex three-dimensional structure of the tissue, some tubules will be cut such that the lumen is not visible, making the process of separating tubules more difficult. For example, Figure 2a shows a connected component resulting from the segmentation, which consist of four tubules, one of which does not show a lumen. Here we suggest a method which discards the tubule that does not have a lumen and separates the correct tubules. In short, the method looks for the narrowing in between tubules that hints at where the connected component should be cut. If we have a propagating curve,



(a) (b) (c)

Fig. 2. (a): One connected component created by the initial segmentation; (b): Evolving curve from lumen; (c): Plot of length of curve at every distance.

starting from the border of the lumen (Figure 2b), and measure the length of this curve at every step (Figure 2c), we will find a strongly decreased length at the narrowest point in between two tubules. The location of this minimum is where the connected component should be cut.

We compute the evolving curve inside a connected component using the geodesic distance transform [18], taking the lumen as the seed. The geodesic distance transform computes the distance between each pixel in the object and the nearest seed point. The distance is computed along the shortest path within the object. The distance along a path is given by the distance between consecutive pixels on that path; we use a

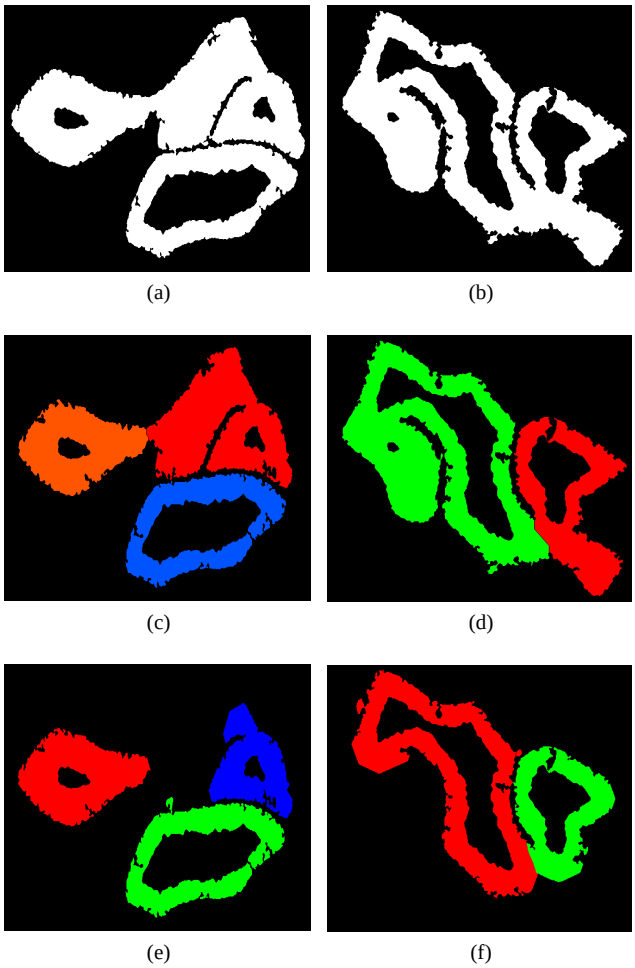


Fig. 3. (a) and (b) are two examples of connected components, (c) and (d) are results of watershed method, (e) and (f) are results of proposed method.

5x5 neighborhood with optimal weights according to [19]. Computing the geodesic distance transform is a minimum-cost path problem, efficiently solved using Dijkstra's algorithm [20]. Points which are at the same distance from the lumen form an evolving curve; the length of this curve at every time step is easily obtained by the histogram of the distance transform. Figure 2b shows this evolving curve for one of the lumens inside the object. The length of the curve increases until it reaches the border of object, then it starts to decrease, and as soon as the curve enters another tubule, the length of curve increases again. We will thus have a clear minimum where the tubules connect to each other. For each of the three lumens in the connected component in Figure 2a, we find the border of the corresponding tubule using this method.

Figure 3, compares our suggested method with watershed for two different examples. In the first example (Figure 3a) the watershed method finds as many tubules as lumens (three), leaving the fourth tubule attached to one of the others. Our method instead discards the tubule that does not show a lumen (Figuer 3e). It is also evident from the second example (Figure 3b) that our method not only is capable of separating

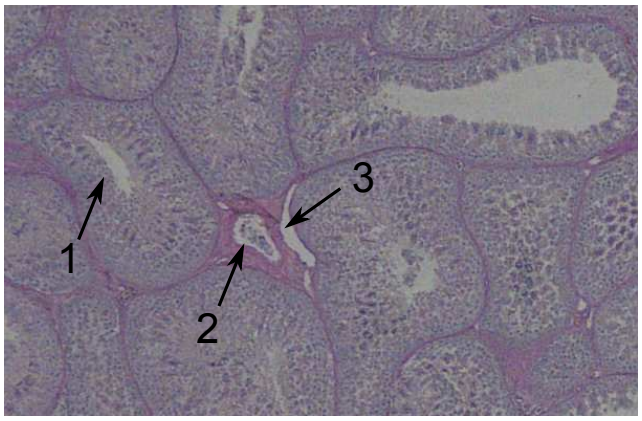


Fig. 4. (1) Correct lumen, (2) and (3) are false detected lumens.

connected tubules as the watershed method does, but also can nicely discard the redundant part of the tubules (Figure 3f). The separation of tubules might not be still optimal in some cases but we can get a reasonable and trustworthy result for the majority of the cases. Another advantage of our method is discarding the false positive lumens, which will be discussed in the next section.

C. Removing False Positive Lumens

As mentioned before, we first find the bright area and then use them as the initialization of the evolving curve. There might be some false positives when detecting lumens. This would be large areas of high intensity, such as tears in the tissue. These bright areas have the same features as lumens, and it is a challenging task to distinguish them from the correct lumens. The result of the tubule segmentation is strongly affected by these wrongly detected lumens. Figure 4 shows two types of false detected lumens. We know that the correct lumens are located in the middle of the tubule and is surrounded with epithelial cells. The bright areas are used as seeds for the geodesic distance transform as described in section II-B. The geodesic distance curve, can not evolve from the seeds which are completely outside the objects (tubules). So the evolving curve for these wrongly detected lumens will not have any minimum and we can discard them. Bright areas outside of tubules that are connected to a tubule (number 3 in Figure 4), can be distinguished by density function of the distance between border of the lumens and border of tubules. All bight areas are assigned to its closest tubule. A false positive like number 3 partly shares the same border with its corresponding tubule and the mode of the density of the distance between their borders becomes zero. For the correct lumen like number 2, the mode is not zero.

III. Results

In this section, we apply the suggested algorithm to segment tubular structure in testicular tissue. Segmented tubules can be used later for studying the morphology of tissue. Exposure to substances like xeno-estrogens during embryonic development

or early in life may affect the morphology and function of testicular tissue. One suggested quantity to study the morphology of tissue is the height of the epithelium. Epithelial height as shown in Figure 1 is the distance from the border of the testicular tubule to the border of its lumen. We use our algorithm to delineate the borders of the tubule and the lumen.

We obtained white field microscope photographs from testicular tissue using the following procedure. Healthy, sexually mature roosters were euthanized and transverse tissue slices (about 2 mm thick) were cut from testis. The tissue was fixed in modified Davidsson fluid for 24 hours at 4 °C then dehydrated and embedded in paraffin wax. The samples were cut in 5 μm sections and stained with Periodic Acid Shiffs (PAS). The sections were photographed using a x10 objective lens.

A sample image is shown in Figure 5a. Using k-means clustering we find all bright areas (Figure 5b). The result of k-means clustering is used for initialization of the level set. As an output of level set, we get a nice and smooth segmented lumen (Figure 5c), which is used later for initialization of distance transform. The epithelial tissue of the tubules, and the interstitial tissue in between the tubules, have a nearly-identical response to the standard staining process, and their color values are mostly the same.For these kind of images, in which the color of different objects are similar, usually color spaces like $L^*a^*b^*$ give a better contrast than RGB color space. This is because $L^*a^*b^*$ separates the brightness from the color components. The $L^*a^*b^*$ color space consists of a brightness channel L^*, a color channel a^* indicating where the color falls along the red-green axis, and a color channel b^* indicating where the color falls along the blue-yellow axis. Figure 5d shows a^* channel of the image. As it is seen in this figure, the interstitial tissue outside the tubules, is quite distinct in channel a^*. To reduce the noise we convolve the image with a Gaussian kernel with variance of 1. By choosing a proper threshold, we can find most of the borders of the tubules. The result of thresholding is shown Figure 5e. We are only interested in the tubules with the lumens. Also, often groups of tubules are connected to each other and labeled as one object. We use the distance transform, as described in Section II-B, to separate tubules and to remove the regions without a lumen. Figure 5f shows that many of the tubule cross-sections with awkward shapes, caused by sectioning these highly-bent tubules, have been cut or removed by the algorithm, leaving only tubule cross-sections suitable for measurement. We remove the false positive lumens based on the features we described in Section II-C. The final segmentation result is shown in Figure 6.

We have applied this method to 36 images, delineating a total of 1855 tubules. The images are taken from four different healthy mature roosters (9 images per animal). A tubule like the one shown in figure 1, where a lumen is surrounded by epithelial cells, is recognized by biologists as the correct tubule. Tubules cut in such a way that the lumen is not visible, or the lumen is cut more than once, are considered incorrectly presented, as they are unsuitable for measurement.

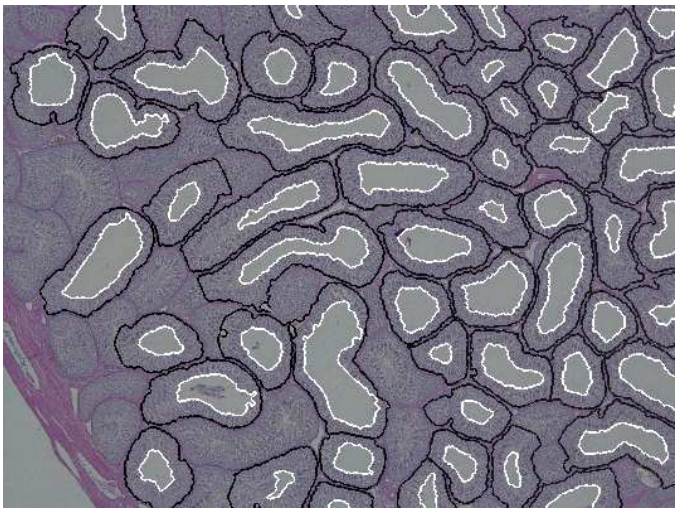


Fig. 6. Border of the tubules in black and borders of the lumens in white, obtained based on the suggested algorithm.

89% of regions segmented by the algorithm are correctly presented and are delineated satisfactorily. About 72% of incorrect delineations are caused by errors in the detection of lumen. Although, we remove some of the wrongly detected lumens, in some of the images where the tissue is severely damaged we still have false positive lumens. When the border of the segmented regions is very jagged, we sometimes obtain strong minima in the evolving curve length before the curve reaches the border. As a result, the delineation will be wrong as well. 28% of incorrect delineations are due to this effect. In spite of this, many tubules with jagged borders are delineated correctly, as can be seen in the top right corner of the image in Figure 5f, many tubules with jagged borders are delineated correctly.

IV. Conclusion

We proposed a method based on the geodesic distance transform for separating clustered tubules from each other after an initial segmentation. The method uses the lumens as an initialization for the distance transform. In general, lumens are easy to detect, as they are big bright areas in the image. However, tears and some similar issues can have the same appearance, which make it hard to distinguish them from the correct lumens. We applied the level set method for segmenting lumens. Two features based on location of the bright areas was used to remove the false positive lumens. However, for the images where tissue is severely damaged, we need to find better features. 72% of incorrect delineations of tubules are caused by areas incorrectly marked as lumen.

Once tubules are correctly delineated, a whole host of new measures becomes available. We have used the method presented here to obtain a measure of the epithelial height in testicular tissue. Measures related to the shape, size, cell count, etc. would be trivial to implement. Our method should also be applicable for analyzing tissues with similar structures like kidney and prostate.

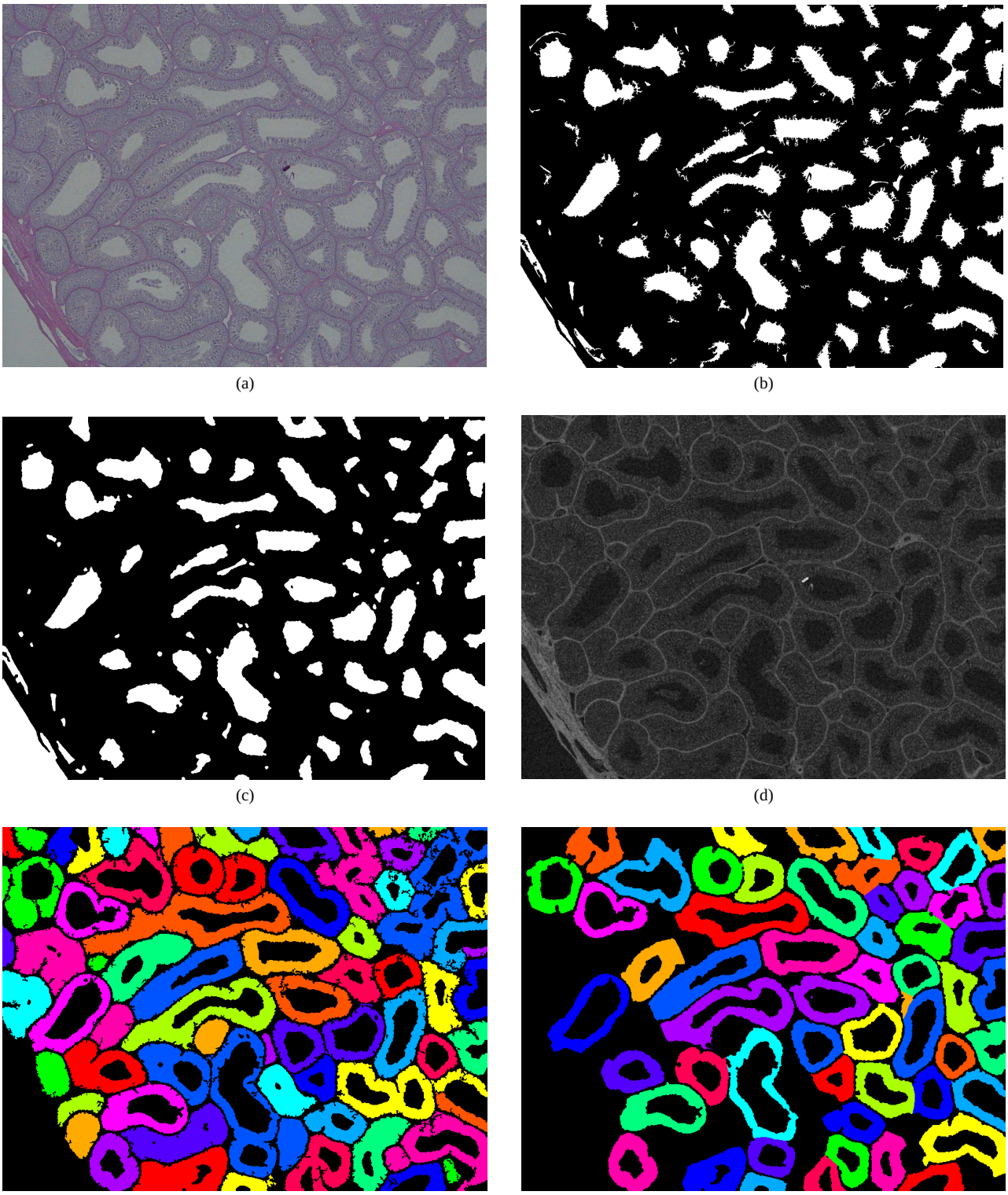


Fig. 5. (a): Thin section of testis. (b): Segmented bright area. (c): Potential lumens as an output of level set method. (d): Channel a^* of the image. (e): Result of threshoding (f): Result of the suggested method.

REFERENCES

[1] T. Mouroutis, S. J. Roberts, and A. A. Bharath, "Robust cell nuclei segmentation using statistical modelling," *Bioimaging*, vol. 6, pp. 79–91, 1998.

[2] K. M. Lee and W. N. Street, "An adaptive resource-allocating network for automated detection, segmentation, and classification of breast cancer nuclei topic area: Image processing and recognition," *IEEE Transactions on Neural Networks*, vol. 14, pp. 680–687, 2003.

[3] P. Bamford and B. Lovell, "Unsupervised cell nucleus segmentation with active contours," *Signal Processing*, vol. 71, pp. 203–213, 1998.

[4] D. L. Weaver, D. N. Krag, E. A. Manna, S. P. H. T. Ashikaga, and K. D. Bauer, "Comparison of pathologist-detected and automated computer-assisted image analysis detected sentinel lymph node micrometastases in breast cancer," *Modern Pathology*, vol. 16, p. 115963, 2003.

[5] D. Glotsos, P. Spyridonos, P. Petalas, D. Cavouras, P. Ravazoula, P. A. Dadioti, I. Lekka, and G. Nikiforidis, "Computer-based malignancy grading of astrocytomas employing a support vector machine classier, the who grading system and the regular hematoxylin-eosin diagnostic staining procedure," *Analytical and Quantitative Cytology and Histology Anal Quant Cytol Histol*, vol. 26, p. 7783, 2004.

[6] C. Wahlby, I. M. Sintorn, F. Erlandsson, G. Borgefors, and E. Bengtsson, "Combining intensity, edge and shape information for 2d and 3d segmentation of cell nuclei in tissue sections," *Journal of Microscopy*, vol. 215, p. 6776, 2004.

[7] V. R. Korde, G. T. Bonnema, W. Xu, C. Krishnamurthy, J. RangerMoore, K. Saboda, L. D. Slayton, S. J. Salasche, J. A. Warneke, D. S. Alberts, and J. K. Barton, "Using optical coherence tomography to evaluate skin sun damage and precancer," *Lasers in Surgery and Medicine*, vol. 39, p. 68795, 2007.

[8] M. Gurcan, T. Pan, H. Shimada, and J. H. Saltz, "Image analysis for neuroblastoma classication: Hysteresis thresholding for cell segmentation," *IEEE Engineering in Medicine & Biology Society*, vol. 9, p. 48447, 2006.

[9] L. Y. adn P. Meer and D. J. Foran, "Unsupervised segmentation based on robust estimation and color active contour models," *IEEE Transactions on Information Technology in Biomedicine*, vol. 1, pp. 475–89, 2005.

[10] A. Hafiane, F. Bunyak, and K. Palaniappan, "Fuzzy clustering and active contours for histopathology image segmentation and nuclei detection," *Lect Notes Comput Sci*, vol. 5259, pp. 903–914, 2008.

[11] A. Basavanhally, E. Yu, J. Xu, S. Ganesan, M. Feldman, J. E. Tomaszewski, and A. Madabhushi, "Incorporating domain knowledge for tubule detection in breast histopathology using O'Callaghan neighborhoods," *SPIE Medical Imaging*, vol. 7963, no. 1, p. 796310, 2011.

[12] S. Naik, S. Doyle, M. Feldman, J. Tomaszewski, and Madabhushi, "Gland segmentation and computerized gleason grading of prostate histology by integrating low-high-level and domain specific information," *Proceedings of MIAAB*, 2007.

[13] S. Petushi, F. U. Garcia, M. M. Haber, C. Katsinis, and A. Tozeren, "Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer," *BMC Med Imaging*, vol. 6, 2006.

[14] A. Hermansson, "Effects on the reproductive system in domestic fowl (gallus domesticus) after embryonic exposure to estrogenic substances." Ph.D. dissertation, Swedish University of Agricultural Sciences, 2007.

[15] M. Orazizadeh, L. S. Khorsandi, and M. Hashemitabar, "Toxic effects of dexamethasone on mouse testicular germ cells," *Andrologia*, vol. 42, pp. 247–253, 2010.

[16] M. J. Golalipour, R. Azarhoush, S. Ghafari, A. M. Gharravi, S. A. Fazeli, and A. Davarian, "Formaldehyde exposure induces histopathological and morphometric changes in the rat testis," *Folia Morphology (Warsz)*, vol. 66, no. 3, pp. 167–171, 2007.

[17] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *IEEE Trans. Image. Process.*, vol. 19, no. 12, pp. 3243–3254, 2010.

[18] P. Soille, *Morphological Image Analysis: Principles and Applications*, 2nd ed. Berlin: Springer, 2003.

[19] B. J. H. Verwer, "Local distances for distance transformations in two and three dimensions," *Pattern Recognition Letters*, vol. 12, pp. 671–682, 1991.

[20] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematlk*, vol. 1, pp. 269 – 271, 1959.