Semantic Segmentation of Microscopic Images Using a Morphological Hierarchy

Cristian Smochina[1], Vasile Manta[1], and Walter Kropatsch[2]

[1] Faculty of Automatic Control and Computer Engineering,
"Gheorghe Asachi" Technical University of Iasi, Romania
smochina.cristian@yahoo.com, vmanta@cs.tuiasi.ro
[2] Pattern Recognition and Image Processing Group, Institute of Computer Graphics and
Algorithms, Vienna University of Technology, Austria
krw@prip.tuwien.ac.at

Abstract. The objective of semantic segmentation in microscopic images is to extract the cellular, nuclear or tissue components. This problem is challenging due to the large variations of these components features (size, shape, orientation or texture). In this paper we present an automatic technique to robustly identify the epithelial nuclei (crypt) against interstitial nuclei in microscopic images taken from colon tissues. The relationship between the histological structures (epithelial layer, lumen and stroma) and the ring like shape of the crypt are considered. The crypt inner boundary is detected using a closing morphological hierarchy and its associated binary hierarchy. The outer border is determined by the epithelial nuclei, overlapped by the maximal isoline of the inner boundary. The evaluation of the proposed method is made by computing the percentage of the mis-segmented nuclei against epithelial nuclei per crypt.

Keywords: Crypt segmentation, Morphological hierarchy, Biomedical imaging, Pathology, Microscopy.

1 Introduction

In diagnostic pathology, the pathologists give a diagnostic after a set of biological samples (tissues stained with different markers) are viewed and many specific features of the objects of interest (size, shape, colour or texture) have been analysed. This complex diagnostic process is an important part in clinical medicine but also in biomedical research and can be enhanced by providing the pathologists or the biologists with quantitative data extracted from the images. The image processing techniques are of special interest because they allow large scale statistical evaluation in addition to classical eye screening evaluation and are used in both sections of the pathology: cytology (the study of cells) and histology (anatomical study of the microscopic structure of tissues) [1]. The microscopic image segmentation is considered a hard task due to the following problems also pointed out in [2]:

− Low contrast and weak boundaries on out-of-focus nuclei. Also some nuclei structures can appear as artefacts in the non-uniformly stained slices.

− Different grey values for the background cased by the non-uniform illumination.
− The physical structure of the cells and the way of sectioning determine a non-uniform distribution of material inside the nucleus (lower intensities within nuclei).
− Considerable variation of object features like shape and/or size and/or orientation and different nuclei distribution within the epithelial layer.

This paper is organized as follows. The last part of this section points out the goal of this study. The inner boundaries of the crypts are detected using the morphological hierarchy and the lumen reconstruction (section 2), while in section 3 the outer borders are detected using the maximal isoline. The results are evaluated in section 4 and discussed and concluded in section 5.

1.1 State of the Art

Many studies from the literature cover fields like microscopy, biomedical engineering and imaging, bioinformatics or pattern recognition and introduce techniques for solving the mentioned problems [2]. Beside the nuclei segmentation attempts [1], [2], also the segmentation of histological structures like gland or crypt is addressed. In [3] a threshold is used to identify the gland seeds which are grown to obtain the nuclei chain. In [4] the pixel labelling to different classes is performed using a clustering approach based on the textural properties.

An object-graphs approach is described in [5] where the relationship between the primitive objects (nucleus and lumen) is considered. The prostate cancer malignancy is automatically graded (Gleason system) in [6] after the prostate glands are detected.

1.2 Aim of the Study

The basic functional unit of the small intestine is the crypt (crypt of Lieberkühn) [7] and it comprises two main structures of interest: the lumen and the epithelial layer (Fig. 1a). The epithelial layer contains epithelial nuclei and surrounds the lumen which is an empty area. The interstitial cells on the other side form heterogeneous regions (stroma) placed between crypts. The stroma areas (Fig. 1a) contain isolated cells with non-regular shape and without particular patterns of arrangement.

This work provides specific techniques to segment the crypts from fluorescence images of colorectal cancer tissue sections. We used 8 bit greyscale images (Fig. 1a) containing nuclei labelled with DAPI, a fluorescent stain that binds strongly to DNA [8] and acquired using a TissueFAXS slide scanner (TissueGnostics GmbH, Austria).

The main motivation for segmenting crypts is to provide the pathologists with quantitative data regarding the areas covered by epithelial nuclei. One alternative approach is to segment each nucleus and to analyze the structures that they form. Since this approach can encounter additional problems, our objective is to directly find the boundaries of the crypts, i.e. to delineate the area covered by the epithelial nuclei without dealing with the individual nuclei (Fig. 2). In [5] also the high level information are preferred against the *local* one but their approach uses different images type (hematoxylin-and-eosin stained images) which provide more biological details (also cell cytoplasm is available) than our DAPI stained nuclei images.
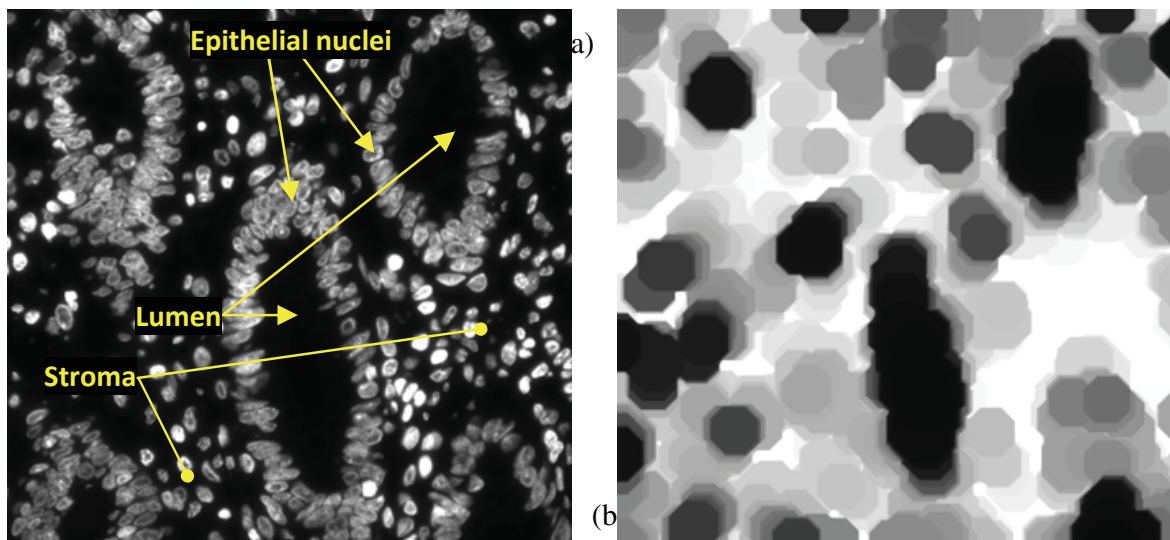
Fig. 1. a) Fluorescence image with crypts from a colon tissue section. b) The image from the top level of the morphological hierarchy. The black regions indicate the lumen.
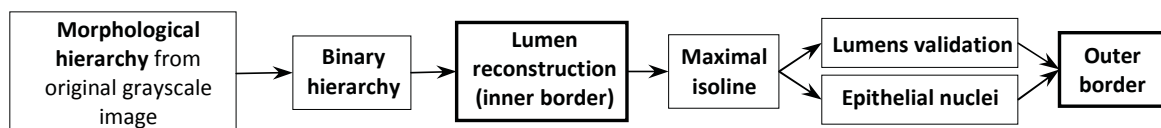

Fig. 2. Overview scheme of the proposed technique

2   Lumens Segmentation

Without considering the relations between the crypt components (epithelial layer and lumen), the stroma and the background, the low level cues will not be able to separate the regions having a particular meaning [9]. A way must be found to keep only the important information and to remove the unnecessary details. In order to detect these regions, the role of *local* information (pixel grey values or gradient) is very important but not sufficient; also *global* information like the region's size and relation with the other region types must be included [9].

A rough assumption about the nuclei distribution over different region types can be made. The lumen does not contain nuclei and appears like a big round black area surrounded by a 'ring' with variable thickness. This ring contains a high density of touched/chained epithelial cells. The exceptional cases appear when the lumen gets to be in touch with the stroma area due to missing cells that 'break' this ring.

By applying the morphological closing operation [10] on the grey image, the nuclei closer than the size of the structure element (SE) will be connected. The epithelial nuclei and those from the stroma area can be connected by relating the SE's size to the size of the lumen region (the connection should not pass over the lumen). For that we build the hierarchical image decomposition similar to the morphological pyramid [11]: the upper levels are obtaining by applying a morphological operator on the base image. The difference consists in lack of sub-sampling step [12].

2.1   Building the Morphological Hierarchy

Let I denote the input grey scale image and $\bullet$ denote the morphological closing operation. The SE ψ_k is a two-dimensional disk of diameter $2k+1$. The hierarchical morphological representation $\Pi_\bullet$ consists of L levels. The first one is the original grey scale image $\Pi_\bullet^1 = I$ and each level $\ell > 0$ is given by $\Pi_\bullet^\ell = \Pi_\bullet^1 \bullet \psi_{2\ell}, \ell = \overline{1,L}$.

   The closing operation smoothes the objects' boundary and removes the dark holes smaller than the SE. Since the size of the SE increases according to the level of the hierarchy, in the lower levels only the small gaps will be filled, while the bigger ones will be closed in the upper levels. To prevent the SE from growing too large, the maximum number of levels is established by limiting the SE's size so that it covers maximally 2-3 nuclei (in our experiments the SE ψ_{50} gives 25 levels). The lumens should 'survive' till the top level (Fig. 1b) and should be easier highlighted; also the gaps from the crypts (do not exceed the size of 1-2 nuclei) should be filled.

2.2   Lumen Reconstruction

The proper reconstruction of each lumen based on the found regions from the top level must be done by analyzing the lower levels of the hierarchy where more details are present. A binary hierarchy Π_{bw} is build in which each level represents the result of the thresholding applied on the corresponding level from the $\Pi_\bullet$. The hierarchy Π_{bw} consists also of L levels and each level l is given by

$$\Pi_{bw}^l = \Pi_\bullet^l < c \cdot thr_{\text{Otsu}}(\Pi_\bullet^l), l = \overline{1,L} \tag{1}$$

where $thr_{\text{Otsu}}(\cdot)$ computes the threshold for an image using the Otsu's method [13] and $0 < c \le 1$ (0.5 in our experiments).

   Our goal is to find for each partition (ancestor) from the top of Π_{bw} the corresponding partition (descendent) from a bottom level which properly identifies the lumen. Each level l of the Π_{bw} contains np_l unconnected regions $\Pi_{bw}^l = \{P_1^l, P_2^l, P_3^l, ..., P_{np_l}^l\}$. A vertical relation between partitions of successive levels can be established: each partition of a level is included in a partition from the below level.

$$\forall P_p^l (2 \le l \le L, 1 \le p \le np_l), \exists p'(1 \le p' \le np_{l-1}), \text{ such that } P_p^l \subset P_{p'}^{l-1} \tag{2}$$

   This inclusion is valid due to the reduction of the black regions caused by the increasing of SE's size used in the closing operation. According to Eq. 2, for each partition of any level of Π_{bw}, the corresponding partition from the base level can be found so that the inclusion rule is validated.

   The base level Π_{bw}^1 will not give for sure the proper regions because the global threshold is applied on the original image where the main structures are not properly highlighted. For each partition of the top level $P_r^L, 1 \le r \le np_L$, the corresponding

descendent from the base level P_r^1 is obtained. By bottom-up analyzing level by level, a certain level l_r is chosen such that the ancestor region ($P_r^{l_r}$) of P_r^1 from level l_r (having as ancestor the P_r^L) checks the following rules:

- solid($P_r^{l_r}$) > $min\_solidity$. The function solid(·) computes the proportion of the pixels from the convex hull that are also in the region; $min\_solidy = 0.8$ in our experiments.

- distEuclid(centroid($P_r^{l_r}$), centroid($P_r^{l_r-1}$)) < max_dist. The function centroid(·) returns the centroid of a region and the distEuclid(·) computes the Euclidean distance between these two centroids. This rule ensures stability by checking the distance between the regions centroids from two successive levels.

The border of the found lumens actually describes the inner border of the crypt (Fig. 3a). The false positive (FP) results are eliminated by a validation rule in 3.1.
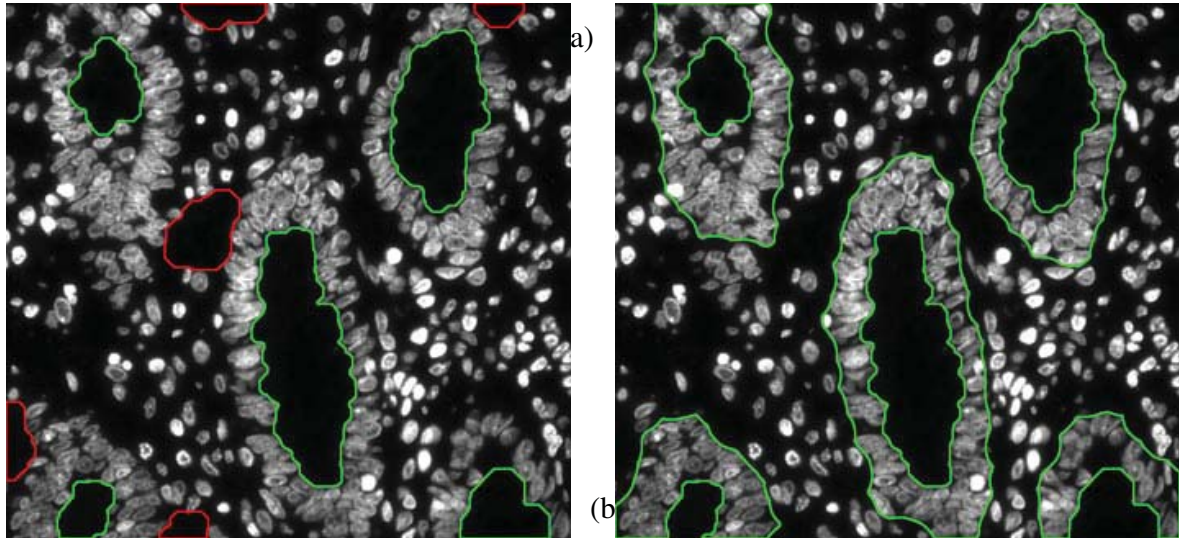

Fig. 3. a) The true positive (green curves) and the FP (red curves) lumen boundaries. b) The green curves indicate the inner and the outer boundaries of the crypts.

3   Crypt's Outer Border

The epithelial layer is differentiated from the stroma areas by considering the nuclei distribution: the crypt's nuclei are packed tightly together while those from the stroma areas are wide spread with considerable distances between them. The isolines of the inner boundary are used to eliminate the FP lumens and to detect the outer border which delineates the epithelial nuclei. Each isoline contains pixels situated on the same distance from the inner boundary.

The maximum distance d_{max} (60 in our experiments) is related to the average width of the epithelial layer. Fig. 4b displays the smoothed signal containing the intensities sum for each isoline. The maximum value (green square) indicates the

isoline which gives the maximum sum of pixels intensities. This maximal isoline (depicted with blue in Fig. 4a) is used for two purposes: to validate the detected lumens and to mark the epithelial nuclei (used to get the outer boundary).

3.1 Lumen Validation

The area covered by nuclei can be identified by subtracting a highly blurred version from the original image: $N_{bw} = (I - G * I) > 0$, where G is a big Gaussian filter (201 by 201 in our experiments) and $*$ denotes the convolution operation. Considering the high nuclei concentration around the lumen, there should be only few situations in which the maximal isoline crosses over the background in N_{bw} (Fig. 4a) i.e. situations of big distances between epithelial nuclei or in case of crypt breaks.
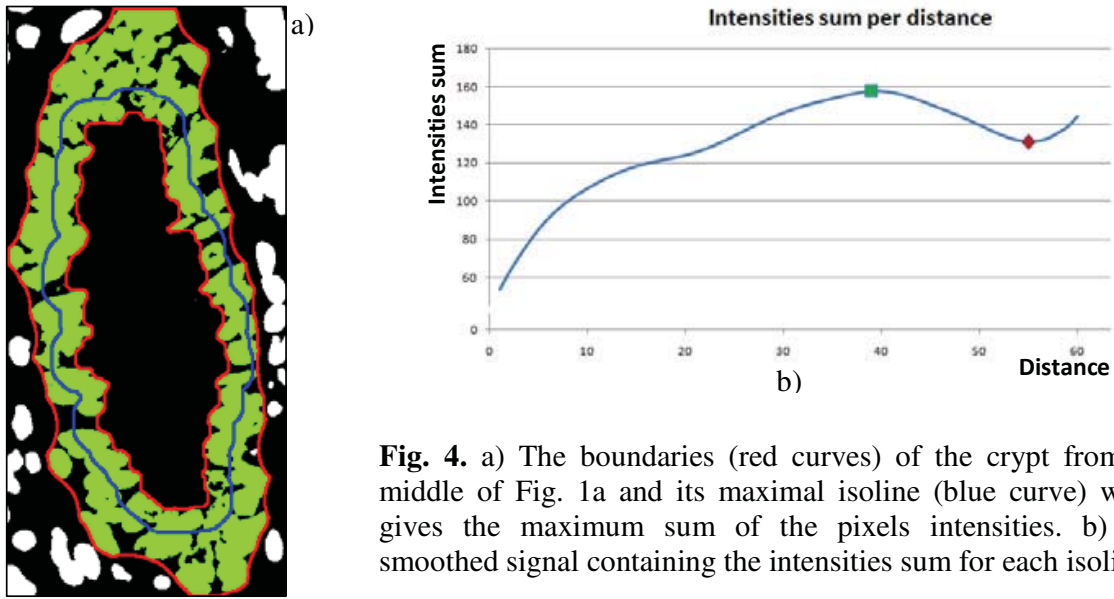

Fig. 4. a) The boundaries (red curves) of the crypt from the middle of Fig. 1a and its maximal isoline (blue curve) which gives the maximum sum of the pixels intensities. b) The smoothed signal containing the intensities sum for each isoline.

Based on this, the following rule is proposed to validate the lumen results from 2.2: if the portions of the maximal isoline overlapping the background in N_{bw} are not considerable high compared to those overlapping the nuclei than the found boundary does not delimit a lumen area.

$$r = \frac{card(MI \cap \sim N_{bw})}{card(MI \cap N_{bw})} \tag{3}$$

The binary image *MI* contains the maximal isoline, $card(\cdot)$ is the cardinality function and $\sim$ gives the complement of a binary image. The FP boundaries ($r > r_{\min}$) are depicted with red in Fig. 3a and the true positive (TP) inner boundaries ($r \le r_{\min}$) with green ($r_{\min} = 0.3$ in our experiments).

3.2 Outer Border Detection

A region from N_{bw} is marked as part of a crypt iff it is overlapped by the band formed by the inner boundary and the maximal isoline. The epithelial nuclei are

depicted with green in Fig. 4a. The outer border of the crypt (Fig. 4a the red curve outward) represents the exterior perimeter of the morphological closing applied on the found epithelial nuclei with a SE covering 2-3 nuclei.

4   Results

We tested the proposed segmentation technique on different datasets of images from tissues labelled with DAPI; some results are show in Fig. 3b and Fig.5. The results confirmed that the proposed method could efficiently segment the crypts with a high degree of accuracy.
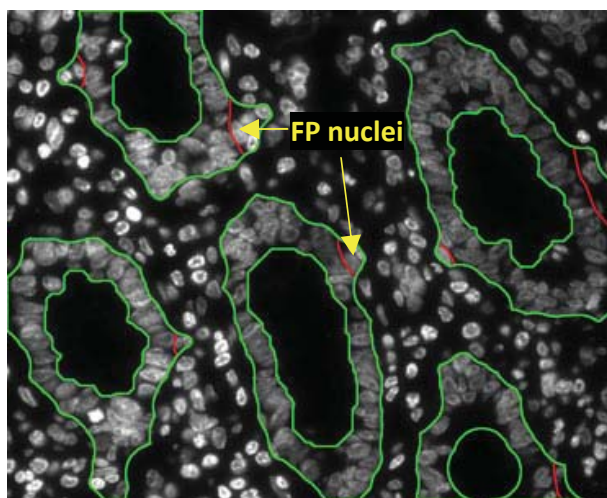

Fig. 5. The green curves indicate the inner and the outer crypt boundaries. The red portions delimit the FP nuclei.

A more rigorous evaluation must be done by comparing the results against the ground truth segmentations. Since a database with reference segmentations for this type of images does not yet exist, a pathology specialist has been asked to validate a set of results. The segmentation quality is established by visual inspecting the number of the mis-segmented nuclei per crypt. A number of 87 crypts have been analyzed resulting in 284 over segmented nuclei. Considering an average of 55 nuclei per crypt, the over-segmented nuclei represent 5.93% from the total crypt's nuclei (an average accuracy of 94.07% per crypt).

5   Conclusions

A new automatic technique for robust crypt segmentation based on hierarchical structures is presented in this paper. A closing morphological hierarchy is used to identify the lumen positions and a binary hierarchy provides enough details for proper lumen reconstruction. The maximal isoline is used to eliminate the false positive lumens and to validate the epithelial nuclei belonging to crypts.

A significant implication of the current work consists of the top-down approach. Firstly the 'obvious' areas (lumens) are detected followed by a more detailed analysis to proper delimit the crypts. The morphological closing operator has been chosen to build the hierarchical representation due to the patterns of nuclei arrangements.

This technique uses a coarser-to-fine approach and can be easily extended on any image with human body cell nuclei from different tissues types (e.g. prostate, breast or lung) but also in any other field in which the objects of interest have the features considered in designing this method. This study will be continued by analysing the topological properties of the graph associated to the tissues components. Considerable

effort will be spent to obtain a database with ground-truth segmentations and to find rigorous evaluation criteria of the results.

References

1. Ta, V.T., Lezoray, O., El Moataz, A., Schupp, S.: Graph-based tools for microscopic cellular image segmentation. Pattern Recognition 42(6), 1113–1125 (2009)
2. Nattkemper, T.W.: Automatic segmentation of digital micrographs: A survey. Journal Studies in Health Technology and Informatics 107(2), 847–852 (2005)
3. Wu, H.S., Xu, R., Harpaz, N., Burstein, D., Gil, J.: Segmentation of intestinal gland images with iterative region growing. Journal of Microscopy 220(3), 190–204 (2005)
4. Farjam, R., Soltanian-Zadeh, H., Jafari-Khouzani, K., Zoroofi, R.A.: An image analysis approach for automatic malignancy determination of prostate pathological images. Clinical Cytometry 72B(4), 227–240 (2007)
5. Gunduz-Demir, C., Kandemir, M., Tosun, A.B., Sokmensuer, C.: Automatic segmentation of colon glands using object-graphs. Medical Image Analysis 14(1), 1–12 (2010)
6. Naik, S., Doyle, S., Feldman, M., Tomaszewski, J., Madabhushi, A.: Gland segmentation and computerized Gleason grading of prostate histology by integrating low-, high-level and domain specific information. In: 2nd MICCAI Workshop Microscopic Image Analysis with Appl. in Biology, Piscataway, NJ, USA, (2007)
7. Humphries, A., Wright, N.A.: Colonic Crypt Organization and Tumorigenesis: Human Colonic Crypts. Nature Reviews Cancer 8(6), 415–424 (2008)
8. Morikawa, K., Yanagida, J.: Visualization of individual DNA molecules in solution by light microscopy: DAPI staining method. Japanese Biochemical Society 89(2), 693–696 (1981)
9. Kropatsch, W.G., Haxhimusa, Y., Ion, A.: Multiresolution Image Segmentations in Graph Pyramids. In: Kandel, A., Bunke, H., Last, M. (eds.) Applied Graph Theory in Computer Vision and Pattern Recognition, vol. 52, pp. 3–41. Springer, Wien (2007)
10. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd edn. Addison-Wesley, Reading (1992)
11. Haralick, R., Lee, J., Lin, C., Zhuang, X.: Multi-resolution Morphology. In: First IEEE Conference on Computer Vision, London (1987)
12. Haralick, R.M., Zhuang, X., Lin, C., Lee, J.S.J.: The digital morphological sampling theorem. IEEE Trans. Acoust., Speech, Signal Processing 37, 2067–2090 (1989)
13. Otsu, N.: A Threshold Selection Method from Gray-Level Histograms. IEEE Transactions on Systems, Man, and Cybernetics 9(1), 62–66 (1979)