Separation of Benign and Malignant Glands in Prostatic Adenocarcinoma

Sabrina Rashid[1], Ladan Fazli[2], Alexander Boag[3], Robert Siemens[3], Purang Abolmaesumi[1], and Septimiu E. Salcudean[1]

[1] Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada
[2] The Vancouver Prostate Center, University of British Columbia, Vancouver, BC, Canada
[3] Kingston General Hospital, Kingston, ON, Canada

Abstract. This paper presents an analysis of the high resolution histopathology images of the prostate with a focus on the evolution of morphological gland features in prostatic adenocarcinoma. Here we propose a novel technique of labeling individual glands as malignant or benign. In the first step, the gland and nuclei objects of the images are automatically segmented. Individual gland units are segmented out by consolidating their lumina with the surrounding layers of epithelium and nuclei. The nuclei objects are segmented by using a marker controlled watershed algorithm. Two new features, Number of Nuclei Layer (N_{NL}) and Ratio of Epithelial layer area to Lumen area (R_{EL}) have been extracted from the segmented units. The main advantage of this approach is that it can detect individual malignant gland units, irrespective of neighboring histology and/or the spatial extent of the cancer. The proposed algorithm has been tested on 40 histopathology scenes taken from 10 high resolution whole mount images and achieved a sensitivity of 0.83 and specificity of 0.81 in a leave-75%-out cross-validation.

Keywords: Nuclei layer, epithelial layer, prostatic adenocarcinoma.

1 Introduction

Prostate cancer is one of the most frequently diagnosed cancers and ranks second among the cancer related deaths of men worldwide [1]. Analysis of the histopathology specimens of prostate is an important step for prostate cancer diagnosis and treatment planning.

The tissue features of these histopathology images are the key indicators of prostate cancer. Among the different types of prostate cancer, the most common one is the prostatic adenocarcinoma, cancer pertaining to the gland units of the prostate. Pathologists determine the extent of this cancer by carefully evaluating the changes in the gland morphology. The gland is the main histopathological structural unit in prostate. Fig. 1 shows the structure of a normal gland unit. It mainly comprises a lumina of irregular shape, a layer of epithelial cells, and

K. Mori et al. (Eds.): MICCAI 2013, Part III, LNCS 8151, pp. 461–468, 2013.

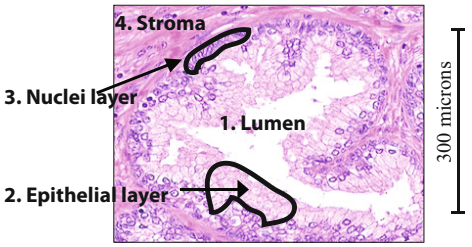

Fig. 1. Illustration of the histopathology components associated with a gland unit: 1. Lumen, 2. Epithelial layer, 3. Nuclei, and 4. Stroma.

nuclei surrounding the lumina. The unit is supported by a surrounding fibromuscular stroma. When the slides are stained using a Hematoxylin and Eosin (H&E) solution, the nuclei turn dark blue and the epithelial layer and stroma turn into different shades of purple to pink.

The recent literature on computerized diagnosis of prostate cancer quantizes the morphological and architectural features associated with the gland units for cancer detection and grading. The most commonly used features on analyzing histopathology specimens are related to the size and shape of gland lumina, nuclei shape and density [2], [3], [4], [5]. Though gland size and shape do contain information about the abnormality of prostate tissue, this feature is not exclusive to cancerous tissue only. In case of other prostate anomalies such as, atrophy and benign prostatic hyperplasia the gland size and shape resembles that of cancerous ones [6]. Apart from gland-based features, some approaches exploit overall image features such as, energy and entropy of multiwavelet coefficient of the image [7], fractal dimension [8] and so on. But these features are not specific to each gland and do not capture the gland specific features that are clinically used for cancer classification.

Therefore, more decisive features are needed for effective separation of individual glands. Here, we propose two novel features based on which glands can be classified: i) Number of Nuclei Layer (N_{NL}), and ii) Ratio of Epithelial layer area to Lumen area (R_{EL}). Fig. 2 illustrates the change of appearance between benign and malignant glands. In benign glands there are multiple layers of nuclei surrounding the gland unit with relatively thinner epithelial layer compared to lumen area. On the other hand, in malignant units there are usually a single layer of nuclei surrounding the gland unit with a thick epithelial layer. Our proposed features N_{NL} and R_{EL} quantize these two properties of malignant glands. In a recent literature, Nguyen et al. [3] achieved an accuracy of 0.79 in classification of benign and malignant glands by exploiting region-specific features such as percentage of nuclei pixels, gland shape, crowdedness etc. In comparison to that, our proposed features are strictly gland specific and involve i) pixel labeling, ii) segmentation of each nuclei in gland, and iii) finding the number of layers of nuclei for each gland from angle-dependent histograms. The advantage

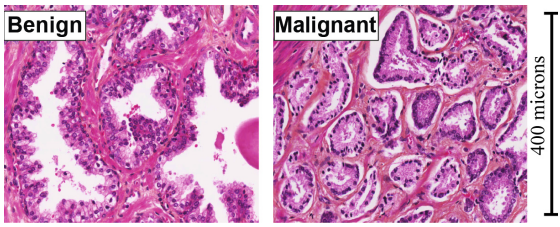

Fig. 2. Visual comparison between benign and malignant prostate glands

of this technique is that it can detect a malignant gland irrespective of the region properties. In cases where malignant glands are present in close proximity of benign glands, this approach might provide a more sensitive cancer annotation compared to the approaches that use region-dependent image features [3]. The proposed technique has been evaluated on 40 histopathology scenes extracted from 10 whole mount images with a resolution of 0.5 μm per pixel and achieved 0.83 and 0.81 of sensitivity and specificity, respectively, in a leave-75%-out cross-validation experiment.

This paper is organized as follows: the methodology of the complete gland classification algorithm is presented in Section 2. In Section 3 the experimental results of the proposed algorithm are presented. Finally, Section 4 presents concluding remarks and suggestions for future work.

2 Methodology

2.1 Segmentation of the Gland

The gland segmentation algorithm has been partially adopted from another work of Nguyen et al. [9]. In the first step, each pixel in the image is categorized into one of these four categories: i) Gland lumen, ii) epithelial layer, iii) nuclei, and iv) stroma. Small training patches of each class have been selected to train the classifier for pixel labeling. The classification is based on the color information of these histological objects in the two chromaticity layer 'a*' and 'b*' of the L*a*b color space.

For pixel labeling, we have used linear discrimination analysis instead of the Voronoi tesselation based approach from [9]. The main drawback of Voronoi tesselation approach is that when the number of training samples is large, the classification time for each testing data point is very high compared to that of linear discriminant analysis [10]. Therefore, when the number of testing samples are very large the reported nearest neighborhood approach will be very expensive to compute. After having the labeled image, the complete gland units are constructed by iteratively consolidating lumen objects with the surrounding epithelial layer and nuclei pixels [9] (see Fig. 3c).

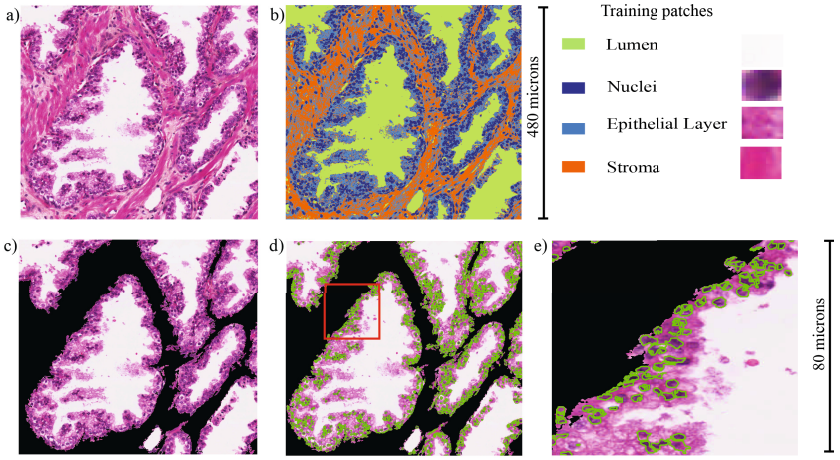

Fig. 3. Gland segmentation. a) A sample histopathology scene, b) Labeled image, c) Segmented gland units after consolidation of lumina with surrounding epithelial layer and nuclei objects, d) Segmented nuclei objects overlaying on the segmented gland units, and e) enlarged view of the segmented nuclei objects.

2.2 Segmentation of the Nuclei Units

Segmentation of nuclei is performed by employing a marker controlled watershed algorithm [11] followed by a Support Vector Machine (SVM) based object classification. The segmentation function used in the watershed algorithm is the gradient image of the inverted grayscale image of the input scene. This gradient image is modified by placing regional minima in the marked pixels of foreground and background objects of the image. The foreground markers are determined by finding the regional minima in the image after the morphological opening and closing by reconstruction operations. To determine the background marker, the same operation is performed with the inverted image. Along with the nuclei objects, the segmented objects often include some other histological objects, i.e., crystalloids inside glands, darkly stained stroma/epithelial object due to nonuniform staining and so on.

The non-nuclei objects are then filtered out by employing a SVM classifier. The features used in this classification are the mean intensity, entropy, and the standard deviation of the segmented objects. The classifier is trained using a linear kernel and manually selected training samples of nuclei and non-nuclei objects. Fig. 3d) and e) depicts the segmented nuclei objects that are part of the gland unit. Since our main focus is to extract the features related to gland morphology, the segmented nuclei units that float in the stromal region are not considered for further analysis.

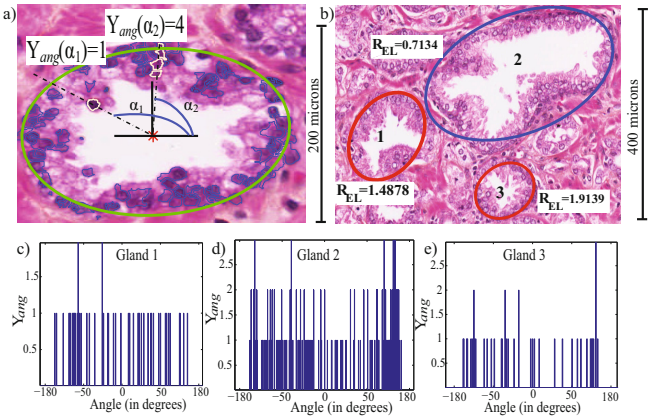

Fig. 4. a) Graphical illustration of Y_{ang} calculation. b) Sample histopathology scene with a single benign (marked with blue ellipse) and two malignant gland units (marked with red ellipses). c), d), and e) illustrate the different appearances of the histograms (Y_{ang}) of benign and malignant glands.

2.3 Extraction and Classification of the Features

Number of Nuclei Layer, N_{NL}: To determine the number of nuclei layers pertaining to each gland, at first the segmented nuclei objects are paired with the corresponding gland unit that minimizes the distance between the centroid of the nuclei and the gland lumen boundary. For each of the combined gland-nuclei object, an ellipse is fit around it. The angular location of each of the nuclei is evaluated by calculating the angle of the connecting line of the gland centroid and corresponding nuclei centroid (see Fig. 4a). Then the feature N_{NL} is evaluated from the histogram Y_{ang} of angular locations of nuclei. Customized bin spacing has been utilized to account for glands of different sizes. The bin spacing for the histogram is evaluated as $360/P_g$, where P_g is the perimeter of the corresponding ellipse surrounding the gland. Then the N_{NL} is evaluated by counting the total number of instances where multiple nuclei have same angular location in the histogram and then normalizing it by dividing by P_g. Mathematically, $N_{NL} = \frac{1}{P_g}|\{n|Y_{ang}(n) >= 1\}|$. Fig. 4c-e illustrates the different nature of histogram, Y_{ang} in case of benign and malignant glands. As can be observed from the figure, the benign histogram provides more instances of multiple nuclei having same angular location.

Ratio of Epithelial Layer area to Lumen Area, R_{EL}: This feature is evaluated by simply taking the ratio of the epithelial layer area to lumen area of the gland. In case of malignant glands, fast multiplication of cells lead the epithelial layer to invade more in to gland lumina. As a result, the ratio gets larger in case of malignant gland units.

After the feature extraction we choose optimum thresholds on the features, $\tau_{N_{NL}}$ and $\tau_{R_{EL}}$ for the classification of benign and malignant glands. We classify a gland (G_i) as malignant when the parameters fulfill the following criteria, $Label_{G_i} = \{Malignant | N_{NL}(G_i) < \tau_{N_{NL}}, R_{EL}(G_i) > \tau_{R_{EL}}\}$. These threshold parameters are tuned by performing Receiver Operator Characteristics (ROC) analysis in a leave-75%-out experiment (discussed in the following section).

3 Experiments and Results

The proposed algorithm has been evaluated on 40 different histopathology scenes containing a total of 2145 glands. These scenes have been extracted from 10 whole mount histopathology images obtained from eight radical prostatectomy patients. These whole mount histopathology images are digitized at 20× magnification (0.5 μm per pixel) with an Aperio scanner. Cancerous regions in these images are annotated by an expert pathologist. These annotations are used as the gold standard to evaluate the performance of the proposed algorithm.

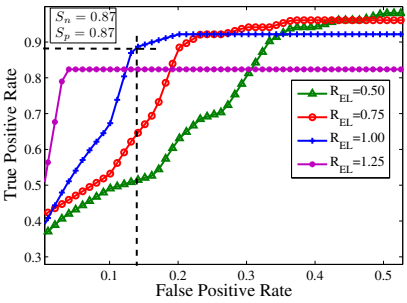

Fig. 5. a) ROC of our algorithm for four discrete values of R_{EL}. At the optimum operating point the algorithm achieves, sensitivity $S_n = 0.87$ and specificity $S_p = 0.87$.

The performance of the algorithm is influenced by the choice of the parameter value of R_{EL} and N_{NL}. We tune the parameters by performing ROC analysis on randomly selected 25% of the glands in the dataset. The ROC curve of the classifier is generated by varying the parameter N_{NL} as $\{0, 0.08, 0.16, ..., 4\}$. To determine the effect of varying R_{EL} on the classifier performance the following operation has been performed: for each choice of R_{EL} in the set $R_{EL} = \{0, 0.25, ..., 2.5\}$

Table 1. AUC obtained by our algorithm for different parameter values of R_{EL}

R_{EL}	0.25	0.5	0.75	1	1.25	1.5	1.75	2	2.25	2.5
AUC	0.84	0.88	0.89	0.87	0.81	0.81	0.81	0.81	0.80	0.77

individual ROC curve by varying N_{NL} has been generated. We choose the thresholds $\tau_{N_{NL}}$ and $\tau_{R_{EL}}$ corresponding to the optimum operating point in the ROC curve. Then we use these thresholds on the remaining 75% of the data for gland classification. By repeating the above leave-75%-out experiment 10 times we achieved 0.83 ± 0.007 of sensitivity and 0.81 ± 0.005 of specificity. When we performed similar analysis on the entire dataset we achieved 0.87 of sensitivity and 0.87 of specificity at the optimum operating point. The corresponding values of $\tau_{R_{EL}}$ and $\tau_{N_{NL}}$ at this point are 1 and 2.24, respectively. Table 1 lists the Area Under the Curve (AUC) obtained for different values of R_{EL}. Fig. 5 illustrates the ROC curve for four different R_{EL} values .

Fig. 6 illustrates the performance of the proposed algorithm on three sample histology scenes, one entirely benign scene, one entirely malignant scene obtained from marked cancerous region by the pathologist, and one scene comprising both malignant and benign glands in close proximity of each other. Experimentally classified benign and malignant units are marked by blue and yellow ellipses, respectively. In all the examples, strong agreement between the pathologist's marking and experimental classification of glands corroborates the effectiveness of the proposed algorithm.

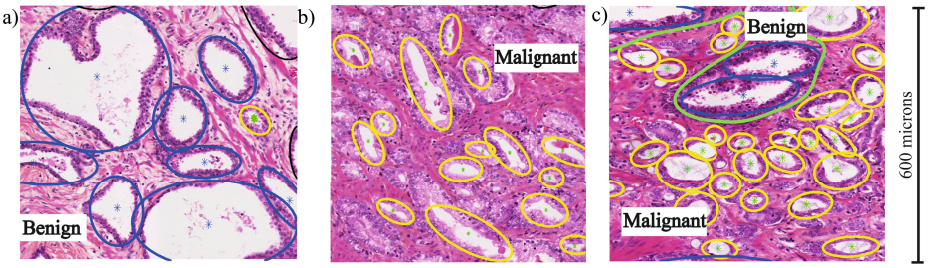

Fig. 6. Application of the proposed technique on three sample histology scenes. The yellow and blue ellipses are used to denote the malignant and benign gland units respectively. Sample scenes containing a) only benign glands, b) only malignant glands and c) both the benign and malignant glands. The green annotation mark by pathologist denotes the separation of benign and malignant glands. All the images are shown in the same scale of magnification.

4 Conclusion

In this paper, we have proposed a technique for the classification of individual benign and malignant glands based on two novel features, Number of Nuclei layers and Ratio of Epithelial layer area to Lumen area. To the best of our knowledge, this is the first work to quantify nuclei layers associated with each gland based on angular histogram. This is also the first report of individual gland labeling as malignant or benign without relying on the surrounding histology information. Since most reports on automatic cancer annotation are region based, cases with very small cancerous area might not be diagnosed by those approaches.

The application of individual gland-based technique proposed here will lead to a more sensitive cancer annotation and thus improved diagnosis of early stage prostate cancer. In the future, we plan to implement this technique on entire whole mount images. Moreover, we plan to investigate the relationship between the proposed features and long term disease progression.

Acknowledgment. Funding from the NSERC-CIHR Collaborative Health Research Projects, the BC Innovation Council NRAS Program, and a travel grant from the Institute for Computing, Information and Cognitive Systems (ICICS) at the University of British Columbia are gratefully acknowledged.

References

1. Bohring, C., Squires, T.: Cancer statistics. CA Cancer J. Clin. 43, 7–26 (1993)
2. Monaco, J.P., Tomaszewski, J.E., Feldman, M.D., Hagemann, I., Moradi, M., Mousavi, P., Boag, A., Davidson, C., Abolmaesumi, P., Madabhushi, A.: High-throughput detection of prostate cancer in histological sections using probabilistic pairwise markov models. Medical Image Analysis 14(4), 617 (2010)
3. Nguyen, K., Sarkar, A., Jain, A.K.: Structure and context in prostatic gland segmentation and classification. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part I. LNCS, vol. 7510, pp. 115–123. Springer, Heidelberg (2012)
4. Clark, M.D., Askin, F.B., Bagnell, C.R.: Nuclear roundness factor: a quantitative approach to grading in prostate carcinoma, reliability of needle biopsy tissue, and the effect of tumor stage fore usefulness. The Prostate 10, 199–206 (1987)
5. 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: Proc. of 2nd Workshop on Micro. Image Anal. with Applications in Biology (2007)
6. Epstein, J.I., Netto, G.J.: Biopsy interpretation of the prostate. Lippincott Williams & Wilkins (2007)
7. Khouzani, J.K., Zadeh, S.H.: Multiwavelet grading of prostate pathological images. In: Proc. SPIE, vol. 4628, pp. 1130–1138 (2002)
8. Huang, P.W., Lee, C.H.: Automatic classification for pathological prostate images based on fractal analysis. IEEE Transactions on Medical Imaging 28(7), 1037–1050 (2009)
9. Nguyen, K., Sabata, B., Jain, A.K.: Prostate cancer grading: Gland segmentation and structural features. Pattern Recognition Letters 33, 951–961 (2011)
10. Krzanowski, W.: Principles of multivariate analysis. Oxford Uni. Press (1996)
11. Meyer, F.: Topographic distance and watershed lines. Signal Processing 38(1), 113–125 (1994)