Title	Automatic detection of malignant prostatic gland units in cross-sectional microscopic images
Author(s)	Xia, T; Yu, Y; Hua, J
Citation	The 17th IEEE International Conference on Image Processing (ICIP 2010), Hong Kong, China, 26-29 September 2010. In Proceedings of the 17th ICIP, 2010, p. 1057-1060
Issued Date	2010
URL	http://hdl.handle.net/10722/139999
Rights	International Conference on Image Processing Proceedings. Copyright © IEEE.

AUTOMATIC DETECTION OF MALIGNANT PROSTATIC GLAND UNITS IN CROSS-SECTIONAL MICROSCOPIC IMAGES

Tian Xia *Yizhou Yu* *Jing Hua*

University of Illinois at Urbana-Champaign Wayne State University

ABSTRACT

Prostate cancer is the second most frequent cause of cancer deaths among men in the US. In the most reliable screening method, histological images from a biopsy are examined under a microscope by pathologists. In an early stage of prostate cancer, only relatively few gland units in a large region become malignant. Discovering such sparse malignant gland units using a microscope is a labor-intensive and error-prone task for pathologists. In this paper, we develop effective image segmentation and classification methods for automatic detection of malignant gland units in microscopic images. Both segmentation and classification methods are based on carefully designed feature descriptors, including color histograms and texton co-occurrence tables.

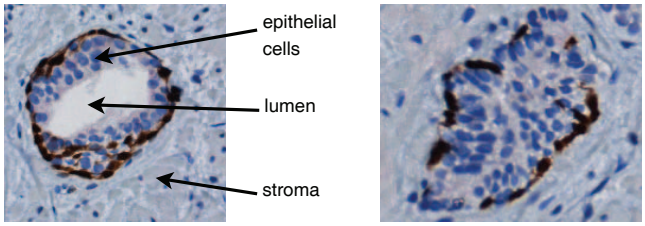
Index Terms— Prostate Glands, Histological Images, Classification, Segmentation

1. INTRODUCTION

Prostate cancer is the second most frequent cause of cancer deaths among men in the US. Almost one-third of American men over 50 years old will be diagnosed with prostate cancer during their life times. Nevertheless, if diagnosed in a sufficiently early stage, prostate cancer patients have a high probability to survive. Several screening methods exist, the most reliable one being biopsy. Sample tissues from a biopsy are usually stained by Hematoxylin and Eosin (H&E) and sliced into cross sections for examination under a microscope.

Diagnosis of prostate cancer is performed by examining the glandular architecture in histological images of the specimen. Normal prostate tissue, illustrated in Fig. 1(a), consists of gland units surrounded by fibromuscular tissue, called stroma. Each gland unit consists of layers of epithelial cells located around an empty tubular region, named the lumen. When cancer occurs, epithelial cells replicate in an uncontrolled way to fill up the lumens as depicted in Fig. 1(b). In a more serious state, even the stroma disappears, giving way to replicated epithelial cells. Therefore, the condition of gland units is the most important indicator of prostate cancer.

Cancerous development in gland units is usually spatially varying. Some regions may progress faster than others. In an early stage of prostate cancer, only relatively few gland units in a large region become malignant. Discovering such sparse malignant gland units in cross sections of the specimen under a microscope is a labor-intensive and error-prone task for pathologists. It would be very easy for them to miss sparse malignant units in a sea of benign ones. Nevertheless, the screening of malignant gland units is crucial for early diagnosis of prostate cancer. In this paper, we propose a method to automatically detect malignant gland units in cross-sectional microscopic images.



(a) normal gland unit (b) malignant gland unit

Fig. 1. Examples of prostatic gland units in cross-sectional microscopic images.

Achieving this task is challenging for the following reasons. First, an accurate segmentation of gland units, especially malignant gland units, is required in order to correctly collect features from them for classification. However, a gland unit is highly inhomogeneous, consisting of epithelial cells, some background tissue and the lumen. In addition, it may have an irregular shape and may not have a clear boundary. Second, criteria for classifying a gland unit should be partially based on the spatial distribution of epithelial cells within the unit. But it is not always possible to segment individual epithelial cells especially when the microscopic image does not have a sufficient resolution.

Our solution for malignant gland unit detection makes use of two cascaded classifiers. The first classifier performs pixel-level classification. It is trained to discern pixels inside gland units from those outside. The classification result provides a noisy initial segmentation of gland units. A refinement to the initial segmentation by improving spatial coherence is performed through a graph cut algorithm. Features within every segmented gland unit are then extracted and fed to the second classifier, which determines whether a gland unit is malignant or not. An important class of features we use are based on textons [1, 2], whose spatial distribution serves as a representation of the cellular structures within a gland unit. There exist two separate training stages for the two classifiers both of which share the same set of training images. The first classifier is trained simply using a set of labeled pixels in the training images. However, the training data for the second classifier needs to be prepared using the first classifier, which segments out individual gland regions in the training images. A human expert then labels the segmented gland regions. Experimental results confirm the effectiveness of our solution.

2. RELATED WORK

There exists much related work on prostatic image classification and cancer grading. Pitts *et al.* [3] investigated the application of gray level co-occurrence matrix techniques for the interpretation of

prostate cancer lesions and identifying corresponding features on the color images of the section. Jafari-Khouzani and Soltanian-Zadeh [4] used features based on multiwavelets combined with a k-nearest neighbor classifier to classify each image into grades 2 through 5. The classification distance metric is optimized using simulated annealing. Doyle *et al.* [5] introduced a multiresolution scheme where pixel-wise Bayesian classification is performed at each image scale to obtain corresponding likelihood values. Starting at the lowest scale, they apply the AdaBoost algorithm to analyze pixels with a high probability of malignancy at subsequent higher scales. Tabesh *et al.* [6] proposed an automatic two-stage system for prostate cancer diagnosis and Gleason grading. The color, morphometric, and texture features are extracted from prostate tissue images. Linear and quadratic Gaussian classifiers were then used to classify images into tumor/nontumor classes, and further into low/high grades for cancer images. Most recently, Huang and Lee [7] proposed two feature extraction methods based on fractal dimension to analyze variations of intensity and texture complexity in regions of interest. Each image can be classified into an appropriate grade by using Bayesian, k-NN, and support vector machine (SVM) classifiers, respectively. A review of computer technologies in detection and staging of prostate carcinoma can be found in [8].

Most of the aforementioned work performs image classification and grading without explicitly segmenting out gland units. Therefore, such methods tend to produce incorrect results when sparse malignant gland units are distributed among a large number of benign ones. In our method, we perform high-quality gland unit segmentation using state-of-the-art techniques, such as graph cut. In addition, the classification of gland units is based on local texture pattern analysis using textons, which have proven to be useful for texture description.

3. GLAND UNIT SEGMENTATION

Our gland unit segmentation is composed of two stages, an initial binary classification followed by a graphcut based refinement. We perform texture classification on microscopic images because the micro structures of tissues and cells resemble a texture image with repetitive patterns. Since texture features for classification need to be defined over local spatial neighborhoods, we first group nearby pixels sharing similar color and intensity attributes into small patches called superpixels using the algorithm in [9].

Oriented filter banks have proven to be an effective tool to characterize textures. We apply the oriented filter bank in [2] to three color channels separately. The filter bank has 36 elongated filters at 6 orientations, 3 scales, and 2 phases, 8 center-surround difference of Gaussian filters, and 4 low-pass Gaussian filters. Thus, every pixel has a 144-component filter response vector after filtering. We compute the mean and standard deviation of each component within a superpixel to obtain a 288-component vector.

A normalized histogram is also computed for every color channel of every superpixel. We use 51 bins for each color channel, and the 153-component vector is then concatenated with the 288-component vector. It is reduced to a shorter vector by principal component analysis. This shortened vector, in juxtaposition with the mean and standard deviation of each color channel of the superpixel, forms the texture descriptor for every supervoxel. In our experiments, we have used a descriptor length around 60.

We use AdaBoost [10] with decision trees as our classifier. The above descriptors are fed to our classifier as feature vectors. We limit the maximum depth of each decision tree to 4 to maintain a good generalization capability.

A microscopic image for prostatic cross sections usually contains a large number of gland units. And we typically use multiple training images. To avoid the tremendous work required to prepare the training data by manually segmenting all gland units in the training images, we adopt an interactive training approach [11]. In this approach, the user starts with a few scribbles on one image indicating a few gland regions as well as the background, as shown in Fig. 2(b). Superpixels covered by these scribbles serve as the initial positive and negative training data on which the first tentative classifier is trained and applied to the unlabeled data. If the tentative result is not satisfactory, the user inputs a few more scribbles to further guide the training process. In practice, this iterative procedure runs for a few times before the classifier performs sufficiently well on all microscopic images.

3.1. Graphcut Based Refinement

As the above classifier labels every patch independently, there may exist sporadic misclassifications in the middle of a gland region as in Fig. 2(c). We enhance the spatial coherence of segmentation results by applying a graph cut algorithm. The refined segmentation result is shown in Fig. 2(d).

The idea of using graph cut for segmentation [12] is to treat image segmentation as a binary labeling problem. Specifically, the image induces a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ represents the set of pixels and $\mathcal{E}$ represents the set of edges between every pair of neighboring pixels. Every pixel also has links to two virtual nodes, the source representing the foreground and the sink representing the background. The labeling problem is to assign a unique label l_i for each node $i \in \mathcal{V}$, i.e. $l_i \in \{\text{foreground}(=1), \text{background}(=0)\}$, so that the following objective function is minimized.

$$E(L) = \sum_{i \in \mathcal{V}} R(l_i) + \lambda \sum_{(i,j) \in \mathcal{E}} B(l_i, l_j), \tag{1}$$

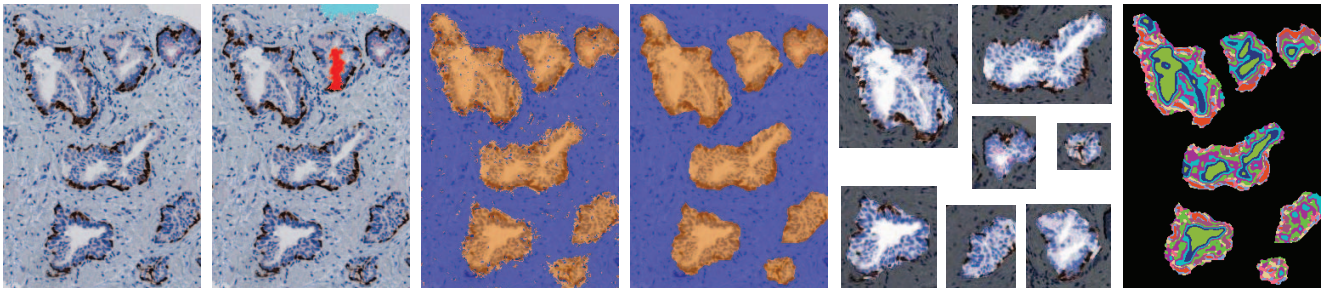
where $R(l_i)$ represents the cost when the label of node i is l_i, and $B(l_i, l_j)$ denotes the cost when the labels of adjacent nodes i and j are l_i and l_j respectively. If we set the weight of edge (i, j) to $B(l_i, l_j)$, the weight of the link between node i and the source to $R(l_i = 1)$ and the weight of the link between node i and the sink to $R(l_i = 0)$, finding an optimal solution of (1) is equivalent to finding a cut with minimal cost in the graph. More details can be found in [12] and [13].

We would like to make use of the interactively trained texture classifier in the graph cut algorithm. instead of using a global color distribution to set the source/sink link cost as in [13], we use the signed confidence value (voting result), $\sum_t \alpha_t h_t(p)$, where $\alpha_t h_t(p)$ is the signed label of a pixel p predicted by a base classifier in AdaBoost. Specifically, we set $R(l_i = 0) = \max(0, -\ln \max(0, 0.5 + 0.5 \sum_t \alpha_t h_t(p)))$ and $R(l_i = 1) = \max(0, -\ln \max(0, 0.5 - 0.5 \sum_t \alpha_t h_t(p)))$. We adopt the conventional scheme of assigning edge cost where the color difference between adjacent pixels is used.

4. GLAND UNIT CLASSIFICATION

4.1. Gland-Level Feature Descriptors

We treat every connected component in the previously segmented foreground as a gland unit (Fig. 2(e)). For every detected gland unit, we build a descriptor that characterizes discriminative features of malignant gland units. Such a descriptor will be used for both training and classification. As malignant gland units and benign ones differ in their color distributions, textural and structural features,



(a) original image (b) user scribbles (c) initial pixelwise classification (d) graph cut refinement (e) segmented gland units (f) texton distribution

Fig. 2. (a)-(d) shows the prostatic gland unit segmentation procedure. (e) shows gland units (highlighted regions) segmented from the background (darkened regions). (f) illustrates the distribution of 10 textons. Each color represents a distinct texton label.

we incorporate such information when building descriptors. More specifically, we generate a color descriptor, a texton descriptor, and a texton co-occurrence descriptor. They are combined into one single feature vector for every gland unit.

Color distributions are well described by histograms. We compute three normalized histograms for every gland unit, one for each of the three color channels. 51 bins are used for each histogram. Thus, the color descriptor of a gland unit is a 153-component vector.

It is well-known that pixelwise histograms are a type of first-order statistics that is inherently ambiguous. That is, randomly swapping pairs of pixel colors within a gland unit would not affect its histograms. Nevertheless, swapping may destroy the shape and structure of local patterns, such as epithelial cells, making the resulting image unrecognizable as a gland unit. Important visual cues that distinguish a malignant gland unit from a benign one include the percentage of area covered by epithelial cells, the spacing among nearby epithelial cells in the interior of a gland unit as well as along its boundary, and the existence of the lumen. Descriptors more informative about the existence and spatial distribution of certain local patterns would be more useful for our intended classification. Thus, we choose to build descriptors for the frequencies of local patterns as well as descriptors reflecting the spacing among nearby local patterns.

Since a texture can be considered as a repetitive spatial arrangement of local texture patterns (textons), we start with sorting out typical local patterns by building a texton library for gland units. Since it is possible to reconstruct a local pattern from its responses to a bank of redundant oriented filters, a local pattern can be uniquely represented by its filter responses. We adopt the same oriented filter bank [2] used in Section 3, which leads to a 144-component filter response vector at every pixel. Following [2], we cluster these filter response vectors using the K-means algorithm and the resulting cluster centers represent a small set of prototype response vectors, $i.e.$ textons. An example of texton distribution can be seen in Fig. 2(f).

We form two descriptors based on textons, a texton histogram and a texton co-occurrence descriptor. Every pixel within a gland unit is first associated with a texton whose filter response vector is most similar to the one at the pixel. We mark every pixel with the label of its associated texton. A normalized texton histogram is then computed. The number of bins in the histogram is equal to the number of distinct textons. This histogram holds the information regarding the existence and frequency of typical local patterns.

Inspired by the grey-level co-occurrence matrix [14], we propose another descriptor based on texton co-occurrence frequencies. Recall that every pixel is associated with a texton label, the texton co-occurrence descriptor measures how often a pair of texton labels at a certain distance occur in a gland unit. Within the gland unit, we tabulate the number of occurrences of all possible pairwise texton combinations at specified distances. Each row of the table represents a specific pair of texton labels, and each column corresponds to a certain distance between pairs of texton labels. Thus, each cell in the table records the number of times a specific texton pair, separated by a given distance, appears in the image. The table is then flattened into a single vector. In our experiments where 10 textons are trained, we normally choose 6 distance values ranging from 1 to 24 pixels (2,6,10,15,20,24), resulting in a descriptor vector of length 330. To make descriptors from different-sized gland units comparable to each other, we normalize them by making the summation of all elements in the vector equal to one. Unlike conventional co-occurrence matrices, our descriptor is rotation-invariant because we do not maintain the orientation of the displacement between a texton pair.

Finally, we concatenate all three normalized descriptors into the same vector and perform PCA on such concatenated vectors to reduce their dimensionality to 80. The vector of PCA coefficients resulting from dimension reduction serves as the final feature vector fed into a classifier.

4.2. Training and Classification

Gland units in cross-sectional microscopic images are present to pathologists for labeling. These labels (both malignant and benign ones) as well as their corresponding feature vectors are used for the training of a binary classifier.

Once a classifier has been trained, undiagnosed gland units in new cross-sectional images can be readily segmented and classified. Note that being a flexible framework, any binary classification technique can apply here. In this paper, we use support vector machines. We have also experimented with decision tree classifiers enhanced by the AdaBoost algorithm [10], which render satisfactory classification accuracy as well. The discriminative power of the decision trees is greatly improved by the boosting mechanism. As a result, the classifier only requires shallow decision trees of moderate accuracy.

5. EXPERIMENTS

All experiments are carried out on a 3.8GHz Pentium 4 processor. We have obtained 27 microscopic images from which 267 gland units are successfully extracted using the segmentation algorithm described in Section 3. Among them, 159 are benign (positive data) and 108 are malignant (negative data). Parameter specifications are given in Section 3 and 4.1.

Feature sets	Classifiers	Accuracy (%)
color histograms + texton histogram + texton co-occurrence	SVM	86.0 ± 1.9
	boosted trees	85.3 ± 2.2
color histograms + filter response + GLCF	SVM	82.1 ± 2.2
	boosted trees	82.3 ± 2.3
filter response + GLCF	SVM	79.7 ± 2.5
	boosted trees	80.2 ± 2.2
color histograms + texton histogram	SVM	81.5 ± 1.9
	boosted trees	81.0 ± 2.4
color histograms + filter response	SVM	79.8 ± 2.3
	boosted trees	80.6 ± 1.8
color histograms + GLCF	SVM	78.5 ± 1.9
	boosted trees	77.6 ± 2.1
color histograms	SVM	73.4 ± 2.2
	boosted trees	74.6 ± 2.4

Table 1. Classification results based on descriptors with different feature combinations.

We perform a 10-fold cross validation using this dataset and calculate confidence intervals at a 95% confidence level. We experimented with both boosted trees and support vector machines as classifiers. 11 boosted trees of depth 4 are used for the Adaboost algorithm. For SVMs, we have chosen an RBF kernel where parameters C and γ are automatically tuned using the LIBSVM package [15].

Classification results are summarized in the first row of Table 1. To demonstrate the discriminative power exhibited by our descriptors, we have compared them with descriptors with other feature combinations. In addition to the three features introduced in Section 4.1, we have extracted for every gland unit two more features. We compute the mean and standard deviation of per-pixel responses to an oriented filter within a gland unit, and define a filter response feature as the concatenation of such mean and standard deviations for all filters within the filter bank used in Section 3. The other is a rotation-invariant gray-level co-occurrence feature (GLCF), which is in the same form as our texton co-occurrence feature except that textons are replaced with 10 evenly divided gray-scale intervals.

Results of typical feature combinations are given in Table 1. As expected, color histogram alone as a descriptor has the least accurate prediction. In combination with filter responses, GLCF, or the first-order texton statistics, it yields better results ranging from 77% to 82%. However, all these combinations have none or incomplete texton statistics, and therefore miss local patterns and structures crucial for malignant gland unit discrimination. They are outperformed by our descriptor which includes both first- and second-order texton statistics.

Comparisons with other recently adopted descriptors [6, 7] on our dataset have also shown that our descriptors have the most discriminative power. For example, multiwavelets [6], Gabor energy+entropy+magnitude [7], and fractal dimensions [7] achieved an accuracy of 68.1%, 75.3% and 79.0%, respectively.

6. CONCLUSIONS

We have introduced effective image segmentation and classification methods for automatic detection of malignant gland units in microscopic images. Both segmentation and classification methods are based on carefully designed feature descriptors, including color histograms and texton co-occurrence tables. The framework is designed to be flexible. Techniques for gland unit segmentation, feature extraction, and binary classification can all be tailored or substituted to meet the specific needs of the input data.

Acknowledgments

This work was partially supported by National Science Foundation (IIS 09-14631) and National Natural Science Foundation of China (60728204/F020404).

References

[1] B. Julesz, "Textons, the elements of texture perception and their interactions," *Nature(London)*, vol. 290, pp. 91–97, 1981.

[2] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *Intl. Journal Computer Vision*, vol. 43, no. 1, pp. 29–44, 2001.

[3] D.E. Pitts, B.S. Premkumar, A.G. Houston, R.J. Badain, and P. Troncosa, "Texture analysis of digitized prostate pathological cross-section," in *SPIE Proceedings of Medical Imaging: Image Processing*, 1993, vol. 1898, pp. 456–470.

[4] K. Jafari-Khouzani and H. Soltanian-Zadeh, "Multiwavelet grading of pathological images of prostate," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 6, pp. 697–704, 2003.

[5] S. Doyle, A. Madabhushi, M. Feldman, and J. Tomaszeweski, "A boosting cascade for automated detection of prostate cancer from digitized histology," in *Medical Image Computing and Computer Assisted Intervention*, 2006, vol. 9, pp. 504–511.

[6] A. Tabesh, M. Teverovskiy, H.Y. Pang, V.P. Kumar, D. Verbel, A. Kotsianti, and O. Saidi, "Multifeature prostate cancer diagnosis and gleason grading of histological images," *IEEE Transactions on Medical Imaging*, vol. 26, no. 10, pp. 1366–1378, 2007.

[7] P.-W. Huang and C.-H. Lee, "Automatic classification for pathological prostate images based on fractal analysis," *IEEE Transactions on Medical Imaging*, vol. 28, no. 7, pp. 1037–1050, 2009.

[8] Y. Zhu, S. Williams, and Zwiggelaar R., "Computer technology in detection and staging of prostate carcinoma: a review," *Medical Image Analysis*, vol. 10, no. 2, pp. 178–199, 2006.

[9] Y. Li, J. Sun, C.-K. Tang, and H.-Y. Shum, "Lazy snapping," *ACM Transaction on Graphics*, vol. 23, no. 3, pp. 303–308, 2004.

[10] R.E. Schapire, "The boosting approach to machine learning: An overview," in *MSRI Workshop on Nonlinear Estimation and Classification*, 2002.

[11] T. Xia, Q. Wu, C. Chen, and Y. Yu, "Lazy texture selection based on active learning," *The Visual Computer Journal*, vol. 26, no. 3, 2010.

[12] Y. Boykov and M.P. Jolly, "Interactive graph cuts for optimal boundary and region segmentation of objects in n-d images," in *Intl. Conf. Computer Vision*, 2001, vol. I, pp. 105–112.

[13] C. Rother, A. Blake, and V. Kolmogorov, "Grabcut–interactive foreground extraction using iterated graph cuts," *ACM Transaction on Graphics*, vol. 23, no. 3, pp. 309–314, 2004.

[14] R.Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.

[15] Chih-Chung Chang and Chih-Jen Lin, "LIBSVM: a library for support vector machines," 2001, Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.