

Automated Gland Segmentation and Classification for Gleason Grading of Prostate Tissue Images

Kien Nguyen and Anil K. Jain¹

Department of Computer Science and Engineering
Michigan State University
East Lansing, U.S.
nguye231@msu.edu, jain@cse.msu.edu

Ronald L. Allen

BioImage Inc
Sunnyvale, U.S.
Ron.Allen@bioimage.com

Abstract—The well-known Gleason grading method for an H&E prostatic carcinoma tissue image uses morphological features of histology patterns within a tissue slide to classify it into 5 grades. We have developed an automated gland segmentation and classification method that will be used for automated Gleason grading of a prostatic carcinoma tissue image. We demonstrate the performance of the proposed classification system for a three-class classification problem (benign, grade 3 carcinoma and grade 4 carcinoma) on a dataset containing 78 tissue images and achieve a classification accuracy of 88.84%. In comparison to the other segmentation-based methods, our approach combines the similarity of morphological patterns associated with a grade with the domain knowledge such as the appearance of nuclei and blue mucin for the grading task.

Keywords-Gleason grading, gland, carcinoma, prostate

I. INTRODUCTION

According to the 2005 United States Cancer Statistics (USCS) [1], prostate cancer is the most prevalent among the top ten cancer types. Among every 100,000 men, about 14,000 are afflicted with this type of cancer. Gleason grading [2] is a well-known method pathologists use to diagnose prostate cancer. It classifies prostatic carcinoma into five grades based on the structure of glandular patterns in the tissue. As the population grows older there is a significant increase in the volume of prostate screening tests. The shortage of pathologists in the face of the increased volume is a strong driver to develop an automated procedure for Gleason grading. Our approach to computing the Gleason grade of the tissue sample requires that we classify glands in the tissue as one of the grades. The classification will then be used to compute the final Gleason grade of the tissue sample. We have developed a classifier with three-classes corresponding to the patterns associated with benign (grades 1 and 2), grade 3 carcinoma and grade 4 carcinoma types. Grade 5 is not considered at this time. We first segment glands from the tissue pattern, extract gland features and

use these features to classify the pattern. An example of a benign tissue pattern is shown in Figure 1. A gland unit consists of the boundary formed by layers of epithelial cells which have the epithelial nuclei (blue dots) adjoining epithelial cytoplasm (purple regions) and the central white region, called lumen. The pink region surrounding the gland is called stroma. The benign patterns are characterized by (i) large and well-separated glands, (ii) large and branchy lumen components, and (iii) thick boundaries. Grade 3 carcinoma patterns, on the other hand, (Figure 2 (a)) have (i) smaller glands with thinner nuclei boundary and (ii) small and round lumen components which are sometimes filled with light blue mucin. In grade 4 carcinoma patterns (Figure 2 (b)) there is a loss of gland unit structure. Glands fuse with each other to create a mass of glands containing multiple lumen components.

Most of the previous studies related to prostate cancer grading simply used generic texture features to classify a tissue pattern as benign or cancerous. The tissue structure and the glands were not effectively analyzed. Diamond et al. [3] used texture features to classify regions of interest into stroma and prostatic carcinoma (PCa). However, large size of lumen area was used to detect benign regions. A classification accuracy of 79.3% was reported in classifying each region as benign, stroma or PCa. Doyle et al. [4] extracted 594 features at three different image scales for classifying a tissue pattern as cancerous or non-cancerous. Features used were based on first-order statistics, texture and Wavelet coefficients. An accuracy of 88% was achieved using AdaBoost.

Naik et al. [5] present a segmentation-based approach. Eight shape features for each of the lumen and the gland inner boundary are computed. An SVM classifier then determines whether the tissue is benign, a grade 3 carcinoma or a grade 4 carcinoma tissue. With a dataset containing 44 images, they reported the following results for the following three two-class classifications: grade 3 carcinoma versus benign (86.35%), grade 4 carcinoma versus benign (92.9%) and grade 3 carcinoma versus grade 4 carcinoma (95.19%). No results were presented for the three-class classification.

¹Anil Jain's research was partially supported by WCU (World Class University) program through the National Research Foundation of Korea funded by the Ministry of Education, Science and Technology (R31-2008-000-10008-0).

In addition to low level image features, we believe the domain knowledge about gland structures, the nuclei components, and mucin should be used in the classifiers to improve the sensitivity and selectivity of the classification. In high-grade carcinoma, the gland units degenerate and it is very difficult to separate individual glands. Therefore, we introduce a new terminology called *gland mass*, a generalization of gland unit, to represent areas of indistinguishable proliferating glands.

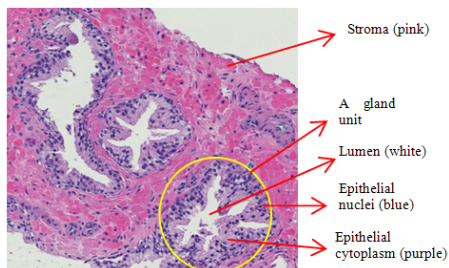


Figure 1. Benign pattern & gland structure.

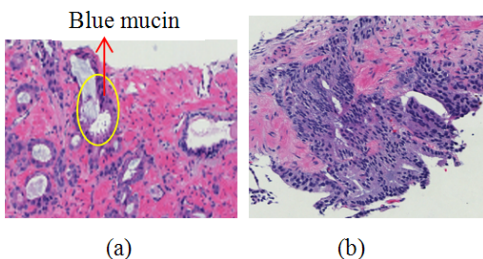


Figure 2. Carcinoma patterns (a) Grade 3; (b) grade 4.

II. GLAND SEGMENTATION

As the shape and position of the three main parts of a gland (nuclei, cytoplasm and lumen) are diverse, model-based approaches such as Active Shape Model [6] or Active Appearance Model [7] are not applicable. Hence, we first identify the three gland components based on color and then unify them to segment the full gland. We propose a three-stage algorithm consisting of: (1) pixel classification, (2) gland boundary extraction and (3) complete gland construction.

A. Pixel classification

We classify each pixel into 5 different classes: stroma (pink), lumen and non-tissue area (white), epithelial nuclei (dark blue), epithelial cytoplasm (purple) and blue mucin (light blue). To enhance the separation of different colors, we convert the original RGB color system to La*b* color system where the L channel corresponds to illumination

and a* and b* channels correspond to the color-opponent dimensions. We use the nearest neighbor classifier to classify each pixel (say at (x,y)) $L(x,y)$ into five classes $\{S,L,N,C,M\}$ where S,L,N,C,M correspond to stroma, lumen, nuclei, cytoplasm, and blue mucin, respectively. Approximately 200 samples from each class are used for training this classifier.

B. Gland boundary extraction

The gland boundary is modeled using the domain knowledge of how typical glands are structured. Since the gland boundary consists of nuclei intermixed with cytoplasm, we need to unify cytoplasm pixels and nuclei components of the same gland together to construct the gland boundary. Lumen is not involved in this step because it is located mostly in the center of the gland and does not touch the gland boundary. Therefore, we first find the connected components of nuclei pixels. Let W_i be a window of size $S_1 \times S_1$ centered at the nuclei component N_i . A cytoplasm pixel c_j and a nuclei component N_i are unified, denoted as $\text{Unified}(c_j, N_i)$ provided (i) $c_j \in W_i$ and (ii) $\text{card}\{L(x,y) = N, (x,y) \in W_i\} > T_1$ where T_1 is the nuclei density threshold. Two nuclei components N_i and N_j are unified if one of the following two conditions hold: (i) there exists a cytoplasm pixel c_k such that $\text{Unified}(c_k, N_i)$ and $\text{Unified}(c_k, N_j)$ hold true or (ii) there exists a chain of nuclei components $N_{i+1}, N_{i+2}, \dots, N_{j-1}$ such that $\text{Unified}(N_i, N_{i+1})$, $\text{Unified}(N_{i+1}, N_{i+2}), \dots$, and $\text{Unified}(N_{j-1}, N_j)$ hold true. The parameter S_1 is chosen basically based on the size of a typical nuclei and a typical distance between nuclei in benign glands. The results of this step are gland boundary segments, where each segment is created by nuclei and cytoplasm which are unified by the algorithm. Figure 4a shows the result of this step for the input image in Figure 1.

C. Complete gland construction

The components of nuclei-cytoplasm (NC) on the gland boundary obtained in the previous step may be separated. So, we use a lumen expansion procedure to unify the lumen with interior cytoplasm (cytoplasm between the boundary and the lumen) and the NC components.

For each lumen component L_m (created by lumen and blue mucin pixels), let (x_0, y_0) denote its centroid and let $B = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ denote the n points on its boundary. A point (x_i^j, y_i^j) is said to be the expanded version of (x_i, y_i) at iteration j if

$$x_i^j = (x_i - x_0) \times (1 + j \times \text{sf}) + x_0$$

$$y_i^j = (y_i - y_0) \times (1 + j \times \text{sf}) + y_0$$

where $j \geq 0$ and $j \leq \text{JMAX}_i$ and sf is the pre-defined scaling factor. The smaller the sf , the slower but finer the expansion. Let W_i^j centered at (x_i^j, y_i^j) with size $S_2 \times S_2$ be the neighborhood of that expanding point. The size

S_2 is selected based on the scaling factor sf so that W_i^j and W_i^{j+1} do not overlap. The expansion ends at iteration $JMAX_i$ when one of the following three conditions is met:

- (i) $\exists(x, y) \in W_i^{JMAX_i}$ such that $L(x, y) = S$.
- (ii) $\exists NC_k$ such that $W_i^{JMAX_i} \cap NC_k \neq \emptyset$.
- (iii) $d\{(x_0, y_0), (x_i^{JMAX_i}, y_i^{JMAX_i})\} > T_2$, where T_2 is a size constraint of the gland.

At each iteration $j < JMAX_i$, $\forall(x, y) \in W_i^j$ and $L(x, y) = C$, set $Unified\{(x, y), Lm\}$. This is to unify all cytoplasm pixels in the window to the lumen.

At iteration $j = JMAX_i$, $\forall NC_k$ such that $W_i^{JMAX_i} \cap NC_k \neq \emptyset$, set $Unified(NC_k, Lm)$. This is to unify nuclei-cytoplasm components to the lumen.

Figure 3 illustrates the lumen expansion procedure. Unlike [5] where each segment is always a gland unit, each output segment of the proposed method can be a gland unit or a mass of gland (usually in grade 4 carcinoma). However, the features we extract for each segment are applicable for both a gland mass and a gland unit. In the final result (Figure 4b), we apply a size constraint to remove very small segments.

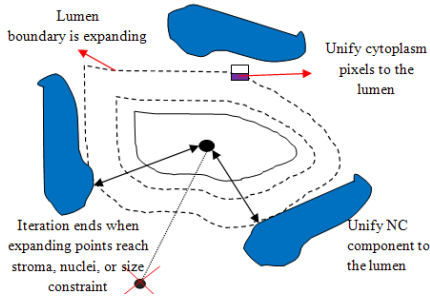


Figure 3. Lumen expansion procedure.

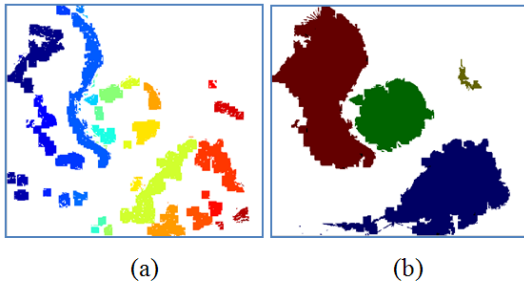


Figure 4. Gland segmentation. (a) gland boundary extraction; (b) complete gland construction.

III. GLAND FEATURE EXTRACTION

For each segment, we extract a set of 15 features (Table 1), average them over all the segments in the pattern to get a 15 dimensional feature vector for each pattern.

Table I
FEATURES EXTRACTED FROM SEGMENTS

Feature type	Feature description
Ten lumen features	Average, variance, max of lumen area; average, variance, max of lumen perimeter; average and variance of lumen roundedness; number of lumen and ratio of lumen area to total segment area.
Two nuclei features	Nuclei density and ratio of nuclei area to total segment area.
Two gland size features	Average and variance of the distances from the lumen center to the nuclei boundary.
Blue mucin feature	Ratio of blue mucin area to total segment area.

IV. RESULTS AND DISCUSSION

Dataset: The dataset was created by selecting sub-images of benign, grade 3 and grade 4 carcinoma patterns from 52 10x whole-slide images with average size of approximately 90,000 x 45,000 pixels. The average size of each sub-image is approximately 501 x 526 pixels. There are 30 sub-images of benign patterns, 28 sub-images of grade 3 carcinoma and 20 sub-images of grade 4 carcinoma. This grading was confirmed by a pathologist which makes the result of this experiment clinically acceptable. Figures 1 and 2 show example images from the dataset.

Experimental Setup: A number of classifiers (SVM, Multilayer Perceptron and K-Nearest Neighbor) were used in our experiments. The mean and variance of the classification performance over 10 different runs of 10-fold cross validation for the best parameters are reported. Weka [8] software was used for all the classifiers.

Three-class classification: For the three-class classification task using the complete feature set (15 features), the best result (87.5% accuracy with a variance of 1.11%) was obtained using SVM (linear kernel, $C = 10$). After performing a best-first-search feature selection to get a subset of 9 features (6 lumen features and nuclei density, blue mucin ratio, gland size variance features) the best result was 88.4% accuracy with 6.2% variance using Multilayer Perceptron (16 hidden nodes).

Classifier fusion: Different combinations of classifiers were tested with the selected subset of features. Applying the max-rule to the scores of SVM (linear kernel, $C = 10$) and Multilayer Perceptron (16 hidden nodes) converted to probabilities, resulted in the best performance with 88.8% accuracy and 1.8% variance.

Baselines: To compare our classification results with [4] and [5], we also solved the four two-class classification problems. The best accuracies (variance) we obtained were:

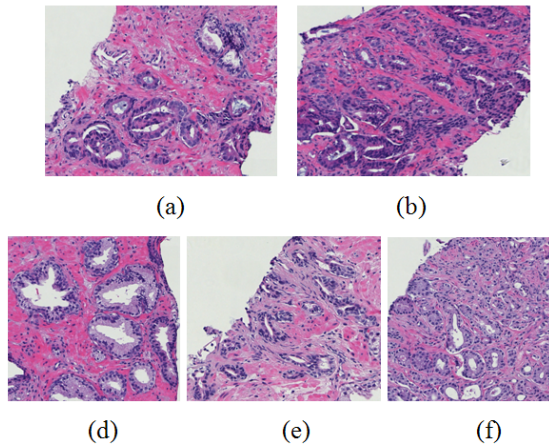


Figure 5. Classification examples. Misclassifications: (a) Grade 3 is classified as grade 4; (b) grade 4 is classified as grade 3. Correct classifications: (d) Benign; (e) grade 3 carcinoma; (f) grade 4 carcinoma.

97.75% (1.35%) for Benign vs Grade 3, 94% (3.55%) for Benign vs Grade 4, 87.3% (0.43%) for Grade 3 vs Grade 4 and 98.58% (0.16%) for Benign vs Carcinoma (grade 3 and grade 4) with 100% sensitivity and 96.67% specificity.

These results show that the proposed algorithm does very well in separating cancer versus non-cancer tissues, which is a crucial decision in diagnostic application. However, the discrimination between grade 3 versus grade 4 is not as good. This is because the dataset contains some patterns that are at the boundary of grade 3 and grade 4 with some glands standing alone and some glands fused with others. These tissues are in the intermediate stage of growth from grade 3 to grade 4 which are very common and cause difficulty for even pathologists. Two examples of misclassifications between these two classes are shown in Figures 5(a) and (b).

Features selected using best-first search confirm that nuclei and blue mucin are important parts of a gland and need to be incorporated in gland segments. In cases where there is a need for a deeper investigation in the prostate tissue, nuclei can also be used for recognizing special structures such as paraganglia, seminal vesicles, perineural indentation and eosinophilic crystalloids [9].

Three examples of correct classifications are also presented in Figures 5(d), (e), and (f).

V. SUMMARY

Compared to non-segmentation-based methods, our approach has two main advantages: (i) the extracted glands not only provide cancer grading information but also facilitate other tasks such as gland retrieval, (i.e. when a pathologist has difficulty in grading a gland, he can search for a similar gland that has been graded), (ii) pathologists can obtain

several biopsies of the same region in the prostate and create images by different staining methods such as H&E and IHC to improve the grading result. In this case, segmented glands can serve as landmarks to register those images together. Other features based on distribution of nuclei and lumen within a gland will be explored more thoroughly to discriminate grade 3 and grade 4. We also plan to create a different feature vector for the pattern by concatenating features of all the glands instead of averaging features to avoid any information loss. Finally, the classification information will be integrated across the whole slide to grade the tissue specimen.

REFERENCES

- [1] U.S. Cancer Statistics Working Group, *United States Cancer Statistics: 1999-2005 Incidence and Mortality Web-based Report*, Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute, 2009, Available at: www.cdc.gov/uscs.
- [2] D. F. Gleason, *Histologic grading and clinical staging of prostatic carcinoma*, in *Urologic Pathology: The Prostate*, M. Tannenbaum, Ed. Philadelphia, PA: Lea and Febiger, pp. 171-198, 1977.
- [3] J. Diamond, N. Anderson, P. Bartels, R. Montironi, and P. Hamilton, *The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia*, *Human Pathol.*, vol. 35, pp. 1121-1131, 2004.
- [4] S. Doyle, A. Madabhushi, Feldman, M., and Tomaszewski, J., *A boosting cascade for automated detection of prostate cancer from digitized histology*, *MICCAI*, pp. 504-511, 2006.
- [5] S. Naik, S. Doyle, A. Madabhushi, et al, *Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information*, *MIAAB Workshop*, 2007.
- [6] T. Cootes, C. J. Taylor, D. H. Cooper, and J. Graham, *Active shape models - their training and application*, *Computer Vision and Image Understanding*, 61(1):38-59, January 1995.
- [7] T. F. Cootes, G. J. Edwards, and C. J. Taylor, *Active appearance models*, in *5th European Conference on Computer Vision*, vol. 2, pp. 484+, 1998.
- [8] Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten, *The WEKA Data Mining Software: An Update*, *SIGKDD Explorations*, vol. 11, Issue 1.
- [9] D. M. Ramnani, *Normal Prostate and Seminal Vesicles*, *Web-Pathology, visual survey of surgical pathology*, 2008, Available at: <http://www.webpathology.com/case.asp?case=14>.