

Structure and Context in Prostatic Gland Segmentation and Classification

Kien Nguyen¹, Anindya Sarkar², and Anil K. Jain¹

¹ Michigan State University, East Lansing, MI 48824, USA

² Ventana Medical Systems, Inc., Sunnyvale, CA 94085, USA
{nguye231,jain}@cse.msu.edu, anindya.sarkar@ventana.roche.com

Abstract. A novel gland segmentation and classification scheme applied to an H&E histology image of the prostate tissue is proposed. For gland segmentation, we associate appropriate nuclei objects with each lumen object to create a gland segment. We further extract 22 features to describe the structural information and contextual information for each segment. These features are used to classify a gland segment into one of the three classes: artifact, normal gland and cancer gland. On a dataset of 48 images at 5× magnification (which includes 525 artifacts, 931 normal glands and 1,375 cancer glands), we achieved the following classification accuracies: 93% for artifacts v. true glands; 79% for normal v. cancer glands, and 77% for discriminating all three classes. The proposed method outperforms state of the art methods in terms of segmentation and classification accuracies and computational efficiency.

1 Introduction

In detecting prostate cancer on a digitized tissue slide, the pathologist relies on: (i) structural information; glands in a cancer region (cancer glands) appear to have structural properties (e.g. nuclei abundance, lumen size) different from glands in a normal region (normal glands) and (ii) contextual information; cancer glands typically cluster into groups and are of similar shape and size¹, while shape and size of normal glands vary widely. These two sources of information can be observed in Fig. 1b. Hence, a reasonable approach to assist a pathologist in finding cancer regions includes segmenting out glandular regions, examining their structural and contextual information and finally classifying them.

The cancer detection problem in prostate tissue images has been studied in the literature. Monaco et al. [2] (Table 1) segmented glands and classified individual glands into normal or cancer by (i) using gland size feature to assign initial gland labels and (ii) applying a probabilistic pairwise Markov model (PPMM) to update gland labels. On the other hand, Nguyen et al. [3] and Doyle et al. [4] detected cancer regions by classifying individual image patches and image pixels, respectively, by using a combination of cytological and textural features [3],

¹ It was also mentioned in [1] that cancer glands tend to appear close to other cancer glands, which is a biological motivation for this contextual information.

or textural features alone [4]. Unlike [2,3,4], other studies focused on the tissue image classification problem, i.e. they classified every prostatic tissue image (which is a region in the whole slide tissue) into normal or cancer or into different cancer grades. This problem was addressed by either segmentation-based approaches [5,6] (Table 1) or texture-based approaches [7,8,9].

Table 1. Studies on gland segmentation and gland feature extraction reported in the literature. Note that each study used a different database.

Study	Segmentation algorithm	Gland feature	Use of context	Final goal	Dataset
Monaco et al. [2]	Region growing	Gland size	PPMM method	Detect cancer regions	40 images at 1.23×
Peng et al. [5]	Region growing	Gland size	No	Classify tissue images	62 images at 100×
Naik et al. [6]	Level set	Shape features of lumen and gland	No	Classify tissue images	44 images at 40×
Proposed method	Nuclei lumen association	Structural features	Contextual features	Classify glands	48 images at 5×

Similar to [6], we also address artifacts² (Fig. 1c) in this paper by including them into the gland classification step, leading to a three-class classification problem: artifact, normal gland and cancer gland. The contributions of the paper include: (i) The proposed segmentation algorithm is computationally efficient and is able to successfully segment out appropriate gland regions, (ii) we explore features related to nuclei which are distinctive for classification, (iii) we introduce robust features to capture contextual information of a gland.

2 Gland Segmentation

Gland Structure: A gland consists of nuclei, cytoplasm and lumen (Fig. 1c, [6], [9]). Hence, a gland segmentation algorithm should capture these components in the results. We achieve this goal by employing the following two steps.

Tissue Component Identification: For each image, we perform k-means clustering algorithm ($k = 4$) in the *RGB* color space of 10,000 randomly selected pixels (Fig. 1d) to find 4 cluster centers. By finding the nearest cluster center, each pixel in the image is assigned a label corresponding to one of the 4 tissue components (stroma, nuclei, cytoplasm and lumen). We apply a connected component algorithm [10] on nuclei pixels and lumen pixels to generate nuclei objects and lumen objects, respectively, which are used for segmentation. Since the colors of nuclei and lumen are quite salient, the k-means algorithm on *RGB* space is sufficient to identify them despite the intensity variation among images.

² The term “artifact” denotes broken tissue areas, and has been used in [9].

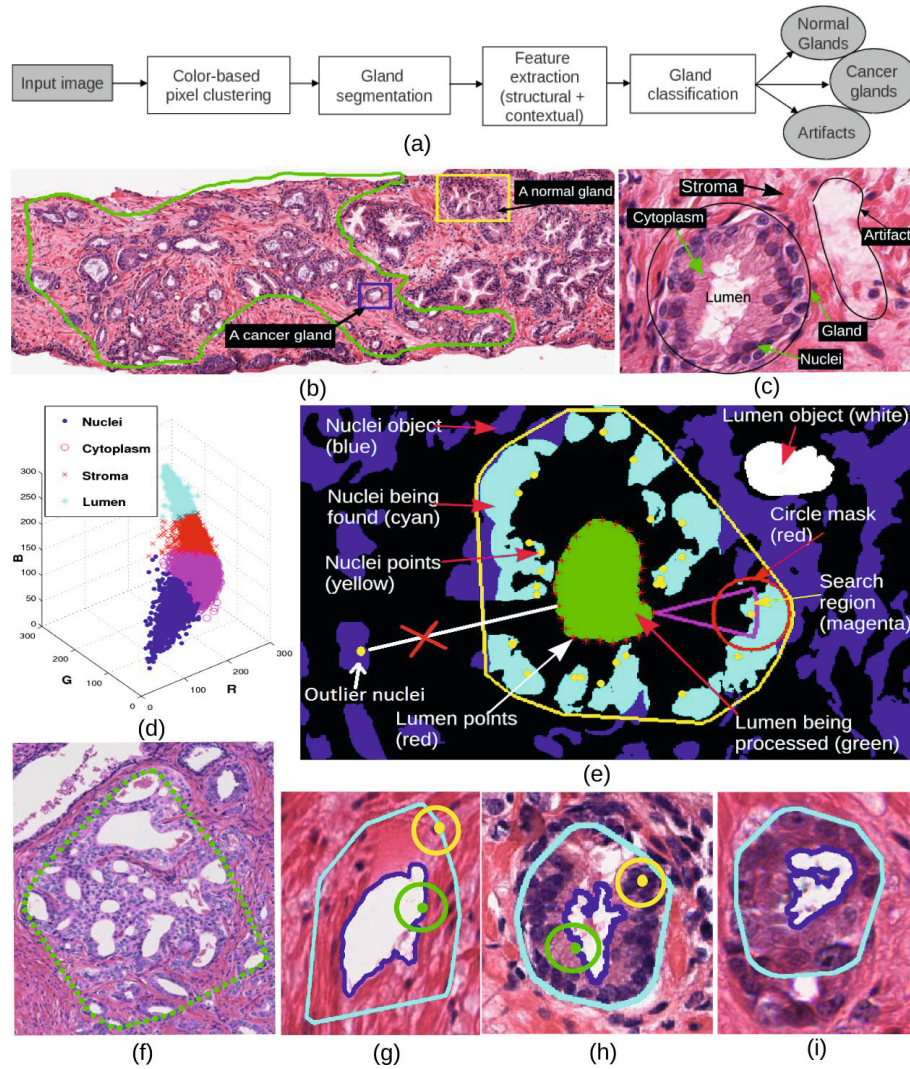


Fig. 1. The proposed method for gland segmentation and classification. (a) Flowchart. (b) Input image showing the cancer glands in a cancer region annotated by a pathologist (**green contour**); normal glands are present in the region outside the green contour. (c) A gland with basic components (nuclei, cytoplasm and lumen) and an artifact. (d) Clustering result of the image pixels in the *RGB* color space. (e) The gland segmentation process, where the segmentation result is depicted by a convex hull enclosing the detected nuclei. (f) Glands are assigned into a group (dotted contour) to compute contextual features. Segmentation results for three classes of interest: (g) Artifact, (h) normal gland and (i) cancer gland. Green and yellow circles in (g), (h) denote the neighborhood of a lumen point and a nuclei point, respectively.

Segmentation Algorithm. The proposed algorithm (Fig. 1e), which is referred to as nuclei-lumen association (NLA) algorithm, associates appropriate nuclei with each lumen to create a gland segment. Nuclei are searched along the normal direction of the lumen boundary contour. The algorithm has three steps:

1. Given n points on the lumen boundary (n may vary from 30 to 3000 depending on the lumen size), by considering the trade-off between a sparse (for computational efficiency) and dense set (adequate search coverage), we sample $n/3$ points uniformly, and refer to them as lumen points.
2. A search region, of a conical shape, centered at each lumen point, is expanded to find nuclei. A circle mask is used to limit nuclei regions to be merged to the gland. This step is repeated for all lumen points.
3. A pruning procedure, based on the median absolute deviation (MAD), is applied to remove outlier nuclei and generate smoother segmentation boundary³. Besides gland segments, the algorithm also produces a set of points located at the detected nuclei, referred to as the nuclei point set. The nuclei point set and lumen point set are used for feature extraction. Although some non-gland segments created by artifacts are present (Fig. 1g), we do not detect them at this step. Instead, we will identify them in the classification procedure.

3 Gland Classification

The differences in structures of the three classes (artifact, normal gland, cancer gland) are as follows. An artifact (Fig. 1g) does not have cytoplasm surrounding the lumen and has very few associated nuclei. Nuclei on the boundary of a normal gland (Fig. 1h) are more abundant and have darker blue color than a cancer gland (Fig. 1i). Lumina of cancer glands commonly appear more circular and smaller than normal glands (Fig. 1b). Based on these differences, we extract the following four sets of *structural features*, including 19 features, for each gland:

1. **Set 1 (8 nuclei features):** For each nuclei point (NP), we compute the mean (μ) and standard deviation (σ) of L, a, b color bands⁴ in the neighborhood of NP (Ω_{NP}). In addition, we compute μ and σ of percentage of Ω_{NP} that contains nuclei pixels, i.e. nuclei abundance on the gland boundary and its variation. Ω_{NP} is a circular region centered at NP (yellow circle in Fig. 1g, 1h), and has a radius R_{NP} . Since cancer glands usually have one nuclei layer (NL) on the boundary, while normal glands have more than one NL (mostly 2 NL), we choose $R_{NP} = 10$ pixels, which corresponds to the 2-NL thickness (the diameter of a nucleus is

³ $MAD = \text{median}_i(|d_i - \text{median}_j(d_j)|)$, where d is the distance between a lumen point and a nuclei point. A nuclei point k with $d_k > 3\sigma + \text{mean}(d)$ ($\sigma = 1.48MAD$) is considered an outlier and being discarded. All pixel distances (used in the paper) can be converted to physical distances when the magnification is known.

⁴ The Lab color space, which separates luminance and chrominance, is suitable to describe the color intensity of the tissue components.

approximately 5 pixels in $5\times$ images). Hence, Ω_{NP} is sufficient to capture most nuclei of the gland, while excluding nuclei of neighbor glands.

2. Set 2 (6 cytoplasm features): For each lumen point (LP), we compute μ and σ of L, a, b color bands in the neighborhood of LP (Ω_{LP}). Ω_{LP} is a circular region (green circle in Fig. 1g, 1h), which is centered at LP and excludes lumen area. The radius of Ω_{LP} is the distance between LP and the corresponding NP. Hence, in a true gland segment, Ω_{LP} mostly contains cytoplasm.

3. Set 3 (3 lumen shape features): Area, solidity (ratio of the lumen area to its convex hull area) and circularity ($(4\pi\text{area})/\text{perimeter}^2$) of the lumen.

4. Set 4 (2 global features): μ and σ of the distance between a LP and a NP.

To explore contextual information, we first assign gland segments into groups (Fig. 1f) by using the connected component (CC) algorithm in graph theory. Let $\{Lu^i\}_{i=1}^n$ denote the n lumen objects used to represent n gland segments, and let $(Lu_{x_o}^i, Lu_{y_o}^i)$ denote the centroid of Lu^i . A graph is built where each node is a gland. If $\|(Lu_{x_o}^i - Lu_{x_o}^j, Lu_{y_o}^i - Lu_{y_o}^j)\| < t_d$,⁵ there is an edge connecting Lu^i and Lu^j . Each CC is considered a group of gland. So groups are disjoint and the grouping is unique. Once groups are formed, we compute the following 3 *contextual features* for each gland segment Lu^i (which belongs to group O):

1. **Neighborhood crowdedness:** $|O|$ or the no. of elements in O .

2. **Shape similarity:** $\frac{1}{|O|} \sum_{j=1}^{|O|} \|LV_i - LV_j\|$, where LV denotes the 3-dimensional lumen shape feature vector described above.

3. **Size similarity:** $\frac{1}{|O|} \sum_{j=1}^{|O|} \frac{\min(|Lu^i|, |Lu^j|)}{\max(|Lu^i|, |Lu^j|)}$, where $|Lu|$ denotes the lumen size.

Finally, each gland segment can be represented by a full feature vector of dimensionality 22 (19 + 3). We use a SVM classifier (linear kernel, $C = 1$) with this feature vector to classify the gland segment⁶.

4 Experiments

a. Data Set: The dataset includes 48 images at $5\times$ magnification (average image size is $900 \times 1,500$ pixels), which come from 20 patients. Glands in images of the same patient still have very large variability in structures. Given the pathologist's annotation on each image, we manually label 525 artifacts, 931 normal glands and 1,375 cancer glands to form the (ground truth) gland dataset. We also implemented the methods in [2] and [6] to compare them with the proposed method. Since all three methods perform segmentation by starting at the same

⁵ We choose $t_d = 65$, which minimizes the cross validation error in classifying the 3 different gland types. However, when minimizing the training error on an independent training set, we also obtain $t_d = 65$.

⁶ The code for gland segmentation and feature extraction, along with the details of the experiments can be found at www.cse.msu.edu/~nguye231/GlandSegClass.html

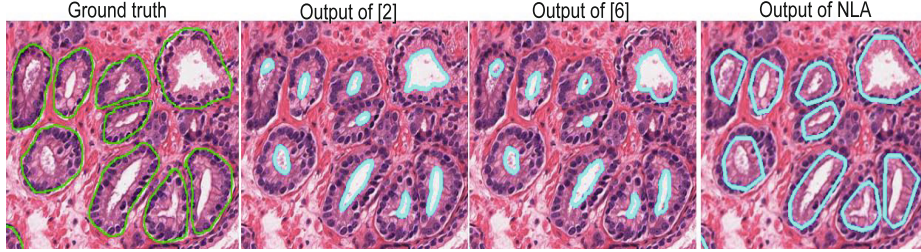


Fig. 2. Comparison of the NLA algorithm with [2] and [6] for gland segmentation. Glands are segmented more completely by the NLA algorithm

Table 2. Gland classification accuracies (s.d.) for the method in [6], method in [2], SVM-SF, PPMM-SF, PPMM-SCF, and SVM-SCF by cross validation

Classification problem	Method in [6]	Method in [2]	SVM-SF	PPMM-SF	PPMM-SCF	SVM-SCF
Artifact v. true gland	0.78 (0.09)	-	0.93 (0.03)	0.93 (0.06)	0.93 (0.04)	0.93 (0.04)
Normal v. cancer	0.67 (0.13)	0.68 (0.13)	0.75 (0.07)	0.73 (0.11)	0.75 (0.11)	0.79 (0.08)
All three classes	0.54 (0.12)	-	0.74 (0.06)	0.75 (0.10)	0.73 (0.08)	0.77 (0.07)

lumen objects (identified in section 2), and use the same ground truth (which is not affected by lumen objects), the comparison is unbiased.

b. Gland Segmentation Evaluation: We manually select 309 glands whose boundaries are well-defined, and create segmentation ground truth for each gland (G_i^0) by outlining its area. We use the Jaccard Index (JI) to evaluate the output of a segmentation algorithm. Given a gland segment G_i^m produced by the algorithm m for the i^{th} gland (with ground truth G_i^0), JI is computed as $J(G_i^0, G_i^m) = |G_i^0 \cap G_i^m| / |G_i^0 \cup G_i^m|$. Higher JI values (range is [0,1]) indicate better segmentation results. *The average JI value per gland segment obtained by the algorithms in [2], [6] and the proposed NLA algorithm are 0.31, 0.43 and 0.66, respectively.* Since the NLA algorithm aims at detecting nuclei surrounding the lumen (while the algorithms in [2] and [6] mostly detect lumen and cytoplasm), it segments more complete gland regions than [2] and [6] (Fig. 2). The computational complexity of the three segmentation algorithms is measured by their running time on a 369×1213 image, containing 64 glands. *The total computation time of [2], [6] and the NLA algorithm for the image are 242.0s, 256.5s and 2.7s, respectively (all algorithms, implemented in Matlab, were run on a 2.93GHz machine with 16GB memory).* While the NLA algorithm processes pixels by only considering their labels (see Fig. 1d and section 2), the algorithms in [2] and [6] perform complicated operations on the grayscale intensities of the pixels. More precisely, the level set algorithm [6] needs to iteratively evolve a zero level curve by minimizing both the internal energy and the external energy of the curve,

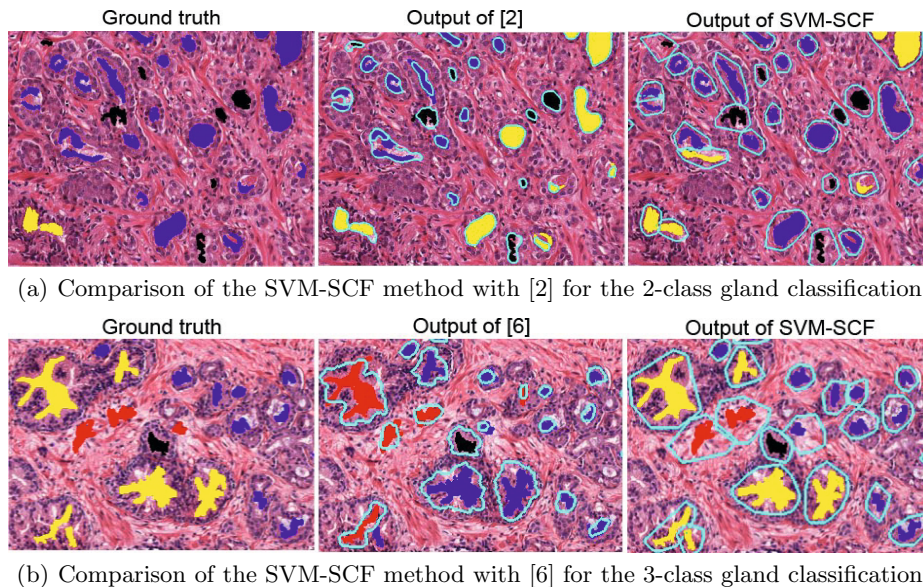


Fig. 3. Gland classification comparison. Cyan contours denote segmentation results, and color of the lumen corresponds to gland label (black, red, yellow and blue denote non-labeled glands, artifacts, normal glands and cancer glands, respectively)

and the region growing algorithm [2] needs to, at every step, add one pixel to the growing region and recompute the boundary strength of the region. Hence, the NLA algorithm is better than [2] and [6] in both accuracy and complexity.

c. Gland Classification Evaluation: We perform a 10-fold cross validation on the gland dataset, and report the average classification accuracy. First, we solve the two 2-class classification problems, i.e. (i) artifacts v. true glands and (ii) normal glands v. cancer glands. Next, we perform the 3-class classification by combining the previous two 2-class classification problems in a hierarchical fashion. In Table 2, besides the methods in [6], [2] and the proposed method (denoted by SVM-SCF, i.e. applying SVM classifier on the structural-contextual features (SCF)), we also report the results of the SVM-SF method (applying SVM classifier on the structural features (SF)), the results of the PPMM-SF method (applying the PPMM (Table 1) on the SF), and the results of the PPMM-SCF method (applying the PPMM on the SCF). Since artifacts were not addressed in [2], we only report normal v. cancer result for this method.

From Table 2, we can see that: (i) The SVM-SCF method obtains the highest accuracy, (ii) the superior performance of PPMM-SF over [2] shows that the SF are robust, and (iii) it is better to use SCF with SVM classifier than with PPMM. A drawback of the PPMM is that it requires a density estimation $p(y|x)$ (y is the feature vector, and x is the class label), which is difficult when y is high dimensional like the SF or SCF. Here, we address this “curse of dimensionality”

problem (which was not discussed in [2]) by testing three different methods to estimate $p(y|x)$, i.e. (i) parzen window density estimator (PWDS), Naive Bayes model with assumption that each feature follows a (ii) Gaussian distribution, and (iii) Gamma distribution. The best results among them are obtained by PWDS, which are reported in Table 2 (columns 5 and 6).

The Role of Segmentation: To evaluate how segmentation results affect classification results, we extract the proposed SF of the gland segments resulted from the level set algorithm in [6], and perform classification using these features. *The results obtained are lower than those when using SF with the proposed NLA algorithm (column 4 of Table 2). This shows that good segmentation results are necessary for good classification results, as we discuss in section 4b that the NLA algorithm performs better than the level set algorithm in [6].*

Feature Weight: By applying linear SVM to compute weights for all 22 features, we observe that the nuclei abundance (on the gland boundary) feature receives the highest weight. This shows the important role of nuclei in classification.

Classifier Selection: We also conduct experiments using other classifiers such as: Neural Network, KNN, and Adaboost. However, the accuracies obtained by those classifiers were lower than SVM. Moreover, performing classification by a linear SVM is fast while the time-consuming training part is an offline process.

5 Conclusions and Future Work

We have presented a novel method to segment and classify glands in a prostate histology image. The proposed method outperforms state of the art methods for both gland segmentation and classification. The detected cancer glands can facilitate the Gleason scoring task performed by either a pathologist or an automated system. Since the proposed segmentation algorithm relies on lumen, it is not applicable to regions with occluded lumina. However, occluded lumina are seldom observed in prostate tissue images. In our future work, we will address this limitation. Moreover, we will improve the accuracy of the normal v. cancer glands classification by further research on contextual information, and textural information of the glands.

References

1. Kumar, V., Abbas, A., Fausto, N.: Robbins and Cotran Pathologic Basis of Disease. Saunders (2004)
2. Monaco, J., Tomaszewski, J., Feldman, M., Hagemann, I., Moradi, et al.: High-throughput detection of prostate cancer in histological sections using probabilistic pairwise markov models. *Medical Image Analysis* 14, 617–629 (2010)
3. Nguyen, K., Jain, A., Sabata, B.: Prostate cancer detection: Fusion of cytological and textural features. *Journal of Pathology Informatics* 2, 2–3 (2011)

4. Doyle, S., Feldman, M., Tomaszewski, J., Madabhushi, A.: A boosted Bayesian multi-resolution classifier for prostate cancer detection from digitized needle biopsies. *IEEE Trans. Biomed. Eng.* 59, 1205–1218 (2012)
5. Peng, Y., Jiang, Y., Eisengart, L., et al.: Segmentation of prostatic glands in histology images. In: *IEEE Int. Symp. on Biomedical Imaging*, pp. 2091–2094 (2011)
6. Naik, S., Doyle, S., Madabhushi, A., Tomaszewski, J., et al.: Automated gland segmentation and Gleason grading of prostate histology by integrating low-, high-level and domain specific information. In: *MIAAB Workshop*, Piscataway, NJ (2007)
7. Yoon, H.J., Li, C.C., Christudass, C., Veltri, R., et al.: Cardinal multiridgelet-based prostate cancer histological image classification for Gleason grading. In: *IEEE Int. Conf. Bioinformatics and Biomedicine*, pp. 315–320 (2011)
8. Khouzani, K.J., Zadeh, H.S.: Multiwavelet grading of pathological images of prostate. *IEEE Trans. Biomed. Eng.* 50, 697–704 (2003)
9. Tabesh, A., Teverovskiy, M., et al.: Multifeature prostate cancer diagnosis and gleason grading of histological images. *IEEE Trans. Med. Img.* 26, 1366–1378 (2007)
10. Haralick, R., Shapiro, L.: *Computer and robot vision*, vol. 2, pp. 28–48. Addison-Wesley (1992)