

# Tree-structured grading of pathological images of prostate

Reza Farjam,<sup>a</sup> Hamid Soltanian-Zadeh,<sup>a,b</sup> Reza A. Zoroofi,<sup>a</sup> Kourosh Jafari-Khouzani<sup>b,c</sup>

<sup>a</sup>Control and Intelligent Processing Center of Excellence,

Electrical and Computer Engineering Department, University of Tehran, Tehran 14395-515, Iran

<sup>b</sup>Image Analysis Lab, Radiology Department, Henry Ford Health System, Detroit, MI 48202, USA

<sup>c</sup>Computer Science Department, Wayne State University, Detroit, MI 48202, USA

E-mails: reza\_dsp@yahoo.com, hamids@rad.hfh.edu, zoroofi@ut.ac.ir, kjafari@rad.hfh.edu

## ABSTRACT

This paper presents a new algorithm for Gleason grading of pathological images of prostate. Structural features of the glands are extracted and used in a tree-structured (TS) algorithm to classify the images into five Gleason grades of 1 to 5. In this algorithm the image is first segmented to locate the glandular regions using texture features and a *K*-means clustering algorithm. The glands are then labeled from the glandular regions. In each stage of the proposed TS algorithm, shape and intensity-based features of the glands are extracted and used in a linear classifier to classify the image into two groups. Despite some proposed methods in the literature which use only texture features, this technique uses the features like roundness and shape distribution, which are related to the structure of the glands in each grade and are independent of the magnification. The proposed method is therefore robust to illumination and magnification variations. To evaluate the performance of the proposed method, we use two datasets. Data set 1 contains 91 images with similar magnifications and illuminations. Data set 2 contains 199 images with different magnifications and illuminations. Using leave-one-out technique, we achieve 95% and 85% accuracy for dataset 1 and 2, respectively.

**Keywords:** Tree-structured classification, Gleason grading, prostate cancer, texture analysis.

## 1. INTRODUCTION

Cancer is the second common cause of death after cardiovascular diseases [1]. Prostate cancer is the most prevalent cancer diagnosed among men over the age of 50 with 25% of patients' death from the disease [2]. Early detection of cancer is very important for survivable treatment planning of prostate cancer. However, 33% of patients have advanced diseases on initial diagnosis [3]. In prostate cancer prognosis, patient is first examined using clinical tests such as measuring prostate specific antigen, digital rectum examination, CT, MRI, or trans-rectal ultrasound scan [4]-[7]. If a cancer is then suspected, biopsy specimens of prostate tissue are taken, stained and observed by pathologists for any evidence of cancer. In the case of cancer, pathologists use a grading system to determine the level malignancy of cancer by assigning each cancerous tissue a grade describing the aggressiveness of the disease.

The most common grading system currently used by pathologists is Gleason grading system (GGS) [8]. In this system, each cancerous tissue is assigned one of five grades 1 to 5 with higher grades indicating higher malignancy. In this system, the architecture of the glands determines the malignancy of cancer. In low grades where cancer is not advanced yet, the glands are differentiated, approximately of the same size and equally spaced. In high grades where cancer is advanced, the glands are erupted and merged. Thus, they are not as well differentiated as in low grades.

Histological grading of the PIP is very important for treatment planning of prostate cancer. It is also very subjective due to inter- and intra-observer differences among the pathologists. Furthermore, it is a time-consuming and in some cases a difficult process. Hence, automatic grading of PIP is of interest.

So far, many attempts have been directed towards analysis of microscopic images for cancer [9]-[14]. An artificial neural network ensemble-based system is proposed in [9] for automatic identification of lung cancer cells from the biopsy images. The proposed system is a two-level ensemble architecture such that the first level is used to judge whether a cell is normal or cancerous and the second is used to deal with the cells that are judged as cancerous cells. Automatic method

for grading of the urinary bladder tumors is also presented in [10]. In this paper, 36 morphological and textural features describing cell nuclei are used as the inputs of an artificial neural network. The system assigns each tumor one of three grades. An automatic system for the analysis of cells' nucleus in the biopsy cancerous mammary tissues is presented in [11]. In this article, biopsy images are enhanced and segmented using morphological transformations. Ultimate erosion is then used to separate cells' nucleus in contact.

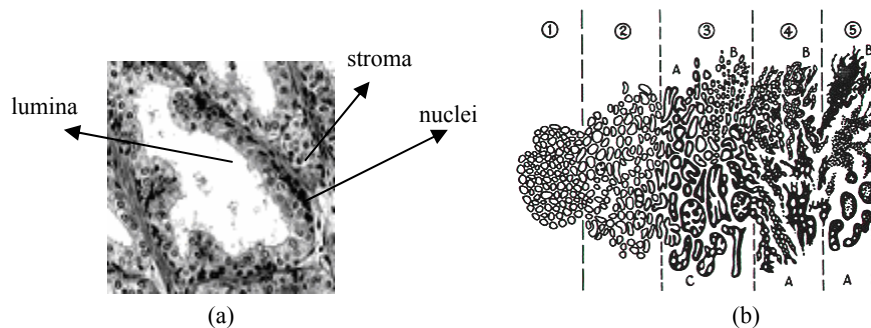
To the best of our knowledge, few attempts have been made towards analysis of the prostate biopsy images. Stotzka *et al.* [12] proposed a method to distinguish between the moderately and poorly differentiated samples, but did not consider the benign cases. Their work is based on extracting textural features that describe the arrangement of nuclei in the image. The nuclear roundness factor analysis (NRF) is proposed in [13] to predict the behavior of the low-grade samples. Since this technique requires manual nuclear contour tracing, it is time-consuming and tedious. Furthermore, the NRF analysis cannot be applied to high-grade samples because the monotonic relationship between NRF and grade is lost in high grades. In [14], the energy and entropy features of the multi-wavelet coefficients of the images are computed. The most discriminative features are then selected using a simulated annealing algorithm. Using a  $k$ -nearest neighbor ( $k$ -NN) classifier the samples are graded. In this work, it is assumed that the images have similar illuminations and magnifications.

In our proposed approach, a Tree-Structure (TS) algorithm is used for Gleason grading of PIP. First, a texture-based method is used to segment the glandular regions of the image. The segmented regions and the texture features are then given as inputs to the TS algorithm. This algorithm contains five branches. In each branch of the proposed tree-structured (TS) algorithm, shape and intensity-based features of the glands are extracted and used in a linear classifier to classify the image into two groups. Despite some proposed methods in the literature which use only texture features, this technique uses the features like roundness, shape distribution, etc., which are related to the structure of the glands in each grade and are independent of the magnification. The proposed method is thus robust to illumination and magnification variations. To compare with other techniques, the specimens are also categorized using the texture features proposed in [14]. Experimental results show the efficiency of the proposed method.

The rest of the paper is organized as follows. In Section 2, we briefly discuss Gleason grading system. In Section 3, extraction of texture features and segmentation of glandular regions are presented. Five different stages of the TS algorithm are explained in Section 4. Experimental results are presented in Section 5. We conclude in Sections 6.

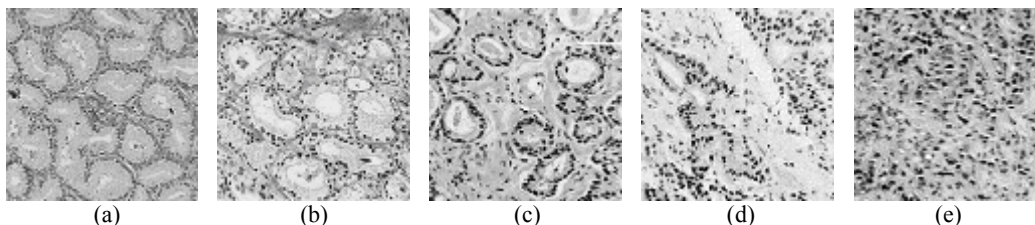
## 2. GLEASON GRADING SYSTEM

Gleason grading is based on the structure of the glands. Fig. 1(a) illustrates the structure of a normal prostate gland. As shown, each gland consists of three main parts: lumina, stroma, and nucleus. Nuclei are the dark areas with low homogeneity and high variance, while stroma and lumina are the brighter regions with high homogeneity and low variance. In a normal gland, lumina is located at the center and usually has an irregular shape. Also, it is the brightest region in the gland. Stroma, surrounds lumina with nuclei floating in it. In a normal gland, these regions are arranged in a way that the gland is a round mass.



**Figure 1:** (a) Structure of a normal prostate gland. Each gland consists of three main parts: lumina, stroma and nucleus. Nuclei are the dark areas with low homogeneity and high variance while stroma and lumina are the bright regions with high homogeneity and low variance. (b) Conceptual diagram of the Gleason grading system. In this system the cancerous specimens are assigned one of five grades from 1 to 5 based on the aggressiveness of cancer.

A normal prostate tissue consists of a large number of glands of approximately the same size and equally spaced. When cancer infects the prostate tissue, it affects the structure of the glands and the distances between them. In Fig. 1(b), a conceptual diagram of the Gleason grading system is presented. As shown, the architecture of the glands determines the grade. In grade 1 and 2, the glands are well differentiated. In grade 1, they are approximately of the same size and equally spaced. But, in grade 2 they have difference sizes. Also, the distances between them are slightly increased. In grade 3, the glands are moderately differentiated. Here, the distances between the glands are too high. In grade 4, and 5, the glands are erupted and merged, and they are not differentiated. In grade 4, nuclei are still connected to each other, but in grade 5 they are separated and floating in the stroma irregularly. In Fig. 2, examples of the five grades of the GGS are presented.



**Figure 2:** Example of the five grades of the Gleason grading system. (a-e). Gleason grades 1 to 5, respectively.

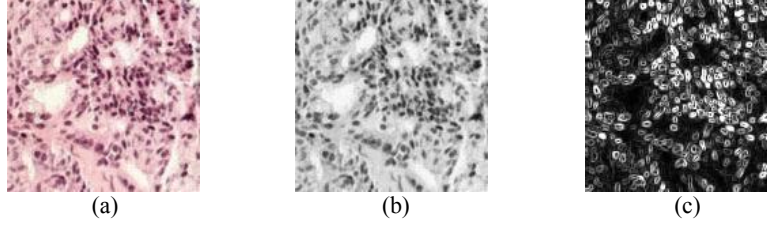
Urologists apply suitable treatments for prostate cancer based on the malignancy of cancer. In low-grade (benign) specimens which cancer is still restricted in the prostate tissue, urologists often remove the prostate by surgical operation (Radical Prostatectomy) [15]. In high-grade (malignant) samples, cancer may metastasize in other organs. In these cases, surgical operation is not sufficient. Therefore, urologists apply other therapies like hormone therapy, radiotherapy and chemotherapy to control the disease [15], [16].

### 3. SEGMENTATION OF THE GLANDULAR REGIONS

To evaluate the architecture of the glands, we need to separate them from the image. To this end, color image of the biopsy sample  $I$  is first converted to gray scale image  $I_{gs}$ . This is because color may vary based on the staining agent and does not play an important role in locating the glandular regions. A variance filter is then applied to  $I_{gs}$  to enhance the nuclei. The output is called  $I_v$ , (see Fig. 3.c). Wavelet-based texture features are then extracted from  $I_{gs}$  and  $I_v$ . A  $K$ -means clustering algorithm is applied to them to segment the image into stroma, lumina, and nuclei. The number of clusters in the  $K$ -means clustering algorithm is set to 2: one for lumina and stroma, and the other for nuclei. The glandular regions are then located from the resulting segmented images. Details of the above steps are explained in the following subsections.

#### 3.1. Variance filter

In practice, biopsy samples are stained by solvents with different illuminations and colors. Differences between the solvents may affect the image color and the segmentation process of the color images. To avoid this problem, we first convert the color image  $I$  into gray scale image  $I_{gs}$  [17], and then apply a variance filter to it to get image  $I_v$ . This filter assigns each pixel, the variance of its neighbor intensities. Since the regions containing nuclei have higher variance than the regions containing stroma and lumina, the application of the variance filter to the gray scale image makes these regions brighter than the other regions. Fig. 3 shows an example of the above process.



**Figure 3:** Application of a variance filter to a prostate biopsy image. (a) Original cancerous image of grade 4, (b) Gray scale version ( $I_{gs}$ ) of (a), (c) Resulting image ( $I_v$ ) after the application of the variance filter to (b). Note that the variance filter enhances the intensity of the nuclei.

### 3.2. Wavelet transform

Wavelet transform (WT) is the decomposition of signal with a family of orthogonal bases obtained through translation and dilation of a kernel function  $\psi(t)$  known as the mother wavelet. The mother wavelet is constructed from a scaling function  $\varphi(t)$  which satisfies the following difference equation [18].

$$\varphi(t) = \sqrt{2} \sum_k h(k) \cdot \varphi(2t - k), \quad \forall k \in Z \quad (1)$$

The mother wavelet  $\psi(t)$  is related to the scaling function via:

$$\psi(t) = \sqrt{2} \sum_k g(k) \cdot \varphi(2t - k), \quad \forall k \in Z \quad (2)$$

In the above equations,  $h(k)$  and  $g(k)$  are low-pass and high-pass filters, respectively [18]. Applying WT to the image  $I$  in the first level of decomposition creates four images with the half length of the original image ( $I_{LL}, I_{LH}, I_{HL}, I_{HH}$ ). We employ the WT to create texture features as explained in the next section.

### 3.3. Texture features

The three regions of PIP explained in Section 2 can be distinguished by texture patterns. Stroma and lumina are structural textures with high illuminations while nuclei are statistical patterns with high variance. In this section, we briefly explain how the roughness information of the image is employed to characterize these areas. We used our proposed texture features [19] for segmenting the glands in the image. The roughness information of the image is extracted using:

$$f(x, y, s) = \exp\left[-\frac{x^2 + y^2}{2s^2}\right]. \quad (3)$$

The first derivative of function  $f$  in direction  $\theta$  is computed as [20]:

$$f'_\theta(x, y, s) = f'_x \cos(\theta) + f'_y \sin(\theta). \quad (4)$$

where  $f'_x, f'_y$  are the first derivatives of  $f$  with respect to  $x$ , and  $y$ , respectively. The features are calculated in the following steps [19].

- Calculate one level of WT of the image to get  $I_{LL}, I_{LH}, I_{HL}$ , and  $I_{HH}$ .
- Convolve  $f'_\theta$  with each of the four components ( $I_{LL}, I_{LH}, I_{HL}, I_{HH}$ ) in a number of directions. For each pixel of these components, consider a symmetric neighborhood and convolve  $f'_\theta$  with this window. In this paper, the window size is 5x5.

$$F_{\theta, w, (LL, LH, HL, HH)}(k, l) = f'_\theta * I_{\text{neigh}, (LL, LH, HL, HH)}(k, l). \quad (5)$$

where  $(k, l)$  and  $I_{\text{neigh}, (.)}(k, l)$  represent the pixel  $(k, l)$  and its symmetric neighborhood in each component. In the above,  $w$  refers to the wavelet space.

- Compute power of  $F_{\theta, w, (.)}(k, l)$  as:

$$\widehat{F}_{\theta, w, (.)}(k, l) = \langle F_{\theta, w, (.)}(k, l)^2 \rangle \quad (6)$$

where  $\langle f(\cdot) \rangle$  denotes the average of function  $f(\cdot)$ .

- Compute average of  $\widehat{F}_{\theta, w, (.)}(k, l)$  with respect to  $\theta$  ( $\widehat{F}_{w, LL}(k, l)$ ,  $\widehat{F}_{w, LH}(k, l)$ ,  $\widehat{F}_{w, HL}(k, l)$ ,  $\widehat{F}_{w, HH}(k, l)$ ).

$$\widehat{F}_{w, (.)}(k, l) = \frac{1}{N_{\theta}} \sum_{\theta} \widehat{F}_{\theta, w, (.)}(k, l) \quad (7)$$

where  $N_{\theta}$  is the number of directions. Note that  $\widehat{F}_{w, LL}$ ,  $\widehat{F}_{w, LH}$ ,  $\widehat{F}_{w, HL}$ ,  $\widehat{F}_{w, HH}$  are matrices with the half-length of the original image.

- Apply the inverse WT to each of the above components assuming the other components are zero ( $\widehat{F}_{LL}$ ,  $\widehat{F}_{LH}$ ,  $\widehat{F}_{HL}$ ,  $\widehat{F}_{HH}$ ).  $\widehat{F}_{(.)}$  is a matrix with the same size as the original image.
- Apply the following operators to obtain the feature set  $\{F_1, F_2, F_3, F_4\}$ :

$$\begin{aligned} F_1 &= \widehat{F}_{LL}, & F_2 &= \sqrt[4]{\widehat{F}_{LL} \cdot \widehat{F}_{LH} \cdot \widehat{F}_{HL} \cdot \widehat{F}_{HH}} \\ F_3 &= \widehat{F}_{LL} + \widehat{F}_{LH} + \widehat{F}_{HL} + \widehat{F}_{HH} \\ F_4 &= \sqrt{\frac{\widehat{F}_{LL}}{\widehat{F}_{LH} + \widehat{F}_{HL} + \widehat{F}_{HH}}}, & F_5 &= \widehat{F}_{HH} \end{aligned} \quad (8)$$

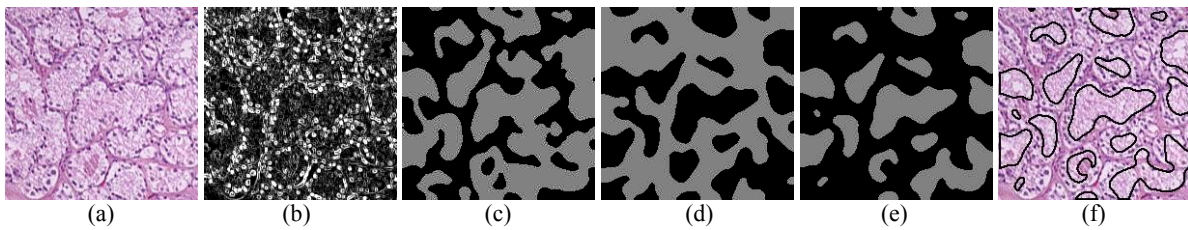
We proposed feature  $F_1$  for highly structural textures (in these textures, high frequency components are less important), features  $F_2$ ,  $F_3$ , and  $F_4$  for fairly structural or statistical textures, and feature  $F_5$  for highly statistical textures (in these textures, high frequency components are very important) [19]. Features  $F_1, F_2, F_3, F_4$  and  $F_5$  are extracted in three scales  $s = 1, 2, 3$  in Equation (3). Thus, a feature set containing 15 elements is obtained for each pixel.

### 3.4. K-means clustering

For segmenting glandular regions in the image, we employ the  $K$ -means clustering algorithm [21] as follows.

- Calculate the first level of wavelet decomposition of the images  $I_{gs}$  and  $I_v$ .
- Apply the feature extraction method explained in Section 3.3 to the resulting images. Since stroma and lumina do not have considerable high frequency components, we omit the features  $F_2$ ,  $F_3$ ,  $F_4$ , and  $F_5$  from its feature set. Therefore, a feature set containing 3 elements is obtained for stroma and lumina.
- Apply the  $K$ -means clustering to the resulting feature set of the image  $I_{gs}$  to segment stroma and lumina. Assume the feature space has two clusters: one for regions containing stroma and lumina and the rest of the image (nuclei).
- Apply the  $K$ -means clustering to the resulting feature set of the image  $I_v$  to segment nuclei with the same assumptions.
- Obtain the glandular regions by excluding the regions containing nuclei from the regions containing stroma and lumina.

In Fig. 4, an example of the glandular regions segmentation is shown.



**Figure 4:** Segmenting the glandular regions in a prostate biopsy image using the method explained in Section 3. Regions containing stroma and lumina, and areas containing nuclei are segmented separately. The glandular regions are then obtained by removing the nuclei regions from the stroma and lumina regions. (a) A benign sample, (b) Resulting image after applying variance filter to gray scale of (a), (c) Segmented image representing areas containing stroma and lumina, (d) Segmented image representing areas containing nuclei, (e) Glandular regions obtained by finding bright pixels of (c) that are not bright in (d), (f) Overlaying the boundaries of (e) on (a).

#### 4. TREE-STRUCTURED CLASSIFICATION

Segmented glandular regions along with the texture features described in Section 3.3 are given as inputs to a TS classification system which grades the PIP in five stages. The output of this algorithm is a grade between one and five. Fig. 5 shows a block diagram of the proposed method. In this section, we describe the different stages of the TS classification system.

##### 4.1. Labeling glandular regions

To extract the glandular features, we need to label the segmented glands. To this end, we consider the connectivity of each pixel in the segmented image showing the glandular regions. We put all connected pixels into a category and assign each category a label (color or intensity). Fig. 6 shows an example of the labeled images. Glandular features of the labeled regions are extracted using the methods proposed in the remainder of this section.

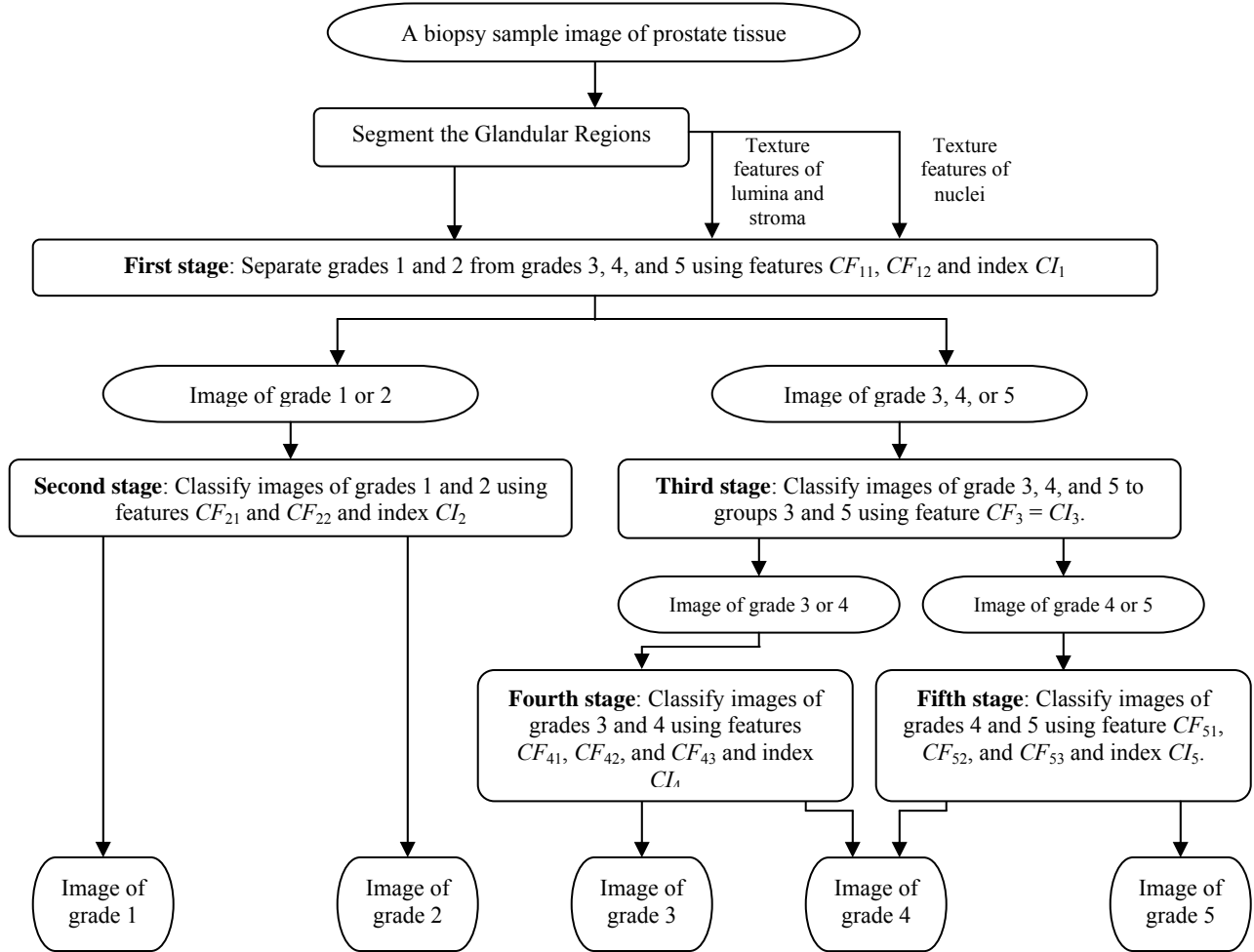
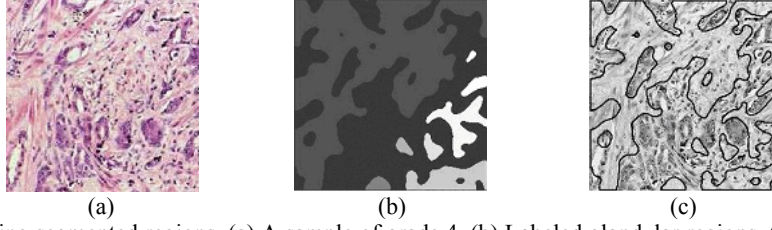


Figure 5: A block diagram of the five stages of TS algorithm.





**Figure 6:** Labeling segmented regions. (a) A sample of grade 4, (b) Labeled glandular regions, (c) Overlaying the boundaries of (b) on (a). Connectivity of the bright pixels in the segmentation map is used to label the regions.

## 4.2. Tree-Structured classification system

In each stage of the TS algorithm a number of features are extracted. The  $j$ th feature in stage  $i$  is shown as  $CF_{ij}$ . The features in each stage are then combined to create a single feature  $CI_i$  called cancer index. A linear classifier is then used to classify the images into two groups.

### 4.2.1. First stage

In the first stage of the TS algorithm, images of grades 1 and 2 (group C1) are separated from the images of grades 3, 4, and 5 (group C2). To this end, variance and roundness of the glandular regions are computed as desirable features.

#### 4.2.1.1. Variance of glandular regions

The glands in low-grade (grades 1 and 2) specimens of PIP are approximately of the same size, while in high-grade samples (grades 3, 4, and 5) they are merged and have different sizes. Furthermore, in low-grade cases the glands are morphologically similar and approximately have the same illumination. In the high-grade cases, not only the glands are dissimilar but also they have different illuminations. We consider this in the definition of the first feature. Assume  $L_i$  is the label of all pixels in the  $i$ th gland.  $SB(i)$ , size of the mentioned gland considering the illumination is obtained by:

$$SB(i) = \sum_{k=1}^N \delta[L(k) - L_i] \cdot I_{gs}(k), \quad i = 1, 2, \dots, M \quad (9)$$

where  $\delta$  is a discrete delta function,  $N$  is the number of pixels in the image,  $k$  represents the  $k$ th pixel,  $I_{gs}(k)$  is the intensity of the  $k$ th pixel in the gray scale image,  $L(k)$  is the label of  $k$ th pixel, and  $M$  is the number of the glands in the segmentation map. The first feature,  $CF_{11}$ , is defined by:

$$CF_{11} = \frac{\text{variance}(SB)}{[\text{mean}(SB)]^2}. \quad (10)$$

The higher  $CF_{11}$ , the higher grade of the specimen.

#### 4.2.1.2 Roundness factor

The glands in low-grade specimens are approximately round, while in high-grade cases they are erupted and have irregular patterns. Hence, we consider the roundness of the glands as the second feature in the first stage. Assuming  $S(i)$  is the area of the  $i$ th gland,  $r(i)$  is the radius of a circle with the area of  $S(i)$ , and  $SR(i)$  is the perimeter of the  $i$ th gland, the roundness factor ( $R_n$ ) of a gland is computed via:

$$R_n(i) = \frac{SR(i)}{S(i) \cdot 2} \cdot r(i) \quad (11)$$

For a circle,  $R_n$  is 1. The higher the roundness of a gland, the closer  $R_n$  to 1. To include both of the roundness and size features, we calculate the following parameter for each gland.

$$S_{Rn}(i) = S(i) \cdot \exp(-|1 - R_n(i)|) \quad (12)$$

The closer  $R_n$  to 1, the closer  $S_{Rn}$  to  $S$ . To include all glands, we compute the following.

$$S_t = \sum_{i=1}^M S(i) \quad , \quad S_{tRn} = \sum_{i=1}^M S_{Rn}(i) \quad (13)$$

Ultimately, we define the second feature of the first stage as:

$$CF_{12} = \frac{S_t - S_{tRn}}{S_t} \quad (14)$$

In low-grade specimens with the round glands,  $CF_{12}$  is small. In high-grade cases where the glands have irregular patterns  $CF_{12}$  is large. We normalize features  $CF_{11}$  and  $CF_{12}$  to  $[0,1]$  [21]. Then, we combine them to compute the following index, which is proportional to the grade of the specimen:

$$CI_1 = CF_{11}^2 + CF_{12}^2 \quad (15)$$

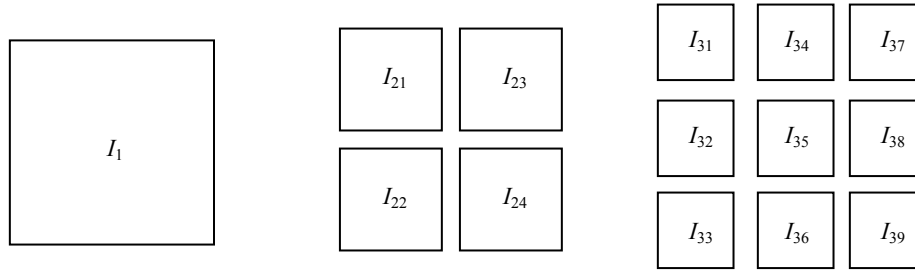
where  $CI_1$  is the cancer index in the first stage of TS algorithm. In low-grade cases where the glands are round and of approximately the same size,  $CI_1$  is low. In high-grade cases, where the glands have irregular patterns and are of different sizes,  $CI_1$  is high. Thus, we apply a linear classifier to  $CI_1$  to classify the images in the first stage.

#### 4.2.2. Second stage

In the second stage of the TS algorithm, images of grades 1 and 2 are separated from C1. Here, we use distribution of the energy in the texture feature space and variance of energy of the segmented regions as effective features.

##### 4.2.2.1 Distribution of energy in the texture feature space

In images of grade 1, the glands are arranged uniformly. But in grade 2, the distances between the glands increase such that their energy distribution changes noticeably. Therefore, we compute energy distribution of the glands as effective feature. Since the glands are better shown in  $I_{gs}$ , we use its feature space to compute this feature.



**Figure 7:** Representation of division of an image into equal parts for calculation of energy distribution of the glands.

- In the  $i$ th stage, the image is divided into  $i^2$  equal regions (Fig. 7). Energy of each region in the texture feature space is then computed separately:

$$E_i(n) = \sum_{k=1}^3 \sum_{j=1}^{N_i} F_{1s=k}(j)^2, \quad n = 1, 2, \dots, i^2 \quad (16)$$

where  $E_i(n)$  and  $N_i$  are the energy and number of pixels in the  $n$ th regions of  $i$ th stage, respectively.  $F_{1s=k}$  is also the  $k$ th texture feature computed for  $I_{gs}$

- In the  $i$ th stage, we define  $A_i$  as an  $i^2$ -element matrix such that its elements are  $E_i(n)$ s.
- In the  $i$ th stage, we compute  $h_i$  proportional to the homogeneity of  $A_i$  as (17). Here,  $H$  denotes the homogeneity.

$$h_i = \frac{H(A_i)}{i^2}, \quad (17)$$

- Finally,  $CF_{21}$  which describes the energy distribution of the glands, is computed as follows:

$$CF_{21} = \sum_{k=1}^Q h_k \quad (18)$$

In the above,  $Q$  is the number of all stages. In image of grade 1 where the glands are arranged uniformly,  $CF_{21}$  is small. But in images of grade 2 where the distribution of the glands changes considerably,  $CF_{21}$  is large.



#### 4.2.2.2 Variance of energy of the segmented glandular regions

In images of grade 1, the glands have very similar structure. But in images of grade 2, the glands do not have similar structure as in grade 1. To verify this point, we study the variance of energy of the segmented glandular regions in the image. Energy of  $i$ th gland in the texture feature space is computed as:

$$E_g(i) = \frac{1}{S(i)} \sum_{k=1}^3 \sum_{j=1}^N \delta[L(j) - L_i] \cdot F_k(j)^2, i = 1, 2, \dots, M. \quad (19)$$

In this equation,  $N$  is the number of pixels in the image and  $F_k$  denotes the  $k$ th feature of the texture feature space. Ultimately, variance of energy of the glands in the texture features is computed as follows:

$$CF_{22} = \frac{\text{variance}(E_g)}{[\text{mean}(E_g)]^2} \quad (20)$$

Like the previous section, we normalize  $CF_{21}$  and  $CF_{22}$  to  $[0, 1]$ . Then, we combine them as:

$$CI_2 = CF_{21} \cdot CF_{22} \quad (21)$$

and apply linear classifier. In image of grade 1 where the glands are arranged uniformly and are similar,  $CI_2$  is small. But, in grade 2 where the glands are not arranged as uniformly as grade 2 and are dissimilar,  $CI_2$  is large.

#### 4.2.3 Third stage

In the third stage of the TS algorithm, images of grade 3, 4, and 5 (group C2) are classified into two groups: groups of grade 3 (C21) and grade 5 (C22). In this step, some of grade 4 images are classified to C21 and the rest to C22. In grade 3, we still see some glands in the image with nuclei connected to each other. In grade 5, on the other hand, we see no glands in the image and nuclei are distinct from each other. We use these facts to classify the grades 3, 4, and 5 into two mentioned groups. Here, we use a feature describing the entropy of lumina, stroma, and nuclei to classify the images. In this stage, we use the feature space of  $I_{gs}$  as follows:

$$CI_3 = CF_3 = - \sum_{k=1}^3 \sum_{i=1}^N \left[ \frac{F_{1k}(i)}{F_{2k}(i)} \right] \cdot \log \left[ \frac{F_{1k}(i)}{F_{2k}(i)} \right] \quad (22)$$

where  $N$  is the number of pixels in the image and  $F_{1k}$  and  $F_{2k}$  are the  $k$ th feature explained in section 3.3 and correspond to  $I_{gs}$  and  $I_v$ , respectively. In this stage also, we apply the linear classifier to the computed feature to categorize the samples.

#### 4.2.4 Fourth stage

In grade 3, we still see round glands, but in grade 4, the glands are erupted and have irregular patterns. We use this property to separate grades 3 and 4 from C21 in the fourth stage of TS algorithm. To this end, we compute ratio of energies of round segmented regions to irregular ones. We combine this feature with those describing entropy of lumina, stroma, and nuclei, respectively.

##### 4.2.4.1 Ratio of energies of round regions to irregular regions

In the first stage of the TS algorithm, we computed a feature ( $CF_{12}$ ) which describes the roundness of the glandular regions. While training the system, a threshold ( $Tr$ ) is obtained for this term. Here, we use this threshold to determine if a gland has round or irregular shape. Hence, ration of round regions to irregular regions is obtained as follows:

$$\begin{aligned} E_{gr} &= \sum_{i=1}^M u[Tr - (1 - \exp(-|1 - R_n(i)|))] \cdot E_g(i) \cdot S(i), \\ E_{gi} &= \sum_{i=1}^M u[(1 - \exp(-|1 - R_n(i)|)) - Tr] \cdot E_g(i) \cdot S(i), \\ CF_{41} &= \frac{E_{gi} - E_{gr}}{E_{gi}}. \end{aligned} \quad (23)$$

##### 4.2.4.2 Entropy features

When cancer advances in the prostate tissue, it increases the level of entropy of the image. Here, we compute features describing the entropy of lumina and stroma ( $CF_{42}$ ), and nuclei ( $CF_{43}$ ), respectively. To this end, we use the feature sets of  $I_{gs}$  and  $I_v$ . These values are obtained as follows:

$$CF_{42} = -\sum_{k=1}^3 \sum_{i=1}^N F_{1k}(i) \cdot \log[F_{1k}(i)], \quad (24)$$

$$CF_{43} = -\sum_{k=1}^{15} \sum_{i=1}^N F_{2k}(i) \cdot \log[F_{2k}(i)]. \quad (25)$$

Like the previous sections, we normalize these three features to  $[0, 1]$ . Then we combine them as follows to generate  $CI_4$ .

$$CI_4 = CF_{41} + CF_{42} + CF_{43} \quad (26)$$

We finally apply the linear classifier to the resulting index. In grade 3 that the cancerous samples have lower entropy and there are still round glands,  $CI_4$  is small, but in grade 4 image  $CI_4$  is large.

#### 4.2.5 Fifth stage

In grade 4, the glands are not merged completely and have different levels of energies. We use these points to separate grades 4 and 5 from C22 in the fifth stage of TS algorithm. Here, we use variance of energy of the segmented regions and features describing entropy of lumina, stroma, and nuclei. Since we compute similar features in the previous sections, here we use some of them as:

$$CF_{51} = -CF_{22}, \quad CF_{52} = CF_{42}, \quad CF_{53} = CF_{43}. \quad (27)$$

$$CI_5 = CF_{51} + CF_{52} + CF_{53} \quad (28)$$

We apply the linear classifier to  $CI_5$  to categorize the specimens.

## 5. EXPERIMENTAL RESULTS

To evaluate the performance of the proposed technique, we use two datasets. The first dataset contains 91 images with grades 1, 2, 3, 4, and 5. The numbers of samples in these grades are 4, 15, 20, 25, and 27 respectively. These images have similar magnifications and illuminations. They are captured with magnification of  $\times 100$ . The second dataset contains 199 images with Gleason grades 1, 2, 3, 4, and 5. The numbers of samples in these grades are 11, 28, 44, 49, and 67, respectively. These images are captured with different magnifications and illuminations.

The leave-one-out technique is used to evaluate the performance of the proposed system in each stage as well as in overall. To train the system, thresholds are calculated for each classifier in all stages. The thresholds are obtained by minimizing the classification error in each stage. Since the numbers of samples in different grades of both datasets are not the same, we trained the classifier such that a balance occurs between the considered samples (as an example, there are 11 and 28 samples in grades 1 and 2 of the second dataset, respectively. In the second stage that these grades are classified, the threshold is derived when the optimum classification errors are obtained for both grades). We computed texture features obtained by Haar, Daubechies (Db) 3, Db 6, Symlet (Sym) 2, Coiflet (Coif) 2, Bi-orthogonal (Bior) 1.1, Bi-orthogonal 3.3, Reverse Bi-orthogonal (Rbio) 1.1, and Reverse Bi-orthogonal 3.3, respectively and employed them in TS algorithm. Table 1 shows the accuracy percentages of the proposed stages (Stg) and the overall system (OS) using these wavelet basis.

To compare the performance of the proposed system with the other works, we classified the images using multi-wavelet features that are proposed in [14]. We computed energy and entropy of the multi-wavelet coefficients of the images in the first and second levels of decomposition. We extracted these features for repeated row and critically sampled preprocessing. We used the multiwavelets GHM, CL, SA4, BiGHM2, and BiH32. To evaluate the error rate of the features, we used a  $k$ -NN classifier and the simulated annealing as described in [14]. The results are presented in Table 2.

**Table 1:** Accuracy percentages of the proposed stages (Stg) and the overall system (OS) using different wavelets.

Wavelet base	First dataset						Second dataset					
	Stg1	Stg2	Stg3	Stg4	Stg5	OS	Stg1	Stg2	Stg3	Stg4	Stg5	OS
Haar	98	100	100	97	96	95	92	97	95	89	91	84
Daubechies 3	98	100	100	97	96	95	94	97	95	88	88	83
Daubechies 6	97	100	100	95	96	94	93	97	95	86	92	84
Symlet 2	97	100	100	93	96	92	91	97	95	86	88	81
Coiflet 2	97	100	100	97	96	95	93	97	95	89	90	85
Biorthogonal 1.1	98	100	100	97	96	95	92	97	95	89	91	84
Biorthogonal 3.3	98	100	100	97	96	95	91	97	95	87	88	82
Reverse Biorthogonal 1.1	98	100	100	97	96	95	92	97	95	89	90	83
Reverse Biorthogonal 3.3	98	100	100	95	96	94	92	97	95	88	89	82

**Table 2:** Accuracy percentages of energy and entropy features of multi-wavelet method proposed in [4] for grading of PIP. The results are obtained for either repeated row (rr) or critically sampled (cs) preprocessing and in first and second levels of decomposition.

Dataset		First dataset										Second dataset									
Level of decomposition		1 <sup>st</sup> level					2 <sup>nd</sup> level					1 <sup>st</sup> level					2 <sup>nd</sup> level				
k		1	3	5	7	9	1	3	5	7	9	1	3	5	7	9	1	3	5	7	9
Multi wavelet (rr)	GHM	94	91	86	82	79	91	92	86	82	76	68	65	64	62	57	69	66	64	62	59
	CL	94	92	87	84	78	94	90	88	86	79	69	67	64	63	61	71	67	65	63	62
	SA4	95	93	89	86	80	95	93	88	85	78	73	67	65	62	60	71	66	65	64	62
	BiGHM2	95	92	87	85	77	94	91	87	82	77	72	66	64	63	60	69	67	65	63	61
	BiH32	93	90	86	83	78	93	91	87	85	78	69	66	64	64	61	69	65	64	62	61
Multi wavelet (cs)	GHM	92	91	83	83	78	95	92	90	87	80	67	65	64	62	59	71	65	64	62	60
	CL	94	93	89	85	78	94	92	89	82	77	70	67	65	63	61	69	67	61	63	61
	SA4	94	91	89	85	79	95	92	87	83	76	72	66	63	62	60	71	66	65	64	62
	BiGHM2	90	92	92	85	82	93	95	94	85	80	67	65	63	62	60	69	66	63	63	61
	BiH32	93	94	93	90	84	92	92	93	90	84	71	69	66	63	62	72	69	67	64	62

Our proposed system grades the PIP with accuracies of about 95% and 85% for the datasets 1 and 2 respectively. The most important advantage of the proposed method compared with the other techniques is that it is based on the physical/biological concepts used by the pathologists. The results of the comparison study suggest that when the images are captured in similar conditions, their characteristics do not have considerable differences. However, when the images are captured in different conditions, the image characteristics may vary noticeably such that each sample has its own characteristics. Thus, classifying the images by searching the most similar samples ( $k$ -NN) may lead to undesirable results. Table 2 shows that the accuracy decreases as  $k$  increases. This may indicate that even when the images are captured with similar magnification and illumination, the pathological samples in the same group may not have the same energy and entropy features. This is attributed to the subjectivity of the PIP. Thus, for training the system using  $k$ -NN classifier and energy and entropy features, a relatively large dataset containing various samples may be needed.

## 6. CONCLUSION

A TS algorithm is proposed and evaluated in this paper for automatic grading of the pathological images of the prostate. Texture features are computed and employed in  $K$ -means clustering to segment glandular regions of pathological image of the prostate. The segmented glandular regions and the texture feature space are then given as the inputs to the TS algorithm for automatically grading the PIP. Performance of the system is evaluated using two datasets containing 91 and 199 images, respectively. These datasets have images with similar and different magnifications and illuminations, respectively. Accuracies of about 95% and 85% have been achieved for the first and second datasets, respectively.

## REFERENCES

1. V. Kumar, R.S. Cotran, and S.L. Robbins, *Basic pathology*, Philadelphia, PA: Saunders, 1997.
2. C. Bohring and T. Squires, "Cancer statistics", *CA Cancer J. Clin.* **Vol. 43**, pp. 7-26, 1993.
3. C. Mettlin, G. Jones, and G. Murphy, "Trends in prostate care in the United States, 1974-1990: Observations from the patient care Evaluation Studies of the American College of Surgeons Commission on cancer", *CA Cancer J. Clin.* **Vol. 43**, pp. 83-91, 1993.
4. C.M. Coley, M.J. Barry, C. Fleming, and A.G. Mulley, "CLINICAL GUIDELINE: PART1: Early detection of prostate cancer: Part I: Prior probability and effectiveness of tests", *Ann Intern Med.* **Vol. 126**, No. 5, pp. 394 – 406, 1997.
5. B.S. Kramer, M.L. Brown, Ph.C. Prorok, A.L. Potosky, and J.K. Gohagan, "Prostate cancer screening: What we know and what we need to know", *Ann Intern Med.* **Vol. 119**, No. 9, pp. 914–923, 1993.
6. F. Cornud, X. Belin, T. Flam, Y. Chretien, S. Deslignieres, F. Paraf, JM. Casanova, N. Thiunn, O. Helenon, B. debre, and JF. Moreau, "Local staging of prostate cancer by endorectal MRI using fast spin-echo sequences: prospective correlation with pathological findings after radical prostatectomy", *Br J Urol.* **Vol. 77**, No. 6, pp. 843 – 850, 1996.
7. M. Huncharek, and J. Muscat, "Serum prostate-specific antigen as a predictor of staging abdominal/pelvic computed tomography in newly diagnosed prostate cancer", *Abdom Imaging.* **Vol. 21**, No. 4, pp. 364-367, 1996.
8. J. Rosai, and L.V. Ackerman, *Ackerman's surgical pathology*. St. Louis, MO: Mosby, 1996.
9. Z.-H. Zhou, Y. Jiang, Y.-B. Yang, Sh.-F. Chen, "Lung cancer identification based on artificial neural network ensembles", *Artificial Intelligence in Medicine*, **Vol. 24**, No. 1, pp. 25-36, 2002.
10. D.K. Tasoulis, P. Spyridonos, N.G. Pavlidis, D. Cavouras, P. Ravazoula, G. Nikiforidis, and M.N. Vrahatis, "Urinary bladder tumor grade diagnosis using on-line trained neural networks", V. Palade, R.J. Howlett, and L.C. Jain (Eds): *KES, Lecture note in artificial intelligence*, No. 2773, pp. 199-206, 2003.
11. E.M. Marroquin, E. Santamaria, X. Jove, and J.C. Socoro, "Morphological analysis of mammary biopsy images", *in Electro technical Conf. MELECON 96 , 8<sup>th</sup> Mediterranean*, No. 2 pp. 1067-1070, 1996.
12. R. Stotzka, R. Manner, P. H. Bartels and D. Thompson, "A hybrid neural and statistical classifier system for histological grading of prostate lesions", *Analytical Quantitative Cytol. Histol.* **Vol. 17**, No. 3, pp. 204-218, 1995.
13. M.T.D. Clark, F.B. Askin, and C.R. Bagnell, "Nuclear roundness factor: A quantitative approach to grading in prostate carcinoma, reliability of needle biopsy tissue, and the effect of tumor stage on usefulness", *The Prostate*. No. 10, pp. 199-206, 1987.
14. K. Jafari-Khouzani and H. Soltanian-zadeh, "Multiwavelet grading of pathological images of prostate", *IEEE Trans. Biomed. Eng.* **Vol. 50**, No. 6, pp. 697-704, 2003.
15. M. B. Garnick, "Prostate Cancer: Screening, Diagnosis, and Management", *Ann Intern Med.*, **Vol. 118**, No. 10, 1993, pp. 804 – 818.
16. D. J. Vaughn, "Hormone Therapy for Advanced Prostate Cancer", *Ann Intern Med.*, **Vol. 132**, No. 7, pp. 584 – 585, 2000.
17. W.K. Pratt, *Digital Image Processing (third edition)*, John Wiley & sons Inc, 2001.
18. C.S. Burrus, R.A. Gopinath, H. Guo, *Introduction to wavelet and wavelet transforms*, Prentice-Hall, 1998.
19. R. Farjam, R.A. Zoroofi, and H. Soltanian-zadeh, "Unsupervised texture segmentation using roughness in wavelet domain", *2<sup>nd</sup> Int. IEEEGCC Conf.*, No. 284, Bahrain, Nov. 2004.
20. D. Charalampidis and T. Kasparis, "Wavelet-based rotational invariant roughness features for texture classification and segmentation", *IEEE Trans. Image Processing.* **Vol. 11**, No. 8, pp. 825 -837, 2002.
21. C.H. Chen, L.F. Pau, and P. S. P Wang, *The handbook of pattern recognition and computer vision (2nd Edition)*, World Scientific Publishing Co, 1998.