

Computer-aided Prostate Cancer Diagnosis from Digitized Histopathology: A Review on Texture-based Systems

Clara Mosquera-Lopez*, *Member, IEEE*, Sos Agaian, *Senior Member, IEEE*
Alejandro Velez-Hoyos, and Ian Thompson

Abstract—Prostate cancer (PCa) is currently diagnosed by microscopic evaluation of biopsy samples. Since tissue assessment heavily relies on the pathologists level of expertise and interpretation criteria, it is still a subjective process with high intra- and interobserver variability. Computer-aided diagnosis (CAD) may have a major impact on detection and grading of PCa by reducing the pathologists reading time, and increasing the accuracy and reproducibility of diagnosis outcomes. However, the complexity of the prostatic tissue and the large volumes of data generated by biopsy procedures make the development of CAD systems for PCa a challenging task. The problem of automated diagnosis of prostatic carcinoma from histopathology has received a lot of attention. As a result, a number of CAD systems, have been proposed for quantitative image analysis and classification. This article aims at providing a detailed description of selected literature in the field of CAD of PCa, emphasizing the role of texture analysis methods in tissue description. It includes a review of image analysis tools for image preprocessing, feature extraction, classification, and validation techniques used in PCa detection and grading, as well as future directions in pursuit of better texture-based CAD systems.

Index Terms—Prostate cancer, Gleason grading, histopathology image analysis, pattern recognition, computer-aided diagnosis (CAD), texture-based CAD systems.

I. INTRODUCTION

PROSTATE cancer (PCa) is the second most frequent cause of cancer behind skin cancer in the United States. The American Cancer Society estimates that about 233,000 new cases will be diagnosed, and 29,480 men will die of prostate cancer in 2014 [1]. Although the lifetime risk of a man being diagnosed with clinically apparent PCa is around 11%, and the lifetime risk of dying of prostate cancer is 3.6%, the age is an important risk factor that increases the probabilities of a man to be diagnosed with prostate cancer to 50% [2]. Today, microscopic analysis of needle biopsy tissue sections remains the gold-standard method for cancer detection and grading [3]–[5]. Other screening methods such as prostate specific antigen

(PSA) and digital rectal exam (DRE) are often used to identify patients who need a biopsy and general suspicious areas in the prostate gland, respectively. However, they are only indicators of prostate cancer risk, but not definitive diagnosis methods. For example, PSA test yields to low sensitivity and specificity; and detection by palpation through DRE is also limited to relatively large and superficial lesions [6], [7].

The prostate samples collected during the transrectal ultrasound (TRUS) or magnetic resonance (MR) guided biopsy [8], [9] are processed in order to produce glass microscopic slides or digitized slides. From the visual or computerized analysis standpoint, one of the most important steps in the tissue processing chain is staining. In general, Hematoxylin and Eosin (H&E) are used for prostatic tissue staining in order to highlight diagnostically important histological and textural features. Due to chemical reactions between stains and tissue components, the nuclei are stained blue, whereas the cytoplasm and extracellular matrix have varying degrees of pink staining [10].

Currently, pathologists visually assess histopathology slides using conventional microscopes, camera-equipped microscopes or computers running software tools developed for use with modern digital pathology equipment such as whole-slide scanners. The Gleason grading system is the standard for PCa diagnosis. Since 2003, this system has been endorsed by the World Health Organization (WHO) [11] and it has been widely adopted for pathologists around the world. The system proposed by Dr. Donald Gleason assigns a grade from 1 to 5 depending on the architectural pattern of the glands of the prostate tumor [12]–[14]. Fig. 1 presents illustrative examples of Gleason grades 3 to 5, which are the most common grades. Gleason grade 1 refers to well-differentiated glands that resemble normal tissue, whereas Gleason grade 5 represents poorly or nondifferentiated glands. Therefore, high grades (grade 4 and 5) are closely related to a more aggressive disease, whereas low grades correlate with a more favorable patient outcome. In contemporary clinical practice, prostate-tumor grading starts with pattern 3 [15]–[17] because Gleason patterns 1 and 2 are rare and may lead to diagnostic errors (i.e., reporting these grades usually reflects undergrading and does not correlate with radical prostatectomy) [18], [19]. A pathology report includes a Gleason score, which corresponds to the sum of the two most predominant Gleason patterns within a histopathology image: predominant grade + secondary grade. When a tumor has only one histologic pattern, the

Manuscript received November 27, 2013; revised March 27, 2014; accepted July 1, 2014. This work has been partially supported by CTRC P30 Cancer Center Support Grant (CA054174) from the National Cancer Institute. Asterisk indicates corresponding author.

Clara Mosquera-Lopez and Sos Agaian are with the Department of Electrical and Computer Engineering, University of Texas at San Antonio, San Antonio, TX, 78249 USA (e-mail: claram.mosquera@gmail.com).

Alejandro Velez-Hoyos is with the Department of Pathology, Pablo Tobon Uribe Hospital, Medellin, Colombia.

Ian Thompson is with the Cancer Therapy and Research Center, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA.

primary and secondary pattern are given the same number. The Gleason score on a biopsy is a powerful indicator for prostate cancer prognosis that correlates with all of the important pathologic parameters at radical prostatectomy, prognosis after radical prostatectomy, patient outcome radiotherapy, and many molecular markers [20]–[22].

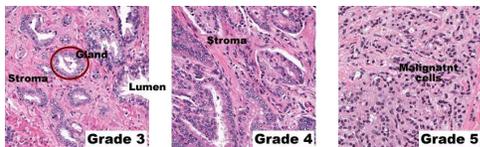


Fig. 1. Examples of Gleason grades 3 to 5

Assessment of prostatic tissue specimens by pathologists is a decisive step in the diagnosis of prostate cancer towards the selection of the best treatment option for a patient. Cancer detection and grading by pathologists from histopathology images is a time consuming and error-prone procedure. Various aspects may affect the accuracy of a pathology report including the pathologist’s experience and fatigue, variability in the interpretation and application of the grading criteria, and complexity of tissue samples. It has been reported in recent studies that intra- and interobserver reproducibility of the Gleason grading system ranges from 60 to 90% [23]. However, high variability in grading occurs when distinguishing between tangentially sectioned Gleason pattern 3 glands and the poorly formed gland subset of pattern 4 [24]. Moreover, comparisons between the assigned grade in the needle biopsy and the grade of the matched whole prostate gland reflect undergrading of the needle biopsy specimen in 42% of the cases and overgrading in 15% of the cases [23]. Undergrading is a serious problem in circumstances where the treatment of a low-Gleason-score tumor (containing grades 1 and 2) would vary from a treatment of a Gleason score 5 or 6 [25]. In order to improve the overall accuracy of prostate cancer diagnosis, two or more pathologist readings are preferred, which increase the cost and time of diagnosis significantly. Ideally, two pathologists should perform independent analyses of biopsy specimens, and a third pathologist would be necessary in the case that the first two readers have different diagnostic impressions as presented in Fig. 2.

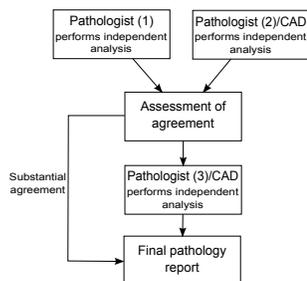


Fig. 2. Ideal scheme for accurate diagnosis of PCa

CAD techniques could offer a cost-effective alternative to image interpretation and double reading as a means of reducing errors and decreasing observational oversight [26].

A CAD system is not intended to replace a pathologist. Instead, it could act as a second reader or decision support tool in clinical practice providing more precise quantitative information about suspicious areas in order to produce accurate and more complete pathology reports. CAD systems, for example, can automatically measure the extent of cancerous areas, percentage of tissue occupied by each Gleason grade, and generate localized cancer maps for visualization. Some important rewards of using computerized systems to diagnose disease are the speed, reproducibility and consistency of the diagnostic methods because the performance of computers is not affected by fatigue, perceptual errors, or variability in classification criteria. This review focuses on an elaborated discussion on feature extraction, feature selection, classification and validation methods used in texture-based CAD systems for prostate cancer detection and grading. Texture features are used extensively in data-driven models because they are simple to extract, useful to describe histopathology images; and when extracted locally, robust to geometric and illumination changes, as well as partial occlusions [27]. In PCa CAD, texture features accurately represent normal tissue and all Gleason grades at a pixel-, tile-, or image-level, whereas tissue-structure-based features require tissue segmentation before feature extraction, and might not be accurate in classifying regions where the gland, lumen, or stroma areas exceed the image (or image patch) under study, or regions of high-grade carcinoma where some of the tissue structures (e.g. glands) cannot be differentiated. The main objective of this paper is to emphasize the potential of intelligent computer systems to be used in clinical practice to help pathologists to analyze and classify prostate cancerous tissue.

The rest of this paper is organized as follows: Section II presents the general pipeline of computer-aided diagnosis of prostate cancer. Section III discusses systems and methods for automated detection and grading of prostatic carcinoma from histopathology using texture-based features. Finally, Section IV concludes the paper with focus on future research directions in the field of CAD for prostate cancer.

II. COMPUTER-AIDED DIAGNOSIS OF PROSTATE CANCER

After approximately a decade of research on quantitative analysis of prostate cancer histopathology, several approaches have been proposed to automatically detect, classify, and grade prostatic carcinomas. The general components used in most of the existing CAD systems are a preprocessing unit, a feature extraction and feature selection unit, and a classification block as illustrated in Fig. 3. The image preprocessing steps are intended to remove irrelevant background noise, enhance diagnostically important details of the images, segment important objects within the image, standardize color and image scale. Image scale standardization in this context refers to adjusting the magnification of the image such that it matches the magnification of the images used for training the classifier. Scale variations affect a variety of image features such as object size, topology, and texture. Next, the feature extraction unit derives relevant properties of the tissue by measuring the size and estimating the shape of segmented tissue structures, or

by computing color and texture features. The subset of features that best differentiate Gleason grades are then selected using several methods for feature selection, and this subset is used for classification. The classification block detects the presence of cancer or determines the malignancy level of the detected disease.

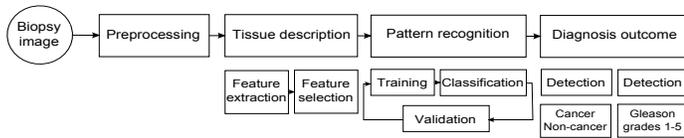


Fig. 3. General architecture of CAD systems for PCa

Based on the feature extraction methods used for tissue description, there are two main classes of computerized recognition systems:

- **Tissue-structure-based CAD systems** employ feature vectors derived from measurements of the size, shape, and spatial arrangement of gland units, lumen, epithelial cytoplasm, epithelial nuclei and other tissue structures to distinguish among different classes.
- **Texture-based CAD systems** use measurements of spatial variations in pixel intensities in order to characterize the pattern of Gleason grades. The properties of a texture can be characterized as fine, coarse, smooth, rippled, molled, and irregular or lineated [28]. Texture analysis can be performed either in spatial or transform domain.

One important advantage of using texture features over tissue-structure-based features is their ability to perform better classifying high-grade prostatic tumors, which are characterized by minimal glandular differentiation. For instance, in images of Gleason grades 4 and 5 some basic elements of the tissue, such as lumen, are absent or can be occluded by blue mucin or cytoplasm [29]; therefore, accurate morphometric measurements can not be obtained. Several combinations of morphological, cytological and texture features have also been explored to distinguish among Gleason grades. Fig. 4 depicts the classification of existing CAD systems for prostate cancer diagnosis base on the feature extraction methods used for classification.

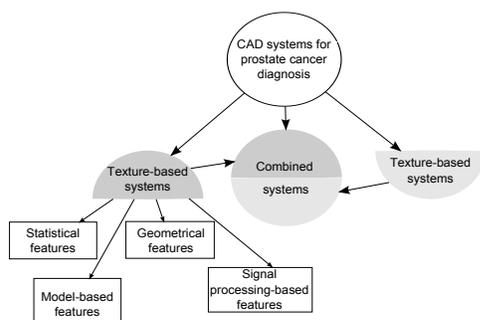


Fig. 4. Classification of CAD systems for PCa according to the feature extraction methods

Texture-based CAD systems, which are the focus of this work, use a variety of texture analysis methods for feature extraction that can be grouped into four general categories:

statistical, geometrical, model-based, and signal processing methods [30]. Statistical methods represent texture indirectly by non-deterministic properties that govern the distribution and relationship among pixel values of an image [31]. Examples of commonly used statistical features include autocorrelation function, gray-level co-occurrence matrix and histograms. Geometrical methods consider a texture as an entity composed of texture primitives and analyze the placement rule that describe the texture. Graph-based features extracted from Voronoi tessellation or Delaunay triangulation are examples of geometrical texture features. Model-based methods describe texture using mathematical models such as autoregressive model, random field models, and fractal analysis. Finally, signal processing approaches extract features from filtered images. Transform-based textural analysis is an important approach that falls in this group. Transform-based methods have the ability of decomposing a signal using predefined basis or data-driven basis, and providing important frequency and spatial information. There exist many transform-based methods; however, the most popular ones are base on the Fourier transform, Gabor and other wavelet transforms. Table I categorizes the CAD systems for PCa diagnosis according to the texture analysis methods and the features that have been used to describe prostate histopathology images.

The development of new algorithms for CAD of prostate cancer is an active research field. Several new implementations based on modifications or combinations of the above texture analysis methods are proven to be effective detecting and grading prostatic carcinomas. However, a great challenge for CAD systems is still to be accurate enough to be used in clinical practice. Selected works on texture-based CAD systems will be described in detail in Section III.

III. SYSTEMS AND METHODS FOR COMPUTER-AIDED PROSTATE CANCER DIAGNOSIS

Most of the published works on CAD of PCa have been focused on analysis of pre-selected regions of interest (ROI) from digitized histopathology, and only few of them have addressed the problem of disease detection from digitized whole-slide biopsy images. Detecting the spatial extent and presence of disease is still a challenging problem due to the large amount of data that is generated by using new scanning technologies. For instance, in order to analyze a digitized version of a single core of a prostate biopsy digitized at 40x magnification, a CAD system must process more than 200 million pixels, and the analysis of a prostate biopsy containing 12-20 samples requires processing more than 2.5 billion pixels of data. Therefore, the development of fast algorithms for description and accurate classification of digitized whole-slide images is still an open research problem in the field of CAD of PCa. More research has to be done on automatic detection of cancerous regions, because relying on preselected ROIs limits the general usability of the automated grading algorithms in clinical practice [32].

Ideally, a CAD system for prostate cancer should be a two-stage system. The first step accurately localizes the cancerous regions, and then the second stage performs more detailed

TABLE I
SUMMARY OF TEXTURE ANALYSIS METHODS USED IN CAD OF PCA.

Category	Subcategory	Description	Features
Statistical	Histogram-based approach	Summarizes information about pixel intensities distributions, but fails to capture spatial relationship among pixels.	First-order statistics [29], [32]–[34], color channels histograms [35], 2D HSV color histograms [36]
	Autocorrelation function	Captures the amount of regularity, fineness and coarseness of the texture.	Autocorrelation coefficients
	Run-length matrix (GLRLM)	Contains information about spatial relationships between groups of pixels having similar gray level values.	GLRLM features [33]
	Co-occurrence matrix (GLCM)		Haralick features [29], [32], [34], [37]–[40]
Geometrical	Graphs	Capture the placement rule that describe a texture.	Statistics from nuclear-based graphs (Delaney, Voronoi, and Minimum Spanning Tree) [36], [41]–[43], network cycle [41]
Model-based	Auto regressive model	Captures the local interaction between pixels. It is assumed that a pixel gray level is a weighted sum of the neighboring pixels.	Set of weights
	Random field models	Model a texture image as a probability model or as a linear combination of a set of basic functions.	Probabilistic pairwise Markov model (PPMM) [44]
	Fractals	Capture the roughness and self-similarity of textures at different scales.	Grayscale fractal dimension [35], [45], color fractal dimension [44], color ratio-based fractal dimension [46], entropy-based fractal dimension [45], [47], fractal code [35]
Signal processing-based	Time-domain filter response	Captures orientation and edge information of the image.	Sobel, Kirsch, gradient, and derivative [29], [32]
	Fourier	Captures dominant orientation and coarseness of a pattern, but fails to provide spatial information.	Statistics and energy of Fourier coefficients' magnitude
	Wavelets	Capture texture complexity at several scales by providing spatial and frequency information.	Gabor wavelet coefficients [29], [32], wavelet energy [46], [48], statistics of wavelet coefficients [35], [49], joint probability of color channels wavelet coefficients [46], energy and entropy of multiwavelet coefficients [40], wavelet-based fractal dimension [47], [48]

analysis on those regions in order to determine the histological grade of the detected tumor. This section describes in detail the techniques based on image textural analysis that researchers have used so far for detection and histological grading of prostate cancer from digitized biopsy images.

A. Image preprocessing

Current high resolution histology imaging systems such as camera-equipped microscopes and whole-slide scanners allow production of high content images. Preprocessing algorithms can be used to reduce the computational cost through image segmentation or multi-scale image decomposition, reduce noise, and enhance diagnostically important details of the tissue. For instance, reduction of computational complexity might be achieved by image subsampling or by wavelet decomposition. Doyle et al. [32] employed a pyramidal decomposition prior to feature extraction in order to obtain images of the same slide at different resolutions. In repetition, Gaussian smoothing is first performed on the full-resolution image followed by subsampling of the smoothed image by a factor of 2 [50]. Then, the low-resolution images are analyzed to roughly locate the cancerous regions of interest, and only these regions go to the higher resolution processing steps. In

case of poor quality input, e.g. severe noise, low intensity contrast with weak edges, and color variations, several preprocessing techniques such as image smoothing, denoising, color normalization or standardization, and enhancement may be applied for image conditioning.

Generic and specific algorithms for noise removal [51], image enhancement [52]–[54], and edge detection may be applied to histopathology images. Those algorithms have proven to have a positive incidence in cancer recognition problems [55]. For instance, Almutashri et al. [48] demonstrated that a logarithmic-ratio edge detection algorithm using visual morphology concepts [56] perform very well in detecting prostatic tissue structures and highlighting edge information of histological details, which are then used to capture pattern similarities between Gleason grades using the fractal dimension measure.

1) *Color normalization and standardization*: In texture-based CAD systems, color normalization and standardization of histopathology images play an important role because the performance of the classification may be adversely affected by color variations. Color variations are caused by variations in staining and scanning conditions due to image acquisition protocols, capturing-device properties, and lighting conditions. Two aspects have made the standardization of

color a challenging problem: the presence of important but subtle diagnostically important details in color images, and the heterogeneity of tissue composition. Several approaches to histopathology color standardization have been proposed so far [57]–[60]. For example, the approach proposed by Yagi [58] is based on the use of standard color filters selected for histology H&E stained slides (i.e., color chart) in order to calibrate and profile imaging devices. The color produced by a particular scanner is mapped to the reference colors using a polynomial transformation, in which the color transformation matrix is constructed according to the parameters of a specific scanner or imaging device. After the polynomial transformation is done, the resulting image is gamma corrected in order to produce the final result. A shortcoming of this method is the need for a specific transformation matrix per scanner. Another approach to color standardization of histopathology images is based on global or local color transference [57], [59], [60]. Mosquera-Lopez and Agaian [46] developed a standardization algorithm that uses fuzzy c-means to segment the reference and input images in meaningful regions and generates a weighting function based on the pixels fuzzy membership index. The membership index modulates the color transference operations. Magee et al. [57] used a multimodal extension to linear normalization in Lab color space, which generates several transforms depending on the image composition (one for each segmented tissue structure present in the image), and the distribution of each class is then mapped to each class of the image being processed. However, none of the aforementioned approaches have used a quality metric to evaluate the performance of the standardization algorithm being used and its impact on the overall quality of the image.

Although several studies have been carried out to develop algorithms for color image standardization, various researchers in the field of CAD of PCa only used color model transformations for image normalization. For instance, Doyle et al. [32] used Red Blue Green (RGB) to Hue Saturation Intensity (HSI) transformation [61] in order to confine color variations to the intensity channel of the HSI color space instead of affecting all three RGB channels [62].

B. Texture-based tissue description

In order to identify prostatic patterns for cancer detection and grading, texture features have been used as discriminative measurements of the tissue samples of different Gleason grades. In general, a large number of features can be generated in CAD systems with the aim of selecting the ones that best derive clinically significant information and efficiently differentiate among classes. In this section a summary of the feature vectors used in both prostate cancer detection and grading is presented.

1) *Prostate cancer detection*: Most of the texture-based CAD systems for PCa detection use statistical texture analysis methods such as first- and second- order statistics for tissue description. Statistical features are easy to extract employing algorithms with relatively low computational complexity. This is important due to the high amount of data that has to be processed when detecting cancer regions from whole-slide digitized images.

Sun et al. [33] utilized a texture analysis method based on the run-length matrix for identifying tissue abnormalities in a sample collected from a radical prostatectomy and digitized at 50x magnification. For PCa detection, they used block processing for classifying each region as normal or PCa according to textural descriptors obtained based on the Gray-level Run-length Matrix (GLRLM) [63]–[65]. By definition, a set of consecutive collinear pixels with the same gray level, in a given direction, constitute a gray level run. The number of pixels in the run is called run length. Then, each element of the GLRLM is given by:

$$GLRLM(i, j | \theta) = \beta_{ij} \quad (1)$$

where β_{ij} is the total number of occurrences of runs of length j at gray level i , in a given direction θ .

Once the GLRLM has been computed, several features can be extracted to model the coarseness of tissue texture for further analysis and classification.

Sun et al. [33] computed four different GLRLM, one for each of the following angular directions: $\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. The GLRLM features were extracted from the matrix resulting of the addition of the four aforementioned run-length matrices. The proposed system also used two pixel intensity-based features corresponding to the mean and standard deviation of gray levels of each block. After feature selection, the most discriminative features (i.e., first order statistics of pixel gray level and low gray-level run emphasis (LGRE), high gray-level run emphasis (HGRE), and run percentage (RPC) from GLRLM) were used for classification. The performance of the system was estimated using k-fold crossvalidation. The reported accuracy, sensitivity, and specificity of the diagnosis system are 89.5%, 90.48% and 89.49%, respectively.

Another set of statistical features commonly used in automated prostate cancer detection is derived from the gray-level co-occurrence matrix (GLCM) [29], [32], [37], [39]. This feature set were proposed by Haralick [28] and comprises 14 features numbered from f_1 to f_{14} : angular second moment, contrast, correlation, sum of squares or variance, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, information measures of correlation and maximal correlation coefficient. The GLCM can be defined as follows:

$$GLCM(i, j | d, \theta) = \alpha_{ij} \quad (2)$$

where α_{ij} represents couples of pixels having i and j , respectively, as grey levels and separated by a distance d in a direction angle θ .

The GLCM reveals certain properties about the spatial distribution of pairs of pixels sharing gray levels in the texture image.

Usage of fractal theory is the most used model-based texture analysis methods in PCa recognition tasks. Most of the proposed features are based on the computation of fractal dimension using the box-counting algorithm [66]. The fractal dimension D of an image is defines as follows [67], [68]:

$$D = \frac{\log(N_r)}{\log\left(\frac{1}{r}\right)} \quad (3)$$

where N_r is the number of non-overlapping copies of a bounding set, each of which has been scaled down by a ratio of r .

The discriminant power of the color fractal dimension was explored by Yu et al. in [44] in conjunction with a probabilistic pairwise Markov Model (PPMM). Yu et al. [44] proposed a method that uses an extended color fractal algorithm for the computation of the fractal dimension on a per pixel basis taking into account the context of hyper-rectangles, as opposed to only hyper-cubes as proposed initially by Ivanocici and Richard [69]. The color fractal dimension, that captures color and textural information of the tissue, is modeled as a mixture of gamma distributions for cancer and benign pixels, whereas the spatial dependencies between pixels are incorporated using a Probabilistic Pairwise Markov Model (PPMM) [70]–[72] after bayesian classification. The system was tested using Leave-one-out crossvalidation on a data set of 27 H&E stained histological sections from radical prostatectomies digitized at 40x magnification. The performance metric used in this study is the area under de ROC curve. It was demonstrated that the introduction of the Markov model produces an increment of the measured AUC from 0.790 to 0.831. The relatively high AUC achieved by this system using a single feature show that integrating the color channels when describing the tissue instead of processing them separately might be helpful to capture significant interchannel information.

In addition, besides using features obtained form statistics and fractal analysis of the image in the spatial domain, many features can also be extracted using other signal processing-based methods such as filtering and frequency transformations. In automated classification of prostate cancer research, several spatial domain filters have been employed to capture image texture properties. Most of the used methods concentrate in measuring edge density for modeling texture coarseness. For example, Sobel filters in the x - and y -, and two diagonal axes, Kirsch filter, gradients in the x - and y -axes, difference of gradients, and diagonal derivative were used by Doyle et al. [32] and Nguyen et al. [29] for cancer detection.

Cancer detection systems cannot be compared directly because each research group use different datasets annotated by different pathologists and different experimental protocols. However, from the results of the selected texture-based CAD systems it can be observed that despite the ease of computation of first-order statistics, they have a limited impact in the accuracy of cancer detection. Reported results in [33] show that only the mean and standard deviation of pixel intensities were selected as a discriminative features, while measures related to linear spatial relationship between texture primitives obtained using the run-length approach were selected as more important features. Moreover, in the multiresolution approach proposed by Doyle et al. [32], it was demonstrated that first-order statistics of pixel intensities do not play a determinant role in PCa detection, whereas co-occurrence features and Gabor filter response were the most effective variables in distinguishing cancer pixels from normal pixels at all resolutions

levels. According to this observations, it can be concluded that quantifying the spatial relationship between groups of pixels is more beneficial in PCa recognition than obtaining plain statistics of image gray levels.

2) *Prostate cancer Gleason grading*: Unlike PCa detection, the CAD systems for PCa Gleason grading do not use pixel- or tile-based classification approach. In general, homogeneous regions (i.e., cancerous regions containing a single Gleason pattern) are used for feature extraction and classification. The texture-based features used in PCa grading are more varied that the ones used in detection (and several categories of texture analysis methods can be mixed) because most of the research work have been concentrated in the recognition of the histological grade of pre-segmented cancerous images or sub-images. In existing PCa grading systems, first-order statistics and co-occurrence features are less common, and fractal analysis and wavelet features are preferred.

Following the fractal-based approach, several systems have been proposed [45], [47], [48]. Huang et al. [45] developed a system that analyzes the texture complexity of histological images using two fractal measurements: fractal dimension calculated by using the conventional differential box counting method, and entropy-based fractal dimension. In order to compute the entropy-based fractal dimension (EBFD), an $M \times M$ image is partitioned into blocks of size $s \times s$. Then, the entropy of the pixels of each block is computed as:

$$e_r = - \sum_{k=1}^{N_g} p_k \log_2(p_k) \quad (4)$$

In equation (4), k indexes the gray level of the pixels in a given block, p_k is the probability of gray level k , and N_g is the total number of gray levels in a block. The total contribution of all blocks for a scale down ration $r = \frac{s}{M}$ is given by:

$$E_r = \sum_i (e_r(i))^2 \quad (5)$$

The EBFD features of an image can be estimated using least-squares linear fitting for $\log(E_r)$ versus $\log\left(\frac{1}{r}\right)$.

The accuracy of the system using both fractal dimension measurements and various classifiers (i.e., Bayesian classifier, k-NN and SVM), that is around 95%, was estimated using cross-validation methods.

In the cases where wavelet transform is employed to describe histopathology images, the most common extracted features include statistics (mean and standard deviation), energy, and entropy of wavelet coefficients and low resolution images. In order to perform 2D wavelet transform, a scaling $\phi_{j,m,n}(x,y)$ and a wavelet function $\psi_{j,m,n}^i(x,y)$ are required. The wavelet transform coefficients of an image at the j^{th} decomposition level can be computed as follows:

$$W_\phi(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \phi_{j,m,n}(x,y) \quad (6)$$

$$W_{\psi^i}(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \psi_{j,m,n}^i(x,y) \quad (7)$$

Equation (6) is used to compute low-frequency components, and equation (7) is used to compute detail coefficients with $i = \{H, V, D\}$ representing horizontal, vertical, and diagonal edge information.

Jafari-Khouzani and Soltanian [40], [73] used multi-wavelet transform to represent cancerous images. They computed the energy and entropy of multi-wavelet coefficients of each resulting sub-matrix along with textural features extracted from a co-occurrence matrix to classify a data set of 100 images into Gleason grades 2 to 5. The maximum reported accuracy of the was 97%. Yoon et al. [49] developed a computer-aided classification system where textural features were extracted from cardinal multiridgelet transform (CMRT) [74] to differentiate images of Gleason 3 from Gleason 4. SVM with a Gaussian kernel was used for the classification task and the accuracy of the system is 93.75% using Leave-One-Out crossvalidation method. Almutashri et al. [48] presented a method for automatic classification of prostate cancer biopsy images by combining energy features from wavelet transform and wavelet-based fractal dimension. Experimental results showed average classification accuracy (with one-vs.-all SVM classifiers) of 95% in a set of 45 images of Gleason grades 3, 4, and 5. Mosquera-Lopez and Agaian [46] presented a system for classification of prostatic carcinomas of grades 3 to 5 by using statistics of the distribution of wavelet energy within cancerous patches, joint probability of wavelet coefficients obtained by wavelet decomposition of the channels of color images and measurements of color fractal dimension. The average accuracy of the system was estimated by cross-validation yielding a 97% of correct classification rate.

Although grading systems perform multiclass classification tasks, the overall accuracy of texture-based grading systems is higher than the accuracy achieved in PCa detection. The feature sets used in grading systems are high-dimensional spaces. In this context, the systems using statistics and energy of wavelet coefficients and fractal measures tend to be more compact and concise because a few variables can accurately capture differences and similarities among Gleason patterns. It was demonstrated by Huang and Lee [45] that the performance of fractal dimension-based features is statistically better or at least equivalent to the performance of feature sets based on multiwavelet, Gabor, and GLCM texture analysis methods when recognizing Gleason grades. From the observations of Huan and Lee and the accuracy of CAD systems using fractal analysis, it can be stated that self-similarity properties are more important in Gleason grading than in cancer detection.

Other approaches to prostate tissue description consist of segmenting tissue structures and extracting features based on individual properties of each one of them. Mainly, nuclear and glandular features have been considered important characteristics in the detection of prostate cancer and its severity [75]–[82]. Tissue structure-based systems exploit the correlation between the size, shape and arrangement of histological structures within the histology image with Gleason grades, but geometrical texture features are also computed based on the location of structures of interest. CAD systems that use geometrical features after segmenting important tissue structures are also considered texture-based systems. For example,

Naik et al. [42] used centroids of segmented nuclear structures to create Voronoi, Delaunay and minimum spanning tree graphs in order to capture the spatial arrangement of nuclei in pathological images represented by the area and edges length features as well as the nuclear density. They classified H&E stained images into cancer, non-cancer and cancer confounders groups using statistics from the constructed graphs. Wetzel et al. [43] constructed spanning tree graph to connect segmented cell nuclei over a tumor image and quantify architectural arrangement of cells. The proposed system was able to correctly match the grade value of pathological images in 80% of the test cases.

Several combinations of feature vectors coming from morphometric and textural image analysis are found in literature. Roula et al. [34], [83] investigated the accuracy of utilizing Haralick features, gland area and nuclear area extracted from multi-spectral microscopy image to separate stroma, benign prostatic hyperplasia (BPH), prostatic intraepithelial neoplasia (PIN) and prostatic carcinoma obtaining an average classification error of 5.57%. Diamond et al. [39] used morphometric and Haralick texture features to identify stroma, normal, and cancerous regions in samples of prostatic tissue from whole-mount radical prostatectomy. Classification of noncancerous regions was performed using morphometric characteristics of the histology image under the assumption that normal tissue exhibit larger areas of associated lumen. On the other hand, classification of stroma and cancerous tissue was done by looking at Haralick features. In this study, 79.3% of the sub-regions of interest were correctly classified. Tabesh et al. [35], [84] integrated object- and image-level features describing the color, texture and morphometric characteristics of histopathology images. Object-level features include statistics of the intensity of the segmented tissue structures, whereas image-level features consider color channel histograms, fractal measurements and statistics of wavelet coefficients. The developed system achieve an accuracy of 96.7% classifying tumor and non-tumor images and a maximum accuracy of 81% in Gleason grading classification tasks.

C. Feature selection and dimensionality reduction

Before classification, features selection and feature space dimensionality reduction are key steps. Features selection is performed in order to identify the most discriminative features, which not only should contain more information about patterns, but also present small intra-class variance and high inter-class variance in order to enhance the class separability and consequently the system classification performance. There are three main aspects that affect the performance of a classification system: the quality of the selected features, samples sizes, and classifier complexity [85], [86]. The selection of a good set of features consists of finding a subset of characteristics with a specified size that satisfies a certain restriction on an evaluation measure or optimizes an evaluation metric. In general, the selection of features is a tradeoff between size of the feature space and the value of its evaluation measure [87]. Examples of evaluation measures include classification accuracy, consistency, information, dependence and distance. Feature selection

and dimensionality reduction methods are often used to reduce the computational complexity of pattern recognition systems and more importantly to control the curse of dimensionality phenomenon. Such a phenomenon requires the number of training data points to be an exponential function of the feature space dimension to achieve a good classification performance [88]. Therefore, it is important to ensure that the number of features is smaller than the number of training patterns in order to prevent the classification performance to be adversely affected (peaking phenomenon) [85], [89]–[92].

There are various approaches to the problem of feature selection. A list of the most common feature selection methods includes exhaustive search, branch-and-bound search, best individual features, sequential forward selection (SFS), sequential backward selection (SBS), sequential forward floating search (SFFS), and sequential backward floating search (SBFS) [85]. In the field of CAD systems for PCa various feature selection methods have been used, but they are not clearly reported in most of the published papers. For instance, Huang and Lee [45] used SFFS to rank the most discriminant fractal-based features. Sun et al. [33] used a piecewise linear network method [93] for selecting the most discriminative features among run-length features. The selection algorithm utilizes a piecewise linear orthonormal least-squares (OLS) procedure combined with floating search to select the most useful features in a computationally efficient manner since only one data pass is required to complete the selection.

Tahir and Bouridane [94] presented a round-robin tabu search (RR-TS) algorithm for selecting the features that best discriminate each possible pairs from the following group of considered classes: stroma, benign prostatic neoplasia, prostatic intraepithelial neoplasia, and prostatic carcinoma. The method uses tabu search [95], [96] to find the subset of features that satisfy a fuzzy logic rule in which the number of features must be minimized as well as the number of incorrect predictions. For the description of tissue from multi-spectral images, the proposed classification system employs first order statistics, second order statistics obtained using Haralick features, and morphology information. An important characteristic of RR-TS is the use of different features to solve each specific binary classification problem, which improve the overall classification accuracy. The reporter classification error across all considered classes is smaller than 2%. Later, Bouatmane et al. [37] extended the round-robin algorithm for sequential forward feature selection.

On the other hand, dimensionality reduction techniques, as indicated by its name, aim to reduce the dimension of the feature space in general by mapping the feature vector onto a lower-dimensional space through some coordinates transformation. Linear methods [97]–[99] have been used in CAD systems for prostate cancer. Examples of the linear methods include principal component analysis (PCA) [83], linear discriminant analysis (LDA) [34], and independent component analysis. As an unsupervised data analysis tool, PCA finds orthogonal eigenvectors along which the greatest amount of variability in the data lies. However, the projection of feature points to the principal component directions may not separate the data well for classification. In contrast, LDA

is a supervised learning tool, which incorporates data label information to find the projections that maximize the ratio of between-classes variance and within-classes variance [100].

Although less common, nonlinear dimensionality reduction techniques have been also employed. These methods overcome a major limitation of linear dimensionality reduction methods, which assume that the geometrical structure of the high-dimensional feature space is a linear relationship. For instance, Sparks and Madabhushi [101], [102] successfully used manifold learning to reduce the dimensionality of the feature space, but preserving non-linear relationship between object instances.

With multiple classes of features extracted from large size histology images, the resulting vast quantity of data can be prohibitive for feasible analysis, even with current high performance computing machines. Feature selection and dimensionality reduction techniques are useful in generating a compact and non-redundant subset of features that improves interpretability and classification generalization, especially in prostate cancer classification problems where images of the same Gleason grade present multiple variations.

D. Classification

In general, detection and grading of prostate cancer using automated systems is done by supervised pattern recognition. In the context of pattern recognition, a pattern is a vector of features describing the properties of an object or class. Supervised classification approaches require the use of annotated samples (features with their respective class label) to train the classifier and to determine the decision boundaries (classes separation) in a given feature space. A classifier is a mathematical function that takes an input vector of features and assigns it to one of K classes considered in a classification problem.

Once an appropriate subset of features is selected, several classification methods can be used for prostate cancer diagnosis. Such classification methods include k nearest neighbors (k -NN), Bayesian classifier, Support Vector Machine (SVM), neural networks, Markov random field (MRF) classifier, Gaussian classifier, and classical linear discrimination (CLD), among others. A list of commonly used classifiers in CAD for PCa is presented in Table II. Fig. 5 shows an example of the output of a cancer detection system using a boosted Bayesian classifier [32]. In this figure, the red spots are more likely cancer pixels.

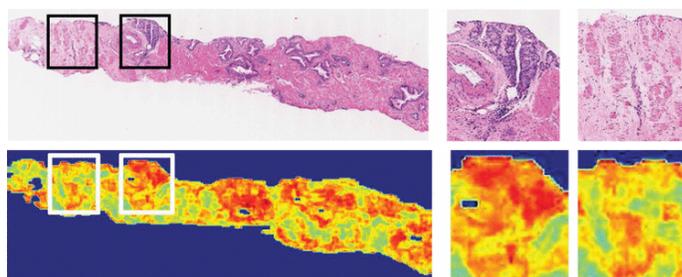


Fig. 5. Example of the output of an automated system for PCa detection (red or darker regions represent cancerous tissue) [32]

TABLE II
CLASSIFICATION METHODS.

Classification method	Description
k-Nearest Neighbor [103]	Assigns pattern to the majority class among k nearest neighbors given a distance metric.
Bayes classifier [104]	Assigns pattern to a class with the maximum posterior probability.
Logistic classifier [104]	Finds the maximum likelihood rule for sigmoidal posterior probability.
Linear discrimination [85]	Finds a linear combination of features which characterizes or separates two or more classes by minimizing the sum of squares error function.
Multilayer perceptron [85]	Optimizes weights of multiple layers of nodes in a directed graph using a nonlinear activation function.
Support vector machine [105]–[107]	Finds a separating hyperplane, such that the margin between classes in the feature space is maximized.

Various studies [35], [45], [109] have published comparative performance analysis among various classification procedures in order to demonstrate which one is better for prostate cancer diagnosis under specific circumstances. A large comparative study on machine learning techniques for prostate cancer diagnosis was conducted by Alexandratou et al. [110]. In their work, 16 supervised machine learning algorithms were compared based on their performance. Classification problems regarding cancer detection (tumor vs. non-tumor), low- vs. high-grade recognition; and the multi-class problem Gleason grading were addressed. Thirteen Haralick texture characteristics were calculated based on grey level co-occurrence matrix of microscopic prostate tissue. For the best-performing algorithm in each case the accuracy obtained was 97.9% for cancer detection, 80.8% for lowhigh grade discrimination, and 77.8% for accomplishing both detection and Gleason grading. Logistic regression and sequential minimal optimization for training a support vector machine were among the top scoring algorithms in each classification problem.

Multi-classifier ensemble [111] is another strategy used for increasing the accuracy of recognition systems. The diversity of data and feature vectors related to PCa histopathology creates a perfect scenario for classifiers ensembles. Multi-classifiers systems may be beneficial for computer-aided prostate cancer diagnosis (detection and grading) given the complexity of the recognition tasks. In general, different feature sets, different training sets, and different classification methods can be integrated into a system to improve the overall accuracy of the labeling because some of classifiers are better at resolving one aspect of the labeling problem, whereas another method may be superior in a different respect [112].

Multiclassifier systems can be grouped into three general categories: parallel, cascading or serial combinations, and hierarchical (tree-like). In the parallel architecture, the predictions of independent classifiers are combined using a fusion rule such as average and voting. In the cascading architecture, individual classifiers are invoked in a linear sequence such that the possible classes are reduced after each classification stage. For the sake of efficiency, inaccurate but cheap classifiers (low computational and measurement demands) are considered first,

followed by more accurate and expensive classifiers. In the hierarchical architecture, individual classifiers are combined into a structure, which is similar to that of a decision tree classifier [85].

While learning ensembles have been successfully employed for various pattern recognition tasks, they have found limited application in problems related to medical image analysis and CAD [113]. Few studies in the field of CAD for prostate cancer from histopathology images have used combination of several classifiers. Doyle et al. [114]–[116] presented a cascaded multi-class pairwise classifier to grade regions of interest of prostate tissue biopsies. The proposed classifier incorporates domain knowledge to partition the multi-class problem into several binary-class tasks, reducing the intra-class heterogeneity that causes errors in one-vs.-all multi-class approaches. In their cascaded approach, successive classifications are performed, beginning with the most broad (i.e., cancer detection) and proceeding to increasingly granular separations (pattern 3+4 vs. pattern 5 and epithelium + atrophy vs. stroma), and finally classifying the most similar classes within each group (pattern 3 vs. pattern 4 and epithelium vs. stroma, light gray line). This reduces classification error by ensuring that the separations are performed between dissimilar classes. The reported simulation results show that the cascaded approach only outperforms the one-shot classification (OSC) and one-vs.-all (OVA) [117] schemes in terms of the positive predictive value (PPV). The average accuracy for Gleason 3, 4 and 5 are 77%, 76% and 95% respectively, and the average accuracy across all classes is 89%.

Another classifier ensemble approach used in CAD of PCa is AdaBoost [118], [119]. AdaBoost is an adaptive and iterative technique that takes advantage of weak hypotheses to construct a strong classification function, which is a weighted linear combination of the responses of weak classifiers and reveals which features are most salient at each resolution level. At each iteration, the AdaBoost algorithm selects a weak classifier $h_t(\cdot)$, which minimizes the following mathematical expression:

$$Z_t = \sum_i D_t(i) \exp(-y_i h_t(x_i)) \quad (8)$$

D_t is the weight on example i at round t , $y_i \in \{-1, 1\}$ is the target label of the example, x_i is the example, and $h_t(\cdot)$ is a binary classifier. After every round the weights are updated as follows:

$$D_{t+1}(i) = \frac{D_t(i) \exp(-y_i h_t(x_i))}{Z_t} \quad (9)$$

If each extracted feature is used to build a weak classifier as in the work presented by Doyle et al. [32], the final classifier will give more weight to the features that best distinguish between classes, intrinsically performing feature selection.

Nguyen, Jain and Sabata [29] also used an ensemble of two SVM classifiers to detect cancer regions using a fusion of cytological and texture features. Each classifier is trained using a feature set and the final label of a patch of a whole-slide image is assigned to the cancerous class based on the

TABLE III
SELECTED CAD SYSTEMS FOR PCA DETECTION.

Author(s)	Features	Dataset	Classification method	Performance measure
Doyle et al. (2012) [32]	First-order statistics, filter response (Sobel, Kirsch, gradient, derivative), co-occurrence features, and Gabor features	100 whole-slide images at 40x optical magnification	Boosted Bayesian classifier	AUC: 0.84, 0.83, and 0.76 for the lowest, medium, and highest image resolution, respectively TPR: 78% FPR: 6%
Nguyen et al. (2011) [29]	Combination of cytological features and texture features including first-order statistics, second-order statistics, and Gabor features	17 whole-slide images digitized at 20x magnification. 6 images were used for training, and 11 for testing	SVM with RBF kernel	
Bouatmane et al. (2011) [37]	Haralick features, glandular area, and nuclear area	592 textures multi-spectral images of size 128 x 128 examined at 40x magnification	k-NN one-vs.-all binary classifiers plus round-robin (RR) sequential forward-feature selection	CCR: 99.83%
Yu et al. (2011) [44]	Color fractal dimension and probabilistic pairwise Markov Model (PPMM)	27 radical prostatectomy specimens digitized at 40x magnification	Markov random field (MRF) classifier	AUC: 0.831
Sun et al. (2009) [33]	Run-length matrix features: low gray-level run emphasis, high gray-level run emphasis, run percentage; mean and standard deviation	9 subimages of a tissue sample from a radical prostatectomy at 50x magnification	Multilayer perceptron (MLP)	CCR: 89.5%

multiplication of the posterior probabilities for cancer and non-cancer classes. The criterium for assigning the cancerous label to an image patch can be expressed mathematically as follows:

$$\prod_{i=1}^n p(f_i(x_i) = 1) > \prod_{i=1}^n p(f_i(x_i) = 0) \quad (10)$$

,where $p(f_i(x_i) = 1)$ and $p(f_i(x_i) = 0)$ denote the probability that the used SVM classifier f classifies the feature vector x_i as normal or cancer, respectively. Otherwise, the patch under study is considered a normal patch. Simulation results, on a data set of 17 whole-slide images which conform two independent sets for training and testing, show a maximum true positive rate TPR = 78% at a false positive rate FPR = 6%. Experiments also show that the combination of cytological features and texture-based features yields to significant improvements in terms of TPR with respect to systems using each the feature set alone.

DiFranco et al. [120] developed a system for cancer detection in digitized images of radical prostatectomy. The system uses an ensemble of a random forest classifier, and SVM with linear and RBF kernel implementing different training models based on a subset of features (obtained from the co-occurrence matrix and spatial filter responses). Each image patch in the data set is classified on a different training model on each pass of the analysis, resulting in multiple predictions for each image tile, which are combined based on a majority voting scheme. In their ensemble approach, AUC values of 95.5%, 95.1% and 94.8% using RBF SVM, random forest and linear SVM respectively were achieved.

A two-stage multiclassifier system was presented by Greenblatt et al. [121] for the assignment of Gleason grades (3, 4, and 5) to H&E biopsy images. The system proceeds in two phases: initial grade assignment, and classification refinement. A quaternion neural network trained using quaternion

wavelet and local binary pattern features performs multiclass-classification. The classification outcome of the neural network is refined by binary linear SVM classifiers if two or more classes have similar probabilities or if the most probable classes are not contiguous Gleason patterns. The system accuracy is 98.89% across all the considered Gleason grades. This classification approach can be generalized to more classification stages using a tree-like structure using different suitable classifiers and be applied to a variety of recognition problems involving multiple classes.

E. System performance assessment

Evaluating classification performance is important for several reasons: when building classifiers, the parameters used for classification can be tuned. For example, at this point, several tests should be done in order to choose predictor variables or features, to estimate parameters, to explore data transformations, and so on. When evaluating given classifiers, it can be determined whether they are good enough for the purpose or whether they provide sufficient improvement over an existing method to merit switching [122]. One of the most used methods for estimating classification performance is cross-validation. There are three different cross-validation methods, namely k-fold, hold-out, and leave-one-out. It is accepted that hold-out cross-validation yields a better estimation of the generalization performance of system, whereas leave-one-out cross-validation is considered to be almost unbiased, but with large variance [123], [124].

In order to obtain reliable performance estimation of a given classifier, it is recommended to have a large number of validation iterations. Cross-validation can be performed on a per-image basis considering that all images in the database are independent, even if they belong to the same patient since each sample is randomly collected from different locations within the prostate. In addition, cross-validation can be performed on

TABLE IV
SELECTED CAD SYSTEMS FOR PCA GRADING.

Author(s)	Features	Dataset	Classification method	Performance measure
Mosquera-Lopez et al. (2013) [46]	Wavelet energy distribution, joint probability of wavelet coefficients, and color ratio-based fractal dimension	71 images of size 512 x 512 of Gleason grades 3 (30), 4 (30), and 5 (11).	SVM linear kernel	CCR: 97%
Mosquera-Lopez et al. (2012) [36]	HSV color features, glandular and nuclear features, and architectural features from Delaunay triangulation	71 images of size 512 x 512 of Gleason grades 3 (30), 4 (30), and 5 (11).	SVM linear kernel	CCR: 95%
Khurd et al. (2011) [41]	Network cycle features and graph features.	25 images of Gleason grade 3 and 50 images of Gleason grade 4. Images are size 1392 x 1040 and were acquired at 10x magnification	SVM	AUC: 0.995
Almuntashri et al. (2011) [48]	Haar wavelet energy and wavelet-based fractal dimension	45 images of size 512 x 512 of Gleason grades 3 (15), 4 (15), and 5 (15).	SVM	CCR: 95%
Yoon et al. (2011) [49]	Variance and entropy of cardinal multiridgelet transform (CMRT) coefficients	42 images of size 768 x 768 of grade 3 carcinoma (16) and grade 4 carcinoma (26).	Gaussian kernel SVM	CCR: 93.75%
Khurd et al. (2010) [108]	Basic texture elements	75 images at magnification 10x of size 1392 x 1040 of grade 3 carcinoma (25) and grade 4 carcinoma (50).	Random forest with SVM	CCR: 94%
Tai et al. (2010) [47]	Wavelet-based fractal dimension, classical fractal dimension and entropy-based fractal dimension computed from each wavelet subband	1000 pathological images	SVM	CCR: 86.3%
Huang et al. (2009) [45]	Differential box counting fractal dimension and entropy-based fractal dimension	205 images	Bayesian, k-NN, and SVM	CCR: 94.6%
Naik et al. (2008) [42]	Glandular shape and size features as well as Voronoi, Delaunay, and minimum spanning tree graph features	44 subimages of benign tissue (17), grade 3 carcinoma (16), and grade 4 carcinoma (11)	SVM	CCR: 91.48%
Alexandratou et al. (2008) [38]	Haralick features	50 samples of histopathological data belonging to Gleason grades 2–5	Multiparameter statistical method of multiple logistic discrimination analysis	CCR: 87%
Diamond et al. (2004)	Glandular and nuclear area as well as Haralick features	Images of size 100 x 100 digitized at 40x magnification	Not specified	CCR: 79.3%
Jafari-Khouzani et al. (2003) [40]	Energy and entropy features calculated from multiwavelet, and co-occurrence matrix features.	100 images at magnification 100x belonging to Gleason patterns 2–5	k-NN	CCR: 97%
Roula et al. (2002) [34]	Haralick features, global variance of pixels, glandular area, and nuclear area.	10 multispectral images of whole mount sections from radical prostatectomy	Supervised Classical Linear Discrimination (CLD)	CCR: 94%
Wetzel et al. (1999) [43]	Glandular features and Nuclear-based spanning tree and Delaunay features	54 prostate cases digitized at 10x magnification	Content-based image retrieval (CBIR) system	CCR: 80%

a per-patient fashion. In this case, a single patient could not have images in more than one group, ensuring that training images are from different patients than testing images [32].

1) *Performance indicators*: Evaluation of system performance is carried out by comparing performance indicators. The most common performance indicators used in CAD systems for PCa are: accuracy or correct classification rate (CCR), sensitivity or true positive rate (TPR), specificity or true negative rate (TNR), positive predictive value (PPV), negative predictive value (NPV), and combinations of those mentioned metrics [125], [126]. Although the overall correct classification have been the most used indicator to measure the performance of PCa classification systems, this indicator generally avoids the classification level of each class in the results. Therefore, a more reliable analysis should consider

other indicators such as sensitivity and specificity, which measure the proportion of positive cases of a given class which are correctly identified as such and the proportion of negative cases of a given class which are correctly identified as such, respectively. Assessing the sensitivity and specificity is particularly important when classifying unbalanced datasets like PCa histopathology. In this cases, the overall accuracy alone might not capture classification errors in the minority class (or classes).

Another system performance indicator often used in prostate cancer detection is the area under the ROC curve. A ROC curve is defined as a plot of the false positive rate (FPR), on the vertical axis, against the TPR, on the horizontal axis. A good classification rule is reflected by a ROC curve which lies in the upper left triangle of the square [127], [128]. The area

under the ROC curve (AUC) measures classifier discrimination capability; that is, the ability of the classifier to correctly separate classes. An AUC of 1.0 represents a perfect classifier, and an AUC of 0.5 represents a worthless test.

Finally, with all system components described using examples, a summary of the developed systems for computer-assisted prostate cancer detection and grading is presented chronologically indicating the main characteristics of each system in Tables III and IV, respectively. The column performance measure in those tables refers to the measures used to assess the CAD systems for PCa detection and grading. However, the studied systems cannot be fairly compared based on the reported performance indicators because of the differences in the experimental setup.

IV. FUTURE DIRECTIONS

Digital pathology has become a useful and valuable tool in clinical and research pathology. This transition started approximately in the 1990s when new technologies for digitizing glass slides were available. The availability of digitized data and the great amount of prostate biopsies performed each year create the opportunity for the development of CAD systems for diagnosis of prostate cancer. In the context of texture-based systems, ongoing research should focus on two main aspects: development of new feature extraction methods and analysis of the correlation between texture features and semantic concepts.

A. Potential feature extraction methods

This subsection describes several methods for feature extraction which has the potential of increasing the accuracy and robustness of CAD systems. Current feature extraction techniques process color histopathology images either in monochrome using only intensity information or as separate color channels, which lead to a loss of valuable inter-channel information. In order to process histopathology as a vector field, the pixels of the image can be treated, for example, as quaternion numbers [121], and new or extensions of texture analysis methods can be used to capture intra- and inter-channel variations. For instance, an extended version of the co-occurrence matrix $\varepsilon - CCM$ can be formulated in terms of the distance ε between color vectors in an appropriate color space.

Regarding signal processing-based feature extraction algorithms, we foresee a lot of potential in histogram-based features obtained using wavelet coefficients (either using real- or complex-valued wavelets), as well as descriptors from hyper-complex compact representations of sub-band's coefficients. For instance, a quaternion representation of the j^{th} wavelet decomposition level is $W_j(m, n) = A(m, n) + H(m, n)i + V(m, n)j + D(m, n)k$.

On the other hand, the quality of the features obtained from the wavelet domain can be improved by using directional wavelet transforms. The commonly used wavelets in PCa diagnosis are real-valued, and they can only capture limited directional information (horizontal, vertical, and diagonal edges). For example, in the sub-band of the highest frequency components, different directions are mixed, and they cannot be

easily separated for image analysis. The availability of more varied texture information obtained from different angular directions allows the development of new multi-scale rotation invariant texture features. Another possible solution to capture more information about the orientation of texture features of prostatic tissue is to use complex wavelets sub-bands. This approach not only generates more wavelet sub-bands oriented at $\pm 15^\circ$, $\pm 45^\circ$, and $\pm 75^\circ$; but also provides additional phase information for further texture analysis.

In addition, the definition of existing and newly-developed texture analysis methods should be generalized to be able to process shapes other than rectangular sections of whole-slide images. This is important because of the high variance of tissue sections shape. Moreover, tumors of small extent located at the boundaries of tissue sections that often cannot be detected by pathologists under a microscope could be prompted by an automated system using a shape-dependant tissue description approach which will improve the fineness of the diagnosis outcome.

B. Analysis of the correlation between texture features and semantic concepts

Although some complex criteria used by pathologists to evaluate tissue samples are often difficult to formulate in computational or mathematical forms, a study of the correlation of features derived from texture analysis methods can be useful to formulate models that relates low-level abstract tissue descriptors to interpretable high-level concepts. These high-level concepts or semantic features may capture biological clues often used by expert pathologists when assessing prostate tissue under a microscope. For instance, low-level properties such as nuclear texture, color density, and gray-level distribution may capture the presence of nucleolus in a nucleus. Also, the fractal dimension of cancerous glands can be also correlated with variations in their shape and size, which is an important indicator of disease progression. Semantic-level features require a large amount of annotated data since each important biological concept should be represented in the training data [129]. Unfortunately, annotated prostate tissue slides are difficult to find. One possible solution to overcome this problem is to use semi-supervised or deep learning algorithms in order to model high-level abstractions in data based on few annotated examples.

C. CAD systems evaluation framework

Although the problem of creating a reproducible performance analysis methodology is not exclusive to texture-based CAD systems for PCa diagnosis, a system evaluation framework should be created in order to assess the accuracy and other statistical parameters of the developed recognition systems. Promising results can be envisioned from the systems and methods discussed in this paper. However, it is not straightforward to evaluate and numerically compare these studies solely based on their reported results [130] because each system is built and tested under different conditions (e.g. different datasets and ground truth annotations). In addition, the reported performance is done using different metrics. In

order to compare several CAD systems for prostate cancer diagnosis or other purposes, it is not enough to have a better indicator measured on a test set, or in a cross-validation or other comparison based on sample data. It is necessary to carry out statistical tests, so that we can be confident that any differences represent genuine underlying differences in performance and are not mere random sampling effects [122]. One resource that might help in the generation of the mentioned evaluation framework is a benchmark annotated database that contains histopathology images of different prostate cancer cases from a large number of patients and examined by different pathologists. Such a dataset will help researchers not only to train their CAD systems, but will also allow comparisons of performance among developed systems in order to identify the most distinguishing features and objectively reduce the set of features to a number of variables that works with the majority of systems' configurations.

V. CONCLUSIONS

Computer-aided prostate cancer diagnosis has become an exciting area of research in the past decade. Several efforts have been done in the development of accurate systems for automated detection and grading of prostatic disease from digitized biopsy images. Most of the developed systems have used texture analysis methods to model the presence and aggressiveness of PCa. Despite of the lack of biological interpretability, texture features are often used because they can be easily extracted at a low computational cost, they are useful, and sometimes sufficient to achieve high classification accuracy in PCa detection and grading. Although several systems as discussed in this paper have been proposed so far, the development of new or extensions of existing monochrome texture analysis methods for tissue description that process color channels as a vector field (allowing for simultaneous analysis of all color data) and the use of arbitrary processing grid shapes are challenging open problems. Solving these issues might further improve the robustness of existing CAD systems. Developed texture-based systems should clearly demonstrate that the accuracy of interpretation of biopsy images with CAD is better than the one without CAD. Therefore, there is a long way to go before CAD systems for PCa become available commercially and widely used in clinics and screening centers.

ACKNOWLEDGMENT

The authors thank the two anonymous reviewers whose comments and insights helped to improve the quality of this manuscript.

REFERENCES

- [1] A. American Cancer Soc., "What are the key statistics about prostate cancer?" Aug 2013.
- [2] D. Tindall and P. J. Scardino, "State of research for prostate cancer: Excerpt from the report of the prostate cancer progress review group," *J. Urology*, vol. 57, no. 4, pp. 28–30, 2001.
- [3] B. Djavan and M. Margreiter, "Biopsy standards for detection of prostate cancer," *World J. Urology*, vol. 25, no. 1, pp. 11–17, 2007.

- [4] A. Angelucci, G. Pace, P. Sanita, C. Vicentini, and M. Bologna, "Tissue print of prostate biopsy: A novel tool in the diagnostic procedure of prostate cancer," *Diagnostic Pathology*, vol. 6, no. 1, p. 34, 2008.
- [5] V. Ravery, S. Dominique, X. Panhard, M. Toubanc, and L. Boccon-Gibod, "The 20-core prostate biopsy protocol - a new gold standard?" *J. Urology*, vol. 179, no. 2, pp. 504–507, 2008.
- [6] M. Moradi, P. Mousavi, and P. Abolmaesumi, "Computer-aided diagnosis of prostate cancer with emphasis on ultrasound-based approaches: A review," *Ultrasound Med. Biol.*, vol. 33, no. 7, pp. 1010–1028, 2007.
- [7] I. M. Thompson, D. K. Pauler, P. J. Goodman, C. M. Tangen, M. S. Lucia, H. L. Parnes, L. M. Minasian, L. G. Ford, S. M. Lippman, E. D. Crawford, J. J. Crowley, and J. Coltman, C. A., "Prevalence of prostate cancer among men with a prostate-specific antigen level 4.0 ng per milliliter," *N. Engl. J. Med.*, vol. 350, pp. 2239–2246, 2007.
- [8] T. Franiel, C. Stephan, A. Erbersdobler, E. Dietz, A. Maxeiner, N. Hell, A. Huppertz, K. Miller, R. Strecker, and B. Hamm, "Areas suspicious for prostate cancer: Mrguided biopsy in patients with at least one transrectal us-guided biopsy with a negative finding multiparametric mr imaging for detection and biopsy planning," *Radiology*, vol. 259, no. 1, pp. 162–172, 2011.
- [9] P. A. Pinto, P. H. Chung, A. R. Rastinehad, A. A. Baccala, J. Kruecker, C. J. Benjamin, S. Xu, P. Yan, S. Kadoury, and C. Chua, "Magnetic resonance imaging/ultrasound fusion guided prostate biopsy improves cancer detection following transrectal ultrasound biopsy and correlates with multiparametric magnetic resonance imaging," *J. Urology*, vol. 186, no. 4, pp. 1281–1285, 2011.
- [10] A. H. Fisher, K. A. Jacobson, J. Rose, and R. Zeller, "Hematoxylin and eosin staining of tissue and cell sections," *CSH Protocols*, vol. 2008, no. 5, 2008.
- [11] D. J. Luthinger and M. Gross, "Gleason grade migration: Changes in prostate cancer grade in the contemporary era," *PCRI Insights*, vol. 9, pp. 2–3, 2006.
- [12] G. T. Mellinger, D. Gleason, and J. Bailar, "The histology and prognosis of prostatic cancer," *J. Urology*, vol. 97, no. 2, pp. 331–337, 1967.
- [13] D. F. Gleason and G. T. Mellinger, "Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging," *J. Urology*, vol. 111, no. 1, p. 58, 1974.
- [14] D. F. Gleason, "Histologic grading of prostate cancer: a perspective," *Human Pathology*, vol. 23, no. 3, pp. 273–279, 1992.
- [15] J. I. Epstein, "Gleason score 2–4 adenocarcinoma of the prostate on needle biopsy: a diagnosis that should not be made," *The Amer. J. of Surgical Pathology*, vol. 24, no. 4, pp. 477–478, 2000.
- [16] R. B. Shah, "Current perspectives on the gleason grading of prostate cancer," *Arch. Pathology Lab. Med.*, vol. 133, no. 11, pp. 1810–1816, 2009.
- [17] J. I. Epstein, "An update of the gleason grading system," *J. Urology*, vol. 183, no. 2, pp. 433–440, 2010.
- [18] J. I. Epstein, W. C. Allsbrook Jr, M. B. Amin, and L. L. Egevad, "The 2005 int. soc. of urological pathology (isup) consensus conf. on gleason grading of prostatic carcinoma," *The Amer. J. of Surgical Pathology*, vol. 29, no. 9, pp. 1228–1242, 2005.
- [19] D. Berney, "Low gleason score prostatic adenocarcinomas are no longer viable entities," *Histopathology*, vol. 50, no. 6, pp. 683–690, 2007.
- [20] L. Egevad, T. Granfors, L. Karlberg, A. Bergh, and P. Stattin, "Prognostic value of the gleason score in prostate cancer," *BJU Int.*, vol. 89, no. 6, pp. 538–542, 2002.
- [21] L. True, I. Coleman, S. Hawley, C.-Y. Huang, D. Gifford, R. Coleman, T. M. Beer, E. Gelmann, M. Datta, and E. Mostaghel, "A molecular correlate to the gleason grading system for prostate adenocarcinoma," *Proc. of the Nat. Academy of Sciences*, vol. 103, no. 29, pp. 10991–10996, 2006.
- [22] D. E. Baydar and J. I. Epstein, "Gleason grading system, modifications and additions to the original scheme," *Turkish J. Pathol.*, vol. 25, no. 3, pp. 59–70, 2009.
- [23] P. A. Humphrey, "Gleason grading and prognostic factors in carcinoma of the prostate," *Modern pathology*, vol. 17, no. 3, pp. 292–306, 2004.
- [24] J. K. McKenney, J. Simko, M. Bonham, L. D. True, D. Troyer, S. Hawley, L. F. Newcomb, L. Fazli, L. P. Kunju, and M. M. Nicolas, "The potential impact of reproducibility of gleason grading in men with early stage prostate cancer managed by active surveillance: a multi-institutional study," *J. Urology*, vol. 186, no. 2, pp. 465–469, 2011.
- [25] W. M. Murphy, "The current status of the pathology of prostate cancer," *Cancer control*, vol. 5, pp. 500–506, 1998.
- [26] R. A. Castellino, "Computer aided detection (cad): an overview," *Cancer Imaging*, vol. 5, no. 1, pp. 17–19, 2005.

- [27] H. Ji, X. Yang, H. Ling, and Y. Xu, "Wavelet domain multifractal analysis for static and dynamic texture classification," *IEEE Trans. Image Process.*, vol. 22, no. 1, pp. 286–299, 2013.
- [28] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural features for image classification," *IEEE Trans. Syst., Man, Cybern.*, vol. 3, pp. 610–621, 1973.
- [29] K. Nguyen, A. K. Jain, and B. Sabata, "Prostate cancer detection: Fusion of cytological and textural features," *J. Pathology Inform.*, vol. 2, 2011.
- [30] M. Tuceryan and A. K. Jain, *Texture analysis*, 2nd ed. World Scientific Publishing Co., 1998, ch. 2, pp. 207–248.
- [31] V. V. Kumar, U. Raju, P. Premchand, and A. Suresh, "Skeleton primitive extraction method on textures with different nonlinear wavelets," *J. of Comput. Sci.*, vol. 4, no. 7, pp. 603–611, 2008.
- [32] S. Doyle, M. D. Feldman, J. E. Tomaszewski, and A. Madabhushi, "A boosted bayesian multiresolution classifier for prostate cancer detection from digitized needle biopsies," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 5, pp. 1205–1218, 2012.
- [33] X. Sun, S.-H. Chuang, J. Li, and F. McKenzie, "Automatic diagnosis for prostate cancer using run-length matrix method," in *SPIE 7260 Medical Imaging: Comput.-Aided Diagnosis*, 2009.
- [34] M. Roula, J. Diamond, A. Bouridane, P. Miller, and A. Amira, "A multispectral computer vision system for automat. grading of prostatic neoplasia," in *IEEE Int. Symp. on Biomedical Imaging*. IEEE, 2002, pp. 193–196.
- [35] 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 Trans. Med. Imag.*, vol. 26, no. 10, pp. 1366–1378, 2007.
- [36] C. M. Mosquera-Lopez, S. Agaian, I. Sanchez, A. Almuntashri, O. Zinalabdin, A. A. Rikabi, and I. Thompson, "Exploration of efficacy of gland morphology and architectural features in prostate cancer gleason grading," in *IEEE Int. Conf. on Syst., Man, and Cybern. (SMC)*. IEEE, 2012, pp. 2849–2854.
- [37] S. Bouatmane, M. A. Roula, A. Bouridane, and S. Al-Maadeed, "Round-robin sequential forward selection algorithm for prostate cancer classification and diagnosis using multispectral imagery," *Mach. Vision and Applicat.*, vol. 22, no. 5, pp. 865–878, 2011.
- [38] E. Alexandratou, D. Yova, D. Gorpas, P. Maragos, G. Agrogiannis, and N. Kavantzias, "Texture analysis of tissues in gleason grading of prostate cancer," in *Biomedical Optics (BiOS) 2008*. SPIE, 2008, pp. 685 904–685 904–8.
- [39] J. Diamond, N. H. Anderson, P. H. Bartels, R. Montironi, and P. W. Hamilton, "The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia," *Human Pathology*, vol. 35, no. 9, pp. 1121–1131, 2004.
- [40] 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.
- [41] P. Khurd, L. Grady, A. Kamen, S. Gibbs-Strauss, E. M. Genega, and J. V. Frangioni, "Network cycle features: Application to computer-aided gleason grading of prostate cancer histopathological images," in *IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*. IEEE, 2011, pp. 1632–1636.
- [42] S. Naik, S. Doyle, S. Agner, A. Madabhushi, M. Feldman, and J. Tomaszewski, "Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology," in *5th IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro (ISBI)*. IEEE, 2008, pp. 284–287.
- [43] A. W. Wetzel, R. Crowley, S. Kim, R. Dawson, L. Zheng, Y. Joo, Y. Yagi, J. Gilbertson, C. Gadd, and D. Deerfield, "Evaluation of prostate tumor grades by content-based image retrieval," in *The 27th AIPR Workshop: Advances in Comput.-assisted Recognition*. Int. Soc. for Optics and Photonics, 1999, pp. 244–252.
- [44] E. Yu, J. P. Monaco, J. Tomaszewski, N. Shih, M. Feldman, and A. Mudabsushi, "Detection of prostate cancer on histopathology using color fractals and probabilistic pairwise markov models," in *Annu. Int. Conf. of the IEEE Eng. in Medicine and Biology Soc. EMBC*, 2011, pp. 3427–3430.
- [45] P.-W. Huang and C.-H. Lee, "Automatic classification for pathological prostate images based on fractal analysis," *IEEE Trans. Med. Imag.*, vol. 28, no. 7, pp. 1037–1050, 2009.
- [46] C. M. Mosquera-Lopez and S. Agaian, "A new set of wavelet- and fractals-based features for gleason grading of prostate cancer histopathology images," in *IS&T/SPIE Electron. Imaging*. Int. Soc. for Optics and Photonics, 2013, pp. 865 516–865 516–12.
- [47] S.-K. Tai, C.-Y. Li, Y.-C. Wu, Y.-J. Jan, and S.-C. Lin, "Classification of prostatic biopsy," in *Conf. on Digital Content, Multimedia Technology and its Applicat. (IDC)*. IEEE, 2010, pp. 354–358.
- [48] A. Almuntashri, S. Agaian, I. M. Thompson, D. Rabah, O. Zin Al-Abdin, and M. Nicolas, "Gleason grade-based automatic classification of prostate cancer pathological images," in *IEEE Int. Conf. on Syst., Man, and Cybern.*, 2001.
- [49] H. J. Yoon, L. Ching-Chung, C. Christudass, R. Veltri, J. I. Epstein, and Z. Zhen, "Cardinal multiridgelet-based prostate cancer histological image classification for gleason grading," in *2011 IEEE Int. Conf. on Bioinformatics and Biomedicine*.
- [50] P. J. Burt and E. H. Adelson, "The laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. 31, no. 4, pp. 532–540, 1983.
- [51] C. B. Smith, S. Agaian, and D. Akopian, "A wavelet-denoising approach using polynomial threshold operators," *IEEE Signal Process. Lett.*, vol. 15, pp. 906–909, 2008.
- [52] S. Agaian, K. Panetta, and A. M. Grigoryan, "Transform-based image enhancement algorithms with performance measure," *IEEE Trans. Image Process.*, vol. 10, no. 3, pp. 367–382, 2001.
- [53] E. Wharton, S. Agaian, and K. Panetta, "Comparative study of logarithmic enhancement algorithms with performance measure," in *SPIE Electronic Imaging*. International Society for Optics and Photonics, 2006, pp. 606 412–606 412–12.
- [54] S. Agaian, B. Silver, and K. A. Panetta, "Transform coefficient histogram-based image enhancement algorithms using contrast entropy," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 741–758, 2007.
- [55] G. D. Tourassi, R. I. Ike, S. Singh, and B. Harrawood, "Evaluating the effect of image preprocessing on an information-theoretic cad system in mammography," *Academic Radiology*, vol. 15, no. 5, pp. 626–634, 2008.
- [56] S. Agaian, "Visual morphology," in *SPIE Electronic Imaging*. International Society for Optics and Photonics, 1999, pp. 139–150.
- [57] D. Magee, D. Treanor, D. Crellin, M. Shires, K. Smith, K. Mohee, and P. Quirke, "Colour normalisation in digital histopathology images," in *Workshop on Optical Tissue Image Anal. in Microscopy, Histopathology, and Endoscopy (MICCAI)*, 2009.
- [58] Y. Yagi, "Color standardization and optimization in whole slide imaging," *Diagnostic Pathology*, vol. 6, no. 1, p. S15, 2011.
- [59] C. M. Mosquera-Lopez and S. Agaian, "Iterative local color normalization using fuzzy image clustering," in *SPIE Defense, Security, and Sensing*, SPIE, Ed., 2013.
- [60] A. Basavanthally and A. Madabhushi, "Em-based segmentation-driven color standardization of digitized histopathology," in *SPIE Medical Imaging*. Int. Soc. for Optics and Photonics, 2013, pp. 86 760G–86 760G–12.
- [61] K. R. Castleman, *Digital image processing*. New Jersey, NJ: Prentice Hall Inc., 1996.
- [62] H. Chen, X. Jian, Y. Sun, and J. Wang, "Color image segmentation: Advances and prospects," *Pattern Recognition Lett.*, vol. 34, no. 12, pp. 2259–2281, 2001.
- [63] M. M. Galloway, "Texture analysis using gray level run lengths," *Comput. Graph. and Image Process.*, vol. 4, no. 2, pp. 172–179, 1975.
- [64] A. Chu, C. M. Sehgal, and J. F. Greenleaf, "Use of gray value distribution run lengths for texture analysis," *Pattern Recognition Lett.*, vol. 11, no. 6, pp. 415–420, 1990.
- [65] X. Tang, "Texture information in run-length matrices," *IEEE Trans. Image Process.*, vol. 7, no. 11, pp. 1602–1609, 1998.
- [66] N. Sarkar and B. Chaudhuri, "An efficient differential box-counting approach to compute fractal dimension of image," *IEEE Trans. Syst., Man, Cybern.*, vol. 24, no. 1, pp. 115–120, 1994.
- [67] B. B. Mandelbrot, "Self-affine fractals and fractal dimension," *Physica Scripta*, vol. 32, no. 4, p. 257, 1985.
- [68] J. Theiler, "Estimating fractal dimension," *JOSA A*, vol. 7, no. 6, pp. 1055–1073, 1990.
- [69] M. Ivanovici and N. Richard, "Fractal dimension of color fractal images," *IEEE Trans. Image Process.*, vol. 20, no. 1, pp. 227–235, 2011.
- [70] J. P. Monaco, J. E. Tomaszewski, M. D. Feldman, M. Moradi, P. Mousavi, A. Boag, C. Davidson, P. Abolmaesumi, and A. Madabhushi, "Detection of prostate cancer from whole-mount histology images using markov random fields," in *Workshop on Microscopic Image Anal. with Applicat. in Biology*, 2008.
- [71] J. Monaco, J. E. Tomaszewski, M. D. Feldman, M. Moradi, P. Mousavi, A. Boag, C. Davidson, P. Abolmaesumi, and A. Madabhushi, "Probabilistic pairwise markov models: Application to prostate cancer detection," in *SPIE Medical Imaging*. SPIE, 2009, pp. 725 903–725 903–12.

- [72] J. P. Monaco, J. E. Tomaszewski, M. D. Feldman, I. Hagemann, M. Moradi, P. Mousavi, A. Boag, C. Davidson, P. Abolmaesumi, and A. Madabhushi, "High-throughput detection of prostate cancer in histological sections using probabilistic pairwise markov models," *Medical Image Anal.*, vol. 14, no. 4, pp. 617–629, 2010.
- [73] K. Jafari-Khouzani and H. Soltanian-Zadeh, "Automatic grading of pathological images os prostate using multiwavelet transform," in *23rd IEEE Int. Conf. on Eng. in Medicine and Biology*, vol. 3, 2001, pp. 2545–2548.
- [74] H.-J. Yoon and C.-C. Li, "Multiridgelets for texture analysis," in *SPIE Wavelet Applicat. in Ind. Process.*, vol. 7248, 2009, p. 2.
- [75] S. Naik, S. Doyle, M. Feldman, J. Tomaszewski, and A. Madabhushi, "Gland segmentation and computerized gleason grading of prostate histology by integrating low-, high-level and domain specific information," in *MIAAB Workshop*. Citeseer, 2007.
- [76] S. Naik, A. Madabhushi, J. Tomaszewski, and M. D. Feldman, "A quantitative exploration of efficacy of gland morphology in prostate cancer grading," in *IEEE 33rd Annu. Northeast Bioengineering Conf. NEBC'07*. IEEE, 2007, pp. 58–59.
- [77] R. Farjam, H. SoltanianZadeh, K. JafariKhouzani, and R. A. Zoroofi, "An image analysis approach for automatic malignancy determination of prostate pathological images," *Cytometry Part B: Clinical Cytometry*, vol. 72, no. 4, pp. 227–240, 2007.
- [78] C. Wittke, J. Mayer, and F. Schweiggert, "On the classification of prostate carcinoma with methods from spatial statistics," *IEEE Trans. Inf. Technol. Biomed.*, vol. 11, no. 4, pp. 406–414, 2007.
- [79] J. P. Monaco, S. E. Viswanath, and A. Madabhushi, "Weighted iterated conditional modes for random field: Application to prostate cancer detection," in *Workshop on Probabilistic Models for Medical Image Anal. PMMIA (in conjunction with MICCAI)*, 2009.
- [80] V. Anand, G. Krishna, G. Mohandass, R. Hemalatha, and S. Sundaram, "Predicting grade of prostate cancer using image analysis software," in *Trends in Inform. Sciences and Computing (TISC)*. IEEE, 2010, pp. 122–124.
- [81] S. I. Fogarasi, F. M. Khan, H.-Y. H. Pang, R. Mesa-Tejada, M. J. Donovan, and G. Fernandez, "Glandular object based tumor morphometry in he biopsy samples for prostate cancer prognosis," in *SPIE Medical Imaging*. Int. Soc. for Optics and Photonics, 2011, pp. 79 633F–79 633F–8.
- [82] K. Diaz, B. Castaneda, M. L. Montero, J. Yao, J. Joseph, D. Rubens, and K. J. Parker, "Analysis of the spatial distribution of prostate cancer obtained from histopathological images," in *SPIE Medical Imaging*. SPIE, 2013, pp. 86 760V–86 760V–11.
- [83] M. Roula, A. Bouridane, F. Kurugollu, and A. Amira, "A quadratic classifier based on multispectral texture features for prostate cancer diagnosis," in *IEEE 7th Int. Symp. on Signal Process. and its Applicat.*, vol. 2. IEEE, 2003, pp. 37–40.
- [84] A. Tabesh, V. P. Kumar, H.-Y. Pang, D. Verbel, A. Kotsianti, M. Teverovskiy, and O. Saidi, "Automated prostate cancer diagnosis and gleason grading of tissue microarrays," in *SPIE Medical Imaging*. SPIE, 2005, pp. 58–70.
- [85] A. K. Jain, R. P. W. Duin, and J. Mao, "Statistical pattern recognition: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 1, pp. 4–37, 2000.
- [86] K. Ganesan, U. R. Acharya, C. K. Chua, L. C. Min, K. T. Abraham, and K. Ng, "Computer-aided breast cancer detection using mammograms: A review," *IEEE Reviews in Biomedical Eng.*, vol. 6, pp. 77–98, 2012.
- [87] L. C. Molina, L. Belanche, and A. Nebot, "Feature selection algorithms: A survey and experimental evaluation," in *IEEE Int. Conf. on Data Mining*, 2002, pp. 306–313.
- [88] C. M. Bishop, *Neural networks for pattern recognition*. Oxford: Clarendon Press, 1995.
- [89] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Trans. Inf. Theory*, vol. 14, no. 1, pp. 55–63, 1968.
- [90] G. V. Trunk, "A problem of dimensionality: A simple example," *IEEE Trans. Pattern Anal. Mach. Intell.*, no. 3, pp. 306–307, 1979.
- [91] A. Jain and D. Zongker, "Feature selection: Evaluation, application, and small sample performance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 2, pp. 153–158, 1997.
- [92] J. Hua, W. D. Tembe, and E. R. Dougherty, "Performance of feature-selection methods in the classification of high-dimension data," *Pattern Recognition*, vol. 42, no. 3, pp. 409–424, 2009.
- [93] J. Li, M. T. Manry, P. L. Narasimha, and C. Yu, "Feature selection using piecewise linear network," *IEEE Trans. Neural Netw.*, vol. 17, no. 5, pp. 1101–1115, 2006.
- [94] M. A. Tahir and A. Bouridane, "Novel round-robin tabu search algorithm for prostate cancer classification and diagnosis using multispectral imagery," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 4, pp. 782–793, 2006.
- [95] F. Glover, "Tabu search i," *ORSA J. Comput.*, vol. 1, no. 3, pp. 190–206, 1989.
- [96] —, "Tabu search ii," *ORSA J. Comput.*, vol. 2, no. 1, pp. 4–32, 1990.
- [97] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2005.
- [98] A. M. Martinez and A. C. Kak, "Pca versus lda," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 228–233, 2001.
- [99] A. Hyvarinen and E. Oja, "Independent component analysis: algorithms and applications," *Neural Networks*, vol. 13, no. 4, pp. 411–430, 2000.
- [100] L. He, L. R. Long, S. Antani, and G. Thoma, *Computer assisted diagnosis in histopathology*. iConcept Press, 2010, pp. 271–287.
- [101] R. Sparks and A. Madabhushi, "Gleason grading of prostate histology utilizing manifold regularization via statistical shape model of manifolds," in *SPIE Medical Imaging: Comput.-Aided Diagnosis*, vol. 8315, 2012, pp. 83 151J–1.
- [102] —, "Statistical shape model for manifold regularization: Gleason grading of prostate histology," *Comput. Vision and Image Understanding*, 2013.
- [103] K. Beyer, J. Goldstein, R. Ramakrishnan, and U. Shaft, *When is nearest neighbor meaningful?* Springer, 1999, pp. 217–235.
- [104] A. Ng and A. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes," *Advances in Neural Inform. Process. Syst.*, vol. 14, pp. 841–848, 2002.
- [105] C. Cortes and V. Vapnik, "Support vector machine," *Mach. Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [106] C. J. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.
- [107] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*. Cambridge University Press, 2000.
- [108] P. Khurd, C. Bahlmann, P. Maday, A. Kamen, S. Gibbs-Strauss, E. M. Genega, and J. V. Frangioni, "Computer-aided gleason grading of prostate cancer histopathological images using texton forests," in *IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*. IEEE, 2010, pp. 636–639.
- [109] K. Nguyen, A. K. Jain, and R. L. Allen, "Automated gland segmentation and classification for gleason grading of prostate tissue images," in *20th Int. Conf. on Pattern Recognition (ICPR)*. IEEE, 2010, pp. 1497–1500.
- [110] E. Alexandratou, V. Atlamazoglou, T. Thireou, G. Agrogiannis, D. Togas, N. Kavantzias, E. Patsouris, and D. Yova, "Evaluation of machine learning techniques for prostate cancer diagnosis and gleason grading," *Int. J. of Computational Intell. in Bioinformatics and Syst. Biology*, vol. 1, no. 3, pp. 297–315, 2010.
- [111] J. Ghosh, *Multiclassifier systems: Back to the future*. Springer, 2002, pp. 1–15.
- [112] B. Tso and M. P. M., *Pattern recognition principles*, 2nd ed. CRC Press, 2009, pp. 41–75.
- [113] A. Madabhushi, J. Shi, M. Feldman, M. Rosen, and J. Tomaszewski, *Comparing ensembles of learners: Detecting prostate cancer from high resolution mri*. Springer, 2006, pp. 25–36.
- [114] S. Doyle, A. Madabhushi, M. Feldman, and J. Tomaszewski, "A boosting cascade for automated detection of prostate cancer from digitized histology," in *Medical Image Computing and Comput.-Assisted Intervention MICCAI*. Springer, 2006, pp. 504–511.
- [115] S. Doyle, M. Feldman, J. Tomaszewski, N. Shih, and A. Madabhushi, "Cascaded multi-class pairwise classifier (cascampa) for normal, cancerous, and cancer confounder classes in prostate histology," in *IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*. IEEE, 2011, pp. 715–718.
- [116] S. Doyle, M. D. Feldman, N. Shih, J. Tomaszewski, and A. Madabhushi, "Cascaded discrimination of normal, abnormal, and confounder classes in histopathology: Gleason grading of prostate cancer," *BMC Bioinformatics*, vol. 13, no. 1, p. 282, 2012.
- [117] M. Galar, A. Fernndez, E. Barrenechea, H. Bustince, and F. Herrera, "An overview of ensemble methods for binary classifiers in multi-class problems: Experimental study on one-vs-one and one-vs-all schemes," *Pattern Recognition*, vol. 44, no. 8, pp. 1761–1776, 2011.
- [118] H. Drucker, C. Cortes, J. L. D., Y. LeCun, and V. Vapnik, "Boosting and other ensemble methods," *Neural Computation*, vol. 6, no. 6, pp. 1289–1301, 1994.
- [119] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *13th Int. Conf. on Mach. Learning*, 1996, pp. 148–156.

- [120] M. D. DiFranco, G. OHurley, E. W. Kay, R. W. G. Watson, and P. Cunningham, "Ensemble based system for whole-slide prostate cancer probability mapping using color texture features," *Computerized Medical Imaging and Graph.*, vol. 35, no. 7, pp. 629–645, 2011.
- [121] A. Greenblatt, C. M. Mosquera-Lopez, and S. Agaian, "Quaternion neural networks applied to prostate cancer gleason grading," in *IEEE Int. Conf. on Syst., Man, and Cybern. (SMC)*. IEEE, 2013, p. 6.
- [122] D. J. Hand, "Assessing the performance of classification methods," *Int. Statistical Review*, vol. 80, no. 3, pp. 400–414, 2012.
- [123] P. Refaeilzadeh, L. Tang, and H. Liu, *Cross-validation*. Berlin: Springer, 2009.
- [124] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Stat. Surveys*, vol. 4, pp. 40–79, 2010.
- [125] A. Fielding and J. F. Bell, "A review of methods for the assessment of prediction errors in conservation presence/absence models," *Environmental Conservation*, vol. 24, no. 1, pp. 38–49, 1997.
- [126] A. G. Lalkhen and A. McCluskey, "Clinical tests: Sensitivity and specificity," *Continuing Educ. in Anaesthesia, Critical Care, and Pain*, vol. 8, no. 6, pp. 221–223, 2008.
- [127] D. J. Hand and R. J. Till, "A simple generalisation of the area under the roc curve for multiple class classification problems," *Mach. Learning*, vol. 45, no. 2, pp. 171–186, 2001.
- [128] K. Tang, R. Wang, and T. Chen, "Towards maximizing the area under the roc curve for multi-class classification problems," in *Assoc. for the Advancement of Artificial Intell. (AAAI)*, 2011.
- [129] S. Kothari, J. H. Phan, T. H. Stokes, and M. D. Wang, "Pathology imaging informatics for quantitative analysis of whole-slide images," *J. of the Amer. Medical Informatics Assoc.*, vol. 20, no. 6, pp. 1099–1108, 2013.
- [130] M. N. Gurcan, L. E. Boucheron, A. Can, A. Madabhushi, N. M. Rajpoot, and B. Yener, "Histopathological image analysis: A review," *IEEE Reviews in Biomedical Eng.*, vol. 2, pp. 147–171, 2009.



Clara Mosquera-Lopez received her B.S. degree in Electronics Engineering and her M.S. degree in Management of Technology from the Universidad Pontificia Bolivariana at Medellin (Colombia) in 2007 and 2010 respectively. Clara is currently a doctoral student in the Department of Electrical and Computer Engineering at The University of Texas at San Antonio. Her research interests include a broad spectrum of topics in signal processing, and machine learning. At this moment, her research is focused on

the development of image processing algorithms for quantitative analysis of biomedical images and pattern recognition methods applied to computer-aided cancer diagnosis.



Sos Agaian is a Peter T. Flawn Professor of Electrical and Computer Engineering at the University of Texas, San Antonio, and a professor at the University of Texas Health Science Center, San Antonio. Dr. Agaian received his M.S. degree (summa cum laude) in Mathematics and Mechanics from Yerevan State University, Armenia; his Ph.D. in Mathematics and Physics from the Steklov Institute of Mathematics, Russian Academy of Sciences (RAS), Moscow; and his Eng.Sc.D. from the Institute of Control Sciences, RAS. He has authored more than 500 scientific

papers, seven books, and holds 14 patents. He is a Fellow of the International Society for Photo-Optical Instrumentations Engineers, the American Association for the Advancement of Science (AAAS), and Imaging Sciences and Technology (IS&T). He also serves as a foreign member of the Armenian National Academy. He is the recipient of MAESTro Educator of the Year, sponsored by the Society of Mexican American Engineers. The technologies that Dr. Agaian invented have been adopted by multiple institutions, including the U.S. government, and commercialized by industry. He is an Editorial Board Member for the Journal of Pattern Recognition and Image Analysis and an Associate Editor for several journals, including the Journal of Electronic Imaging (SPIE, IS&T) and System (Elsevier). His research interests include multimedia processing, imaging systems, information security, artificial intelligence, computer vision, 3D imaging sensors, fusion, and biomedical and health informatics.



Alejandro Velez-Hoyos is an Associate Professor at the Universidad de Antioquia and the Universidad Pontificia Bolivariana, Medellin, Colombia. He is also part of the Department of Pathology at the Pablo Tobon Uribe Hospital and Dinamica IPS. Dr. Velez-Hoyos received his M.D. from the Universidad Pontificia Bolivariana, and after a residency in pathology, he received his specialization degree from the Universidad de Antioquia. He has authored several dozen scientific papers, six book chapters, and he has edited a textbook on pathology. Dr.

Velez-Hoyos has attended several training sessions in the U.S. on cytology, uropathology, dermatopathology, immuno- histochemistry, and endocrine and soft-tissue pathology. He is the vice president of the Colombian Pathology Society and the vice president of the Colombian Cytology Society. His research interests include methods for tissue sampling, and diagnosis of various types of cancerous lesions through microscopic analysis of biopsy specimens. He actively participates in clinical studies on skin, prostate, and breast cancer, among other diseases.



Ian Thompson is the director of the Cancer Therapy and Research Center and a Professor in the Department of Urology in the School of Medicine at the University of Texas Health Science Center in San Antonio. Dr. Thompson received his undergraduate degree from West Point and his M.D. from Tulane University, New Orleans, LA. After a residency in urology in San Antonio, he completed a fellowship in Urologic Oncology at the Memorial Sloan Kettering Cancer Center in New York. He has published over 400 scientific papers, several dozen

book chapters, and has edited five textbooks in medicine and surgery. He previously served as chair of the Residency Review Committee for Urology and currently serves as vice chair of the Early Detection Research Network of the National Cancer Institute. He previously served as the president of the Society of Urologic Oncology. He currently serves as chair of the Genitourinary Committee of the Southwest Oncology Group, the largest clinical trials organization supported by the National Cancer Institute. Dr. Thompson has served as a Visiting Professor at most major academic institution in the U.S. as well as at many leading cancer centers in Europe, Asia, Central and South America, as well as Australia.