

MAPS: A Quantitative Radiomics Approach for Prostate Cancer Detection

Andrew Cameron, Farzad Khalvati*, *Member, IEEE*, Masoom A. Haider, and Alexander Wong, *Member, IEEE*

Abstract—This paper presents a quantitative radiomics feature model for performing prostate cancer detection using multiparametric MRI (mpMRI). It incorporates a novel tumor candidate identification algorithm to efficiently and thoroughly identify the regions of concern and constructs a comprehensive radiomics feature model to detect tumorous regions. In contrast to conventional automated classification schemes, this radiomics-based feature model aims to ground its decisions in a way that can be interpreted and understood by the diagnostician. This is done by grouping features into high-level feature categories which are already used by radiologists to diagnose prostate cancer: Morphology, Asymmetry, Physiology, and Size (MAPS), using biomarkers inspired by the PI-RADS guidelines for performing structured reporting on prostate MRI. Clinical mpMRI data were collected from 13 men with histology-confirmed prostate cancer and labeled by an experienced radiologist. These annotated data were used to train classifiers using the proposed radiomics-driven feature model in order to evaluate the classification performance. The preliminary experimental results indicated that the proposed model outperformed each of its constituent feature groups as well as a comparable conventional mpMRI feature model. A further validation of the proposed algorithm will be conducted using a larger dataset as future work.

Index Terms—Computer-aided detection, feature models, multiparametric MRI (mpMRI), prostate cancer detection, prostate MRI, radiomics.

I. INTRODUCTION

PROSTATE cancer is the most diagnosed form of cancer in Canadian men, with roughly 23 500 new cases and 4000 deaths in 2014 [1]. Although it is the third leading cause of cancer death in men, prostate cancer has good prognosis if detected early [2]. Thus, effective and reliable screening methods are crucial for high survivability rate.

Conventional methods of screening for prostate cancer include the prostate-specific antigen (PSA) test, transrectal ultrasound (TRUS), and magnetic resonance imaging (MRI). Some of these methods, such as the PSA test, have been criticized for oversensitivity, causing expensive and often painful follow-

up procedures [3]–[5]. Imaging-based methods produce image series which require expert review from an experienced diagnostician, who looks for subtle features to make a diagnosis. When multiple types of imaging are performed, they must be registered to a common coordinate system. Even when registered, it can be challenging to search for features across multiple imaging modalities at once leading to intra and interobserver variability in diagnosis.

The European Society of Urogenital Radiology recently introduced a set of guidelines for interpreting several MRI image series together known as Prostate Imaging, Reporting, and Diagnosis System (PI-RADS). This scoring system aims to improve consistency between diagnosticians by providing a common set of criteria to consider when diagnosing a case. However, as the features laid out in PI-RADS are subjective, it is not immune to interobserver variability [6].

Research on automatic methods for detecting and localizing prostate cancer consists mostly of ensemble classifiers trained on huge numbers of features. Furthermore, the feature calculations are usually performed on all local windows of pixels (e.g., 5×5) that cover the prostate gland. The classification decisions made by such automatic systems can be difficult for a radiologist to reason about, consisting of a binary output or some kind of probability or confidence score. Without a mechanism to reason about why an automatic classifier computes the result it does, clinicians tend to distrust them.

The aim of this paper is twofold. First to incorporate comprehensive quantitative feature sequences or radiomics as a high-dimensional mineable feature space that can be used as both detection and prognostic tools for prostate cancer [7]. To achieve this goal, instead of applying the feature extraction function on all local windows without discrimination, we first identify the candidate regions which might contain cancerous pixels from which the features are extracted. Second, we introduce a framework to help bridge the gap between the quantitative characterization and clinical interpretation of prostate cancer diagnosis. By grouping classifier features together, we aim to create a diagnostic support system that clinicians can leverage to augment their understanding rather than supplant it. Beyond stating that a region is suspected to be cancerous, a system built on this framework can explain that the region has an unusual morphology, for instance. With intuitive features, classification decisions can be more easily justified and explained to medical practitioners, while simultaneously providing more consistently quantifiable measures. In this paper, we introduce a novel strategy for a radiomics feature model for prostate cancer analysis that combines:

- 1) a tumor candidate identification scheme based on multiparametric MRI (mpMRI) and morphology, and

Manuscript received May 12, 2015; revised August 4, 2015; accepted September 25, 2015. Date of publication October 1, 2015; date of current version May 18, 2016. This work was supported by the Ontario Institute of Cancer Research, Canada Research Chairs programs, the Natural Sciences and Engineering Research Council of Canada, and the Ministry of Research and Innovation of Ontario. Asterisk indicates corresponding author.

A. Cameron and A. Wong are with the Department of Systems Design Engineering, University of Waterloo.

*F. Khalvati is with the Department of Medical Imaging, University of Toronto, Toronto, ON M5T 1W7, Canada, and also with the Sunnybrook Research Institute, Toronto, ON M4N 3M5, Canada (e-mail: farzad.khalvati@sri.utoronto.ca).

M. A. Haider is with the Department of Medical Imaging, University of Toronto, and Sunnybrook Research Institute.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBME.2015.2485779

- 2) a comprehensive feature model comprising groups of features evaluating tumor candidate morphology, asymmetry, physiology, and size (MAPS).

By presenting results as a collection of factors that the clinicians are already used to for evaluating prostate cancer, the proposed system should foster user acceptance, and the use of objective quantitative features can help to reduce inter- and intraobserver variability.

II. RELATED WORK

In this section, a background review on prostate cancer clinical workflow and computer-aided detection tools is given.

A. Clinical Workflow

The PSA test measures the concentration of specific markers in the patient's blood of which high levels indicate high risk for prostate cancer. Although PSA is used as a biomarker for prostate cancer, its use has been criticized for having an unacceptably high occurrence of false positives, causing healthy patients to undergo expensive and sometimes invasive tests to determine whether the PSA test result was simply a false positive [8], [9]. The PSA test is, therefore, recommended particularly for patients with a predisposition, such as those with a family history of prostate cancer. Following a positive PSA test, a series of images might be acquired followed by a systematic TRUS-guided biopsy where multiple samples are collected. Although core biopsies are very accurate, they are intrinsically invasive, causing significant discomfort to the patient and exposing him to possibly unnecessary surgical risk. Furthermore, some tumors are isoechoic and respond to ultrasound similarly to the surrounding tissues, making TRUS-guided biopsies even more difficult as the surgeon must attempt to collect tissue samples from invisible tumors. Depending on the results, further biopsies may be indicated, bearing further surgical risk. Following any biopsies, the patient may be treated with focal therapy, chemotherapy, or even a radical prostatectomy, depending on the severity of his cancer.

The European Society of Urogenital Radiology recently introduced the PI-RADS [10], modeled after the BI-RADS system for breast cancer reporting. PI-RADS breaks the decision process into several component scores, providing guidelines on the scoring of the various recommended modalities.

B. Multiparametric MR Imaging

mpMRI has shown considerable promise for diagnosis, especially the combination of T2-weighted MRI (T2w) and diffusion-weighted MRI (DWI). T2w gives the clinicians the ability to see relatively high-resolution imagery of the prostate and surrounding tissue while allowing detection of subtle structural features. Recent research has focused on DWI and the DWI-derived apparent diffusion coefficient, which quantifies the diffusion of water molecules through tissue and has been shown to delineate between healthy and cancerous tissue [11], [12].

Other MRI modalities include dynamic contrast-enhanced MRI (DCE-MRI), in which a contrast agent is introduced and the

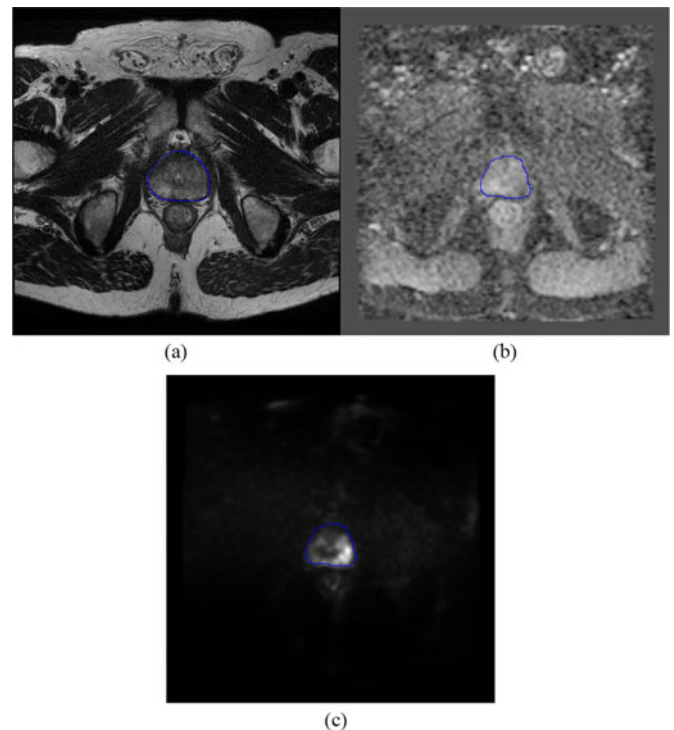


Fig. 1. Sample mpMRI slices of the same prostate gland, using T2w and DWI. The lower-right corner of the prostate gland, outlined in blue, which appears dark on ADC and bright on CDI, was indicated as tumorous by an expert. (a) T2-weighted. (b) DWI: ADC. (c) DWI: CDI.

subject is repeatedly imaged while the contrast agent is absorbed into the tissue. The contrast agent uptake behavior of any given voxel of tissue is modeled and the time-series imagery is used to estimate tissue parameters such as onset time, time to peak, and washout. Magnetic resonance spectroscopy (MRS) is a rarely-used MRI modality, known as magnetic resonance spectroscopic imaging or *in vivo* nuclear magnetic resonance spectroscopy, which provides information about the chemical composition of the tissue. Correlated diffusion imaging (CDI) [13], [14] is a recently developed modality which, similar to ADC, is derived from DWI, and has shown promise in differentiating between healthy and tumorous tissue [15], [16]. Sample mpMRI scans are shown in Fig. 1 where T2w, ADC map, and CDI sample are shown in Fig. 1(a)–(c), respectively. As it is seen, a tumorous region on the lower-right corner of the prostate gland appears dark on ADC and bright on CDI.

The aforementioned PI-RADS system puts guidelines into place for the interpretation of mpMRI by diagnosticians. PI-RADS specifies the use of several different MRI modalities, including T2w, DWI, MRS, and DCE-MRI, each of which is scored, Likert-style, from 1 to 5. Different scoring guidelines are presented for each modality, so while MRS features are relatively straightforward to interpret by reading off the relative heights of two peaks from a spectrogram, other modalities, such as T2w, may have very subtle features. Although these features are difficult to interpret without an understanding of prostate physiology, PI-RADS also provides a scale for performing qualitative analysis of MRS which makes the meaning of the scores for all the modalities clear: a score of 1 indicates that

“clinically significant disease is highly unlikely to be present,” while a score of 5 indicates that “clinically significant cancer is highly likely to be present.”

Unfortunately, despite the wealth of information provided by various MRI modalities, prostate cancer remains difficult to distinguish from other conditions, notably benign prostatic hyperplasia (BPH). The additional tissue growth in the transitional zone (TZ) occurring due to BPH causes similar MRI signal intensity (SI) changes as prostate cancer. Furthermore, attempts to differentiate cancerous tissue from healthy tissue in the central gland have failed to show differences in mean T2w or ADC values [17]. However, since BPH only occurs in the central gland, this issue only arises in the minority of cases which occur elsewhere than the peripheral zone (PZ).

C. Computer-Aided Prostate Cancer Detection

Automatic approaches to prostate cancer detection can be grouped into methods which attempt to segment the prostate into discrete regions of concern, and methods which instead assign some kind of numeric score to each voxel in the prostate. A variety of features and classification strategies is used to produce these scores, whether the scores are region based or voxel based. Features and classifiers will be discussed later, but for region-based methods, an additional step must first be performed, namely the segmentation of the prostate into regions of concern.

1) *Segmenting Tumor Candidates*: The problem of segmenting tumor candidates is not yet well explored; most classifiers built on MRI intensity features did not perform any tumor candidate segmentation, but instead computed voxelwise feature statistics, using an ROC curve analysis, or output from support vector machines (SVMs), RVMs, or MRFs [17]–[26]. Some high-level classifiers leveraged the added detail apparent in stained histology scans to segment different gland structures in the prostate, namely lumen regions, the cytoplasm surrounding the lumen, and the nuclei of the surrounding cells [27], [28]. Classifiers were trained on features computed from these regions to predict the Gleason grade of prostate tissue, but MRI modalities do not have adequate resolution to observe these structures, so these features cannot be leveraged for MRI tumor candidate identification.

2) *Feature Extraction*: Detecting prostate cancer automatically has been a popular area of research and to date, these methods have focused primarily on low-level features. Conventional mpMRI prostate cancer analysis approaches use one or more of the MRI modality values or derived values, including ADC and T2w, with no complementary high-level features [23]. Published values for classification accuracy using these low-level features ranges from 64% to 89% [11], [18], [19], [23], [25], [29].

More sophisticated approaches calculate large numbers of low-level features and then combine them in various ways; low-level features considered in such approaches include statistical features (e.g., local variance) and change-of-basis features, such as Gabor filter coefficients [30], discrete cosine transform (DCT) coefficients, and textural features (both first and second order) such as those derived from cooccurrence matrices (CMs). Chan

et al. [25] reported that compared to using MRI intensity features exclusively, DCT and CM features improved accuracy between 6.2% and 11.0% (accuracy of 72.9% without those features, and 79.1%–83.9% with them). Madabhushi *et al.* [30] reported a positive predictive value of approximately 21% when using a combination of first- and second-order statistical features, gradient features, and Gabor features within an ensemble classification framework. However, since the low-level features lack semantic meaning, the classification decisions made by such approaches are difficult to justify to radiologists and clinicians, who typically use subjective heuristics to diagnose patient cases. Although the latter approach is easy to understand and justify, its subjective nature makes it highly inconsistent.

In contrast to low-level features, which typically characterize some kind of voxelwise physiology characteristic, high-level features often exploit *a priori* knowledge of the problem to quantify high-level domain concepts. High-level features have been used in detecting and localizing prostate cancer; however, their use is still limited, mostly to analysis of histopathology, which necessitates invasive surgery to acquire the tissue samples subsequently scanned at high resolution; for instance, Naik *et al.* [27] demonstrated that classifiers built using high-level features designed using domain knowledge are effective in diagnosing prostate tissue as cancerous and distinguishing cancerous tissue between two grades on the Gleason scale, with classification accuracy in the range of 86%–95%. Notable among these high-level features was the use of region-based features; the added magnification achieved by histopathology scans allows the observation and segmentation of structures that are too small to be apparent via MRI. Naik *et al.* computed features on the segmentation boundaries of these structures, such as smoothness and area, and used these features to perform classification. The use of such region-based features is not currently well explored for MR imagery. These particular high-level features cannot be computed at all for mpMRI, as the structures used are too small to be resolved. Haider *et al.* [23] described a scoring scheme for diagnosing prostate cancer based on high-level features observed in T2w and ADC imagery. However, these features are not easily quantifiable and are intended for use by trained radiologists and not classification algorithms.

As an initial attempt for a radiomics-driven feature model for prostate cancer detection, some preliminary results were presented in [31]. In this paper, we have significantly extended upon the work presented in [31]. First, the previous work only used T2w and ADC for texture feature calculations, whereas this paper also incorporates DWI and CDI for both textural and intensity features. Second, in [31], we only used textural and morphological features, while here, in addition to these two classes of features, we incorporate size and asymmetry features. Using more data and features (more images using DWI and CDI and more features) makes the texture feature model more accurate, in terms of the cancer detection accuracy. Third, only five patients datasets were used in the previous work, whereas 13 patients datasets have been used in this paper allowing for a better validation of the proposed texture feature model.

We have used the proposed radiomics-driven model in this paper for prostate cancer detection. Given that the proposed

model incorporates candidate tumor regions for feature calculations, it can be augmented for prognostic of prostate cancer as well. Studies on lung and head-and-neck cancer patients have confirmed the prognostic power of radiomics features for patient outcome prediction for personalized medicine [32], [33]. However, the prognostic capability of radiomics features has not been fully investigated for prostate cancer and this is a novel approach for identifying prostate tumors phenotypes.

D. Classification

To date, several approaches of classification have been applied in various prostate cancer imaging contexts. Among approaches investigating mpMRI, SVMs are widely used [11], [25], [29], [34], while other methods applied maximum likelihood classification [23], [25], logistic regression [19], or Markov random fields (MRFs) [18], [29]. Approaches attempting classification of histopathology data tended to use more features and, thus, more sophisticated methods of combining those features; for instance, combining many weak Bayesian predictors using an ensemble method such as AdaBoost [24], [30], or using a manifold learning technique, such as graph embedding or local linear embedding to reduce the size of the feature space passed into an SVM classifier [27], [35], [36]. Interestingly, the mpMRI approach with the greatest number of features reported greatest classification accuracy with the comparatively simple Fisher's linear discriminant classifier [25].

Despite the recent introduction of PI-RADS, systematic prostate cancer evaluation schemes are still in their infancy. Although much research has been carried out on the use of various modalities and numerous features for detecting prostate cancer, most of the automatic approaches do not support their decisions with a similar framework as those used by radiologists diagnosing prostate cancer manually. There is no attempt to map automatic classification decisions to the language familiar to diagnosticians in their work. The frameworks which do group features together in this manner, such as those presented by Haider *et al.* [23] or in PI-RADS [10], are meant for interpretation by radiologists. They are not quantifiable, and due to their subjective nature lack consistency in interpretation from one session to the next and from one clinician to the next. The field is ripe for a comprehensive feature model which enables consistent quantifiable feature analysis in a manner that is familiar to diagnosticians. Our proposed feature model has been designed with the aim of providing a quantitative model for PI-RADS scheme to make it a more systematic approach leading to consistent diagnosis. To our best knowledge, this is the first attempt in quantifying the PI-RADS scoring scheme for prostate cancer.

III. PROPOSED MAPS FEATURE MODEL

The proposed model consists of an initial tumor candidate identification scheme followed by the MAPS feature model used to score the candidate regions. The goal of the proposed model is to incorporate high-level radiomics features using candidate tumor regions to construct a high-dimensional feature space that can be mined for different purposes such as detection or prognosis of cancer. In the proposed method, initial identification

of candidate tumor regions is automatically performed using mpMRI and region morphology. After candidate regions are identified by the automatic tumor candidate identification algorithm, numerous features are extracted from the various mpMRI modalities. These features are grouped into high-level categories according to the type of characteristics they describe. The Morphology feature group describes region shape, whether the boundary is smoothly varying or jagged, whether it consists of small protrusions or a focal mass, etc. The Asymmetry feature group describes the extent to which the two halves of the region appear similar when split along its major and minor axes. Tissue characteristics are captured in the Physiology group, which includes mean SI for all the MRI modalities as well as a series of textural features and filter responses. Finally, the Size feature group captures the physical extent of the region. Collectively, these feature groups provide the diagnostician with an improved ability to interpret the system's decisions.

Fig. 2 illustrates a detailed block diagram of the proposed radiomics feature model for a given slice where ten tumor candidate regions have been identified by the model. Four groups of features (Morphology, Asymmetry, Physiology, and Size) are also shown, which are used to construct the radiomics sequences. The tumor candidate identification algorithm, as well as each of the feature groups and all their constituent features, are described in detail in the following sections.

A. Automatic Tumor Candidate Identification

Tumor candidate region identification is the first step in radiomics-driven diagnosis of prostate cancer. An automated system working from straightforward screening procedures can quickly and thoroughly scan all the imagery for a given case, flagging screened regions for review by a radiologist, producing screening results with less interobserver variability. Tumor candidate regions were identified automatically in the proposed system using guidelines for clinical mpMRI prostate cancer screening by a radiologist [23]. Tissue regions satisfying these criteria were grouped into connected regions and analyzed further with the feature model detailed in Section III-B. In particular, diffusion characteristics and morphology were used to automatically identify candidate regions in the proposed system.

As low ADC values are associated with tumorous tissue [23], tissue with ADC values below a particular threshold were automatically identified by the proposed system as possible tumor candidates. Considering that the priority of the candidate identification phase is to highlight regions for consideration, rather than to eliminate them, the threshold value used in [23] was relaxed to capture more regions. All voxels with an ADC value less than $700 \times 10^{-6} \text{ mm}^2/\text{s}$ were flagged as possible tumor candidates. These flagged voxels were then grouped into connected regions, and a second phase was used in the automatic tumor candidate identification algorithm to take morphology into account. In particular, to eliminate single islands smaller than most tumors, all connected regions larger than the voxel resolution, which is usually about 1.6 mm^2 , were considered as the final set of tumor candidate regions.

Sample candidate identification results are shown in Fig. 3. The tendency of the identification scheme to identify many false

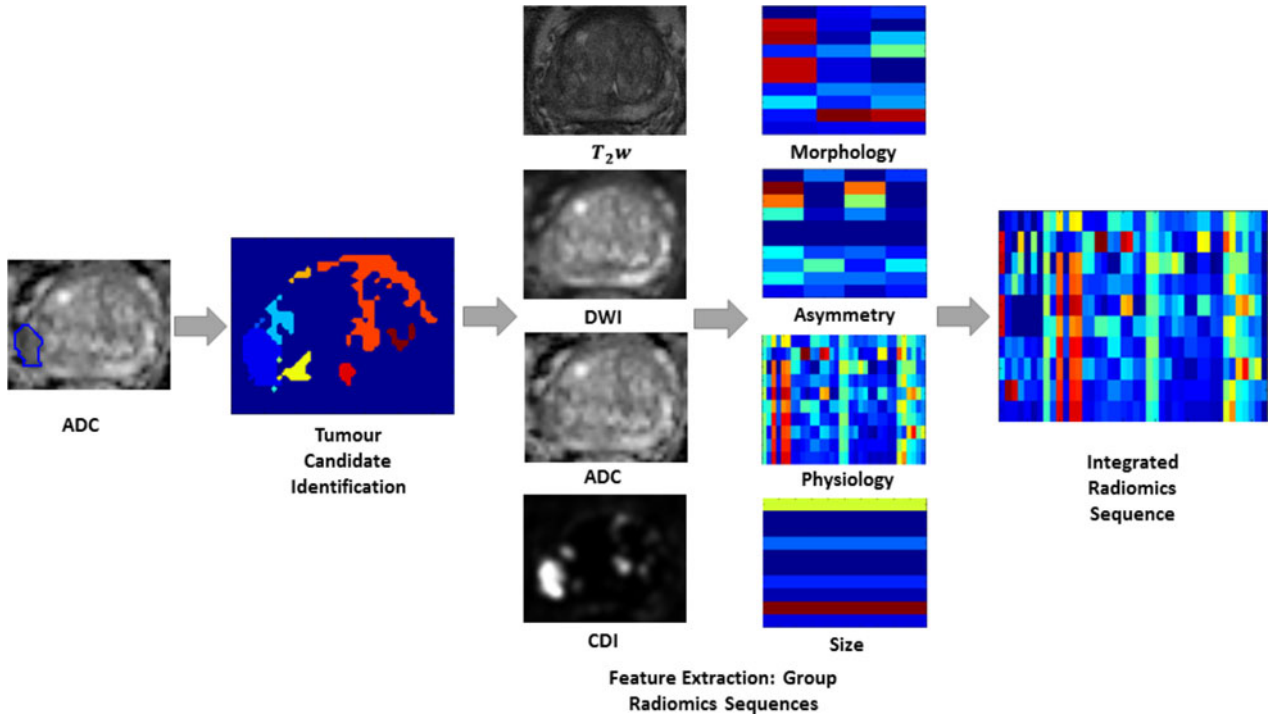


Fig. 2. Detailed block diagram of the proposed radiomics feature model. For each feature group, rows and columns represent tumor candidate regions and feature values, respectively.

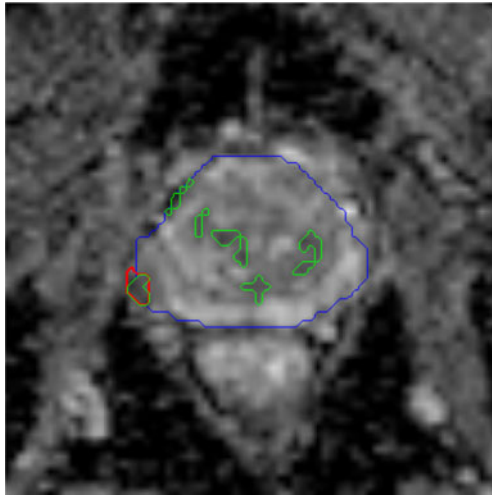


Fig. 3. Sample tumor candidate identification results; prostate gland (blue), true positive regions (red), and regions identified by the tumor candidate identification algorithm (green), overlaid on the ADC image.

positives is evident, although the region labeled as tumorous in the ground truth was successfully identified, so were many other regions. This practice is common in clinical screening, which is effectively the purpose of tumor candidate identification: to identify all suspects so that a more specific test can eliminate the false positives. This helps to reduce the number of false negatives or regions of concern missed by the test that are nevertheless cancerous.

B. MAPS Feature Model

In this section, we discuss the proposed radiomics-based feature model, which provides a high-level description of the tumor

candidate regions. Each feature group in the MAPS model (Morphology, Asymmetry, Physiology, Size) is expounded along with its component features.

1) *Morphology*: The PI-RADS guidelines make mention of morphology in several areas; the appearance of a “focal mass,” for instance, is an indicator for a higher PI-RADS score in several categories. Viewed with T2w, a “bulging” effect in the PZ is highly indicative of cancer, while in the TZ cancerous regions are “usually lenticular or water-drop shaped.” Viewed with DWI, focal regions as well as “linear, triangular, or geographical [regions]” are important biomarkers [10]. In the proposed feature model, a set of features is computed to characterize the morphology of the tumor candidate region. The morphological features used in the MAPS model capture structural information about a candidate region by applying operations which smooth the shape of the region boundary. Regions with little morphological irregularity undergo little change with the smoothing operator, while regions with highly irregular shapes will see a drastic difference. The various smoothing operators used to compute morphology features capture different characteristics about the region shape. In total, three morphological features are computed from the region boundary.

The first feature in this group is the normalized difference in area between the morphological closing of the region and the morphological opening of the region, using an identical disk structuring element for both operations [37]

$$f_1^M = \frac{A_{\text{closed}} - A_{\text{opened}}}{A_{\text{initial}}}. \quad (1)$$

Here, A denotes the area of a region. Peaks and valleys in the border of the region will cause the area to increase after closing, while it will decrease after opening (i.e., $A_{\text{closed}} \geq A_{\text{opened}}$);

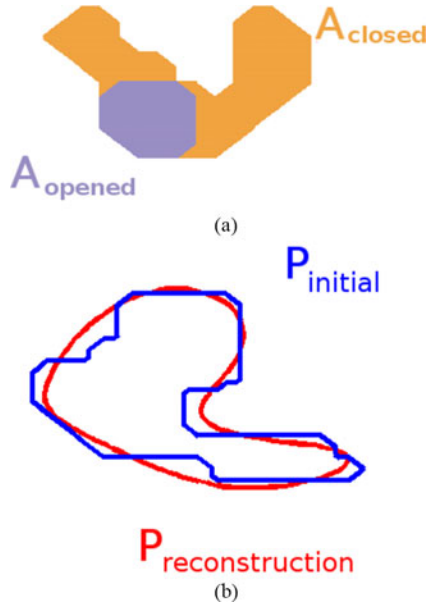


Fig. 4. Illustrations of the morphological features. (a) Large difference in areas between the closing and opening contours results in a high feature value for this region. (b) Perimeter feature illustration, a candidate region contour, shown in blue, has a perimeter similar in length to that of the region's low-frequency Fourier reconstruction, shown in red, resulting in a low feature value for this region.

therefore, regions with very irregular borders will have a greater difference between these two values, and the feature value will be greater. Smaller feature values then correspond to regions with borders which do not feature sharp peaks and valleys. An example region is shown in Fig. 4(a).

The second morphological feature compares the length of the region's perimeter before and after eliminating the high-frequency components in the Fourier space, and normalizing the difference

$$f_2^M = \frac{|P_{\text{initial}} - P_{\text{reconstruction}}|}{P_{\text{initial}}} \quad (2)$$

where each P is the perimeter of the region denoted by its subscript. Since high-frequency components capture rapid changes in the shape of the region, this feature will be greater for regions with rapidly varying boundaries than for those with smooth slowly varying boundaries [37]. An example region is shown in Fig. 4(b).

The third morphological feature is similar to the second, in that it examines the difference between two Fourier reconstructions of the region boundary, one at a low frequency and one at a higher frequency. However, it computes difference in area, rather than perimeter

$$f_3^M = \frac{|A_{\text{initial}} \oplus A_{\text{reconstruction}}|}{|A_{\text{initial}} \cup A_{\text{reconstruction}}|} \quad (3)$$

where \oplus denotes the set symmetric difference, finding voxels which are in the low-frequency reconstruction or the high-frequency one, but not both. This value is normalized by the area of the union of both reconstructions, denoted by \cup .

2) *Asymmetry*: The PI-RADS guidelines suggest that asymmetry in lesions denotes an increased likelihood of cancerousness [10]. The Asymmetry feature group represents the degree

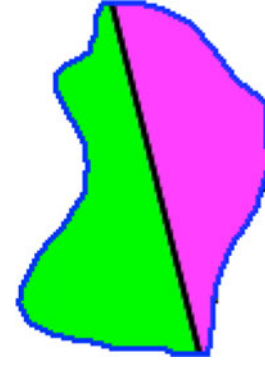


Fig. 5. Asymmetry feature illustration: A sample region contour is shown in blue, with the major axis drawn in black. The difference in area between the green and magenta regions is taken and then normalized in two different ways, producing two different features. The process is then repeated using the minor axis.

of bilateral symmetry exhibited by a candidate region. In order to quantify the region's symmetry, the region is split in half along an axis passing through the region's center of mass. The halves to either side of this axis are then compared by taking the difference in their areas, and normalizing it

$$f_4^A = \frac{A_{\text{large}} - A_{\text{small}}}{A_{\text{normalize}}} \quad (4)$$

where A_{large} and A_{small} represent the areas of the region halves, chosen such that $A_{\text{large}} \geq A_{\text{small}}$, and $A_{\text{normalize}}$ represents the area of the region used to normalize the difference. Four different Asymmetry features (illustrated in Fig. 5) are computed in total by choosing either the major or the minor axis to split the region, and by choosing either the entire region area or the area of the smaller half region as $A_{\text{normalize}}$. As regions may be symmetric about one axis but not the other, using both the major and minor axes to measure region asymmetry allows the feature model to distinguish a greater variety of regions.

3) *Physiology*: Cancerous tissue is often indicated by a number of physiological factors. PI-RADS includes several physiological MRI biomarkers as criteria, including "uniform high SI," "homogeneous low SI," or "heterogeneous [...] adenoma." The Physiology feature group attempts to capture such physiological factors. In addition to raw mpMRI signal values, many derived features are computed from the mpMRI data. Many of these features emulate various parts of the human visual system, such as second-order textural feature images, or Gabor filter responses.

While the features in the other groups can be computed from the region boundary, Physiology features are computed on a per-voxel basis. For each feature, the mean feature score was taken from all the voxels in the region, and the mean value used as the feature score for the region.

a) *Raw mpMRI Signal Intensities* The most directly interpretable features consist of the MRI SI values for the T2w, ADC, DWI, and CDI modalities. Important T2w biomarkers in the PI-RADS scheme include areas in the PZ of "uniform high SI" and "homogeneous low signal [intensity] focus/mass," as well as "areas of [...] homogeneous low SI" in the TZ [10]. DWI measures the ability water to diffuse through tissue, and involves repeated imaging while varying several acquisition parameters, collectively referred to as the b -factor or b -value, to

vary the degree of diffusion allowed to occur during acquisition. The DWI acquisitions at the various b -values are used collectively to estimate the ADC. Unlike DWI, which induces signal decay through diffusion (resulting in dark areas where water can diffuse), the ADC measures the rate of diffusion of tissue, so poorly diffusing areas appear dark. Raw DWI and ADC intensity values are used in PI-RADS where areas of concern are indicated by “focal areas of reduced ADC” and “intense SI on high b -value images” [10]. CDI is not yet widely used, but evidence suggests that CDI SI is effective at delineating tumors from healthy prostate tissue [13].

b) *Textural Features* Texture features were included to capture the different textural characteristics between cancerous and healthy tissue [30]. They play an important role in PI-RADS, which uses properties including uniformity and homogeneity (or heterogeneity) in diagnosis of T2w imaging of the TZ [10]. A series of local statistical features were extracted from the mpMRI images in order to quantify textural characteristics of imaged tissue [30]. Feature values for each region were obtained by averaging texture image values over the region.

For each candidate region, a set of low-level texture features were computed. Texture features quantify the occurrence of various patterns in an image. Whereas raw mpMRI values indicate the signal response from each respective voxel, texture images derived from the MR images indicate how the signal response at any given voxel relates to those around it. For example, the mean texture image can be computed using a $k \times k$ sliding window, where each voxel in the mean texture image consists of the mean value of the k^2 voxels in the $k \times k$ window centered around that voxel in the original image. As four different raw modalities were used (T2w, DWI, ADC, CDI), each texture operation described below constitutes four features in the Physiology group.

Texture features may be classified as first-, second-, or higher-order features, depending on the number of voxels in a particular spatial arrangement being compared in the computation of each class of feature. The first-order features compare a single voxel to some global value, e.g., the median feature. By contrast, second-order features such as homogeneity are computed using a CM, which is a measure of the distribution of pairs of intensities in an image. Higher-order textural features compare the relationship between more than two voxels.

First-Order Features: The first-order textural features in the MAPS model consist of the median, standard deviation, and mean absolute deviation. Each of these three features is computed on the T2w, ADC, DWI, and CDI modalities using a sliding window. Two window sizes were used for a total of 24 first-order texture features. The median and standard deviation feature images are straightforward to compute; each voxel (i, j) in the feature image is the result of evaluating the appropriate function (e.g., median) on $\mathcal{N}_{i,j}$, the set of voxel values belonging to the sliding window around voxel (i, j) in the original image I

$$f_{\text{median}}(i, j) = \text{median}(\mathcal{N}_{i,j}) \quad (5)$$

$$f_{\text{stddev}}(i, j) = \sqrt{\text{var}(\mathcal{N}_{i,j})} \quad (6)$$

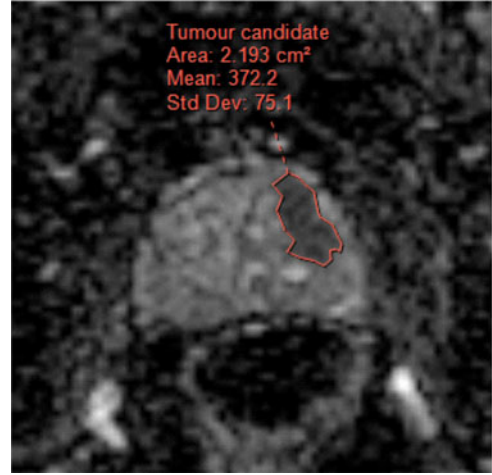


Fig. 6. Size feature (area) illustration for a tumor candidate.

while the mean absolute deviation feature image is simply the absolute value of the voxelwise difference between the original image and the mean feature image

$$f_{\text{absdev}}(i, j) = |I(i, j) - \text{mean}(\mathcal{N}_{i,j})| \quad (7)$$

Second-Order Features: The second-order features in the MAPS model consist of contrast, correlation, and homogeneity, previously used in [30], computed on T2w imagery at two scales each for a total of six features. Computed from a CM, the second-order textures are given by

$$\text{contrast} = \sum_{i,j} |i - j|^2 p(i, j) \quad (8)$$

$$\text{correlation} = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)p(i, j)}{(\sigma_i \sigma_j)} \quad (9)$$

$$\text{homogeneity} = \sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \quad (10)$$

where p is the corresponding value from gray-level CM of the sliding local window.

4) *Size:* Region size is used by clinicians to assess mpMRI; the scheme proposed by Haider *et al.* [23] uses the presence of nodules larger than 5 mm in the TZ as indicative of cancer and further uses size to distinguish between classifications of “Probably cancer” and “Definite cancer,” while PI-RADS suggests that if the size of the region’s interface with the prostate surface is too large, cancer is indicated [10]. The Size feature group contains a single feature measuring the area of the candidate region. Since the resolution of a voxel changes with the modality and with the equipment, counting region sizes by number of voxels would be inconsistent. The size is, therefore, calculated in physical units, such as square millimeters. The size feature is shown in Fig. 6.

The proposed method introduces a novel means of segmenting candidate tumor regions by combining physiological and morphological information. Features are then extracted from the candidate regions in several high-level categories derived from clinical practice: Morphology, Asymmetry, Physiology,

TABLE I
SUMMARY OF FEATURE GROUPS IN PROPOSED MAPS MODEL

Feature group	Number of features	Description
Morphology	3	area regularity (1), perimeter regularity (2)
Asymmetry	4	region bilateral symmetry (4)
Physiology	34	
raw mpMRI	4	mean values for T2w, ADC, DWI, and CDI
textural (first-order)	24	median, standard deviation, and mean absolute deviation at two scales for T2w, ADC, DWI, and CDI
textural (second-order)	6	contrast, correlation, and homogeneity at two scales for T2w
Size	1	size of region
total	42	all features

and Size. These feature groups are summarized along with the number of features in each group in Table I. The proposed radiomics feature model uses all features in each of the feature groups to train the classifier. These feature groups capture information commonly used by radiologists in the diagnosis of prostate cancer using mpMRI data, and group this information in a way familiar to the diagnosticians. Features in the Morphology, Asymmetry, and Size groups describe properties of the region, while features in the Physiology group more traditionally describe properties of the tissue on a finer scale. In order to produce a single feature score for any given region, then, the feature values at the finer scale are averaged, and their mean value is used as the region's feature score. Collectively, this tumor candidate identification scheme and set of feature groups are the MAPS feature model. The proposed model lays out the fundamentals for radiomics-based computer-aided diagnosis and prognosis for prostate cancer. The high-level feature model, MAPS, proposed here can be used for both detection of prostate cancer, as presented in this paper, and prognosis of this disease where patient outcome data are available. In the next section, the performance of this model at detecting prostate cancer will be evaluated and analyzed.

IV. EXPERIMENTAL RESULTS

The proposed MAPS feature model was tested by using actual clinical mpMRI data to extract the MAPS features, which were then used to train automated classifiers. The performance of classifiers using the proposed feature model and various related feature sets was measured against ground truth diagnoses provided by an expert diagnostician confirmed by histopathology data. Various subsets of the entire model were independently evaluated to compare, for instance, the performance of the Morphology features on their own compared to the Asymmetry features. This section describes the clinical mpMRI data, feature extraction, and classifiers training evaluation. Numeric results are presented and discussed in Section IV-C.

A. Data Acquisition Protocol

Thirteen patients were imaged in total where the mpMRI images and expert annotations were compared to the corresponding histopathology data, obtained through radical prostatectomy with Gleason score 7 and above, as ground-truth to confirm the accuracy of the mpMRI markings. Informed consent was obtained from all patients, and approval for this study was obtained from the ethics review board of Sunnybrook Health

TABLE II
SUMMARY OF PATIENT INFORMATION FOR DWI ACQUISITIONS

Age	DFOV (cm ²)	Resolution (mm ³)	TE (ms)	TR (ms)
74	24 × 24	1.667 × 1.667 × 3	61	6692
61	20 × 20	1.5625 × 1.5625 × 3	61	6178
70	20 × 20	1.5625 × 1.5625 × 3	61	6178
53	20 × 20	1.5625 × 1.5625 × 3	61	6693
53	20 × 20	1.5625 × 1.5625 × 3	61	6178
74	24 × 24	1.875 × 1.875 × 3	63	8717
56	20 × 20	1.786 × 1.786 × 3	61	6667
57	24 × 24	1.667 × 1.667 × 3	61	6177
60	20 × 20	1.5625 × 1.5625 × 3	61	6178
75	20 × 20	1.5625 × 1.5625 × 3	61	6178
62	20 × 20	1.5625 × 1.5625 × 3	61	6178
70	20 × 20	1.5625 × 1.5625 × 3	61	6178
60	24 × 24	1.667 × 1.667 × 3	61	6692

Sciences Centre (SHSC), Toronto, ON, Canada. All acquisitions were made at SHSC with a Philips Achieva 3.0 T device with SENSE-Cardiac coil. Patient ages and DWI acquisition parameters are summarized in Table II. Images were reviewed and annotated by a trained radiologist with 18 and 13 years of experience interpreting body MRI and prostate MRI, respectively. Each annotation consists of a contour of the prostate gland as well as a contour around any cancerous regions identifiable by MRI.

Next, tumor candidates were identified by the proposed MAPS tumor candidate identification algorithm. Initially, all voxels with ADC values under $700 \times 10^{-6} \text{ mm}^2/\text{s}$ were segmented as tumor candidates. Following this, only the connected regions larger than the voxel resolution considered as the final set of tumor candidate regions. An affine transformation (scaling, rotation, and translation) was applied to the DWI and ADC imagery as to register it to the T2w imagery according to the patient coordinates in the MRI metadata. The gland and region masks were likewise transformed. This ensured that the various modalities could be meaningfully compared on a voxelwise basis by registering the imagery to a common coordinate system and interpolating to a common resolution. Finally, MAPS feature values were computed for each region, and the resulting feature vectors used to train and evaluate classifiers.

B. Classification and Cross Validation

The features extracted from the clinical mpMRI data were used to train many naïve Bayes classifiers. Naïve Bayes classifiers assume that each feature is independent from one another. Formally, the joint probability of the region label (C) and

feature values (f) is equal to the product of the prior probability and the likelihood of the assumed independent feature values

$$p(C|f_1^M, f_2^M, f_3^M, f_1^A, \dots) = p(C) \prod_i f_i. \quad (11)$$

Although this produces an unsophisticated classifier, in practice they often perform quite well; furthermore, the simplicity of the model implies a lower number of parameters to train, which is helpful considering that the dataset used here had in the order of twice as many training samples as there are features in the model. Ultimately, the classifications performed for the sake of this research were done to compare the performance of different feature groups in a consistent way, rather than to maximize classification performance, so while a reduction in model complexity and classification accuracy may seem to be a tradeoff for the sake of limited training data, it actually serves to facilitate analysis.

In total, 251 regions were used to train and evaluate classifiers. These 251 tumor candidate regions had a median size of 73 mm² (28 pixels) where the actual tumorous regions median size was 122 mm² (46 pixels). In general, the tumor candidate regions included regions which were much smaller or larger than the actual tumorous regions. Of these 251 regions, 53 were cancerous regions identified by an expert diagnostician confirmed by histopathology data. In order to produce noncancerous ground truth regions, the automatic tumor candidate identification algorithm was run and any regions it identified which belonged to the cancerous region labels were removed. Altogether, 198 noncancerous regions were identified in this way and used for evaluation in a cross-validation approach. The cancerous and noncancerous regions that belonged to one patient were grouped together to enable leave-one-patient-out cross validation; to measure the performance for a given patient's data, all cancerous and noncancerous regions belonging to the patient were excluded from the training data and accuracy, sensitivity, and specificity were calculated for that patient data separately. The results of individual patients' data were then averaged to generate the final results.

C. Evaluation and Discussion

Classifier performance was evaluated using three metrics in common use in clinical practice: accuracy, sensitivity, and specificity where accuracy measures the proportion of predictions which are correct, sensitivity measures the proportion of positive cases correctly identified as such, and specificity measures the proportion of negative cases correctly identified as negative. Highly sensitive classifiers are typically used in the clinical process as screens, while highly specific ones are used to weed out false positives from the set of patients identified by the screen. The following section explains the calculation and usage of these metrics. The final section presents and interprets the experimental results.

Several feature groups were evaluated with the scheme described above: each component feature group of the MAPS model and the model as a whole. Results are shown in Fig. 7. In the following, we will discuss the performance of the various

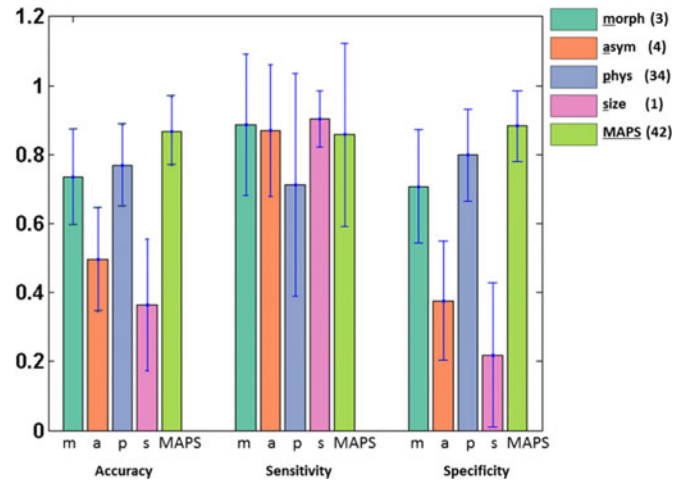


Fig. 7. Classification performance (accuracy, sensitivity, and specificity) of each MAPS feature group independently (Morphology, Asymmetry, Physiology, Size), and the MAPS feature model as a whole using leave-one-patient-out cross validation.

MAPS feature groups relative to one another and to the model as a whole. We will also examine the performance of MAPS compared to commonly used mpMRI feature model (i.e., T2w + ADC).

1) *Performance of the MAPS Feature Groups:* Classification performance of each of the feature groups comprising the MAPS model is shown in Fig. 7. Several trends are worth noting: some feature groups achieve almost consistent scores with each metric, such as the Physiology feature group, which achieves 73%–79% accuracy, sensitivity, and specificity. Other groups achieve a high sensitivity, but low specificity and low accuracy, while the Size group achieves a sensitivity of 91%, its accuracy is only 37% and its specificity is lower still. The Asymmetry group is similar, at about 87% sensitive and 50% accurate. It is logical that the Size feature group, consisting of a single measure of total region area, should fail to accurately classify samples, as there are many factors which can contribute to region size besides cancer, such as haemorrhaging. Likewise, while asymmetry may successfully detect the majority of cancerous cases, it fails to distinguish from healthy cases, producing many false positives.

The Morphology and Physiology feature groups, on the other hand, achieve scores better than a random guess (i.e., better than 50%) across all three performance metrics. The Morphology group achieves an accuracy of 74% and specificity of 71%, though 89% of sensitivity. Morphology would seem to be better at detecting true positives than true negatives; however, considering that there are only three Morphology features, their performance is better than one might expect given that the Asymmetry group has a similar number of features and nearly twice the misclassification rate. Clearly, the inclusion of region shape adds to the diagnostic power.

The unabridged MAPS model performs better than any one of its feature groups (accuracy of 87%). Taken together, each group provides complementary information, so that tissue information captured by the Physiology features and shape information

TABLE III
95% CONFIDENCE INTERVALS ON RADIOMICS FEATURE MODEL PERFORMANCE

Feature set	Accuracy	Sensitivity	Specificity	Average
Morphology	74% \pm 14% [65% 82%]	89% \pm 21% [76% 99%]	71% \pm 17% [60% 81%]	78%
Asymmetry	50% \pm 16% [41% 59%]	87% \pm 19% [75% 98%]	38% \pm 18% [27% 49%]	58%
Physiology	77% \pm 11% [70% 84%]	73% \pm 31% [54% 92%]	79% \pm 14% [70% 87%]	76%
Size	37% \pm 19% [26% 48%]	91% \pm 7% [86% 95%]	22% \pm 21% [10% 35%]	50%
MAPS	87% \pm 10% [81% 93%]	86% \pm 27% [70% 99%]	88% \pm 10% [82% 95%]	87%
MAPS*	87% \pm 2% [86% 89%]	84% \pm 11% [76% 92%]	88% \pm 4% [85% 91%]	86%

All results obtained via leave-one-patient-out cross validation except for the last row (MAPS*) where a tenfold random cross validation (50%–50% training/testing) was used.

captured by the Morphology features bolster each others' shortcomings; the sensitivity of the Size and Morphology groups can be exploited while using the strength of the other groups to reject more false positives. The MAPS feature model achieved sensitivity and specificity scores of 86%, and 88%, respectively. Although Size and Morphology groups give higher sensitivity results compared to MAPS (91% and 89% versus 86%), their accuracy and specificity are significantly lower than those of MAPS (see Table III). When averaging all three metrics (accuracy, sensitivity, and specificity), MAPS produces the best results (i.e., 87%) compared to other feature groups. Table III summarizes the results.

In addition to leave-one-patient-out cross validation, we ran another experiment for all features combined (MAPS) where half the data were selected randomly for training and the remaining was used for testing. A tenfold cross validation was run using this setting and as shown in Table III (row MAPS*), the final results were almost similar to that of leave-one-patient-out approach (accuracy 87%, sensitivity 84% versus 86%, specificity 88%).

To investigate the contribution of each feature group to the final results, we used maximum relevance, minimum redundancy feature ranking technique [38]. The first five top discriminant features included two morphology features (area and perimeter regularity) and three physiology features (mean ADC, mean CDI, and correlation CDI). The first ten top discriminant features included two morphology features (area and perimeter regularity), one asymmetry feature, and seven physiology features (mean T2, mean and median ADC, mean CDI, standard deviation DWI, correlation T2, and correlation CDI). Interestingly, the size feature ranked 21 out of 42 features. This is also reflected in Fig. 7 where Size feature yields low specificity and accuracy.

Fig. 8 shows a sample slice of mpMRI where the tumor has been marked in ADC map. It also illustrates tumor candidate regions picked by the proposed radiomics feature model. Fig. 9 shows sample radiomics maps for different features of the proposed model.

2) *MAPS Versus Conventional mpMRI*: The performance of the MAPS feature model was compared to the most commonly used mpMRI which is noninvasive, the combination of T2w and ADC. For the conventional model, a comprehensive texture feature set was used which included a total of 96 features for each imaging modality: 4 from first-order and 72 from second-order statistical features, 8 from Kirsch edge detection, and 12

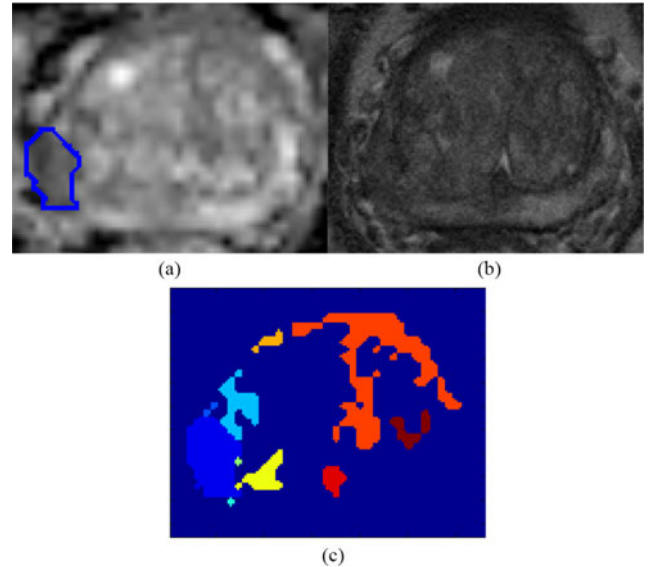


Fig. 8. Sample mpMRI slices of the same prostate gland, using T2w and ADC where the tumor regions is marked. In this example, the radiomics feature model's marking is the same as that of the clinician's. (a) ADC map with tumor delineation. (b) T2w. (c) Tumor candidate regions.

from Gabor filters [16]. The proposed MAPS feature model outperformed the conventional model in all three metrics (see Table IV): accuracy (87% versus 59%), sensitivity (86% versus 79%), and specificity (88% versus 57%). The many false positives produced by the conventional method is a byproduct of the similar T2w and ADC characteristics between healthy and tumorous tissue, suggesting that using low-level features based on T2w and ADC alone is not sufficient for good classification. Indeed, adding more information in the form of additional MRI modalities as well as incorporating high-level features resulted in a much greater improvement, as shown with the performance of the MAPS model. Therefore, in addition to being easier to interpret by radiologists, the MAPS feature model achieves higher classification performance. This paper also outperformed the preliminary results reported in [31] where accuracy and specificity were below 80% and sensitivity was about 60%.

V. FUTURE WORK

Although the initial results are promising, further work is needed to support the use of the MAPS model in the clinical decision-making environment.

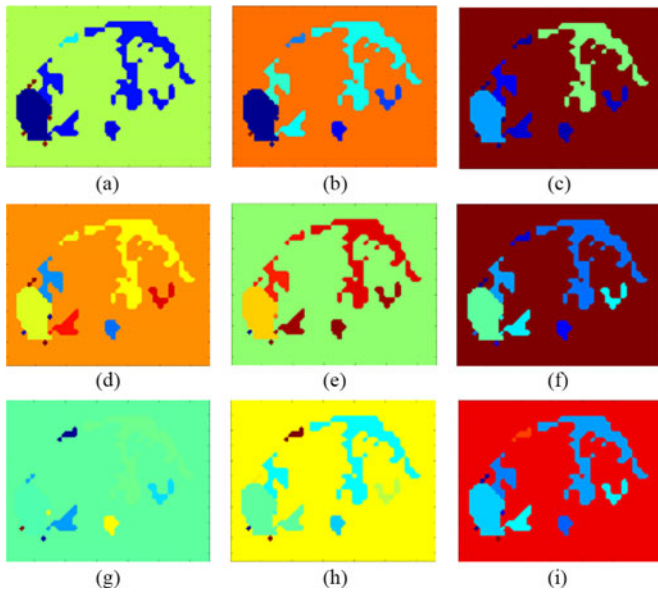


Fig. 9. Examples of radiomics features for a sample slice normalized in $[-1, +1]$. (a) Morphology, (b) asymmetry, (c) size, (d) physiology (mean T2w), (e) physiology (mean ADC), (f) physiology (mean CDI), (g) physiology (texture: contrast), (h) physiology (texture: correlation), (i) physiology (texture: homogeneity).

TABLE IV
PERFORMANCE COMPARISON BETWEEN THE PROPOSED RADIOMICS FEATURE MODEL AND CONVENTIONAL MPMRI VIA LEAVE-ONE-PATIENT-OUT CROSS VALIDATION

Feature Model	Accuracy	Sensitivity	Specificity	Average
MAPS	87% \pm 1%	86% \pm 3%	88% \pm 1%	87%
T2w+ADC	59% \pm 16%	79% \pm 26%	57% \pm 17%	65%

A. Larger Dataset and Feature Selection

The results reported in this paper are preliminary results using a rather small dataset (13 patients) which require a further validation using larger datasets. As future work, we will conduct additional validation using a larger dataset (e.g., 50 patients). Further labeled MRI data would be invaluable in supporting the experimental results shown thus far. Additional data would provide stronger evidence of the performance of the MAPS feature model at detecting prostate cancer. In particular, with enough data, the model could be refined through feature selection in order to discard features shown to have little predictive power.

B. New Features and Changes to Existing Features

Furthermore, there are many additional features that could be considered for inclusion in the model. Although MAPS currently tries to capture many of the features used by radiologists when they are searching for cancer, additional ones can no doubt be designed, which should be evaluated along with the existing features for inclusion into the model. The Size feature group in particular, consisting of a single feature measuring region area, could be expanded with several different measures of size, such

as the major axis length of the region, or the diameter of the largest circle which can be inscribed therein, for example.

C. MAPS for Prognosis of Prostate Cancer

Using the proposed radiomics-driven texture feature model based on tumor candidate regions, we will investigate the prognostic power of MAPS for prediction of patient outcome. This will require additional data on the outcome and survival of the patients in the study.

D. User Testing

Finally, and perhaps most importantly, the MAPS model needs to be evaluated in a clinical decision-making environment. Although initial results show promise for the framework's ability to automatically detect prostate cancer, no amount of automated testing can stand in for user experience testing. In order to validate MAPS as being easy to interpret, it must be evaluated by practitioners. Studies should be run to measure the impact (if any) MAPS has on observer variability, on the accuracy of the diagnostician accuracy, and on the speed with which diagnosticians can review patient cases.

VI. CONCLUSION

We proposed a comprehensive feature model as a quantitative radiomics approach for automatic detection of prostate cancer consisting of several feature groups in order to foster user acceptance and facilitate interpretability of classification results. The initial results are promising where the features in the MAPS model achieved an accuracy of 87% while providing a structured framework for interpretation. The strength of the results presented in this paper is limited due to the size of labeled dataset available for training, and hence, further validation is required using a larger dataset.

REFERENCES

- [1] Canadian Cancer Society's Advisory Committee on Cancer Statistics, *Canadian Cancer Statistics 2014*, Toronto, ON, Canada, 2014.
- [2] Canadian Cancer Society's Steering Committee on Cancer Statistics, *Canadian Cancer Statistics 2012*, Toronto, ON, Canada, 2012.
- [3] F. H. Schrder *et al.*, "ERSPC investigators. Screening and prostate-cancer mortality in a randomized European study," *New Engl. J. Med.*, vol. 360, pp. 1320–1328, 2009.
- [4] G. L. Andriole *et al.*, "PLCO project team. Mortality results from a randomized prostate-cancer screening trial," *New Engl. J. Med.*, vol. 360, pp. 1310–1319, 2009.
- [5] H. U. Ahmed *et al.*, "Systematic review of complications of prostate biopsy," *Eur. Urol.*, vol. 64, no. 6, pp. 876–892, 2013.
- [6] Arumainayagam *et al.*, "Multiparametric MR imaging for detection of clinically significant prostate cancer: A validation cohort study with transperineal template prostate mapping as the reference standard," *Radiology*, vol. 268, no. 3, pp. 761–769, 2013.
- [7] R. Leijenaar *et al.*, "Radiomics: Extracting more information from medical images using advanced feature analysis," *Eur. J. Cancer*, vol. 48, no. 4, pp. 441–446, 2015.
- [8] P. R. Carroll, "Early detection of prostate cancer," *Commun. Oncol.*, vol. 7, pp. 25–27, 2010.
- [9] I. M. Thompson *et al.*, "Prevalence of prostate cancer among men with a prostate-specific antigen level ≤ 4.0 ng per milliliter," *New Engl. J. Med.*, vol. 350, pp. 2239–2246, 2004.
- [10] J. O. Barentsz *et al.*, "ESUR prostate MR guidelines 2012," *Eur. Radiol.*, vol. 22, no. 4, pp. 746–757, Apr. 2012.

- [11] S. Ozer *et al.*, "Prostate cancer localization with multispectral MRI based on relevance vector machines," in *Proc. IEEE Int. Symp. Biomed. Imag., From Nano to Macro*, 2009, pp. 73–76.
- [12] M. D. Blackledge *et al.*, "Computed diffusion-weighted MR imaging may improve tumor detection," *Radiology*, vol. 261, no. 2, pp. 573–581, 2011.
- [13] A. Wong *et al.*, "Correlated diffusion imaging," *BMC Med. Imag.*, vol. 13, p. 26, Jan. 2013.
- [14] A. Wong *et al.*, "Dual-stage correlated diffusion imaging," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2015, pp. 75–78.
- [15] F. Khalvati *et al.*, "A multi-parametric diffusion magnetic resonance imaging texture feature model for prostate cancer analysis," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention, Comput. Diffusion MRI Workshop*, 2014, pp. 79–88.
- [16] F. Khalvati *et al.*, "Automated prostate cancer detection via comprehensive multi-parametric magnetic resonance imaging texture feature models," *Biomed. Central Med. Imag.*, vol. 15, pp. 15–27, 2015.
- [17] K. Hosseinzadeh and S. D. Schwarz, "Endorectal diffusion-weighted imaging in prostate cancer to differentiate malignant and benign peripheral zone tissue," *J. Magn. Reson. Imag.*, vol. 20, no. 4, pp. 654–461, Oct. 2004.
- [18] X. Liu *et al.*, "Prostate cancer segmentation with simultaneous estimation of Markov random field parameters and class," *Med. Imag.*, vol. 28, no. 6, pp. 906–915, 2009.
- [19] D. L. Langer *et al.*, "Prostate cancer detection with multi-parametric MRI: Logistic regression analysis of quantitative T2, diffusion-weighted imaging, and dynamic contrast-enhanced MRI," *J. Magn. Reson. Imag.*, vol. 30, no. 2, pp. 327–334, Aug. 2009.
- [20] P. Kozlowski *et al.*, "Combined diffusion-weighted and dynamic contrast-enhanced MRI for prostate cancer diagnosis—correlation with biopsy and histopathology," *J. Magn. Reson. Imag.*, vol. 24, no. 1, pp. 108–113, Jul. 2006.
- [21] C. K. Kim *et al.*, "High-b-value diffusion-weighted imaging at 3 T to detect prostate cancer: Comparisons between b values of 1,000 and 2,000 s/mm²," *Am. J. Roentgenol.*, vol. 194, no. 1, pp. W33–W37, Jan. 2010.
- [22] K. Katahira *et al.*, "Ultra-high-b-value diffusion-weighted MR imaging for the detection of prostate cancer: Evaluation in 201 cases with histopathological correlation," *Eur. Radiol.*, vol. 21, no. 1, pp. 188–196, Jan. 2011.
- [23] M. A. Haider *et al.*, "Combined T2-weighted and diffusion-weighted MRI for localization of prostate cancer," *Am. J. Roentgenol.*, vol. 189, no. 2, pp. 323–328, Aug. 2007.
- [24] S. Doyle *et al.*, "A boosting cascade for automated detection of prostate cancer from digitized histology," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, vol. 9, no. Pt 2, pp. 504–511, Jan. 2006.
- [25] I. Chan *et al.*, "Detection of prostate cancer by integration of line-scan diffusion, T2-mapping and T2-weighted magnetic resonance imaging; a multichannel statistical classifier," *Med. Phys.*, vol. 30, no. 9, pp. 2390–2398, 2003.
- [26] Shafiee *et al.*, "Apparent ultra-high b-value diffusion-weighted image reconstruction via hidden conditional random fields," *IEEE Trans. Med. Imag.*, vol. 34, no. 5, pp. 1111–1124, May 2015.
- [27] S. Naik *et al.*, "Gland segmentation and computerized Gleason grading of prostate histology by integrating low-, high-level and domain specific information," *Proc. Microsc. Image Anal. With Appl. Biol. Workshop*, 2007, pp. 1–8.
- [28] S. Naik *et al.*, "Automated gland and nuclei segmentation for grading of prostate and breast cancer histopathology," in *Proc. IEEE Int. Symp. Biomed. Imag., From Nano to Macro*, May 2008, pp. 284–287.
- [29] S. Ozer *et al.*, "Supervised and unsupervised methods for prostate cancer segmentation with multispectral MRI," *Med. Phys.*, vol. 37, no. 4, pp. 1873–1883, 2010.
- [30] A. Madabhushi *et al.*, "Automated detection of prostatic adenocarcinoma from high-resolution ex vivo MRI," *IEEE Trans. Med. Imag.*, vol. 24, no. 12, pp. 1611–1625, Dec. 2005.
- [31] A. Cameron *et al.*, "Multi-parametric MRI prostate cancer analysis via a hybrid morphological-textural model," in *Proc. IEEE Annu. Conf. Eng. Med. Biol. Soc.*, 2014, pp. 3357–3360.
- [32] H. J. W. L. Aerts *et al.*, "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature Commun.*, vol. 5, 2014.
- [33] O. Gevaert *et al.*, "Non-small cell lung cancer: Identifying prognostic imaging biomarkers by leveraging public gene expression microarray data—Methods and preliminary results," *Radiology*, vol. 264, pp. 387–396, 2012.
- [34] S. Ruan *et al.*, "Tumor segmentation from a multispectral MRI images by using support vector machine classification," *Imaging, From Nano*, vol. 2, no. 1, pp. 1236–1239, 2007.
- [35] S. Doyle *et al.*, "Automated grading of prostate cancer using architectural and textural image features," in *Proc. IEEE 4th Int. Symp. Biomed. Imag., From Nano to Macro*, 2007, pp. 1284–1287.
- [36] A. Madabhushi *et al.*, "Graph embedding to improve supervised classification and novel class detection: Application to prostate cancer," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2005, pp. 729–737.
- [37] R. Amelard *et al.*, "Extracting morphological high-level intuitive features (HLIF) for enhancing skin lesion classification," in *Proc. IEEE 34th Annu. Int. Conf. Eng. Med. Biol. Soc.*, San Diego, CA, USA, 2012, pp. 4458–4461.
- [38] C. Ding *et al.*, "Feature selection based on mutual information criteria of maxdependency, max-relevance, and min-redundancy," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 8, pp. 1226–1238, Aug. 2005.



Andrew Cameron received the B.A.Sc. degree in systems design engineering from the University of Waterloo, ON, Canada, in 2011 and the M.A.Sc. degree in systems design engineering in 2014.

He is currently working as a Computer Vision Researcher at SportLogiq. His research interests include feature design, detection, and classification.



Farzad Khalvati (S'98–M'06) received the M.A.Sc. and Ph.D. degrees in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2003 and 2009, respectively.

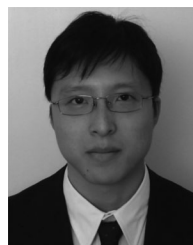
He is currently an Assistant Professor at the Department of Medical Imaging, University of Toronto, Toronto, Canada and a Jr. Scientist at Sunnybrook Research Institute, Toronto, Canada. His research interests include medical image analysis, computer-aided diagnosis and prognosis, image processing, and high-performance image processing.



Masoom A. Haider received the M.D. degree from the University of Ottawa, ON, Canada, in 1986.

In 1994, he completed his residency in radiology from the University of Toronto, ON, Canada, where he is currently a Professor of Radiology in the Department of Medical Imaging. He is also the Chief of the Medical Imaging Department at Sunnybrook Health Sciences Centre, Toronto, ON, Canada. His research interests include abdominal and pelvic MRI with special emphasis in prostate cancer, therapeutic response assessment and functional imaging of

cancer.



Alexander Wong (M'05) received the B.A.Sc. degree in computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2005, the M.A.Sc. degree in electrical and computer engineering from the University of Waterloo, Waterloo, ON, Canada, in 2007, and the Ph.D. degree in systems design engineering from the University of Waterloo, ON, Canada, in 2010.

He is currently the Canada Research Chair in Medical Imaging Systems, the Co-director of the Vision and Image Processing Research Group, and an Assistant Professor in the Department of Systems Design Engineering, University of Waterloo, Waterloo, Canada. His research interests include around imaging, image processing, computer vision, pattern recognition, and cognitive radio networks, with a focus on integrative biomedical imaging systems design, probabilistic graphical models, and biomedical and remote sensing image processing and analysis.