Computer-aided diagnosis
Rampun, Yambu Andrik; Chen, Zhili; Malcolm, Paul; Tiddeman, Bernie; Zwiggelaar, Reyer

Published in:
International Journal for Numerical Methods in Biomedical Engineering

DOI:
10.1002/cnm.2745

Publication date:
2016

Citation for published version (APA):
Computer Aided Diagnosis: Detection and Localisation of Prostate Cancer within the Peripheral Zone

Andrik Rampun*, Zhili Chen**, Paul Malcolm***, Bernie Tiddeman*, Reyer Zwiggelaar*

*Department of Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, UK
**Information and Control Engineering Faculty, Shenyang Jianzhu University, Liaoning 110168, China
***Department of Radiology, Norfolk Norwich University Hospital, Norwich NR4 7UY, UK

SUMMARY

We propose a methodology for prostate cancer detection and localisation within the peripheral zone based on combining multiple segmentation techniques. We extract four image features using Gaussian and median filters. Subsequently, we use each image feature separately to generate binary segmentations. Finally, we take the intersection of all four binary segmentations, incorporating a model of the peripheral zone, and perform erosion to remove small false positive regions. The initial evaluation of this method is based on 275 MRI images from 37 patients and 86% of the slices were classified correctly with 87% and 86% sensitivity and specificity achieved, respectively. This paper makes two contributions: firstly, a novel Computer Aided Diagnosis approach which is based on combining multiple segmentation techniques using only a small number of simple image features. Secondly, the development of the proposed method and its application in prostate cancer detection and localisation using a single MRI modality with the results comparable to the state-of-the-art multi-modality and advanced computer vision methods in the literature. Copyright © 2015 John Wiley & Sons, Ltd.

KEY WORDS: Prostate cancer detection, MRI, Prostate cancer localisation

1. INTRODUCTION

Prostate cancer is the most commonly diagnosed cancer among men and remains the second leading cause of cancer death in men globally. In 2013, there were approximately 240,000 and 40,000 cases reported in the United States and United Kingdom, respectively, and is estimated to reach 1.7 million cases globally by 2030 [1, 2]. In the last decade, prostate cancer screening has been receiving more attention because it can assist in detecting cancer at an early stage before there are any externally detectable symptoms. Statistically, nine out of ten men survive for at least five years if the cancer is diagnosed at the earliest stage [3]. However, early detection of prostate cancer remains a source of uncertainty and controversy [3]. Clinical diagnostic tools such as prostate-specific antigen (PSA) level, digital rectal examination (DRE), transrectal ultrasound (TRUS) and biopsy tests are very popular and globally used despite their inconsistency in producing accurate results [4]. According to Schroder et al. [5], although the use of PSA reduces the rate of death by 20%, the benefit was associated with a high risk of overdiagnosis and overtreatment. It should be noted that the PSA test is not able to predict the aggressiveness of cancer. As a result, slow-growing and non-aggressive prostate cancer is frequently diagnosed in older patients [6]. In terms of TRUS-guided biopsy,
many reports have shown that systematic biopsies do not detect all clinically significant cancers: [6] showed in a large study that nearly a quarter (23%) of detectable cancers were missed [7]. Finally, according to [8] DRE is less effective than the PSA blood test in finding prostate cancer, but it can sometimes find cancers in men with normal PSA levels. One common problem with DRE is if a tumor is located away from the rectal wall, the physician will not be able to palpate it. A recent study concluded that an abnormal DRE had 0.44 sensitivity and 0.68 specificity [9]. On the other hand, although TRUS enables the accurate determination of prostate size and depicts zonal anatomy, its ability to detect cancer tissue is limited with sensitivity and specificity varying between 40-50% [10].

Computer Aided Diagnosis (CAD) of prostate magnetic resonance imaging (MRI) has the potential to improve the accuracy, sensitivity and specificity of clinical diagnostics. According to [11], it could achieve a sensitivity ranging from 61%-81% (average: 71%), specificity 91%-96% (average: 93.5%) and accuracy 84%-92% (average: 88%) while being a non-invasive technique. Moreover, a study conducted in [6] has shown that combining anatomical, functional and metabolic MRI information could achieve on average 83% (75%-92%) sensitivity. Unfortunately, the assessment of prostate MRI requires a high level of expertise and suffers from observer variability [6]. CAD systems can be of benefit to improve the diagnostic accuracy of radiologists, reduce variability and speed up the reading time [6]. The initial goal of CAD is to automatically delineate malignant regions, leading to a reduction of search and interpretation errors, as well as a reduction of the variation between and within observers [6]. CAD has been successfully implemented in different fields of medical imaging such as mammography [12], CT chest [13], CT colonography [14] and brain imaging [15]. Figure 1 shows an example MRI image with the ground truth of the prostate gland, central zone and tumor represented in yellow, green and red, respectively (left image), while the right image shows a simpler schematic overview of the prostate derived from the prostate anatomy proposed in [72] with central zone (CZ), peripheral zone (PZ), and tumor (T). Note that the transitional zone (TZ) is located within the CZ but no definable boundary between these regions is expected on MRI. The ultimate goal of this research is to develop a CAD tool for prostate cancer detection and localisation within the PZ mainly because a) about 80% of the prostate cancers appear in the PZ [10, 16, 17, 18] and b) in general, prostate cancer that arises in the peripheral zone is more aggressive than that which arises in the transitional or central zones [74]. Therefore, in this paper we propose a new method for detecting prostate abnormality within the PZ using four different image features (see Figure 3) extracted using Gaussian and median filters. The main goal of this method is to identify malignant regions (and hence localise them) within the PZ by taking the overlapping binary segmentation from each image feature. This means pixels (or tissues) which are classified in the same malignant cluster in all image features are considered to have the highest probability of being malignant. However, if a pixel is classified as belonging to a benign or normal tissue cluster in one of the image features we considered it to be a benign or normal tissue (this will be explained in more detail in section 4).

The novelty of this method resides in an approach which combines simple features (this is similar to a forest of weak classifiers which together provide strong results), for the first time and applied to prostate T2-W using one modality. To our knowledge, no existing methods in the literature have used the technique of finding cancer regions by taking the overlapping binary segmentation extracted from a small number of image features.
2. RELATED RESEARCH AREA

CAD is a valuable tool and becoming increasingly important in assisting and automating specific clinician’s tasks such as detection, localisation, the study of anatomical structure, treatment planning and computer-assisted surgery [75]. We are focusing on developing a CAD tool for the detection and localisation of abnormal region within the PZ in T2-MRI imaging. CAD can be applied in many different medical imaging applications (such as brain, breast, chest and prostate) using different segmentation techniques. According to [93] the most commonly used segmentation techniques in CAD systems could be categorised into six groups:

1. Contour and shape based (e.g. Active Contour, Level Set, Graph Searching, Atlas-based, Deformable models, etc).
2. Machine learning based (e.g. Support Vector Machine (SVM), k-Nearest Neighbors (k-NN), Fuzzy C-Mean (FCM), K-means, etc)
3. Region based (e.g. Thresholding, Edge-based, watershed, split and merge, etc)
4. Statistical based (e.g. Markov Random Field (MRF), Gaussian Mixture Model (GMM), etc)
5. Multiresolution based analysis (e.g. Discrete Wavelet Transform (DWT), etc)
6. Hybrid and soft computing methods (e.g. Level Set + Artificial neural network (ANN), Fuzzy C-means + DWT, etc)

The most popular segmentation techniques in biomedical imaging fall under the supervised and unsupervised machine learning based techniques, and contour and shape based methods. In this section we will briefly discuss these techniques and their applications in biomedical imaging. For the other techniques and their applications in medical imaging we refer to [11, 12, 13, 14, 15, 96].

The level set technique has been applied to several human organs (e.g. brain, cardiac, prostate, breast, etc). Dubey et al. [77] proposed a semi-automatic segmentation method of MRI brain tumors. Firstly, the method generated a tumor probability map by classifying each voxel into the tumor or background class using intensity-based fuzzy c-means. Subsequently, the tumor probability map was used to locally guide the propagation direction of the level set. Tsai et al. [78] developed a shape-based approach using level sets and demonstrated their method by applying it to the segmentation of cardiac and prostate MRI. The proposed method derived a parametric model for an implicit representation of the segmenting curve by applying principal component analysis to a collection of signed distance representations of the training data. The parameters of this representation were then manipulated to minimize an objective function for segmentation. Liu et al. [79] proposed a method for mass segmentation in mammograms using a level set to improve the initial segmentation performed using a watershed algorithm. On the other hand, Shi et al. [80] used k-means clustering followed by a morphological opening operation for initial mass segmentation. For the level set segmentation a linear discriminant analysis (LDA) classifier with stepwise feature selection was used to merged the extracted features into a classification score.

Recently, Yeo et al. [81], proposed a level-set segmentation method using active contour modelling applied to synthetic and real images (e.g. brain and knee MRI and carotid CT image). The proposed method consisted of an image attraction force, which was used to propagate contours toward object boundaries, and a global shape force, which deforms the model according to the shape distribution learned from a training set. On the other hand, Sachdeva et al. [82] proposed a method which used intensity and texture information (extracted from Gray-Level Co-occurrence Matrices) present within the active contour to overcome weak or diffused edges in an image. In [83], a novel automatic approach to identify brain structures in magnetic resonance imaging (MRI) is presented for volumetric measurements. This approach combines the active contour model with a support vector machine (SVM) classifier. The SVM features are selected according to the structure of brain tissues: gray matter (GM), white matter (WM), and cerebrospinal fluid (CSF). Jones et al. [95] developed an interactive segmentation method by combining both region selection and user point selection. Evaluation results showed on average more than 98% accuracy based on 248 intravascular ultrasound (IVUS) images.

Graph searching techniques have been studied for the segmentation of biomedical images such as brain, knee and glottis [84, 85, 86]. Pedoia and Binaghi [84] proposed a fully automatic 2D
brain segmentation using graph searching technique which consisted of border detection based on
two-dimensional graph searching principles and radial contour detection. In the border detection
phase, polar conversion is performed first, followed by skull and brain boundaries detection. Li
et al. [85] extended the optimal graph-searching techniques to 3D and higher dimensions using a
polynomial-time algorithm for surface segmentation in volumetric images. The method is efficient
and robustly tested on MR arterial walls and IVUS image. Finally, a study in [86] uses the output of
SVM to drive a graph-cuts segmentation, which was initially trained as a local Golgi detector based
on rotationally invariant features. It should be noted that in some of the described segmentation
approaches difference images have been used but in general these rely on multi-modality data or
on some temporal sequences, neither of which have been used in the proposed approach. In the
developed approach we rely on a single modality and the variation in appearance between normal
and abnormal tissues within the peripheral zone.

Despite the relevant achievements obtained, the main limitations of contour and shape based
methods for our research are

1. Most of them require a significant amount of user interaction for initial region selection. In
our case, we want to eliminate (or minimise) user interaction in finding cancerous regions.
2. They work well only if the boundary of the object is well defined within the image: such
methods work well for prostate gland detection [94]. In many prostate T2-W MRI, cancerous
regions are vague both in terms of appearance and shapes, and as such many training samples
would be needed.
3. The results of many of these methods are highly dependent on the initialisation.

Therefore, we applied an unsupervised machine learning based method: FCM clustering as
proposed by Chen and Zwiggelaar [66]. This method incorporates local spatial and intensity
information based on an adaptive local window filter whose weighting coefficients differentiate the
neighbouring pixels within the local window. The method is less sensitive in dealing with different
types of noise and intensity inhomogeneities. Several noise reduction techniques such as median
and Gaussian filters were used to reduce image noise and erosion is used to reduce false positives.

3. MODELLING THE PERIPHERAL ZONE

Pathologically, about 80% of prostate cancers arise in the PZ and the rest are within the CZ [16].
Since the percentage occurrence of cancers in the PZ is high and as these tend to be more aggressive,
we aim to detect prostate abnormality within this region. We did not perform prostate segmentation
because all prostates were already delineated by an expert. It should be noted that Zhu et al. [94]
developed a method to detect the prostate capsule. Based on the schematic overview shown in the
right image of Figure 1 (also proposed in [72]), we defined our 2D prostate model based on Figure
2.

![Figure 2](image)

Figure 2. Prostate gland (black) and the defined PZ below \( y = ax^2 + bx + c \) (green) which goes through
\( v_1, v_2 \) and \( v_3 \).

The generic prostate’s PZ model in this paper is mainly inspired from similar models proposed
by Makni et al. [73] and Liu et al. [91] which used catenary and polynomial curves, respectively.
From a radiological point of view, the prostate is mainly divided into two regions in MR (the PZ and CZ). Therefore, according to Makni et al. [73], when segmenting these regions, expert radiologists tend to follow the rule of ‘imagining’ outlines due to contrast or strong artifacts. The process of distinguishing these regions are heavily relying on a priori knowledge of the most likely locations [73]. Indeed, a more accurate way (probably more complex and time consuming) could be achieved by segmenting the PZ within the prostate gland. However, these approaches are complicated and require high accuracy in distinguishing tissues in the PZ and CZ. In cases where there is no clear boundary between the PZ and CZ, most segmentation based methods suffer from over-segmentation (hence, could lead to many false positives). In contrast, defining quadratic curves is simple and fast.

In our CAD system, we used the quadratic equation \( y = ax^2 + bx + c \) based on three crucial coordinate points of the prostate which are \( v_1, v_2 \) and \( v_3 \), which are determined by the outmost \( x \) and \( y \) coordinates of the prostate boundary: \( x_{\text{min}}, x_{\text{max}}, y_{\text{min}}, y_{\text{max}} \) (see Figure 2). For example, \( x_{\text{min}} \) and \( y_{\text{max}} \) can be determined by taking the minimum \( x \) and maximum \( y \) coordinates along the prostate boundary. Moreover, the \( x \) coordinate of \( v_1 \) and \( v_3 \) are captured from \( x_{\text{min}} \) and \( x_{\text{max}} \) and their \( y \) coordinate is determined by taking the \( y \) coordinate between \( y_{\text{min}} \) and \( y_{\text{max}} \). On the other hand, the \( x \) coordinate of \( v_2 \) is taken from the \( x \) coordinate \( x_{\text{min}} \) and \( x_{\text{max}} \) and its \( y \) coordinate is determined by taking \( \frac{3}{4} \) of the distance from \( y_{\text{min}} \) to \( y_{\text{max}} \). The coefficient (\( \epsilon = 0.75 \)) is selected as it gives balanced results in terms of accuracy, sensitivity and specificity (see Figure 13). Mathematically, these can be represented in equations (1), (2), (3) and (4).

\[
C_p = \left( \frac{(x_{\text{min}} + x_{\text{max}})}{2} \right) \left( \frac{(y_{\text{min}} + y_{\text{max}})}{2} \right) \quad (1)
\]

\[
v_1 = \left( \frac{(x_{\text{min}} + y_{\text{max}})}{2} \right) \quad (2)
\]

\[
v_2 = \left( \frac{(x_{\text{min}} + x_{\text{max}})}{2} \right) \left( \frac{(y_{\text{max}} - y_{\text{min}}) \times \epsilon}{2} \right) \quad (3)
\]

\[
v_3 = \left( \frac{(x_{\text{max}} + y_{\text{max}})}{2} \right) \quad (4)
\]

Once the coordinates of \( v_1, v_2 \) and \( v_3 \) are defined, we can determine the values of \( a, b \) and \( c \) (therefore a final quadratic equation is defined). Finally, by taking every \( x \) coordinate from \( x_{\text{min}} \) to \( x_{\text{max}} \) into a quadratic equation we are able to determine the \( y \) coordinate which will define the PZ’s boundary (the main goal is to analyse the region under the green line in Figure 2). The approximation model is able to capture most of the PZ area, easy to implement and computationally efficient.

4. METHODOLOGY

Figure 3 shows the overview of the proposed methodology. First, we perform Gaussian and median filtering on the original image to obtain \( G_1 \) and \( M_1 \). We extract a probability image from \( G_1 \) and \( M_1 \) using greyscale frequency before we obtain the third feature \( F_1 \) which is the magnitude of \( G_1 \) and \( M_1 \). This means, each element in \( F_1 \) is the sum vector of each component from \( G_1 \) and \( M_1 \). On the other hand, the fourth feature \( F_2 \) is the vector magnitude of probability images from \( G_2 \) and \( M_2 \). Subsequently, we use each feature separately to generate binary segmentations taking the feature space and intensity values into account. We perform erosion on each of the segmentations to remove small false positive regions. Finally, we take the intersection of all four binary segmentations, taking a model of the peripheral zone (see Section 3) into account.

4.1. Preprocessing

Since MRI images often suffer from different types of noise, it is necessary to apply different types of denoising methods before doing any further processing. From a clinical point of view, this is an appropriate step to enhance characteristics of an important region of interest (such as textures and boundaries). Moreover, it does not deform the anatomical locations of tissue regions because this step deals only with noise without affecting the spatial information.

Hendrick [59] reported that one type of noise in MRI images is Gaussian noise. Therefore, Gaussian smoothing is selected, which is also effective in reducing noise that is problematic for
Figure 3. Overview of the proposed methodology. Intensity values are represented by greyscale colours, with the darkest representing the lowest intensity.

image analysis algorithms. For example, image segmentation is often affected by the presence of too many local minima/maxima and inflection points in the data \cite{19}. Studies performed by Barentsz et al. \cite{89} and Viswanath et al. \cite{90} suggested that most cancers shows textural distortions in T2-W images. Litjens et al. \cite{92} captured these characteristics using Gaussian filters. Many previous studies have applied Gaussian filters in denoising MRI images and mammograms \cite{59,60,61}. We are aware that there are more sophisticated denoising methods such as those based on Fourier analysis \cite{87} or anisotropic filtering \cite{88} which could be explored in future work. The 2D Gaussian function is defined as

\[
g(s, t) = \frac{1}{2\pi\sigma^2} e^{-\frac{s^2+t^2}{2\sigma^2}} \quad (5)
\]

where \(s\) is the distance from the origin in the horizontal direction, \(t\) is the distance from the origin in the vertical direction, and \(\sigma\) is the standard deviation of the Gaussian distribution. In the proposed method we used the following parameters: the kernel size \((ks)\) is \(15 \times 15\) and the standard deviation \((\sigma)\) is 3.0. See subsection 4.2 for the selection of the Gaussian parameters \((\sigma\) and \(ks\)) and their resulting variability can be seen in Figure 12.

On the other hand, we used median filtering to preserve the regional boundaries (e.g. tumor regions). It is claimed that using median filtering is much better at preserving sharp edges \cite{30} and in our case we want to preserve the information-bearing structures such as tumor boundaries \cite{10}. The median filter works by replacing the pixel value with the median value in the neighborhood of that pixel. We used a sliding window of \(5 \times 5\) pixels. Other sizes are possible (such as \(3 \times 3, 7 \times 7\) and \(9 \times 9\)) give similar results.

We calculated the probability images using equation (6) and calculate the vector sum using equation (7). Probability images are commonly used to model the expected appearance of an object (e.g. tumor region) in a given reference space. Many studies suggested that prostate cancer appears darker within the PZ and is similar to the tissues outside the prostate gland. By computing a
probability image we are able to quantify the likelihood of every pixel/voxel belonging to specific tissues (e.g. tumor region). This also means each pixel/voxel can be represented by a value of the likelihood of being malignant. The conversion to probability images acts as a normalisation across the feature images and as such these can all be treated in the same framework. We calculate the probability image for each of $G_1$ and $M_1$. This means, for an $f(i,j)$ image, the probability value for the $k^{th}$ grey level is calculated using:

$$P(i,j) = \frac{\#(f(i,j) = k)}{M \times N} \quad (6)$$

where $\#(f(i,j) = k)$ is the number of pixels at the $k^{th}$ intensity level in a $M \times N$ image, and as such each element in $P$ is the probability value for a particular intensity level. Others have exploited probability images for segmentation [62, 63]. To calculate features $F_1$ and $F_2$, we find the sum vector for every corresponding element in $G_1$ and $M_1$ using equation (7). According to [65] by taking the square root of the sum of squares between two corresponding signals produces a good image with little noise and continuous edge marking while another study performed in [64] suggested that equation (7) can improve the signal to noise ratio (SNR). In this case, corresponding signals are the corresponding pixels in $G_1$ and $M_1$.

$$I_n(i,j) = \sqrt{G_2^2(i,j) + M_2^2(i,j)} , n = 1 \text{ or } 2 \quad (7)$$

In total, four features are extracted, namely the Gaussian feature ($G_1$), the median feature ($M_1$), the magnitude Gaussian and median features ($F_1$) and the magnitude of the probability images of Gaussian and median features ($F_2$). Before image segmentation is performed, we applied noise reduction to $F_1$ and $F_2$ to minimise the noise retained/created after being processed using equations (6) and (7). In the proposed method we applied a robust noise reduction method developed by Garcia [20] which is robust in dealing with weighted, missing, and outlying values by using an iterative procedure (which is the case in $G_2$ and $M_2$). Figure 4 shows examples of all extracted features $G_1$, $F_1$, $F_2$ and $M_1$. We can see that the malignant region appears brighter in $F_2$ and darker in the other features.

![Figure 4. Example extracted features, from left to right: $G_1$, $F_1$, $F_2$ and $M_1$. The outline prostate is defined by the yellow line and the malignant region (in the PZ) is indicated by the red arrow.](image)

### 4.2. Gaussian parameters

The selection of the parameters for the Gaussian smoothing function is based on the studies in [23, 24, 25, 26, 27], which indicated that, the standard deviation ($\sigma$) and kernel size ($ks$), are linked. According to the experiments with different Gaussian convolution algorithms conducted in [23], the authors showed that the amount of error (the smaller the error the closer the denoised image is in comparison to the original image) did not change significantly after $\sigma \geq 2$. This means, for many Gaussian algorithms the error is much higher (less accurate) when $\sigma < 2$. On the other hand, for the selection of kernel size ($ks$) several authors [24, 25, 26, 27] suggested that in general, filter size should be $[3\sigma]$ to $[5\sigma]$ and odd [27, 28]. For instance, if $\sigma = 2.5$, the recommended minimum kernel size is $9$ (3 × 2.5 = 7.5 = 8, since it should be an odd number according to [27, 28], the nearest odd value is 9). Similarly, selecting $\sigma = 1$ would suggest the smallest kernel size of 3 × 3. Although there are no quantitative experimental results for optimal Gaussian parameters on medical images such as MRI or ultrasound, their results indicate a general guideline for selecting Gaussian parameters. In the proposed method we used several $\sigma$ values together with several kernel sizes.
and chose the ones that give the highest accuracy, sensitivity and specificity (see Figure 12). The selection of parameters is not the major focus of this study but the development of a novel method of prostate cancer detection and localisation within the PZ is.

4.3. Clustering

In the proposed method, image segmentation is performed using a Fuzzy C-Means (FCM) algorithm as it has been widely applied in a variety of medical image segmentation applications [29, 66]. However, one common problem with FCM is its ability in handling different types of noise and intensity inhomogeneities taking local spatial and intensity information into account. The FCM algorithm assumes that every pixel can belong to multiple classes with varying degrees of membership. The algorithm works by assigning membership to each data point corresponding to each cluster center on the basis of distance between the cluster and the data point. The closer the data point to the cluster center the higher its membership value for that cluster. Let $Y = (y_1, y_2, ..., y_R)$ denote an image with $R$ pixels to be partitioned into $d$ clusters. FCM iteratively minimises the objective function defined as

$$J_{fcm} = \sum_{p=1}^{d} \sum_{q=1}^{R} u_{pq}^m \|y_q - v_p\|^2$$

with the following constraints: $\sum_{p=1}^{d} u_{pq} = 1$ for all $q$ and $0 < \sum_{q=1}^{R} u_{pq} < R$ for all $p$, where $u_{pq}$ represents the membership of pixel $y_q$ to the $p^{th}$ cluster, $y_q$ represents the feature data of the $q^{th}$ pixel, and $v_p$ is the prototype value of the $p^{th}$ cluster centre. The parameter $m$ (equal to 2 in this study) is a weighting exponent on each fuzzy membership that controls the fuzziness of the resulting partition. We segment every image feature into four different classes. We selected four classes based on the number of tissue categories in the prostate: normal (non-neoplastic) prostatic tissue, benign prostatic hyperplasia, high-grade prostatic intraepithelial neoplasia, and prostatic adenocarcinoma [52]. The first two categories are benign tissues, the third one is a risk factor for malignancy (we included this into one of the malignant classes to reduce false negatives) and the last one is malignant. Most cancer regions in the PZ tend to have a dark appearance [16, 31]. Moreover, several studies suggested that prostate cancer tissue tends to appear darker on a T2-weighted MRI image [32, 33, 34]. In fact, radiologists also tend to use darker regions to identify abnormality within the PZ [35]. Since most malignant regions contain lower intensities, cancerous regions could be detected within the prostate by taking the segmented regions that correspond to the first two lowest intensity fuzzy c-means clusters (indicated by the superscript ‘low’ in Figure 5) in $G_1$, $F_1$ and $M_1$. However, since malignant regions in $F_2$ are represented by higher average intensity values, we take segmented regions which correspond to the two highest intensity fuzzy c-means clusters. This process can be represented using the following equation

$$O = G_1^{low} \cap F_1^{low} \cap F_2^{high} \cap M_1^{low}$$

where ‘low’ and ‘high’ are low and high intensity represented in the segmented regions within the prostate and $O$ represents the overlapping region from all four binary segmentations. Figure 5 shows an example of this process.

After selecting the regions of interest (segmented areas which are under the approximate PZ’s boundary (green line in Figure 2)), we combine all binary segmentations and find its overlapping region as shown in Figure 5. Finally, we perform erosion to remove noisy pixels which will be explained in the next subsection. Note that segmented areas above the green line were ignored in this study because we are only interested in detection within the PZ.

4.4. Post processing

By performing erosion on the binary segmentation, we can reduce the number of false positives. The number of pixels removed from the objects in an image depends on the size and shape of the structuring element used. In the proposed method we used a ‘disk’ shaped structuring element with
Figure 5. After performing FCM clustering on every feature, we only take segmented regions which correspond to the first two lowest intensity FCM in $G_1$, $F_1$ and $M_1$, and segmented regions which correspond to the two highest intensity clusters in $F_2$. Note that only segmented areas which are under the green line in Figure 2 (the estimation of the PZ’s boundary) will be taken into account. Segmented areas above the green line were removed.

5. DATABASE DESCRIPTION

Data from 37 patients (range: 40-74 years) with biopsy-proven prostate cancer were included in this study. All patients underwent T2-W MR imaging at the Department of Radiology at the Norfolk and Norwich University Hospital, Norwich, UK. MR acquisitions were performed prior to radical prostatectomy. All patients gave their written consent to participate in this study which was approved by the institutional review board. All images were obtained on a 1.5 Tesla magnet (Sigma, GE Medical Systems, Milwaukee, USA) using a phased array pelvic coil, with a $24 \times 24$ cm field of view, $512 \times 512$ matrix, $3\,mm$ slice thickness, and $0.5\,mm$ inter-slice gap. Each patient has 5 to 12 slices. However, since our current study is focusing only within the PZ, slices with no visible PZ (the whole prostate gland is covered by the CZ) were excluded in this study (e.g. see Figure 6). All images were manually annotated by an expert radiologist (and further validated/confirmed by two independent radiologists) with more than 10 years experience in diagnosing prostate cancer in MRI. In total our database contains 275 slices (135 malignant and 140 normal slices). Each slice contains the annotations of prostate gland, central zone and cancerous regions (if present).

Figure 6. The whole of the prostate gland is fully covered by the CZ. All cases like this were excluded in our study because our current focus is within the PZ.
6. EXPERIMENTAL RESULTS

Data was analysed and classified as to whether the prostate contains cancer. The detection of cancer occurs when there are any retained segmented regions \((G_{1}^{low} \cap F_{1}^{low} \cap F_{2}^{high} \cap M_{1}^{low})\) within the peripheral zone. Subsequently, we compared the result with the ground truth whether the prostate contains cancer regions or not. We use several quantitative measures to evaluate the results such as sensitivity, specificity and accuracy. Each of these metrics can be calculated using the following equations

\[
Sensitivity(\text{Sen}) = \frac{TP}{TP + FN} \tag{10}
\]

\[
Specificity(\text{Spe}) = \frac{TN}{TN + FP} \tag{11}
\]

\[
Accuracy(\text{Acc}) = \frac{TP + TN}{TN + TP + FP + FN} \tag{12}
\]

where \(TP\) and \(FP\) denote the number of true positives and false positives, respectively. Similarly, \(TN\) and \(FN\) indicate the numbers of true negatives and false negatives. Accuracy means the number of correct classified slices (or pixels in voxel based classification) out of the total number of slices. Sensitivity measures the proportion of actual positives which are correctly identified (in this case the percentage of malignant slices which are correctly identified) whereas specificity measures the proportion of actual negatives which are correctly identified (in this study the percentage of normal slices which are correctly identified). The proposed method achieved 86% accuracy (237 samples are classified correctly) and 38 samples data are misclassified with 7% (20 samples) false negative and 6% (18 samples) false positive results. In addition, the method produced 87% sensitivity and 86% specificity. On the other hand, in terms of voxel based classification (only within the PZ) we achieved 0.86 \pm 0.06, 0.83 \pm 0.05 and 0.96 \pm 0.07 accuracy, sensitivity and specificity. Erosion with flexible size of structuring element and regions intersection \((G_{1}^{low} \cap F_{1}^{low} \cap F_{2}^{high} \cap M_{1}^{low})\) reduce the number of false positive and false negative results by \(\approx 20\%\).

6.1. Correct detection (classification)

Figure 7 presents several examples of correct detection/classification. Correct detection means an image (MRI slice) is classified correctly (malignant or normal) regardless of the location of tumor within the PZ. The segmentation results are divided into three different categories; small malignant region, large malignant region and obscure malignant region. For the first category, the proposed method shows its sensitivity dealing with small malignant regions within the PZ as shown in image 3, 9, 10, 16 and 18. In those images, the proposed method managed to segment malignant regions correctly (in red line) despite their small sizes. In the second category, we show results in detecting and localising malignant region in larger areas. This can be seen in images 1, 4, 7, 13, 14, 15, 17, 19 and 20 where cancers are spread quite substantially within the PZ and some within the CZ. Results in Figure 7 show that these regions were segmented within the expert radiologist’s ground truth. Finally (third category), we show results when malignant regions are obscure within the PZ. In image 2, there are three dark regions (left, middle and right) within the PZ and visually it is very difficult to identify which one of those regions is cancerous. As a result, although the proposed method managed to segment the malignant region, there is one false positive region in the middle of the PZ. On the other hand, in image 8 we can visually see that there is no sign of irregularity (the whole PZ looks uniform), which makes the abnormal regions obscured. Other examples of the experimental results can be seen in image 5, 6, 11, 12, 21, 22, 23 and 24.

Figure 8 shows examples of experimental results in normal slices. The PZs in image 25 to 28 show no sign (or small signs) of irregularity which made it easier to identify normal slices.
Figure 7. Malignant slices: prostate capsules are delineated in yellow and central zones and tumors are in green and red, respectively. The detected regions are indicated as the highlighted regions.

Figure 8. Normal slices: prostate capsules are delineated in yellow and central zones in green. The lack of segmented regions on the images indicates no cancer regions present.

6.2. Correct detection with incorrect localisation

This section presents two examples of results where overall classification is correct but the tumor location is incorrect. For localisation, we compare the position of the segmented region based on $(G_1^{low} \cap F_1^{low} \cap F_2^{high} \cap M_1^{low}) \subseteq M_r$, where $M_r$ is a cancerous region within the PZ. In our evaluation, incorrect localisation is when the area of the segmented region is $< 50\%$ within the cancerous region delineated by an expert radiologist. On the other hand, correct localisation means $\geq 50\%$ of the area of the segmented region is within the annotated malignant region. Our method produced 81% (109 slices true positives) correct localisation with respect to the number of malignant
slices (135 samples) which means 6% (8 slices) of malignant cases were classified correctly but tumors were localised incorrectly and the other 18 slices are false negatives. This may have been caused when normal regions have dark or very similar appearance with cancerous regions (low intensity) in the PZ. According to [58] low signal intensity may be seen in the PZ on T2-weighted when blood products may persist after prostate biopsy. Moreover, when the location of the tumor is outside our PZ model (area under the green line in Figure 2). Figure 9 shows examples from our experimental results for correct classification but incorrect localisation.

![Figure 9](image_url)

Figure 9. In both slices (image 29 and 30) the segmentation results show correct overall classification (true positive) but incorrect localisation in comparison to the location of the ground truth.

### 6.3. False positives and negatives

Figure 10 shows four examples of false positive results from four different prostates. In images 31, 32, 33 and 34, there are clearly dark regions (higher probability of cancer) within the PZ which leads to false positive results. Based on the results in Figure 10, we can visually see that irregularity can occur in some normal slices which makes it hard to differentiate between malignant and normal regions. On the other hand, Figure 11 shows examples of false negative results from four different prostates. The malignant regions show obscure irregularity which lead to false negatives.

![Figure 10](image_url)

Figure 10. False positive results from four different prostates.

![Figure 11](image_url)

Figure 11. False negative results from four different prostates. Tumor regions are delineated in red.

### 6.4. Parameters justification

Figure 12 shows the justification of our selected parameters. Based on the varying $\sigma = 2, 3, 5, 7$ and 9, the following kernel sizes $7 \times 7, 9 \times 9, 15 \times 15, 21 \times 21$ and $27 \times 27$ are applied, respectively. The results show that better sensitivity (above 80%) is achieved when $2 \leq \sigma \leq 5$. The sensitivity of the proposed method decreases when $\sigma > 5$ due to the level of smoothing applied to images. For instance, higher value of $\sigma$ would affect (e.g. over-smoothed) the appearance of small malignant regions, hence decreases the sensitivity. On the other hand, the method achieved its highest specificity when $\sigma = 5$. Using $\sigma \leq 5$ still gives similar sensitivity to the other methods in the literature (see Table 1). Although varying the $\sigma$ and kernel size did not change the results significantly, we have shown the quantitative results for the justification of our selected parameters.
Figure 12. Sensitivity, specificity and accuracy using different values of $\sigma$ and $ks$.

Figure 13. Coefficient $\epsilon = 0.75$ produce balanced results, with higher specificity for larger $\epsilon$ values and higher sensitivity for lower $\epsilon$ values.

Figure 13 shows results using different $\epsilon$ values. Our experimental results show that $\epsilon = 0.75$ produced balanced results in terms of accuracy, sensitivity and specificity. A larger $\epsilon$ value (e.g. 0.95) reduces the area under the curve (green line in Figure 2), hence most cancerous tissues were missed which increased specificity but reduced the algorithm’s sensitivity. On the other hand, a smaller $\epsilon$ value (e.g. 0.55) increases the area under the curve (green line in Figure 2). This increased the algorithm’s sensitivity (and false positives) because the larger area leads to a higher chance of cancerous tissues being detected.

7. DISCUSSION

Various methods using different frameworks, modalities and features have been proposed in the literature and our method achieved similar results. Nevertheless, it is extremely difficult to make a quantitative comparison due to:

1. Differences in datasets (different modalities such as T2-weighted (T2-W) MRI, diffusion-weighted (DWI) MRI, dynamic contrast enhanced (DCE) MRI, Magnetic resonance spectroscopy (MRS), etc) and frameworks used in the other studies.
2. Absence of public datasets also makes a quantitative comparison of methodologies in the literature difficult. Each team of researchers has their own datasets which cause huge range of variability in terms of noise and image quality.

3. Studies were conducted within different regions of the prostate. For example, some studies were conducted within the prostate PZ only and some took the whole prostate gland into account.

4. Some evaluation was at volume, slice, regions or voxel level.

However, to have an overall qualitative estimate (therefore our comparisons are subjective due to the differences stated above) of the functioning of our method we compared with some of the previous studies in Table I, where methods are categorised as CAD and non-CAD. Studies classified as CAD are methods which are similar to our method (abnormality/malignancy is automatically determined by computer algorithm). On the other hand, non-CAD are studies which involved human readers which, means abnormality/malignancy determination is performed by radiologists/observers.

Table I. Results are ordered based on performance (ACC and AUC are accuracy and area under the curve, respectively), sensitivity and specificity, respectively (all measured in %). For region, each study is either conducted for the whole prostate (WP) or only for the PZ.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Data size</th>
<th>Performance</th>
<th>Sens</th>
<th>Spe</th>
<th>Category</th>
<th>Modalities</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sung et al. [37]</td>
<td>42</td>
<td>ACC=89</td>
<td>89</td>
<td>89</td>
<td>CAD</td>
<td>DCE</td>
<td>WP</td>
</tr>
<tr>
<td>Niaf et al. [49]</td>
<td>30</td>
<td>AUC=89</td>
<td>-</td>
<td>-</td>
<td>CAD</td>
<td>T2-W+DCE+DWI</td>
<td>PZ</td>
</tr>
<tr>
<td>Vos at al. [6]</td>
<td>34</td>
<td>ACC=89</td>
<td>-</td>
<td>-</td>
<td>CAD</td>
<td>T2-W+DCE</td>
<td>PZ</td>
</tr>
<tr>
<td>Sham et al. [67]</td>
<td>24</td>
<td>F-measure=89</td>
<td>-</td>
<td>-</td>
<td>CAD</td>
<td>T2-W+ADC+DCE</td>
<td>PZ</td>
</tr>
<tr>
<td>Ampeliotis et al. [38]</td>
<td>10</td>
<td>ACC=87</td>
<td>-</td>
<td>-</td>
<td>CAD</td>
<td>T2-W+DCE</td>
<td>WP</td>
</tr>
<tr>
<td>Tiwari et al. [39]</td>
<td>19</td>
<td>ACC=84</td>
<td>-</td>
<td>-</td>
<td>CAD</td>
<td>T2-W+MRS</td>
<td>WP</td>
</tr>
<tr>
<td>Chan et al. [70]</td>
<td>15</td>
<td>AUC=84</td>
<td>-</td>
<td>-</td>
<td>CAD</td>
<td>T2-W+ADC+PD+T2 map</td>
<td>PZ</td>
</tr>
<tr>
<td>Vos et al. [68]</td>
<td>34</td>
<td>ACC=83</td>
<td>-</td>
<td>-</td>
<td>CAD</td>
<td>DCE</td>
<td>PZ</td>
</tr>
<tr>
<td>Rampun et al. [36]</td>
<td>25</td>
<td>ACC=82</td>
<td>81</td>
<td>84</td>
<td>CAD</td>
<td>T2-W</td>
<td>PZ</td>
</tr>
<tr>
<td>Artan and Yetik. [10]</td>
<td>15</td>
<td>ACC=82</td>
<td>76</td>
<td>86</td>
<td>CAD</td>
<td>DCE</td>
<td>PZ</td>
</tr>
<tr>
<td>Tabesh et al. [47]</td>
<td>29</td>
<td>ACC=81</td>
<td>-</td>
<td>-</td>
<td>CAD</td>
<td>Histological images</td>
<td>WP</td>
</tr>
<tr>
<td>Artan et al. [69]</td>
<td>21</td>
<td>ACC=78</td>
<td>74</td>
<td>82</td>
<td>CAD</td>
<td>T2-W+DCE+DCE</td>
<td>PZ</td>
</tr>
<tr>
<td>Puech et al. [71]</td>
<td>100</td>
<td>AUC=77</td>
<td>100</td>
<td>45</td>
<td>CAD</td>
<td>DCE</td>
<td>PZ</td>
</tr>
<tr>
<td>Kim et al. [55]</td>
<td>20</td>
<td>ACC=75</td>
<td>73</td>
<td>77</td>
<td>Non-CAD</td>
<td>DCE</td>
<td>WP</td>
</tr>
<tr>
<td>Han et al. [46]</td>
<td>46</td>
<td>-</td>
<td>96</td>
<td>92</td>
<td>CAD</td>
<td>Ultrasound</td>
<td>WP</td>
</tr>
<tr>
<td>Engelbrecht et al. [44]</td>
<td>36</td>
<td>-</td>
<td>93</td>
<td>-</td>
<td>Non-CAD</td>
<td>T2-W+DCE</td>
<td>PZ</td>
</tr>
<tr>
<td>Shimofusa et al. [45]</td>
<td>60</td>
<td>-</td>
<td>93</td>
<td>-</td>
<td>Non-CAD</td>
<td>T2-W+DWI</td>
<td>WP</td>
</tr>
<tr>
<td>Ito et al. [56]</td>
<td>111</td>
<td>-</td>
<td>87</td>
<td>74</td>
<td>Non-CAD</td>
<td>DCE</td>
<td>WP</td>
</tr>
<tr>
<td>Litjens et al. [11]</td>
<td>188</td>
<td>-</td>
<td>83</td>
<td>-</td>
<td>CAD</td>
<td>DWI+DCE</td>
<td>WP</td>
</tr>
<tr>
<td>Futterer et al. [40]</td>
<td>34</td>
<td>-</td>
<td>83</td>
<td>83</td>
<td>Non-CAD</td>
<td>T2-W</td>
<td>WP</td>
</tr>
<tr>
<td>Reinsberg et al. [41]</td>
<td>42</td>
<td>-</td>
<td>81-93</td>
<td>64-73</td>
<td>Non-CAD</td>
<td>DWI+MRS</td>
<td>WP</td>
</tr>
<tr>
<td>Girouin et al. [42]</td>
<td>46</td>
<td>-</td>
<td>78-81</td>
<td>32-56</td>
<td>Non-CAD</td>
<td>DCE</td>
<td>WP</td>
</tr>
<tr>
<td>Ocak et al. [18]</td>
<td>50</td>
<td>-</td>
<td>73</td>
<td>88</td>
<td>Non-CAD</td>
<td>DCE</td>
<td>WP</td>
</tr>
<tr>
<td>Llobet et al. [43]</td>
<td>303</td>
<td>-</td>
<td>57</td>
<td>61</td>
<td>CAD</td>
<td>Ultrasound</td>
<td>WP</td>
</tr>
<tr>
<td>Schlemmer et al. [54]</td>
<td>28</td>
<td>-</td>
<td>68</td>
<td>-</td>
<td>Non-CAD</td>
<td>DCE</td>
<td>WP</td>
</tr>
</tbody>
</table>

Table I presents the experimental results of 26 different methods/studies (including our method) and their accuracies, sensitivities, specificity and modalities (some authors did not report one or two
of these). Although results may be more or less similar, due to several factors mentioned previously the comparisons in this section needs to be treated with caution. Note that every method or study used a different number of patients, modalities and frameworks. However Table I gives a sense of how well the proposed method is capable of analysing cancerous regions in MRI.

In terms of performance, the methods proposed in [6, 37, 49, 67], achieved the highest result of 89%, all used support vector machines (SVM) as classifier. Sung et al. [37] employed SVM as a classifier to distinguish prostate cancer from non-cancerous tissues based on a set of perfusion parameters. Vos et al. [6] used first order statistics of the scalar values of volume as statistical features. Sham et al. [67] developed a decision support system (DSS) which used SVM to generate cancer probability maps from multiparametric MR images and Niaf et al. [49] reported that SVM produced the best result based on a comparison study of four different supervised learning methods (SVM, linear discriminant analysis, k-nearest neighbours and naive Bayes classifiers) based on a feature set derived from grey-level images such as first-order statistics, Haralick features, gradient features, etc. This was followed by a study in [38] which is 2% below the best performance (and 1% above the result of our proposed method) in Table I. Ampeliotis et al. [38] used probabilistic Neural Networks (PNNs) to classify a set of feature vectors extracted from T2- morphological images and T1-W DCE. Other methods [10, 36, 39] achieved more than 80% accuracy. Artan and Yetik [10] developed a random walker (RW) algorithm with automated seed initialisation to segment cancerous region within the PZ using weighted image features and Rampun et al. [36] developed an algorithm which compared local peaks information between the right and left region of prostate’s PZ by measuring percentage similarity and Ochiai coefficients to determine the presence of abnormality. A method proposed by Tiwari et al. [39] integrated a semi-supervised multi-kernel (SeSMiK) scheme with a graph embedding framework for data fusion and dimensionality reduction for not only prostate cancer detection but also grading. All non-CAD methods did not report their accuracies except Kim et al. [55] which is, with 75%, below the CAD’s average accuracy.

For sensitivity results, Puech et al. [71] reported their CAD software known as ‘ProCAD’ achieved 100% but only at 45% specificity. The developed software allows for the 2D and multislice 2D contouring of suspicious regions based on a seeded region growing algorithm using standard visualization features such as wash-in and wash-out slopes. Han et al. [46] proposed a method which used a combination of image features and clinical features (e.g. location and shape) and performed SVM to classify cancer and non-cancer region, achieved the second best result of 96% in a single modality. Non-CAD methods [44, 45] used perfusion parameters in multimodalities achieved 93% sensitivity the same as the maximum achieved in [41]. Our method achieved comparable result with Sung et al. [37] (89%) and Ito et al. [56] (87%) in a single modality, while Llobet et al. [43] whose method employed a Hidden Markov Model and k-nn classifiers based on texture descriptors extracted from spatial grey-level dependence metrics and grey-level maps achieved the lowest sensitivity (57%). This may have been influenced by evaluation on a large dataset (4944 ultrasound images). Methods in [41] and [42] show variations of 81-93% and 78-81%, respectively. On the other hand, Niaf et al. [49], who used a combination of three different modalities, achieved 82% similar to the method proposed in [11] based on two modalities. Litjens et al. [11] used a SVM classifier to generate a 3D likelihood map which was used to find points of interest using a local maxima detector. Subsequently, a region is segmented around each local maxima and for every region statistics of the voxel features were calculated and used to discriminate malignant and benign regions.

Finally, in terms of specificity the method of Han et al. [46] obtained the highest result of 92% followed by [37] (89%). Other CAD methods which achieved more than 80% are Artan and Yetik [10], Rampun et al. [36] and our reported method. The result reported in [18], which used perfusion parameters, achieved 88% is the highest among Non-CAD methods followed by Futterer et al. [40] (83%), while the method in [42] achieved the lowest specificity varying from 32 to 56%. Nevertheless, these comparisons are subjective as accuracy, sensitivity and specificity are highly influenced by several factors mentioned previously. For example although the method proposed in [43] produced the lowest sensitivity, the evaluation is based on 303 prostates. On the other hand, the method proposed by [38] shows higher accuracy on 10 different cases but has not been tested
on larger datasets. Similarly, although Han et al. [46] achieved the highest sensitivity and specificity (based on Table 1) but was evaluated on 46 ultrasound images (46 patients).

The challenges of developing a CAD system remain open due to its complexity and limitations both in single and multimodalities imaging. In this study, we are aware that many researchers have attempted to improve results in detecting prostate cancer in MR imaging. Engelbrecht et al. [44] and Sung et al. [37] showed the advantages of using perfusion parameters (e.g. wash in and wash out rates) in detecting prostate cancer in DCE. The method proposed in [6] used a multiparametric MR of T1- and T2-weighted imaging showing better results using a single modality. Shimofusa et al. [45] showed a significant improvement in prostate cancer detection using diffusion-weighted imaging in addition to T2-weighted MRI and reported that sensitivity increased from 87% to 93%.

In another study, Reinsberg et al. [41] combined the use of diffusion-weighted MRI and 1H MR Spectroscopy which lead to 93% sensitivity with 73% specificity. The method proposed in [36] used local peak information to detect prostate abnormality by measuring the information difference between left and right PZ regions in a single modality of T2-weighted imaging. Han et al. [46] used clinical knowledge to discriminate the cancer region by location and shape of the region in addition of image features to increase specificity. Moreover, Kim et al. [55] made a comparison between T2-weighted and DCE imaging and reported that DCE has better accuracy (75%) and sensitivity (73%) but 11% lower in specificity (88% in T2-W). Another comparison study was made by Ito et al. [56] between power Doppler ultrasound (PDUS) and DCE imaging and concluded that DCE has higher sensitivity (87%) and specificity (74%) in comparison to PDUS which only achieved 69% and 61%, respectively within the PZ. Llobet et al. [43] who evaluated their method based on 4944 ultrasound images showed similar results with the ones reported on fewer samples by Schlemmer et al. [54]. However, in contrast to the earlier methods, our method is different in the sense that:

1. The proposed method does not need a training phase to be able to discriminate malignant and benign tissues in contrast to the methods in [6, 11, 38, 39, 49, 54, 10].
2. We only used a single modality for abnormality detection which is T2-Weighted MRI. The methods in [44] used multimodality such as diffusion MRI and MR Spectroscopy. Similarly, the method proposed in [6] used a multiparametric MR of T1- and T2-weighted imaging. Engelbrecht at al. [45] suggests that various techniques such as dynamic contrast material enhanced MR imaging, diffusion-weighted imaging, and MR spectroscopy have the potential to improve the detection of prostate cancer. On the other hand [41] combined the use of diffusion-weighted MRI and 1H MR Spectroscopy to get better results in discriminating malignant and normal tissues.
3. The method in [46] used additional clinical knowledge (e.g. location and shape of the region) to discriminate cancer regions in addition of image features while our method only used image features to achieve similar results.
4. Our method used a small number of image features to discriminate malignant and benign regions and produced similar results to the state of art in the literature whereas the methods in [46, 47] used more features.
5. The methods in [18, 37, 44, 45] used various perfusion parameters on a single modality while our method is purely based on image features but still managed to achieve similar results.

The proposed method produced similar accuracy, sensitivity and specificity to the state of art in the literature particularly in single modality T2-Weighted MRI. However, due to various factors mentioned a direct comparison is less appropriate. Several studies [6, 41, 44, 45, 49] have suggested that using image fusion (e.g. combining MRI T2-W with DCE) produces better results in detecting prostate cancer. Nevertheless, other researchers have attempted to make a single predictor (T2-W MRI only) by detecting prostate cancer on the basis of comprehensive analysis of various perfusion parameters, such as in [18] achieved 75% sensitivity and 80% specificity (50 patients). However, without the parameters they achieved higher sensitivity of 94% but much lower specificity of 37%

In another study, Miao et al. [53] reported 76% and 70% sensitivity and specificity, respectively (30 patients) while a study conducted in [42] achieved 50-60% sensitivity and 13-21% specificity in 46 patients. Kim et al. [55] reported 55% sensitivity and 88% specificity (20 patients), and Schlemmer
et al. [54] achieved 79% sensitivity in 28 patients. On the other hand, other investigators [46] have studied the prospect of combining clinical knowledge and image features to detect prostate cancer and achieved similar results. Whether using a single modality, image fusion or using clinical features, none of these methods provide superior results. Therefore developing a CAD tool for prostate cancer detection and localisation remains a challenge.

In our study, one obvious drawback of the proposed method is the risk of classifying correctly with incorrect localisation of the tumor, which could be problematic from a clinical point of view. Secondly, in some cases when the prostate’s peripheral zone is almost non-existent, the proposed method is more likely to produce false positives. This is due to the intensities being very similar between the central gland and the malignant region [58]. Finally, if the prostate’s shape does not conform to the shape of our prostate model a smaller area of the PZ will be analysed which may increase the chance of malignant regions being missed. Therefore, in order to accommodate these limitations for future work we are planning to use a multiparametric approach (e.g. T2-W+DWI+DCE) instead of stand alone T2-W MRI. This means more image features can be extracted which could help to distinguish malignant and benign tissues. In addition, we intend to cover the whole prostate gland instead of only the PZ.

8. CONCLUSIONS

The proposed method specifies regions which have the highest probability to be malignant (see results in Figure 7), hence help radiologists to perform targeted biopsies and potentially improve the accuracy of prostate cancer diagnosis [57].

In conclusion, we have presented a novel method of prostate cancer detection and localisation within the PZ and successfully applied it on 37 patients. Gaussian and median filters together with probability image information show promising potential to be effective texture descriptors to identify cancer regions within the peripheral region. Our idea, which is based on regions intersection and flexible size of erosion’s structuring element, suggest a good potential to reduce false positive and false negative results in the proposed method.

ACKNOWLEDGEMENT

Andrik Rampun would like to thank for the awards given by Aberystwyth University under the Departmental Overseas Scholarship (DOS) and Doctoral Career Development Scholarships (DCDS). This work was funded in part by the NISCHR Biomedical Research Unit for Advanced Medical Imaging and Visualization.

REFERENCES


