A Block-based Approach for Malignancy Detection within the Prostate Peripheral Zone in T2-weighted MRI

Andrik Rampun¹, Paul Malcolm² and Reyer Zwiggelaar¹

¹Department of Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, U.K.
²Department of Radiology, Norfolk & Norwich University Hospital, Norwich NR4 7UY, U.K.

Keywords: Prostate Cancer Detection, MRI, Block-based approach, Grey Levels Appearance.

Abstract: In this paper, a computer-aided diagnosis method is proposed for the detection of prostate cancer within the peripheral zone. Firstly, the peripheral zone is modelled according to the generic 2D mathematical model from the literature. In the training phase, we captured 334 samples of malignant blocks from cancerous regions which were already defined by an expert radiologist. Subsequently, for every unknown block within the peripheral zone in the testing phase we compare its global, local and attribute similarities with training samples captured previously. Next we compare the similarity between subregions and find which of the subregion has the highest possibility of being malignant. An unknown block is considered to be malignant if it is similar in comparison to one of the malignant blocks, its location is within the subregion which has the highest possibility of being malignant and there is a significant difference in lower grey level distributions within the subregions. The initial evaluation of the proposed method is based on 260 MR images from 40 patients and we achieved 90% accuracy and sensitivity and 89% specificity with 5% and 6% false positives and false negatives, respectively.

1 INTRODUCTION

Prostate cancer is the most commonly diagnosed cancer among men and remains the second leading cause of cancer death in men. 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 (Howlader et al., 2013; Chou et al., 2010; PCUK, 2014). In the last decade, prostate cancer screening has been receiving more attention because it can help to detect cancer at an early stage before there are any symptoms. Nowadays, clinical diagnostic tools are very popular and globally used despite their inconsistency (60% - 90%) in producing accurate results (Yu and Hricak, 2000). Many factors causing this inconsistency such as random biopsy tests (hence, higher chance of cancerous tissues being missed), less sensitivity to detect slow-growing and non-aggressive tumors (as a result tumors often detected in the late age (above 70 years old)), inaccurate results (e.g. an elevated level of PSA in the blood does not necessarily indicate cancer), etc.

Computer-aided diagnosis (CAD) of prostate magnetic resonance imaging (MRI) has the potential to improve the results of clinical diagnostic tools. Unfortunately, prostate MRI requires a high level of expertise and suffers from observer variability (Vos et al., 2010). CAD systems can be of benefit to improve the diagnostic accuracy of clinical methods, reduce variability and speed up the reading time of clinicians (Vos et al., 2010). In contrast to segmentation algorithms, detection algorithms only try to decide if tumor is present and output the approximate tumor location instead of providing a complete segmentation. Therefore, the main goal of our research is to develop CAD methods which automatically delineate and localise malignant regions, leading to a reduction of search and interpretation errors, as well as a reduction of the variation between and within observers (Vos et al., 2010). In this paper, we propose a block-based approach to find malignant regions within the prostate peripheral zone. A block-based approach is chosen due to its efficiency in comparison to a pixel by pixel sliding window approach (see section 4). The key idea of this method is, a block or patch is considered being malignant if its grey levels distribution is similar with our training samples. We use specific metrics to measure similarity for malignancy detection which will be explained in section 3.3.
2 PERIPHERAL ZONE MODEL

Pathologically, 80% – 85% of the cancers arise in the Peripheral Zone (PZ) and the rest are within the central zone (Edge et al., 2010). Since the percentage occurrence of abnormality in the peripheral zone is high, we aim to detect prostate abnormality within that region similar to studies in (Artan and Yetik, 2012), (Ito et al., 2003) and (Ocak et al., 2007). We did not perform prostate segmentation because all prostates were already delineated by our expert radiologist (the same in (Artan and Yetik, 2012)). To automatically capture the PZ, we used a generic 2D mathematical model proposed in (Rampun et al., 2014c; Rampun et al., 2013). Figure 1 shows an example of prostate MRI with its ground truth and the generic 2D prostate model.

The prostate’s PZ is defined using the quadratic equation \( y = ax^2 + bx + c \) based on three crucial coordinate points of the prostate which are \( v_1, v_2 \) and \( v_3 \). They are determined by the outmost \( x \) and \( y \) coordinates of the prostate boundary which are \( x_{min}, x_{max}, y_{min}, y_{max} \) (see Figure 1). For example, \( x_{min} \) and \( y_{max} \) can be determined by taking the minimum \( x \) and maximum \( y \) coordinates along the prostate boundary. Moreover, the \( x \) coordinates of \( v_1 \) and \( v_3 \) are captured from \( x_{min} \) and \( x_{max} \) and their \( y \) coordinate is determined by taking the \( y \) coordinate between \( y_{min} \) and \( y_{max} \).

\[
C_p = \left(\frac{(x_{min} + x_{max})}{2}, \frac{(y_{min} + y_{max})}{2}\right) \quad (1)
\]
\[
v_1 = \left( x_{min}, \frac{(y_{min} + y_{max})}{2}\right) \quad (2)
\]
\[
v_2 = \left( \frac{(x_{min} + x_{max})}{2}, y_{min} + \left(\frac{(y_{max} - y_{min}) \times \frac{3}{4}}{4}\right) \right) \quad (3)
\]
\[
v_3 = \left( x_{max}, \frac{(y_{min} + y_{max})}{2}\right) \quad (4)
\]

where \( C_p \) is the central coordinate of the prostate gland. Once the coordinates of \( v_1, v_2 \) and \( v_3 \) are defined, we can determine the boundary of the PZ.

Finally, the PZ region is divided into four subregions \( (R_1, R_2, R_3, R_4) \) according to the prostate anatomy in the European consensus guidelines division of prostate gland (Dickinson et al., 2011).

3 METHODOLOGY

Figure 2 shows the summary of the proposed methodology. In the training phase, we capture region of interests (ROI, also known as a block) from tumor regions. In Figure 2, \( M_{ROI} \) contains many collections of ROIs which basically represents different patterns of grey level distribution within malignant regions. Subsequently, in the testing phase each unknown block \( U \) within the PZ is compared with every block in \( M_{ROI} \). Based on the comparison results, if \( U \) is similar (measured using specific metrics) with any blocks in \( M_{ROI} \), then we assume \( U \) is most likely being malignant.

3.1 Preprocessing

Similar to the study in (Artan and Yetik, 2012), each image MRI is median filtered to reduce noises as well as to preserve sharp regional boundaries (in our case we want to preserve the information-bearing structures such as tumor’s edge boundaries).

3.2 Training

In training phase, we captured 334 number of malignant blocks (334 of ROIs as illustrated in 2) sized
7 × 7 (this size gives the best quantitative experimental results, see experimental results) from 5 patients (50 malignant regions of 50 MRI slices/images). Note that, each malignant region is already delineated by an expert radiologist. For every malignant region in MRI image, we divide it into blocks (ROI) and only take ROI if all of its elements are within the tumor region (red in figure 2). This means ROI is excluded if one or more of the pixels within the block is outside the tumor boundary to ensure reliability in our training samples (we want to make sure that our training samples are purely taken from malignant tissues). At the end of this process, $M_{ROI}$ contains 334 number of malignant blocks as shown in equation (5).

$$M_{ROI}(i) = \{ROI_1, ROI_2, ROI_3, \ldots, ROI_i\} \quad (5)$$

where $i$ is the $i^{th}$ block in $M_{ROI}$.

### 3.3 Testing

In the testing phase, for each unknown block ($U$) we compare its similarity with every block in $M_{ROI}$ (hence there will be 334 comparisons for every $U$). If $U$ is similar with one of the blocks in $M_{ROI}$, we assume that the PZ has higher possibility of being malignant. Note that for blocks size selection, we performed exactly the same in the training phase within the prostate gland and PZ boundaries. On the other hand, we used the following metrics to measure similarity between $U$ and $M_{ROI}$: global similarity, local similarity and attribute similarity which have the minimum and maximum values of 0 and 1.0, respectively.

In all cases we use a default threshold 0.5 which means if the similarity is more than half of its maximum similarity value (1.0), $U$ is considered being malignant. According to (Ali et al., 2005) the human eye is less sensitive in perceiving a change in shape and texture if the similarity percentage between two objects is less than 50% (hence studies in (Ali et al., 2005; Hasan et al., 2009; Hasan et al., 2012) have used a default threshold 0.5 in measuring similarity between two textures). In fact, a study conducted by (Chen et al., 2006) discussed the use of decision threshold (minimum and maximum are 0 and 1, respectively) adjustment in classification for cancer prediction concluded that when the sample sizes are similar (in our case each $U$ and $ROI_i$ are the same size), they suggested that the optimal decision threshold and balanced decision is approximately 0.5 as their experimental results show high concordance with a balance between sensitivity and specificity. When estimating string similarity, (Elita et al., 2007) considers the sequence of common elements have to be at least 50% in both strings to be considered similar.

#### 3.3.1 Global Similarity

Global similarity ($G$) measures the number of elements (#$U$) in $U$ within the range of $i^{th}$ block in $M_{ROI}$. This metric does not concern about the individual grey level value in ROI but the overall range. To calculate the global similarity:

1. Calculate the mean ($ROI_{mean}$) and standard deviation ($ROI_{std}$)

2. Calculate its lower ($ROI_{Lrange}$) and upper ($ROI_{Urange}$) ranges using equation (6) and (7)

3. Calculate $G$ using equation (8)

$$ROI_{Lrange} = ROI_{mean} - ROI_{std} \quad (6)$$

$$ROI_{Urange} = ROI_{mean} + ROI_{std} \quad (7)$$

$$G_i = \frac{\#U \leq ROI_{Urange} - ROI_{Lrange}}{n \times m} \quad (8)$$

where $n$ and $m$ are the size of the block. Equation (8) calculates similarity based on the number of elements in $U$ within the range of $i^{th}$ block in $M_{ROI}$. This means the more elements in $U$ within the range of $i^{th}$ block, the more similar they are (leads to higher possibility being malignant).

#### 3.3.2 Local Similarity

Local similarity ($L$) measures the amount of overlapping information between two blocks. In comparison to global similarity, this metric compares corresponding elements in $U$ and $ROI_i$. For every $U$ we measure its $L$ using equation (9).

$$L_i = \frac{\sum \min\{U(z), ROI_i(z)\}}{\sum \max\{U(z), ROI_i(z)\}} \quad (9)$$

where $z$ represents every single element in a block. In this measure, the higher the overlapping value means the more similar the blocks (between $U$ and $ROI_i$).

#### 3.3.3 Attribute Similarity

Attribute similarity ($A$) measures the number of occurrences of unique overlapping grey levels between two blocks. This means only grey levels appear in both blocks will be taking into account. This metric can be calculated using the following steps:

1. Find unique elements in $U$ and $ROI_i$, let say $U_{unique}$ and $ROI_{i,unique}$

2. Count the number of grey level occurrences in $U_{unique}$ and $ROI_{i,unique}$, let say $f_{U,unique}$ and $f_{ROI_{i,unique}}$

3. Find the overlapping elements in $U_{unique}$ and $ROI_{i,unique}$, let say $U_{unique} \cap ROI_{i,unique}$
4. Count the number of occurrences for every element in $U_{\text{unique}}$ and $ROI_{\text{unique}}$, let say $f'_{U_{\text{unique}}:ROI_{\text{unique}}}$.

5. Use equation (10) to calculate $A_i$

$$A_i = \frac{\sum f'_{U_{\text{unique}}:ROI_{\text{unique}}}}{\sum f'_{U_{\text{unique}}:ROI_{\text{unique}}}}$$

where $f'_{U_{\text{unique}}:ROI_{\text{unique}}}$ represents the number of grey levels occurrences in both $U$ and $ROI_i$ (the same as the number of grey levels in $U$ and $ROI_i$).

### 3.4 Subregion Similarity

In addition to the metrics in section 3.3, we use histogram intersection distance ($d$) to measure the similarity between two subregions’ histograms (e.g., histogram from $R_1$ and $R_2$ in Figure 1) by measuring their distance in the intersection space (Rubner et al., 2000). Previous studies (Rampun et al., 2014b; Rampun et al., 2014a) have shown that the peripheral zone has a higher chance of being malignant if one of the subregions contains a significant number of lower grey levels (e.g., below 120) in comparison to the other subregions. For example, $R_1$ contains 80% lower grey levels whereas $R_3$, $R_5$, and $R_4$ are dominated by upper grey levels (this makes $R_1$ looks darker in T2-Weighted-MRI image in comparison to the other subregions). Several studies suggested that prostate cancer tissue tends to appear darker on T2-weighted MRI images (Garnick et al., 2012; Ginat et al., 2009; Mohamed et al., 2003).

Moreover, radiologists also tend to use regions intensity (e.g., darker region is more likely to be malignant) to identify abnormality within the peripheral zone (Edge et al., 2010). Therefore, the main objective of this metric is to find subregion which has the highest possibility of being malignant. We chose this metric because of its capability to handle partial matches when the areas of two regions are different (Rubner et al., 2000). In our case, every area of subregion is different (due to PZ and prostate boundaries). For each subregion (e.g., $R_1$), we construct grey level histogram (255 bins) by assigning every pixel to its appropriate grey level and normalise it to sum equal to 1. After normalisation, we calculate the histogram’s mean value to roughly identify which subregion has the highest number of lower grey levels. The same process applied to the other subregions ($R_2$, $R_3$, and $R_4$). Subsequently, we take histograms with the lowest and highest histogram mean values and calculate $d$ using equation (11).

$$d = 1 - \left(\frac{\sum \min\{H_{\text{max}}(j), H_{\text{min}}(j)\}}{\sum H_{\text{max}}(j)}\right)$$

where $j$ represents each bin in the histogram and $H_{\text{max}}$ and $H_{\text{min}}$ are subregions’ histograms which have the highest and lowest histogram mean values, respectively. In contrast to the metrics in section 3.3, smaller $d$ indicates higher similarity as it means smaller distance separates two histograms in the intersection distance (hence lower possibility of the PZ being malignant).

### 3.5 Malignancy Detection

An unknown sample/block ($U$) is considered to be malignant if all of the following conditions are true:

1. If its $G$, $L$ and $A$ are greater than 0.5 and
2. If $d > 0.5$ and
3. If the location of $U$ is within the subregion with the lowest histogram mean value

Figure 3 shows the flow chart decision rules for malignancy detection in the proposed method. Note that the selected $R$ is the subregion with the lowest histogram mean value. If the location of the segmented region is not within the subregion which has the lowest histogram mean value, we do not consider $U$ as malignant because previous studies (Rampun et al., 2014b; Rampun et al., 2014a) shows that in many cases cancerous region has lower intensity as previously mentioned in (Garnick et al., 2012; Ginat et al., 2009; Taneja, 2004; Mohamed et al., 2003). Note that the threshold 0.5 is selected based on the studies conducted by (Chen et al., 2006; Ali et al., 2005; Hasan et al., 2009; Hasan et al., 2012; Elita et al., 2007) in...
different applications. Hence, we did not perform experiments to determine a new threshold value due to the extensive experiments have been done in the literature.

4 EXPERIMENTAL RESULTS

For evaluation, our database contains 260 MRI images (118 images are identified as malignant and 142 are benign/normal) from 40 different patients aged 54 to 74 collected from the Norfolk and Norwich University Hospital. The prostates, cancers and central zones were delineated by an expert radiologist on each of the MRI images and all malignant lesions are biopsy proven. Each image was analysed and classified as to whether the prostate contains malignancy (and segment the approximate tumor location) based on the decision rules explained in section 3.5. Subsequently, we compared the result with the ground truth whether the prostate contains malignancy or not. The proposed method achieved 90% accuracy and sensitivity and 89% specificity with 5% and 6% false positives and false negatives, respectively. Note that the term accuracy here means correct classification rate (CCR) which means if an image is classified as containing malignancy and the ground truth is malignant then the image is classified correctly. The first two examples in Figure 4 shows the results of two MR images containing malignant region in each of them.

![Figure 4: Examples of malignant cases.](image1)

The ground truth of malignant regions are within the red lines and green blocks are blocks identified as malignant. For every case, we use a generic PZ mathematical model in section 2 to define the PZ boundary (magenta line) and divide the PZ model into four subregions (partitioned by the blue lines). As we can see in image 1 and 2 the segmented blocks are within the malignant regions indicating correct classification and approximate location. In image 1, a small block segmented within $R_3$ (false positive) which indicates malignancy. However, since the subregion $R_1$ has the lowest average grey level value (and $d > 0.5$), we consider the segmented region in $R_1$ has higher possibility of being malignant and only consider blocks in $R_1$. Similarly in image 2, segmented blocks in $R_3$ are considered having higher possibility of being malignant because subregion $R_3$ has the lowest grey level average value (and $d > 0.5$), hence a segmented block in $R_3$ are considered to have lower possibility of being malignant.

![Figure 5: Examples of cases with two malignant regions.](image2)

Figure 5 shows examples of MR images containing two malignant regions. In both image 3 and 4, sub-region $R_1$ has the lowest grey level mean value and $d > 0.5$. The results show that most malignant blocks were found to be malignant within the cancerous regions (red lines). We also can see some segmented blocks within the second malignant regions in both images. On the other hand, Figure 6 shows examples of segmentation results using different window sizes. In image 5, there is no segmented region using $9 \times 9$ window size but we can clearly see several regions were segmented in image 6 using $7 \times 7$ window. This indicates bigger window size decreases the sensitivity rate and increases the specificity rate.

![Figure 6: Examples of segmentation results using different window sizes.](image3)

We tested the accuracy, sensitivity and specificity of the proposed method using different window sizes of $3 \times 3, 5 \times 5, 7 \times 7, 9 \times 9$ and $11 \times 11$. Based on the results shown in Figure 7, sensitivity increases using a small size of window $3 \times 3$ and gradually decreases as the window size is getting bigger. On the other hand, specificity decreases using a small window and increases when a bigger window size is used (e.g. $9 \times 9$). A balanced results in terms of accuracy, sensitivity and specificity were achieved using a medium window size of $7 \times 7$.

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 the differences in datasets (different modalities such as T2-weighted MRI, diffusion-weighted MRI, etc) and frameworks (e.g. combining different modalities) used in the other studies. In fact, absence of public datasets also makes a quantitative
From the qualitative comparisons in Table 1, the proposed method achieved similar results with the state of the arts in all metrics (accuracy, sensitivity and specificity). Note that some of the authors did not report one or two of their results. In terms of accuracy and sensitivity, our method achieved the best result in comparison to the other methods in Table 1. However, the proposed method of (Sung et al., 2011) achieved similar results in all metrics of 89%. (Vos et al., 2010) and (Ampeliotis et al., 2007) did not report sensitivity and specificity of but their proposed methods achieved similar results 89% and 87%, respectively. For sensitivity, (Han et al., 2008) reported their proposed method achieved 96% which is the highest in Table 1 where the proposed methods of (Engelbrecht et al., 2003) and (Shimofusa et al., 2005) achieved 93% followed by our method 90%. (Llobet et al., 2007) achieved 57% sensitivity based on 303 number of cases where (Litjens et al., 2011) achieved 83% in 188 number of cases. Finally, in terms of specificity, the proposed method of (Han et al., 2008), once again achieved the highest result of 92% where our proposed method achieved 89% similar with the methods proposed by (Sung et al., 2011), and (Artan and Yetik, 2012). (Llobet et al., 2007) once again achieved the lowest result of 61% but evaluated based on the largest dataset in Table 1.

These comparisons are subjective as results are highly influenced by the number of datasets, different modalities and methods frameworks. For example although the method proposed of (Llobet et al., 2007) produced the lowest sensitivity and specificity but the evaluation is based on 303 prostates. On the other hand, although (Han et al., 2008) achieved the highest sensitivity and specificity (based on Table 1 ) but was evaluated with smaller dataset contains 46 ultrasound images (46 patients). Similarly, although the proposed method achieved 90% for both accuracy and sensitivity, these are based on a specific window size (7 × 7).

The performance of the proposed method vary according to window sizes (e.g accuracy 72% and 80% with 3 × 3 and 5 × 5 window size, respectively, see 7). In terms of computationally complexity, the proposed method took approximately 3 hours 30 minutes in both training and testing phases (310 MRI images, Matlab 2013, Windows 7 Intel core i5). On the other hand using a pixel by pixel sliding window approach took approximately 30 hours for the entire experiment with 7 × 7 window size. Using smaller window size (e.g. 5 × 5) took even longer. Therefore we chose block based approach due to its speed performance.

5 DISCUSSIONS

In contrast to the earlier methods, our method is different in the sense that:

1. We only used a single modality for abnormality detection which is T2-Weighted MRI whereas (Engelbrecht et al., 2003) used multimodality such as diffusion MRI and MR Spectroscopy.

2. The method in (Han et al., 2008) used additional clinical knowledge (e.g. shape of the region) to discriminate cancer regions in addition of image features while our method only used the information of grey level distributions between two samples to achieve similar results.
3. The methods in (Ocak et al., 2007; Sung et al., 2011) used various perfusion parameters on a single modality while our method is purely based on the information of grey level distributions to achieve similar results.

This paper makes three contributions:

1. A novel approach of CAD diagnosis method which is based on similarity measure between unknown and sample blocks.

2. The development of the proposed method and its application in prostate cancer detection and localisation using a single MRI modality with comparable results to the state-of-the-art methods in the literature.

3. We introduced a simple approach in measuring similarity based on the occurrences of overlapping grey levels between two samples (section 3.3.3).

6 CONCLUSIONS

The proposed method could help clinicians to perform targeted biopsies which ought to be better and potentially improve the accuracy of prostate cancer diagnosis. In conclusion, we have presented a novel method of prostate cancer detection and localisation within the PZ. In this paper we have showed the potential of grey level information (G, L and A) in predicting malignancy by comparing each unknown sample with malignant samples. In addition to that, we also have used statistical grey levels information (mean value) to find subregion which has the most possibility of being malignant. By combining these information, we achieved promising and similar results with the state of the arts in the literature. Nevertheless, the best results were achieved using $7 \times 7$ window size and achieved lower accuracy and specificity in smaller window sizes.

REFERENCES


ABlock-basedApproachforMalignancyDetectionwithintheProstatePeripheralZoneinT2-weightedMRI


