Prostate Cancer Recognition by Region-based Heterogeneity Analysis of T2w-MRI

Gihan Samarasinghe[1] Arcot Sowmya[1] Daniel A. Moses[2]

[1] School of Computer Science and Engineering,
University New South Wales, Sydney NSW 2052, Australia.
gihan.samarasinghe, a.sowmya@unsw.edu.au
[2] Prince of Wales Clinical School, Prince of Wales Hospital,
Faculty of Medicine,
University New South Wales, Sydney NSW 2052, Australia.
daniel.moses@unsw.edu.au

Technical Report
UNSW-CSE-TR-201808
November 2018

UNSW SYDNEY

School of Computer Science and Engineering
The University of New South Wales
Sydney 2052, Australia

Abstract

Recognition of prostate cancer is important prior to treatment, especially inside the prostate peripheral zone where the majority of tumours occur. A classification framework for recognition of suspicious peripheral zone lesions by region-based heterogeneity analysis in prostate T2w-Magnetic Resonance Images (MRI) is developed and evaluated. The most critical component in the proposed framework is the feature extraction method. Four different novel features are derived for a selected peripheral zone lesion, based on the heterogeneity of the whole peripheral zone on the corresponding 2D MRI slice. When deriving these features, in addition to the relative intensity distribution of regions within the remaining peripheral zone, a distance function that measures the distance of a region from the lesion is taken into account. This guarantees that the adjacent regions, where the influence of the selected lesion is higher, are assigned a lower weight in the computation of intensity discrimination.The developed features were used to build a probabilistic Naive Bayes classifier using 108 peripheral zone lesions in 3.0T T2-weighted MRI datasets for 56 patients, and evaluated against the ground truth provided by an expert radiologist. Quantitative results obtained using 5-fold cross-validation show that the classification performance depends on the distance function used in feature extraction. A second order distance function achieves the best classification results (90.8% sensitivity, 92.3% specificity, 91.5% accuracy, and 91.6 AUC), and significantly outperforms traditional image features found in the literature that are based on true intensities and specific intensity within the peripheral zone.

1 Introduction

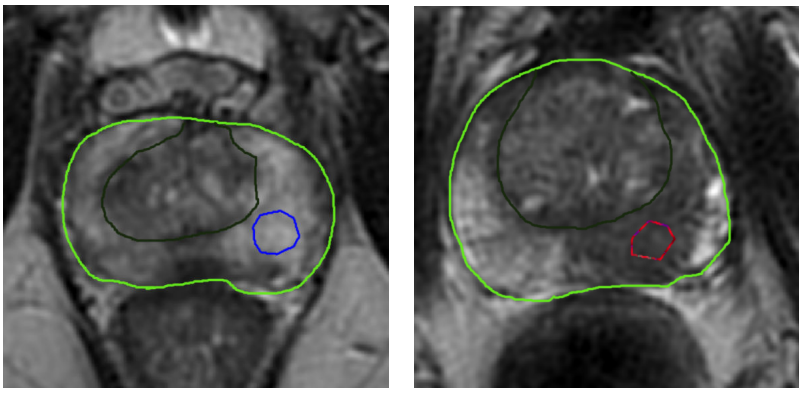
Improving treatments for prostate cancer is critical, and accurate recognition of prostate cancer in terms of detection, diagnosis, staging and tracking is as important as the treatment itself. Traditional physical diagnosis methods including Digital Rectal Examination (DRE) and the Prostate Specific Antigen (PSA) test fail to provide adequate sensitivity and specificity, and needle biopsies are known to cause morbidity [22]. Instead, medical imaging is proving to be a more accurate non-invasive method of recognising prostate cancer [12]. In comparison with other popular imaging techniques used for imaging of the prostate gland, including Transrectal Ultrasound (TRUS) and Computed Tomography (CT), Magnetic Resonance Imaging (MRI) has several advantages, including higher spatial resolution and contrast, better safety in terms of no radiation involvement and availability of multiple MRI modalities [23]. Advances in MRI technology facilitate scanning under higher strength magnetic fields, and modern 3.0T MRI scanners provide higher quality images with higher resolution and contrast, reduced noise and faster acquisition times [9].

Availability of high quality, higher order information with modern MRI technology allows computer based processing to play a significant role in assisting medical experts make better decisions in prostate cancer recognition. MRI based on a variety of different protocols, called multimodal MRI, appears to have great potential for improving prostate cancer diagnosis and treatment planning, improving both sensitivity and specificity at virtually all stages of prostate cancer diagnosis and treatment.

T2w-MRI is the basic and most common MRI modality used in prostate cancer recognition due to its wide availability, lower cost and faster acquisition compared to other more novel MRI modalities. Even though the usage of multi-modalities of MRI in unified CAD algorithms is becoming popular and successful, T2w-MRI alone can provide good accuracy of ROI classification, when used with advanced feature extraction [10]. Many studies have applied various supervised and unsupervised classification algorithms to classify ROIs in T2w-MRI, however the selection of image features is the most important step in achieving better accuracy [13]. Most cancers in the prostate are found in the peripheral zone [17], therefore many studies focus on ROI classification in the peripheral zone. A healthy prostate peripheral zone exhibits homogeneously higher T2w-MRI signal intensity distribution. In the presence of a tumour, however, T2w-MRI signals have lower values [8], which permits the utilisation of region-based statistical intensity features for ROI classification. A tumour ROI and a normal ROI are marked in two example prostate T2w-MRI slices with manually segmented prostate boundary and the peripheral zone are shown in Fig 1.1.

It is also well known that the accuracy of ROI classification in T2w-MRI is affected by inter-patient and inter-scan variations in signal intensity distribution, therefore standardisation of images is important for effective feature extraction [19]. Three methods are traditionally used to address this: statistical modelling of voxels within the peripheral zone, within the whole prostate or even within the whole image including the background [2, 14, 15, 20, 24, 25]; modelling of the intensity variations of neighbouring organs in the MRI dataset [18, 26]; and heterogeneity analysis either in the peripheral zone or in the whole prostate [21].

Deriving features using region-based heterogeneity analysis within the prostate



(a) (b)

Figure 1.1: Two different 2D prostate T2-MRI slices where the prostate boundaries are marked in green, and the peripheral zones are marked in black. A normal ROI is marked in blue in (a), where an ROI is marked within a tumour in (b).

or within the peripheral zone is a sound method for eliminating the bias towards absolute T2w-MRI signal intensity values. Region-based heterogeneity analysis techniques can derive features based on relative signal intensity variation of different regions within an organ, instead of absolute signal intensity values, and is also a good method of bypassing image normalisation. However this approach has not been used much in the literature, and the only study in this category [21] presents an algorithm to detect tumours in the peripheral zone by analysing the heterogeneity of pre-defined sub-regions of the peripheral zone.

This paper presents a novel method of feature extraction in the prostate peripheral zone by analysing the heterogeneity within the whole peripheral zone in the same 2D slice of a T2w-MRI dataset. A set of intensity-based features is derived for a marked ROI on a 2D T2w-MRI slice, where the divergence of the ROI primarily from the remaining peripheral zone is measured, and secondarily from the central gland. Importantly the features are constructed such that relative intensities within a single MRI dataset are taken into account, and the inter-patient and / or inter-scan intensity distribution variations are eliminated from the feature extraction model. The derived heterogeneity-based features are then used in computer-based supervised classification algorithms, and the results achieved are tested against the ground truth marked by an expert radiologist. Also the classification performance with the proposed features are compared against the classification performance with absolute intensity-based features alone.

The rest of the paper is organised as follows. The proposed feature extraction method is described in section II. The experimental design with evaluation and discussion of the results is in section III. Finally, section IV concludes the paper.

2 Method

The implementation of the proposed method is carried out in two major steps: (i) pre-processing of T2w-MRI datasets and (ii) feature extraction. These are discussed now.

2.1 Data Pre-Processing

In order to reduce irregular noise, each MRI dataset is filtered using a three dimensional Gaussian distribution function to obtain a filtered image signal field I_0 by convolving the original MRI signal field I_{mri} with a Gaussian kernel G:

$$I_0(x, y, z) = I_{mri}(x, y, z) * G(x, y, z) \tag{2.1}$$

where

$$G(x, y, z) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2+z^2}{2\sigma^2}} \tag{2.2}$$

The purpose of Gaussian filtering is to eliminate local irregularities (outliers) within smaller windows, therefore after subjective observation of Discrete Fourier Transformation (DFT) outputs of 10 randomly selected ROIs, σ was selected to have value 2. Then as the first step towards eliminating bias towards the absolute T2w-MRI signal intensity of a voxel $I_0(x, y, z)$, the original image voxels within the whole prostate gland (including the peripheral zone and central gland) are normalised and represented as $I(x, y, z)$:

$$I(x, y, z) = I_0(x, y, z).\frac{1}{I_{max}(WG) - I_{min}(WG)} \tag{2.3}$$

where $I_{max}(WG)$ and $I_{min}(WG)$ are the maximum and minimum absolute T2w signal intensities within the whole prostate gland (WG) respectively.

2.2 Feature Extraction and Representation

Four different features are extracted for a selected ROI L within the prostate Peripheral Zone PZ; these features emphasise the signal intensity of the ROI relative to the rest of the peripheral zone and the Central Gland CG, based on the fact that tumours in the peripheral zone show relatively lower T2w-MRI signal intensities. To extract these features, heterogeneity is analysed only on the 2D slice on which the ROI L is marked. To derive all the features discussed below, the scaled intensity model ($I(x, y, z)$ in eqn 2.3) is used. As all the parameters are calculated on the same axial 2D slice on which the ROI is marked, each pixel is notated with respect to the x and y coordinates only (i.e. $p(x, y)$).

Feature I: Distance Weighted Specific Intensity within Peripheral Zone

This feature emphasises the intensity difference of the selected ROI L against the remaining pixels within the peripheral zone on the corresponding 2D T2w-MRI slice. When computing this feature, the intensity difference of each pixel within the rest of the prostate and the average intensity of the selected ROI L is weighted by the Euclidean distance to the particular pixel from the geometric

centroid of the ROI L. The motivation behind defining this feature is that when the selected ROI L is part of a tumour, the tumour and its lower intensity behaviour may spread to adjacent pixels [16]. Therefore it is important to give higher weight to pixels that are farther away but still within the peripheral zone when computing the specific intensity of the ROI L. Initially the unit Euclidean distance of each pixel within the rest of the peripheral zone from the ROI centroid $c_L(x_L, y_L)$ is calculated as:

$$d(x,y) = \frac{1}{D}\sqrt{(x_L - x)^2 + (y_L - y)^2} \tag{2.4}$$

where:

$$D = \sum_{\{\forall(x,y)|(x,y)\in PZ,(x,y)\notin L\}} \left(\sqrt{(x_L - x)^2 + (y_L - y)^2}\right) \tag{2.5}$$

Obtaining the unit Euclidean distance for the pixels guarantees that the bias towards the size of the ROI and the peripheral zone is eliminated. Then for each pixel inside the rest of the peripheral zone, Distance Weighted Specific Intensity ($DWSI$) is calculated as:

$$DWSI(x,y) = (I(x,y) - I_\mu(L)).f(d(x,y)) \tag{2.6}$$

where $I_\mu(L)$ is the average intensity within the ROI L and $f(d(x,y)$ can be a linear or non-linear function of the unit Euclidean distance d. The values computed for distance functions (i) $f(d(x,y)) = d(x,y)$, (ii) $f(d(x,y)) = d(x,y)^2$, and (iii) $f(d(x,y)) = log(d(x,y))$, are represented as grey-levels with respect to a marked ROI, in Fig 2.1. Finally the first feature value for the ROI L is computed as the average of $DWSI$ values for all the pixels inside the rest of the peripheral zone:

$$DWSI_\mu(L) = \frac{1}{N} \sum_{\{\forall(x,y)|(x,y)\in PZ,(x,y)\notin L\}} \Big(DWSI(x,y)\Big) \tag{2.7}$$

where N is the number of pixels within the rest of the peripheral zone. A graphical representation of $DWSI$ when $f(d(x,y)) = d(x,y)^2$ is given in Fig 2.2 and it illustrates the difference for the feature value for a tumour ROI and a normal ROI.

Feature II: Distance Weighted Histogram Match against Identical Lesions within Peripheral Zone

The intensity histogram of an ROI marked on a 2D T2w-MRI slice reflects the T2w signal distribution within the ROI. The intuition behind this feature is the fact that tumours have lower intensity, therefore their intensity histograms are likely to be skewed towards lower intensity levels. To evaluate the intensity distribution of a selected ROI L with respect to the rest of the peripheral zone, the set of all possible similarly shaped and oriented ROIs $\ell = \{\ell_1, \ell_2, ...\ell_n\}$ is obtained by systematically moving the original ROI L inside the peripheral zone on the same axial slice. The corresponding set of unit Euclidean distances ($d =$

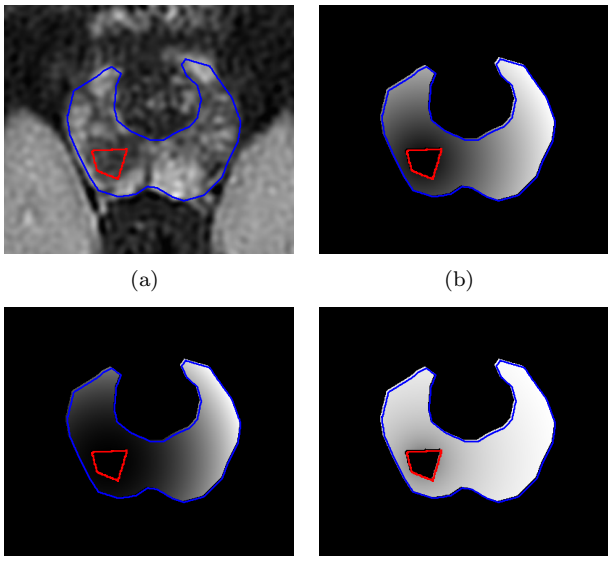


Figure 2.1: Distance function evaluation for each pixel within the rest of the peripheral zone: (a) An ROI is annotated in red and the peripheral zone boundary is annotated in blue on a 2D T2w slice, and the evaluation of the distance function is represented in grey-levels when (b) $f(d(x,y)) = d(x,y)$, (c) $f(d(x,y)) = d(x,y)^2$, and (d) $f(d(x,y)) = log(d(x,y))$.

$\{d(\ell_1), d(\ell_2), ...d(\ell_n)\}$) from the centroid $c_L(x_L, y_L)$ of ROI L to the centroid $c_i(x_i, y_i)$ of each ROI ℓ_i in ℓ is computed as:

$$d(\ell_i) = \frac{1}{\hat{D}}\sqrt{(x_L - x_i)^2 + (y_L - y_i)^2} \tag{2.8}$$

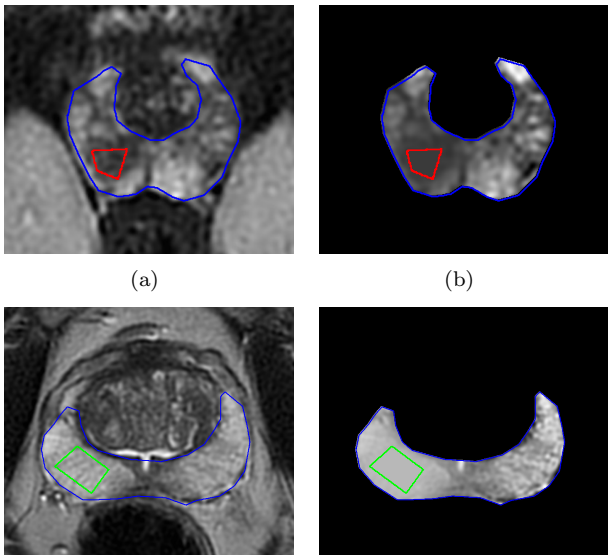
where

$$\hat{D} = \sum_{i=1}^{n} \sqrt{(x_L - x_i)^2 + (y_L - y_i)^2} \tag{2.9}$$

when ℓ_i is the ROI obtained by moving L by $x_L - x_i$ and $y_L - y_i$ signed pixel distances along the x and y axes respectively inside the peripheral zone. Then using the scaled intensity model $I(x,y)$, the grey-level histogram h_L of the original ROI L and the histogram h_i for each ROI ℓ_i in ℓ are obtained. As all the datasets are scaled during initial DICOM data processing to have pixel intensity values between 0 and 255 (8-bit unsigned integer), each histogram h is 8-bit and contains 256 equally distributed bins $[0 - 255]$. Each histogram is normalised such that the sum of frequencies in all bins is 1:

$$\sum_{b=0}^{255} h(b) = 1 \tag{2.10}$$

Then each histogram h_i is subtracted bin-wise from h_L to obtain a signed histogram vector for L relative to each ROI ℓ_i in ℓ:



(a) (b) (c) (d)

Figure 2.2: *DWSI* value for the pixels within the rest of the peripheral zone in grey-levels, where $f(d(x,y)) = d(x,y)^2$: (a) a tumour ROI is annotated in red, (b) scaled grey-levels for DWSI in and around the tumour ROI ($DWSI_\mu(L) = 0.3973$), (c) a normal ROI is annotated in green, and (b) scaled grey-levels for *DWSI* in and around the normal ROI ($DWSI_\mu(L) = -0.2009$).

$$\delta h_i(b) = h_L(b) - h_i(b) \tag{2.11}$$

Based on each histogram difference δh_i a Distance Weighted Histogram Difference feature ($DWHD$) is computed for each l_i as:

$$DWHD(i) = f(d(\ell_i)) \sum_{b=1}^{255} \Big(\delta h_i(b).(255 - b) \Big) \tag{2.12}$$

where $f(d(\ell_i))$ is a linear or non-linear function of $d(\ell_i)$. Finally the second feature value for the ROI L is computed as the average of $DWHD$ values for all ROIs in ℓ:

$$DWHD_\mu(L) = \frac{1}{n} \sum_{i=1}^{n} DWHD(i) \tag{2.13}$$

Intermediate results when computing the feature $DWHD(L)$ for a tumour ROI and a normal ROI are shown in Fig 2.3. Clearly the computed feature value $DWHD(L)$ serves as a logically comparable feature for histograms of a tumour ROI and a normal ROI selected from two different patients.

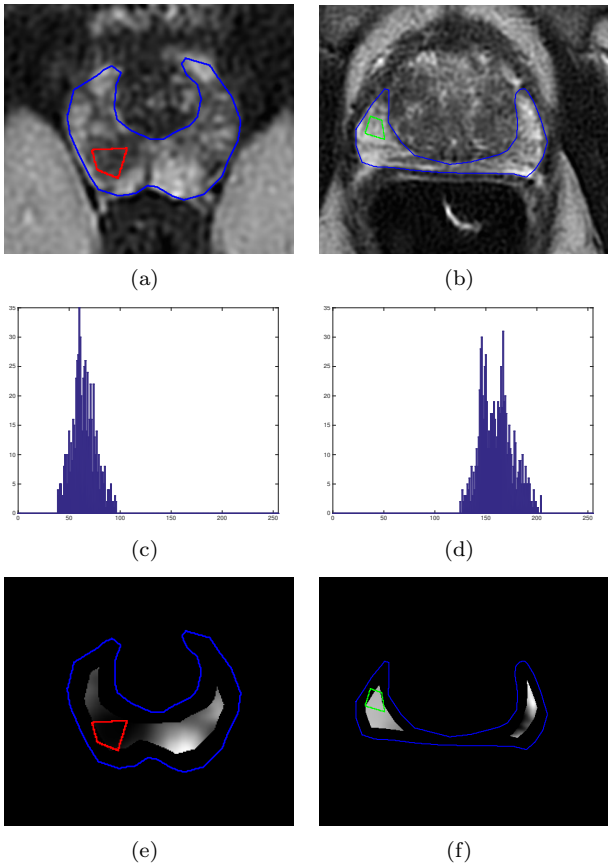


Figure 2.3: $DWHD$ value distribution for centroids of ROIs in ℓ, where $f(d(\ell)) = d(\ell)^2$: (a) tumour ROI is annotated in red, (b) normal ROI is annotated in green, histograms for the tumour ROI and for the normal ROI are presented in (c) and (d) respectively, the calculated $DWHD$ values around the tumour ROI ($DWHD_\mu(L) = 0.3026$) and around the normal ROI ($DWHD_\mu(L) = -0.1184$) are shown in (e) and (f) respectively.

Feature III: Histogram Difference Against the Rest of the Peripheral Zone

The third feature exploits the difference between the intensity distribution of the selected ROI L and the rest of the peripheral zone. The motivation for this feature is the fact that histograms of the tumours are mainly skewed towards the lower end of intensities compared to the rest of the peripheral zone. 8-bit (256 bins) unit histograms H_L for the ROI L, and H_R for the rest of the peripheral zone, are computed exactly as described already in section 2.2. Then the bin-wise signed Difference of Histograms within Peripheral Zone HPZ_δ is computed:

$$HPZ_\delta = \sum_{b=0}^{255} \Big(H_L(b) - H_R(b) \Big) \tag{2.14}$$

To compute the final feature value for L, each bin difference is weighted by the additive inverse of the bin label value:

$$HPZ_\delta(L) = \sum_{b=0}^{255} \Big(HPZ_\delta(b).(255 - b) \Big) \tag{2.15}$$

Step by step formulation of HPZ_δ for a tumour ROI and a normal ROI is illustrated in Fig 2.4.

Feature IV: Intensity Distribution Difference Against the Central Gland

This feature exploits the difference of the intensity distribution of a selected ROI L and the intensity distribution of the Central Gland CG on the same 2D T2w-MRI slice. The central gland, in comparison with the homogeneously higher intensity distribution of the peripheral zone, shows a mixture of low to high intensities [4] in a healthy prostate. Thus before comparing the intensity of ROI L against the central gland, the intensity distribution of the central gland is separated into higher and lower regions. Unsupervised k-means clustering [7] is used to separate the intensity distribution of the central gland into two clusters $J := \{low(CG), high(CG)\}$ with respective cluster means $c := \{low_\mu(CG), high_\mu(CG)\}$.

$$(J, c) = \arg\min_{J,c} \sum_{i=1}^{2} \sum_{\{\forall(x,y)|(x,y)\in CG\}} ||I_{(x,y)} - c_i||^2 \tag{2.16}$$

Then to compute the Intensity Distribution Difference against the Central Gland (ICG_δ) for the ROI L, the mean of the lower intensity cluster of the central gland is subtracted from the mean intensity $I_\mu(L)$ within the ROI L as follows:

$$ICG_\delta(L) = I_\mu(L) - low_\mu(CG) \tag{2.17}$$

2.3 Summary of the Derived Features

The expected behaviour of relative values of each derived feature of the proposed model for tumour ROIs and for normal ROIs is shown in Table 2.1.

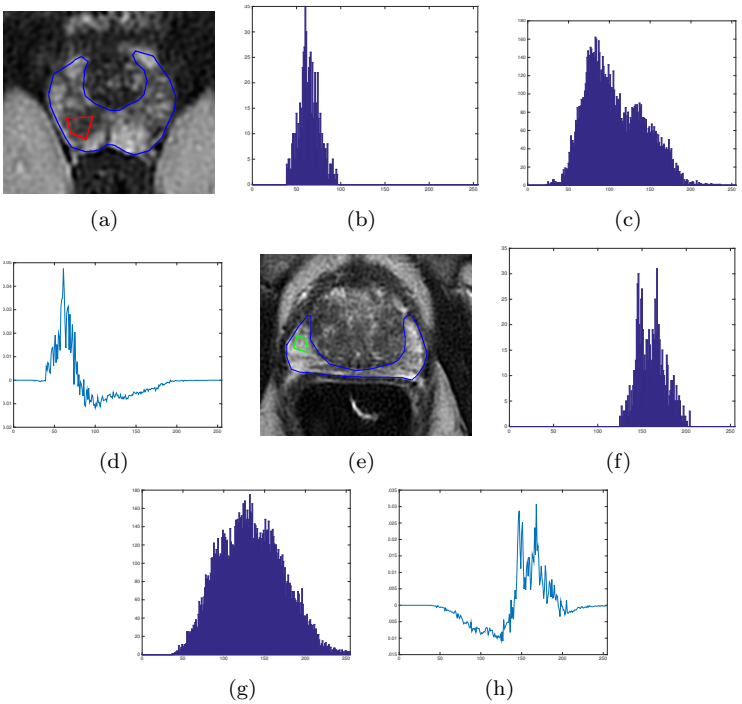


Figure 2.4: Top two rows present formulation of HPZ_δ for a tumour ROI ($HPZ_\delta(L) = 42.619$): (a) ROI belongs to a tumour region, marked in red, (b) histogram for the tumour ROI, (c) histogram for the rest of the peripheral zone, and (d) bin-wise differences of the unit histograms. Bottom two rows show the formulation of HPZ_δ for a normal ROI ($HPZ_\delta(L) = -26.866$) in similar order as in (a), (b), (c) and (d)

Table 2.1: Expected behaviour of relative values of each derived feature

Feature	Expected behaviour in tumour ROIs	Expected behaviour in normal ROIs
$DWSI_\mu(L)$	higher	lower
$DWHD_\mu(L)$	higher	lower
$HPZ_\delta(L)$	higher	lower
$ICG_\delta(L)$	higher	lower

3 Experiments

3.1 Dataset

A T2w-MRI dataset acquired on a General Electric (GE) Healthcare - Discovery MR750w 3.0T MRI scanner, for 56 patients of age ranging from 48 to 73 years, was used to evaluate the proposed feature extraction framework. A senior radiologist affiliated to a tertiary public hospital and private practice supported this research by collecting and anonymising patient data, providing image annotations and verifying the results of experiments. A total of 108 ROIs with 40 positive ROIs and 68 negative ROIs were annotated in T2w-MRI scans of 56 patients. 5-fold cross validation [11] was repeated 5 times to evaluate the experiments. ROIs were split at patient-level in each training / testing iteration. The number of negative and positive ROIs in each patient dataset is different, therefore the composition of negative and positive ROIs within different training / testing iterations within different repetitions vary slightly.

3.2 Experiment I

A first set of experiments was performed to determine the best order of the distance function, in calculation of features $DWSI_{\mu}(L)$ (section 2.2) and $DWHD_{\mu}(L)$ (section 2.2). To evaluate each these features, their feature values were calculated using a range of orders of distance functions, and used in single-feature binary classification. The evaluated distance functions were: $f(d) = \frac{1}{d^2}$, $f(d) = \frac{1}{d}$, $f(d) = 1$, $f(d) = log(d)$, $f(d) = d$, $f(d) = d^2$, $f(d) = d^3$, $f(d) = d^4$, $f(d) = d^5$ and $f(d) = d^6$ for both the features.

3.3 Experiment II

As the main goal of the second set of experiments was to evaluate the proposed feature extraction protocol, generic Naive Bayes Classifier [1] was used as a probabilistic binary classifier. For comparison, two basic intensity features and their combination were used as reference features to train and test the classifier separately:

1. absolute average intensity $I_{0\mu(L)}$ of each marked ROI L:

$$I_{0\mu(L)} = \frac{1}{N_L} \sum_{\{\forall(x,y)|(x,y)\in L\}} I_0(x,y) \tag{3.1}$$

   where N_L is the number of pixels within the ROI L.

2. ratio between average intensity of the ROI L and the whole of the peripheral zone on the corresponding 2D T2w-MRI slice ($I_r(L)$):

$$I_r(L) = \frac{\frac{1}{N_L}\sum_{\{\forall(x,y)|(x,y)\in L\}} I_0(x,y)}{\frac{1}{N_{PZ}}\sum_{\{\forall(x',y')|(x',y')\in PZ\}} I_0(x',y')} \tag{3.2}$$

   where N_L is the number of pixels within the ROI L and N_{PZ} is the number of pixels within the whole peripheral zone.

The reference features are listed in Table 3.1. Each feature set was used to train and test a Naive Bayes Classifier using 5-fold cross validation.

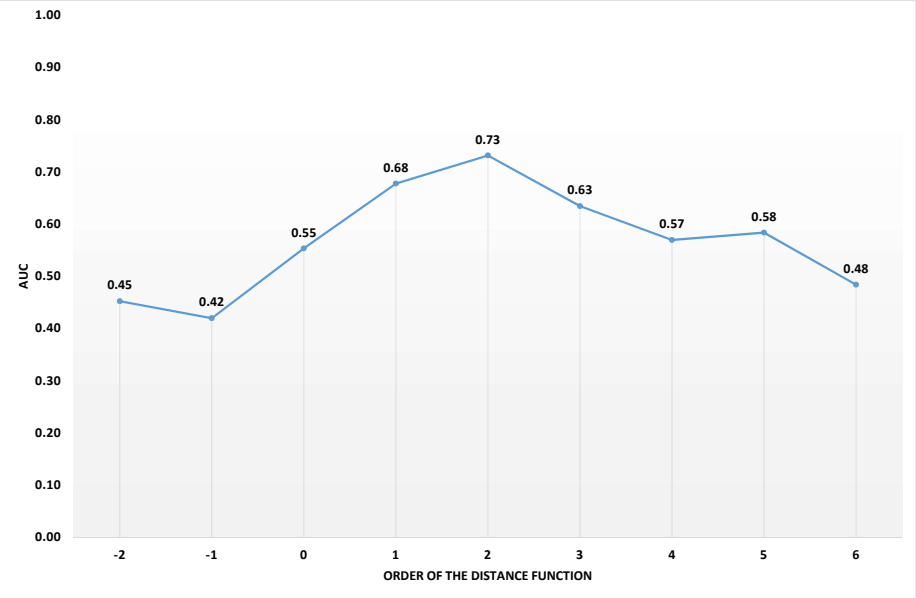


Figure 3.1: AUC against the order of distance function when calculating the feature $DWSI_{\mu}(L)$.

3.4 Results and Evaluation

Area Under the Curve (AUC), when using the computed values of $DWSI_{\mu}(L)$ with different orders of the distance functions in single-feature binary classification, is given in Fig 3.1. Similarly for the feature $DWHD_{\mu}(L)$, AUC against the order of distance functions is presented in Fig 3.1.

By analysing the ROCs, it is clear that the optimal order for the distance function, when calculating both $DWSI_{\mu}(L)$ and $DWHD_{\mu}(L)$, is 2. Performance on classification gradually degrades when increasing / decreasing the order of the distance functions above / below 2. Distance functions of order 2 are named Quadrance in rational geometry [5], and are known to assign progressively greater weights to more distant objects compared to order-1 Euclidean distance. However these results are purely experimental at this stage, with no known clinical reasoning.

Based on the above results, the quadratic distance function was selected ($f(d) = d^2$) for the second set of experiments, where the classification results for the reference feature sets in Table 3.1 were compared with the results of the classification based on the features of the proposed model. ROC curves for the whole proposed feature set and the reference feature sets are shown in Fig 3.3. In Fig 3.4, the sensitivity, specificity, accuracy and AUC of each feature set are summarised, with error-bars indicating the standard deviation of each measure under 5-fold cross validation.

The results show that the proposed region-based heterogeneity-based features perform better than the traditional intensity features in terms of sensitivity, specificity, accuracy and AUC. Sensitivity is generally the most desired

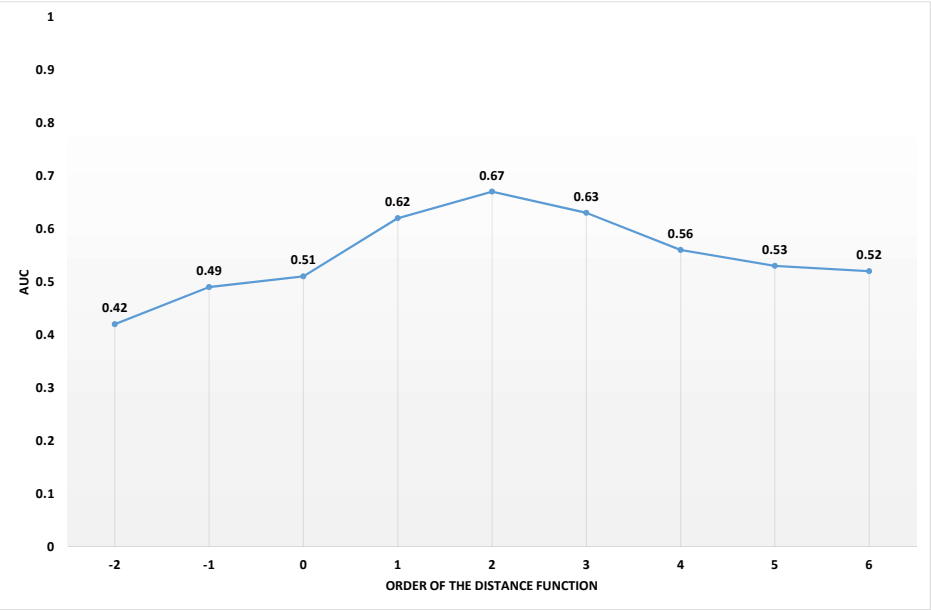


Figure 3.2: AUC against the order of distance function when calculating the feature $DWHD_{\mu}(L)$.

Table 3.1: Summary of the reference feature sets

Feature Set	**Content**
ref-I	$I_{0\mu(L)}$
ref-II	$I_r(L)$
ref-III	$I_{0\mu(L)}$, $I_r(L)$

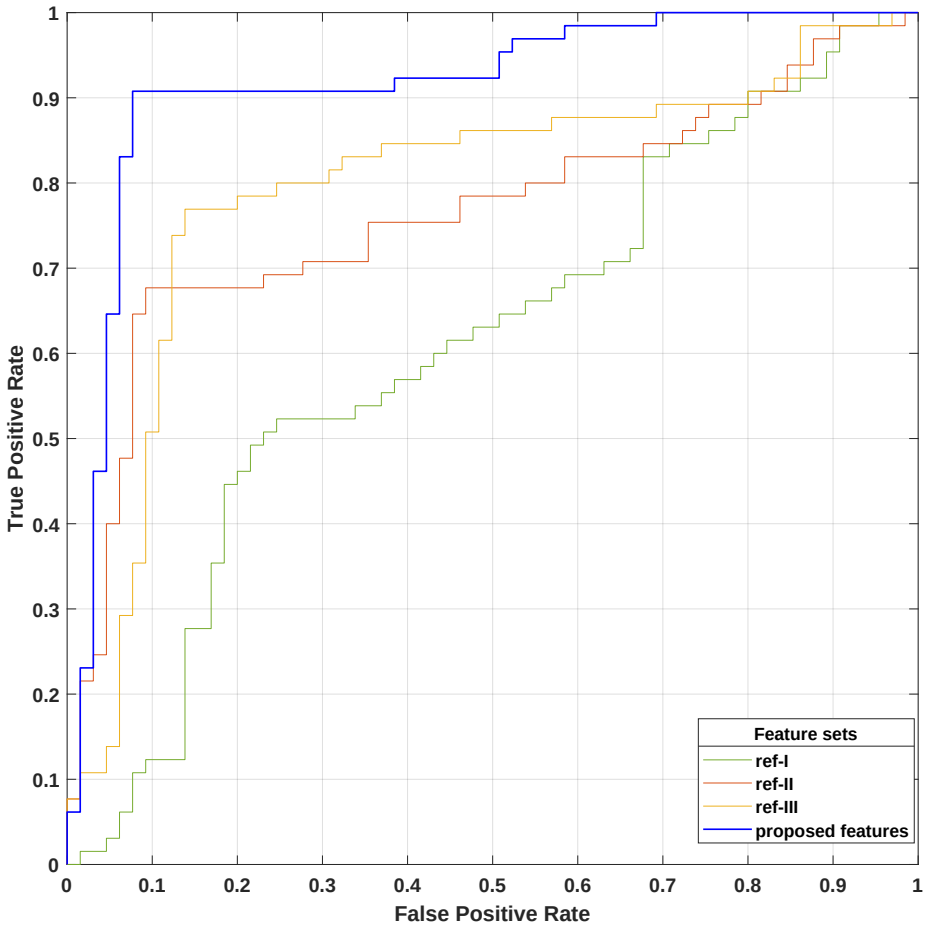


Figure 3.3: ROC curves for classification with the whole proposed feature set vs the reference feature sets.

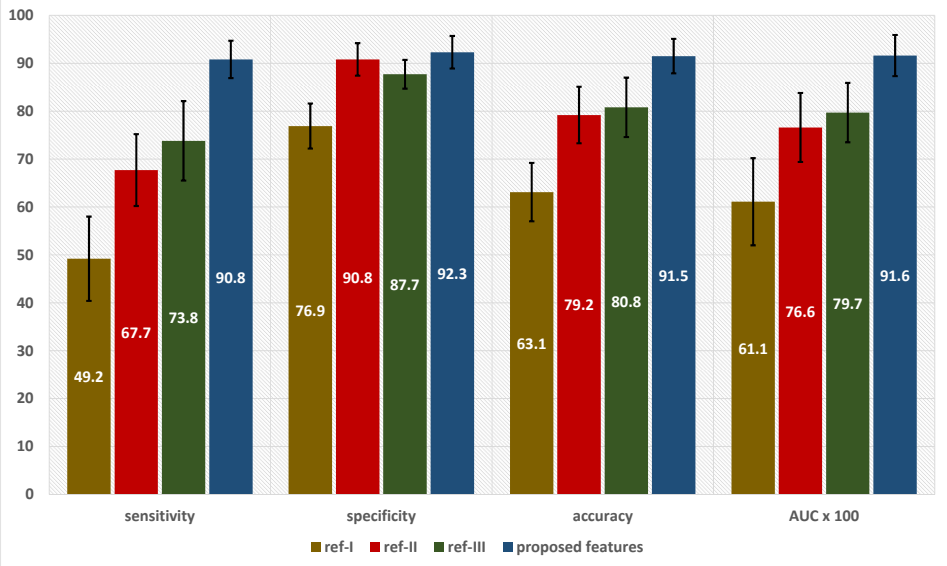


Figure 3.4: Sensitivity, specificity, accuracy and AUC of classifier using different feature sets, where standard deviations for cross-validation iterations are presented as error-bars.

performance measure in CAD frameworks. ROC analysis also shows that the proposed feature extraction framework works well with the Naive Bayes classifier.

For comparison, the proposed feature set was trained on two other classifiers, Support Vector Machine with linear kernel [6], and Random Forest classifier [3]. Summarised results for these classifiers are shown in Fig 3.5, and exhibit similar behaviours to Naive Bayes classifier.

4 Conclusion

A novel feature extraction method for prostate T2w-MRI has been proposed in this paper. Four different novel features are derived for a selected peripheral zone lesion, based on the heterogeneity of the whole peripheral zone on the corresponding 2D MRI slice. When deriving these features, in addition to the relative intensity distribution of the regions within the remaining peripheral zone, a distance function that measures the distance of a specific region from the ROI was taken into account. This guarantees that adjacent regions, where the influence of the selected lesion is higher, are assigned a lower weight in the computation of intensity discrimination.

The developed features were used to train a Naive Bayes classifier. Quantitative results show that the accuracy, sensitivity and specificity depend on the distance function used in feature extraction. The classifier yields the best results of sensitivity 90.8%, specificity 92.3%, accuracy 91.5%, and AUC 91.6 with 5-fold cross validation on the dataset, when a quadratic distance function is used. Two other classifiers trained on the same dataset yield similar results.

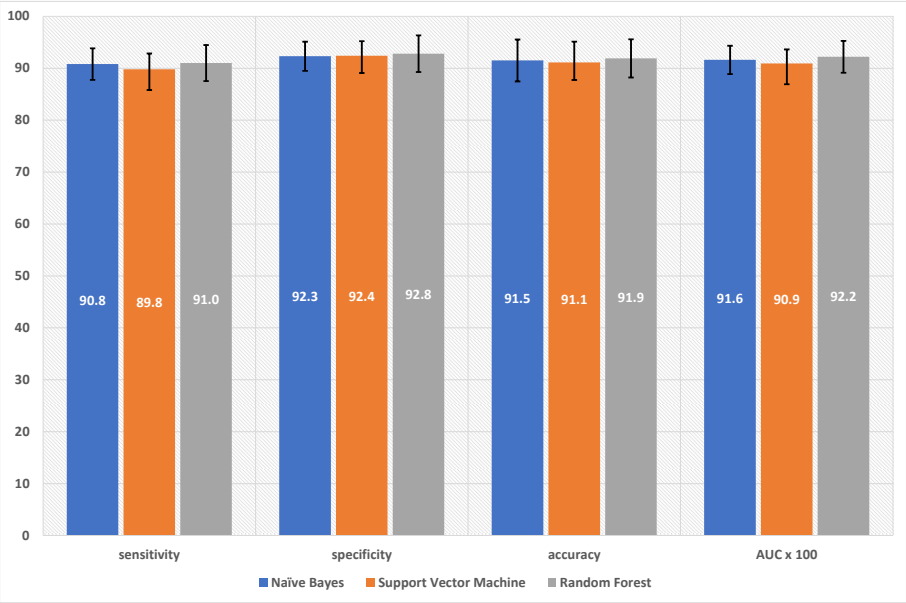


Figure 3.5: Classification results with error-bars for different classifiers trained on the proposed novel features.

Further experiments show that the regional heterogeneity-based features lead to significantly better classification results, compared to traditional image features based on local intensities of ROIs alone.

Bibliography

[1] D.W. Aha, D. Kibler, and M.K. Albert. Instance-based learning algorithms. *Machine learning*, 6(1):37–66, 1991.

[2] Yusuf Artan, Masoom A Haider, Deanne L Langer, and Imam Samil Yetik. Semi-supervised prostate cancer segmentation with multispectral mri. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pages 648–651. IEEE, 2010.

[3] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

[4] D Cheng and C M Tempany. MR imaging of the prostate and bladder. *Seminars in ultrasound, CT, and MR*, 19(1):67–89, February 1998.

[5] Michael Gilsdorf. A comparison of rational and classical trigonometry, 2006.

[6] S.R. Gunn et al. Support vector machines for classification and regression. *ISIS technical report*, 14, 1998.

[7] John A Hartigan and Manchek A Wong. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.

[8] Hedvig Hricak, GC Dooms, John E McNeal, Alexander S Mark, Miljenko Marotti, Antony Avallone, Mark Pelzer, Evelyn C Proctor, and Emil A Tanagho. Mr imaging of the prostate gland: normal anatomy. *American journal of roentgenology*, 148(1):51–58, 1987.

[9] Chan Kyo Kim, Byung Kwan Park, and Bohyun Kim. Localization of prostate cancer using 3t mri: comparison of t2-weighted and dynamic contrast-enhanced imaging. *Journal of computer assisted tomography*, 30(1):7–11, 2006.

[10] Alexander PS Kirkham, Mark Emberton, and Clare Allen. How good is mri at detecting and characterising cancer within the prostate? *European urology*, 50(6):1163–1175, 2006.

[11] Ron Kohavi et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In *Ijcai*, volume 14, pages 1137–1145. Stanford, CA, 1995.

[12] K Yu Kyle and Hedvig Hricak. Imaging prostate cancer. *Radiologic Clinics of North America*, 38(1):59–85, 2000.

[13] G. Lemaître, R. Martí, J. Freixenet, J.C. Vilanova, P.M. Walker, and F. Meriaudeau. Computer-aided detection and diagnosis for prostate cancer based on mono and multi-parametric mri: A review. *Computers in biology and medicine*, 60:8–31, 2015.

[14] Dongjiao Lv, Xuemei Guo, Xiaoying Wang, Jue Zhang, and Jing Fang. Computerized characterization of prostate cancer by fractal analysis in mr images. *Journal of magnetic resonance imaging*, 30(1):161–168, 2009.

[15] A. Madabhushi and J.K. Udupa. New methods of mr image intensity standardization via generalized scale. *Medical Physics*, 33(9):3426–3434, 2006.

[16] John E McNeal. Cancer volume and site of origin of adenocarcinoma in the prostate: relationship to local and distant spread. *Human pathology*, 23(3):258–266, 1992.

[17] John E McNeal, Elise A Redwine, Fuad S Freiha, and Thomas A Stamey. Zonal distribution of prostatic adenocarcinoma: correlation with histologic pattern and direction of spread. *The American journal of surgical pathology*, 12(12):897–906, 1988.

[18] Emilie Niaf, Olivier Rouvière, Florence Mège-Lechevallier, Flavie Bratan, and Carole Lartizien. Computer-aided diagnosis of prostate cancer in the peripheral zone using multiparametric mri. *Physics in medicine and biology*, 57(12):3833, 2012.

[19] László G Nyúl, Jayaram K Udupa, et al. On standardizing the mr image intensity scale. *image*, 1081, 1999.

[20] Sedat Ozer, Deanna L Langer, Xin Liu, Masoom A Haider, Theodorus H van der Kwast, Andrew J Evans, Yongyi Yang, Miles N Wernick, and Imam S Yetik. Supervised and unsupervised methods for prostate cancer segmentation with multispectral mri. *Medical physics*, 37(4):1873–1883, 2010.

[21] Andrik Rampun, Paul Malcolm, and Reyer Zwiggelaar. Computer-aided diagnosis method for mri-guided prostate biopsy within the peripheral zone using grey level histograms. In *Seventh International Conference on Machine Vision (ICMV 2014)*, pages 94451J–94451J. International Society for Optics and Photonics, 2015.

[22] Larissa V Rodriguez and Martha K Terris. Risks and complications of transrectal ultrasound guided prostate needle biopsy: a prospective study and review of the literature. *The Journal of urology*, 160(6):2115–2120, 1998.

[23] Baris Turkbey, Peter A Pinto, and Peter L Choyke. Imaging techniques for prostate cancer: implications for focal therapy. *Nature Reviews Urology*, 6(4):191–203, 2009.

[24] Satish Viswanath, B Nicolas Bloch, Mark Rosen, Jonathan Chappelow, Robert Toth, Neil Rofsky, Robert Lenkinski, Elizabeth Genega, Arjun Kalyanpur, and Anant Madabhushi. Integrating structural and functional imaging for computer assisted detection of prostate cancer on multiprotocol in vivo 3 tesla mri. In *SPIE medical imaging*, pages 72603I–72603I. International Society for Optics and Photonics, 2009.

[25] Satish E Viswanath, Nicholas B Bloch, Jonathan C Chappelow, Robert Toth, Neil M Rofsky, Elizabeth M Genega, Robert E Lenkinski, and Anant Madabhushi. Central gland and peripheral zone prostate tumors have significantly different quantitative imaging signatures on 3 tesla endorectal, in vivo t2-weighted mr imagery. *Journal of Magnetic Resonance Imaging*, 36(1):213–224, 2012.

[26] Marlène Wiart, Laura Curiel, Albert Gelet, Denis Lyonnet, Jean-Yves Chapelon, and Olivier Rouvière. Influence of perfusion on high-intensity focused ultrasound prostate ablation: A first-pass mri study. *Magnetic resonance in medicine*, 58(1):119–127, 2007.