Computer Aided Diagnosis of Prostate Cancer within the Peripheral Zone in T2-Weighted MRI

Rampun, Yambu Andrik; Zheng, Ling; Malcolm, Paul; Zwiggelaar, Reyer

Published in:
Proceedings of the 19th Medical Image Understanding and Analysis Conference

Publication date:
2015

Citation for published version (APA):
Rampun, Y. A., Zheng, L., Malcolm, P., & Zwiggelaar, R. (2015). Computer Aided Diagnosis of Prostate Cancer within the Peripheral Zone in T2-Weighted MRI. In *Proceedings of the 19th Medical Image Understanding and Analysis Conference* (pp. 207-212)

Computer Aided Diagnosis of Prostate Cancer within the Peripheral Zone in T2-Weighted MRI

Andrik Rampun[1]
yar@aber.ac.uk
Ling Zheng[1]
liz5@aber.ac.uk
Paul Malcolm[2]
paul.malcolm@nnuh.nhs.uk
Reyer Zwiggelaar[1]
rrz@aber.ac.uk

[1] Institute of Physics, Mathematics and Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, UK

[2] Department of Radiology, Norfolk Norwich University Hospital, Norwich NR4 7UY, UK

Abstract

Many studies have reported the limitations of computer-aided diagnosis systems using a single T2-W MRI which include weak texture descriptors and an extensive amount of noise. Therefore, researchers have used multiparametric MRI to improve the performances of their methods. We propose a computer-aided diagnosis (CADx) method for prostate cancer within the peripheral zone using a single modality of T2-W MRI and qualitatively compared our results with some of the methods in the literature. The proposed method was tested based on 418 T2-W MR images taken from 45 patients and evaluated using 9-fold cross validation with five patients in each fold. The results demonstrated a comparable performance with CADx systems using multiparametric MRI. We achieved area under the receiver operating curve (A_z) $88\% \pm 9\%$ and $87\% \pm 10\%$ for Random Forest and Naive Bayes classifiers, respectively, while the combined classifier achieved $91\% \pm 7\%$.

1 Introduction

Prostate cancer is one of the major leading causes of death in men. Clinical methods such as transrectal ultrasound (TRUS) biopsy, prostate specific antigen (PSA) blood test and digital rectal examination (DRE) have shown that it is possible to reduce prostate cancer mortality by 20%-30% [5]. However, these methods are associated with several problems such as high risk of overdiagnosis and overtreatment, unable to predict the aggressiveness of cancer and low sensitivity and specificity, as well as being invasive. Currently, magnetic resonance imaging (MRI) is integrated as a second-line modality in most clinical practices as it has the potential to improve sensitivity and specificity. MRI has not been progressed to a first-line modality mainly because it requires substantial expertise from radiologists and reading prostate MRI is time consuming [6].

Automated computer-aided diagnosis (CADx) of prostate cancer could overcome both of these problems. However, developing CAD for prostate using a single modality of T2-W MRI is quite challanging due to weak texture descriptors and an extensive amount of noise [1, 4, 6, 10, 12]. In contrast, multiparametric MRI is costly, time-consuming, and cumbersome for both patient and physician [8]. The aim of this paper is to construct a CADx system for prostate cancer diagnosis within the peripheral zone (PZ) because 80% of prostate cancers arise within this region and prostate cancer that arises within the PZ is more aggressive than that which arises in the central zone (see similar studies in [1, 4, 10, 12]). For this purpose, we extract a set of 217 feature descriptors and perform feature selection before employing six different classifiers namely Naive Bayes (NB), k-nearest neighbours (k-nn), Support Vector Machine (SVM), Linear Logistic Regression (LLR) and Random Forest (RF). An additional combined classifier (CC) using a voting approach with average probability as a combination rule was also considered in this study.

2 Method

Figure 1 shows the pipeline of our method. For every input image, we roughly estimate the area of the prostate's PZ and extract a set of 217 feature descriptors. We normalise (feature scaling) each of the selected features captured in the previous phase. Feature selection was performed to eliminate irrelevant or redundant features and use them to train our classifiers. Finally in the testing phase, for every unseen pixel within the PZ the trained classifiers will decide whether it belongs to malignant or benign class.

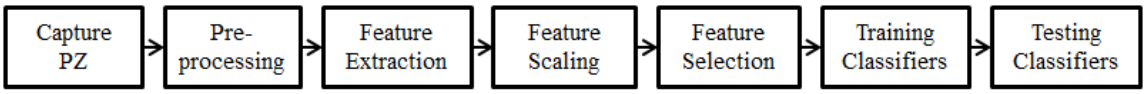

Figure 1: A general overview of the proposed method.

2.1 Capturing the Peripheral Zone

To capture (or segment) the PZ, we employed the mathematical 2D model of Rampun *et al*. [3] which uses a quadratic equation based on the central coordinates of the prostate gland, the left-most and right-most coordinates of the prostate gland boundary. This allows us to model a *priori* general knowledge of radiologists which is similar to the methods of Makni *et al*. [9] and Liu *et al*. [13]. Figure 2 shows example MRI images with the ground truth of prostate gland, central zone (CZ) and tumor (T) represented in red, yellow and green, respectively, while magenta line is the estimated boundary of PZ based on the method in [3]. Note that in this study we did not perform prostate segmentation because all prostates were already delineated by an expert radiologist.

2.2 Pre-processing

MR images typically suffer from two main problems: a) corruption by thermal noise due to receiver coils and b) intensity variations due to different scanning protocols. Following the studies in [1, 14], each image is median filtered to preserve edge boundaries. Subsequently,

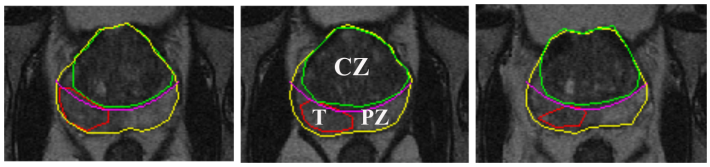

Figure 2: Three different prostate MRI images with ground truth delineated by an expert radiologist and the estimated PZ region indicated under the magenta line.

image intensities were normalised to zero mean unit variance and diffusion anisotropic filtering is applied to remove noise. This three-step pre-processing method has the following advantages a) while suppressing the noise, it simultaneously preserves the edge boundaries b) it standardise image intensities for all patients avoiding dissimilar intensity values for the same tissue types c) it is a robust denoising method without blurring the tumor nodule edges.

2.3 Feature Extraction

In this study, we extracted a set of 217 image features. First and second order statistical features, Tamura texture features and grey-level percentile based features, for each pixel are computed over a local 9×9 window [4].

First order statistical features. Niaf *et al*. [4] used mean, median and standard deviation in their study. On top of that, we extracted mean and median absolute deviation, skewness, kurtosis, the mean of correlation coefficients, local contrast [3], variance and local probability [3] (11 features in total).

Second order statistical features (Grey Level Co-occurrence Matrix (GLCM)). We extracted all features originally suggested by Haralick *et al*. [11] and all features which were further suggested by Soh and Tsatsoulis [18] and Clausi [2]. To maximise the texture information captured from the co-occurrence matrix we considered four orientations ($\theta = 0^\circ, 45^\circ, 90^\circ$ and 90°) with distance d limited to 1. In addition, we calculate the mean, variance and standard deviation for all orientations of each of the features (156 features in total).

Grey-level percentile based features namely percentile 25% and percentile 75%. Vos *et al*. [10] and Niaf *et al*. [4] extracted similar features and found that many cancerous regions have smaller values of percentile 75% (2 features in total).

Tamura texture features [7]. The authors proposed six texture features corresponding to human visual perception: coarseness, contrast, directionality, line-likeness, regularity, and roughness. However, from experiments testing the significance of these features with respect to human perception, it was concluded that the first three features are very important. Therefore we only use the first three features in this study (3 features in total).

Gradient features. There are many operators (e.g. Sobel filter, Kirsch filter, etc.) that could be used to extract these features. In this study we only selected the most discriminate ones according to the results by Niaf *et al*. [4], namely image numerical gradient at 0° and 90° orientations and image magnitude. Secondly, using Sobel operators we extracted image gradient at 0°, 90° and diagonal orientation and image magnitude (7 features in total).

Filter bank features. Most cancers show textural distortions in T2-W MRI. Litjens *et al*. [6] captured these characteristics in features using Gaussian texture bank. However the conventional Gaussian texture bank is a) less sensitive with rotational invariance (hence, rotated versions of cancer textures would be classified as benign unless those rotated versions

were included in the training set) and b) it does not incorperate spots/bars and edges. Therefore, we employed a filter bank proposed by Varma and Zisserman [19] which is rotationally invariant and takes edges and spots/bars into account (38 features in total).

2.4 Feature Scaling and Feature Selection

Since we have 217 texture descriptors, feature selection is necessary to a) reduce over-fitting when building a classifier model and less chance of making decisions based on noise, b) possibly improve accuracy because only the most relevent attributes are selected to build a classifier model and c) reduce time training because fewer features are used in making decisions. Before feature selection is performed, we normalised each selected feature to avoid that absolute values play a role [17]. Following the suggestion in [17], each of the selected features was linearly scaled to the range [-1,+1] ([0,1] is also possible) and the same was applied to the test data.

Subsequently, we employed the CfsSubsetEval [16] attribute evaluator and the GreedyStepwise search method in WEKA [15]. The CfsSubsetEval method measures the value of a subset of features by considering each feature's predictive ability with the degree of redundancy between the other features within the subset, while the GreedyStepwise search method performs a greedy forward or backward search through the feature space.

3 Experiments

Dataset: Our dataset consists of 418 T2-W (1.5T) MR images (512×512 pixels) taken from 45 patients age between 54 to 74 (all patients were biopsy-proven prostate cancer). Each patient has between 6 to 13 slices from top to the bottom slice of the prostate gland. The prostate gland, cancer and central zone were delineated by an expert radiologist on each of the MR images (as shown in Figure 2). All pixels within the radiologist's tumor annotation were extracted as prostate cancer samples (e.g. within the red line region in Figure 2). This area was truncated by the tumor mask, to ensure no pixels outside the tumor region were included into the malignant samples. On the other hand, every pixel outside the tumor region and within the PZ (under the magenta line in Figure 2) were considered as benign samples. Similarly, this region is truncated by the tumor and prostate gland masks to ensure no pixels within the tumor region or outside the prostate gland were included into benign samples. In total 171, 518 samples (instances) were extracted (97, 310 and 74, 208 benign and malignant samples, respectively) with 217 image features (attributes).

Experimental setup: Six classifiers were used in this study: NB, k-nn, SVM (SMO procedure with polynomial kernel), LLR, RF and CC. The CC classifier is a combination of three classifiers which have the 3 best AUC values (based on the results produced in the training phase) from any of the first five classifiers. The voting approach with average probability as a combination rule was employed in WEKA [15]. A stratified nine runs 9-fold cross-validation (9-FCV) scheme was employed. A leave one patient out approach was employed to ensure no samples from the same patient were used in the training and testing phases. On the other hand, we chose 9 folds instead of 10 folds to ensure each fold has the same number of patients (45 patients in our case, hence each fold contains 5 patients). Each classifier was trained and in the testing phase, each unseen instance/sample from the testing data (taken from 5 patients) was classified as benign or malignant. All parameters for all classifiers were left on the default setting in WEKA [15].

Results: The performance evaluation metrics used for this work are classification accuracy (*CA*), which represents the percentage of pixels classified correctly and the area under the ROC curve (AUC), denoted A_z which represents the true positive rate (TPR) against the false positive rate (FPR). Table 1 shows the results of six different classifiers employed in this study where the CC classifier achieved the best A_z=91% (*CA*=84%). In our experiment, selecting the best 2 AUC values produced the same A_z=91%. The RF classifier achieved the best *CA*=87% (A_z=88%), while the NB and LLR classifiers achieved good $A_z$values of 87% and 86%, respectively. Using the CC classifier shows a significant improvement (*p=0.025*) at the 95% confidence level in comparison to the best individual classifier (RF) in Table 1. In comparison with exisitng methods in the literature, Niaf *et al.* [4] reported an A_z=89% using 140 texture descriptors and the method proposed by Vos *et al.* [10] achieved A_z=83% both based on 30 patients. In a smaller number of dataset (15 patients) Yetik and Artan [1] reported *CA*=83% while the method of Viswanath *et al.* [12] (15 MRI images) produced *CA*=76% and A_z=77%. However, we would like to emphasise that direct comparisons are impossible in this study due to a) differences in datasets (e.g. different modalities and data size), b) absence of public datasets and c) evaluation was performed at different levels (e.g. region of interest, voxel, etc).

Metric	NB	*k*-nn	SVM	LLR	RF	CC
$CA(\%)$	71 $\pm$19	70 $\pm$12	78 $\pm$9	78$\pm$9	87 $\pm$5	84 $\pm$8
$A_z(\%)$	87 $\pm$10	68 $\pm$9	76 $\pm$8	86 $\pm$7	88 $\pm$9	91 $\pm$7

Table 1: Performances comparisons of six different classifiers employed in this study.

The top 10 most selected features (80% chance of being selected in every single run) based on the number of selections in 9-FCV (81 runs) are: a bank (edge) of filters [19], a bank (bar/spot) filters [19], a bank (Laplacian of Gaussian) of filters [19], image magnitude of a Sobel filter, image gradient of a Sobel filter (diagonal orientation), numerical image magnitude, numerical image gradient (90°), Tamura texture feature (contrast), image gradient of a Sobel filter (90°), local probability feature [3].

4 Discussions and Conclusions

A study of Litjens *et al.* [6] reported the overall per-voxel performance in a single modality of T2-W MRI as A_z=0.76 (which is similar to our results using a SVM classifier). The most recent study in the literature [6] achieved $A_z = 0.89$ which covered the whole prostate gland using multiparametric MRI. Previous studies [1, 4, 6, 7, 12] showed the feasibility of developing CADx and its prospect in detecting prostate cancers using multiparametric MRI. Although the study conducted in this paper is only based on a single modality of T2-W MRI we obtained promising results similar to those in [1, 4, 6, 10, 12]. On the other hand, Niaf *et al.* [4] reported in their study (preliminary results) that a radiologist has an average A_z=0.83 which is similar to most of the results reported in the literature. This means, a CADx system could be a valuable tool to assist radiologists as a second reader. In conclusion, we have presented a novel CADx method for prostate cancer detection within the PZ using single modality T2-W MRI and performance evaluation shows that it achieved similar results with the state-of-the-art (included the ones based on multiparametric MRI), although the comparison has limitations mainly due to the different evaluation dataset.

References

[1] Y. Artan and I. S. Yetik. Prostate cancer localization using multiparametric mri based on semi-supervised techniques with automated seed initialization. *IEEE Transactions on Information Technology in Biomedicine*, 16(6):2986–2994, 2012.

[2] D. A. Clausi. An analysis of co-occurrence texture statistics as a function of grey level quantization. *Canadian Journal of Remote Sensing*, 28(1):45–62, 2002.

[3] A. Rampun et al. Detection and localisation of prostate cancer within the peripheral zone using scoring algorithm. In *Proc. 16th Irish Machine Vision and Image Processing*, pages 75–80, 2014.

[4] E. Niaf et al. Computer-aided diagnosis of prostate cancer in the peripheral zone using multiparametric mri. *Physics in Medicine and Biology*, 57:3833–3851, 2012.

[5] F. H. Schroder et al. Screening and prostate-cancer mortality in a randomized european study. *New England Journal of Medicine*, 360(13):1320–1328, 2009. doi: 10.1056/NEJMoa0810084. URL http://dx.doi.org/10.1056/NEJMoa0810084. PMID: 19297566.

[6] G. Litjens et al. Computer-aided detection of prostate cancer in mri. *IEEE Transactions on Medical Imaging*, 33(5):1083–1092, 2014.

[7] H. Tamura et al. Textural features corresponding to visual perception. *IEEE Transaction on Systems, Man, and Cybernetics*, 8(6):460–472, 1978.

[8] J. T. Rothwax et al. Multiparametric mri in biopsy guidance for prostate cancer: Fusion-guided. *BioMed Research International*, 2014:1–7, 2014.

[9] N. Makni et al. Computer-aided detection of prostate cancer in mri. *Medical Physics*, 38:6093–6105, 2011.

[10] P. C. Vos et al. Computerized analysis of prostate lesions in the peripheral zone using dynamic contrast enhanced mri. *Medical Physics*, 35(3):888–899, 2008.

[11] R. M. Haralick et al. Textural features of image classification. *IEEE Transactions on Systems, Man and Cybernetics*, (6):610–621, 1973.

[12] S. Viswanath et al. Enhanced multi-protocol analysis via intelligent supervised embedding (empravise): Detecting prostate cancer on multi-parametric mri. In *Proc. SPIE 7963, Medical Imaging 2011: Computer-Aided Diagnosis*, 2011.

[13] X. Liu et al. Automated prostate cancer localization with mri without the need of manually extracted peripheral zone. *Medical Physics*, 38(6):2986–2994, 2011.

[14] Y. Artan et al. Semi-supervised prostate cancer segmentation with multiparametric mri. In *Proc.Int.Symp.Biomed.Imag*, pages 648–651, 2010.

[15] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: an update. *ACM SIGKDD explorations newsletter*, 11(1):10–18, 2009.

[16] M. A. Hall. Correlation-based feature selection for discrete and numeric class machine learning. In *Proc. 17th Proc. Int. Conf. Machine Learning*, pages 359–366, 2000.

[17] C. Hsu, C. Chang, and C. Lin. A practical guide to support vector classification, 2010.

[18] L. Soh and C. Tsatsoulis. Texture analysis of sar sea ice imagery using gray level co-occurrence matrices. *IEEE Transactions on Geoscience and Remote Sensing*, 37(2):780–795, 1999.

[19] M. Varma and A. Zisserman. A statistical approach to texture classification from single images. *International Journal of Computer Vision*, 62:61–81, 2005.