



Published in final edited form as:

Comput Med Imaging Graph. 2016 September ; 52: 44–57. doi:10.1016/j.compmedimag.2016.02.001.

In vivo MRI based prostate cancer localization with random forests and auto-context model

Chunjun Qian^{a,b}, Li Wang^b, Yaozong Gao^b, Ambereen Yousuf^c, Xiaoping Yang^a, Aytekin Oto^c, and Dinggang Shen^{b,d,*}

^aSchool of Science, Nanjing University of Science and Technology, Jiangsu, China

^bDepartment of Radiology and BRIC, University of North Carolina at Chapel Hill, NC, United States

^cDepartment of Radiology, Section of Urology, University of Chicago, Chicago, IL, United States

^dDepartment of Brain and Cognitive Engineering, Korea University, Seoul, Republic of Korea

Abstract

Prostate cancer is one of the major causes of cancer death for men. Magnetic resonance (MR) imaging is being increasingly used as an important modality to localize prostate cancer. Therefore, localizing prostate cancer in MRI with automated detection methods has become an active area of research. Many methods have been proposed for this task. However, most of previous methods focused on identifying cancer only in the peripheral zone (PZ), or classifying suspicious cancer ROIs into benign tissue and cancer tissue. Few works have been done on developing a fully automatic method for cancer localization in the entire prostate region, including central gland (CG) and transition zone (TZ). In this paper, we propose a novel learning-based multi-source integration framework to directly localize prostate cancer regions from in vivo MRI. We employ random forests to effectively integrate features from multi-source images together for cancer localization. Here, multi-source images include initially the multi-parametric MRIs (i.e., T2, DWI, and dADC) and later also the iteratively-estimated and refined tissue probability map of prostate cancer. Experimental results on 26 real patient data show that our method can accurately localize cancerous sections. The higher section-based evaluation (SBE), combined with the ROC analysis result of individual patients, shows that the proposed method is promising for in vivo MRI based prostate cancer localization, which can be used for guiding prostate biopsy, targeting the tumor in focal therapy planning, triage and follow-up of patients with active surveillance, as well as the decision making in treatment selection. The common ROC analysis with the AUC value of 0.832 and also the ROI-based ROC analysis with the AUC value of 0.883 both illustrate the effectiveness of our proposed method.

Keywords

Prostate cancer; MRI segmentation; Random forests; Auto-context model; Cancer localization

*Corresponding author at: Department of Radiology and BRIC, University of North Carolina at Chapel Hill, 130 Mason Farm Road, Chapel Hill, NC 27599-7513, United States. Tel.: +1 919 966 3535; fax: +1 919 843 2641. dgshen@med.unc.edu (D. Shen).

1. Introduction

Prostate cancer is the most commonly diagnosed non-skin cancer and the second leading cause of cancer death among U.S. men [1]. Current clinical practice for the diagnosis of prostate cancer is often based on a transrectal ultrasound (TRUS) biopsy, after the patient shows an elevated serum prostate specific antigen (PSA) level. A large screening trial using PSA and TRUS has shown that it is possible to reduce prostate cancer mortality by 20–30% [2]. However, these studies have also shown that PSA testing in combination with TRUS biopsies has a relatively low specificity. Additionally, cancer is often under-graded in TRUS biopsies [3]. These problems lead to over-diagnosis and overtreatment of patients for prostate cancer [4]. Alternatively, multi-parametric high-contrast magnetic resonance (MR) imaging provides a powerful and noninvasive imaging tool for detecting suspicious cancerous tissues [5], as shown in Fig. 1, which includes T2, diffusion-weighted (DW) imaging (DWI) and apparent diffusion coefficient (ADC) MR images. Since ADC MR image is calculated from DWI and related to the b -value, in our experiments, only one type of ADC MR image – dADC (the regular ADC map) – is used.

Nowadays, MRI is often used as a second line modality after repeating negative TRUS biopsies. One of the reasons why MRI has not yet progressed to the first line modality for prostate cancer diagnosis is mainly the need of high-level expertise of radiologist to read prostate MRI. Such expertise is not widely available [6,7]. Besides, the MRI-based diagnosis suffers from large observer variability due to the difficulty of the task [7]. Additionally, due to the large number of 3D images, reading prostate MRI data is quite time-consuming and labor-intensive.

Automated computer-aided detection and diagnosis (CAD) of prostate cancer could help reduce the above problems and open the door to prostate cancer screening using multi-parametric MRIs. In the recent decade, computer-aided detection and diagnosis of prostate cancer is becoming an active field of research and many methods have been proposed [6,8–16]. However, there are still some limitations for the previous works:

1. Prostate consists of several main regions, namely peripheral zone (PZ), central gland (CG), and transition zone (TZ). In MR images, cancer in the PZ has richer textural information than that in the CG [17], which makes the computer-aided detection relatively easy in the PZ. Therefore, early researches of MRI-based prostate cancer detection mainly focused on identification of cancer only in the PZ, instead of the entire prostate region. For example, Chan et al. [10] were the first to implement the prostate cancer detection method by using multi-parametric MRIs. In their paper, they used the SVM classifier to identify prostate cancer in the PZ by combining information from three different MRI techniques: T2-weighted (T2W), T2-mapping, and line scan diffusion imaging (LSDI). Niaf et al. [8,14] and Vos et al. [9] proposed machine-learning methods for suspicious region of interest (ROI) classification in the PZ. In addition, many techniques have also been developed in the imaging field to facilitate image-guided prostate cancer detection in the PZ. For example, Langer et al. [15] developed a multi-parametric model suitable for prospectively identifying prostate cancer in the PZ

by using MRI. However, many studies have demonstrated that prostate cancer is not limited to the PZ, but also appears in the CG [11,12,16] and TZ [18,19]. But, identifying TZ tumor is a difficult task [20] since the TZ is the site of origin of benign prostatic hyperplasia (BPH), which can have a heterogeneous appearance [21,22].

2. In the literature, prostate cancer is usually detected via a two-stage CAD system. In the first stage, the suspicious cancer ROIs are manually delineated by radiologist. In the second stage, those suspicious ROIs are classified into benign and cancer tissues by using textual information within those ROIs. Most of these two-stage CAD systems are semi-automatic, as the suspicious cancer ROIs often have to be delineated manually by radiologist, which is labor-intensive, time-consuming, and subject to manual errors. For example, Niaf et al. [8,14] proposed a CAD system, aimed to separate cancer ROIs from benign ROIs by using supervised learning methods. In their methods, suspicious cancer ROIs were manually delineated by two persons, i.e., one histopathologist and one researcher, by using histological sections as gold standard. Compared with these two-stage methods, our proposed method directly localizes cancer regions voxel-by-voxel, similar to the detection systems proposed by Tiwari et al. [23] and Sparks and Madabhushi [24], thus is able to detect much smaller cancer regions. This aspect is important to guide prostate biopsy, as well as precisely target the tumor in focal therapy planning. Besides, our method is fully automatic and also computationally efficient.
3. There are still a few works that classify the whole prostate region into benign or malignant tissues, especially in the early diagnosis of prostate cancer. Firjani et al. [13] used DWI and a KNN-based classifier to classify the prostate into benign or malignant regions based on three appearance features extracted from the registered ADC images with different b -values. With their methods, prostate with malignant or benign cancer could be discriminated, but the identified cancer regions were too coarse to guide prostate biopsy. In other words, such method only focused on determining whether cancer exists in the prostate, which cannot be adopted for guiding prostate biopsy that needs more accurate cancer localization.

Litjens et al. [25] proposed an automated zone-specific CAD pipeline for prostate cancer detection. They mainly used an atlas-based method to segment the prostate into PZ and TZ. Then, they detected cancer from each zone. Recently, they [6] again proposed a novel and fully automatic CAD system to detect prostate cancer. Similar to the method in [25], in the first stage of their method, for detecting the candidate regions, they used a multi-atlas-based prostate segmentation, three different voxel classifiers (by integrating various features such as intensity, pharmacokinetic, texture, blobness and anatomical features from multi-parametric MRIs), as well as local maxima detection. Then, in the second stage, the statistics (of voxel-wise features), local contrast, symmetry, and shape features were extracted as candidate representations and further input to a classifier for both cancer ROI identification and segmentation. Finally, with lesion-based FROC and patient-based ROC analysis, their system showed a reasonable performance in prostate cancer detection.

Inspired by these existing methods, T2, DWI, and ADC map are used in our work for prostate cancer localization. As demonstrated in the literature, by combining any two of three multi-parametric MRIs, prostate cancer regions can be more effectively discriminated from non-cancerous regions. For instance, Miao et al. [26] found that the performance of DW imaging in prostate cancer detection was significantly better than that of T2W imaging. Shimofusa et al. [27] compared the prostate cancer detection ability of T2W imaging without *or* with DW imaging. Morgan et al. [28] compared the performance of prostate cancer detection by using T2W alone, and also jointly using T2W and DWI or ADC map. In their works, by analyzing the statistical results, they concluded that, compared with the cases of using T2W without *or* with ADC map, using T2W with DWI had the superiority in cancer detection. All these methods were based on medical statistical analysis, and their results reported in the papers demonstrated that multi-parametric MRIs provided better information about cancerous and normal regions in the prostate, when compared to the methods using a single MRI modality. This was similarly verified by Ozer et al. [29].

To deal with the aforementioned limitations of previous works, in this paper, we propose a novel learning-based multi-source integration framework, similar to the method proposed by Montillo [30], for accurate localization of prostate cancer within the entire prostate. Specifically, in our method, we first employ random forests [31] to train a classifier based on the training subjects with multi-parametric MRIs (T2, DWI, and dADC, as shown in Fig. 1). The trained classifier provides an initial cancer probability map for each training subject. Then, inspired by auto-context model [32,33], the estimated tissue probability map is further used as additional source of information to train the next classifier, which integrates the high-level context features (from the estimated cancer probability map) with the appearance features (from multi-parametric MRIs) for refining cancer classification. By iteratively training the classifiers based on the updated cancer probability map (and the multi-parametric MRIs), we can finally obtain a sequence of classifiers. Similarly, in the application stage, given a target subject, the learned classifiers are sequentially applied to iteratively refine the estimation of cancer probability map by combining the multi-parametric MRI information with the previously-estimated cancer probability map. We have validated our proposed method on 26 patient subjects, with promising performance.

The rest of the paper is organized as follows. A detailed description of source images and our proposed method are introduced in Section 2. Experimental design, results and analysis are then presented in Section 3, followed by the Discussions and Conclusion sections.

2. Materials and method

2.1. Data and image preprocessing

In our work, all patients underwent prostatectomy and preoperative multi-parametric endorectal MR imaging, including all sequences of T2W and DW MR imaging in the axial direction with Philips MR scanners (Achieva; Philips Healthcare, Eindhoven, the Netherlands, a 3T imager) between March 2008 and March 2010, and have not been previously treated before (such as with radiation or hormone). ADC values were calculated from DW MR images by using a linear least squares [34] fitting to the logarithmic form of a mono-exponential DW signal model, as follows:

$$\ln \left(\frac{S_0}{S_b} \right) = b \cdot \text{ADC} \quad (1)$$

where S_0 and S_b are the DW MR signal intensities with diffusion weighting of 0 and b , respectively [35–37]. Here, we use one kind of ADC image – dADC image – as a source image for cancer localization as shown in Fig. 1. Multi-parametric MRI data (T2, DWI and dADC) from 26 patients are included in this study. Among 26 subjects, there are totally 42 tumors, including 29 PZ tumors, 7 CG tumors and 6 TZ tumors.

All MR images were acquired with an endorectal coil (Medrad; Bayer Healthcare, Warrendale, PA) and a phased-array surface coil with the aforementioned Philips MR scanners. Immediately before MR imaging, 1 mg of glucagon (Glucagon; Lilly, Indianapolis, Ind) was injected intramuscularly to decrease peristalsis of the rectum. The data acquisition parameters [38] are shown in Table 1.

The reference standard of prostate cancer foci in the MR images was established through a systematic consensus-seeking correlative review of histological findings and MR images by a genitourinary pathologist (with 8 years of experience in genitourinary pathology) and a radiologist (with 9 years of experience in prostate MR imaging). The pathologist sliced each prostate and identified all distinct tumor foci larger than approximately 5 mm in diameter, and then the radiologist manually outlined the corresponding ROIs of the tumor foci on the MR images according to the tumor regions on the histological slices by the pathologist. Totally, 1–3 cancer delineations on MR images of each subject were cooperatively done manually by radiologist and the pathologist. Note that these cancer ROIs were mostly outlined on T2 images, and some were outlined on DW images.

We resampled all images into the size of 512×512 with the resolution of $0.3125 \text{ mm} \times 0.3125 \text{ mm} \times 3 \text{ mm}$. For the alignment, since the dADC image was generated from DWI, we need only to align DWI and dADC images to T2 image for each subject. Through our experiments, we found that the results of registration affected the final voxel level classification, e.g., better registration leads to better cancer region localization. We also found that the linear registration method can meet requirements in our work due to no large deformation occurring among multi-parametric MR images of each subject. Here, we use the FMRI's linear image registration tool (FLIRT) [39] to implement this procedure. However, since the contrast of the prostate and other tissues in DWI and dADC is not high enough while the boundary of bladder is obvious, we first align T2 image to DWI along with the bladder masks, and then align DWI to T2 image with the prostate masks. Histogram matching [40] was also performed on each type of MRI across different subjects.

2.2. Prostate cancer localization with random forests and auto-context model

In our method, we formulate the localization of prostate cancer as a two-class classification problem. To solve such a classification problem, we proposed a novel learning-based multi-source integration framework by employing random forests [31] and auto-context model [32,33]. For simplicity, let N_{sub} be the total number of the training subjects and let

$\mathbf{I} = \{I_{T2}^i, I_{DWI}^i, I_{dADC}^i, i=1, \dots, N_{sub}\}$ be a set of multi-parametric MRIs. As a supervised learning method, our method consists of both *training* and *testing* stages. The flowchart of training stage is shown in Fig. 2. In the training stage, we will train a sequence of classification forests, each with the input of multi-source images/maps. In the first iteration, the classification forest takes only the multi-parametric MRIs \mathbf{I} as input, and uses the image appearance features extracted from different multi-parametric MRIs for voxel-wise classification. By applying the trained forest in the first iteration, each i th training subject will produce tissue probability maps for prostate cancer I_{PC}^i and non-prostate cancer I_{nPC}^i , respectively, as shown in the second column of Fig. 2. In the later iterations, as inspired by the auto-context model [32,33], the tissue probability maps $\bar{\mathbf{I}} = \{I_{PC}^i, I_{nPC}^i, i=1, \dots, N_{sub}\}$ obtained from the previous iteration will act as additional source information for training. Specifically, the high-level context features are extracted from the tissue probability maps to assist the classification, along with appearance features from the multi-parametric MRIs. Since context features are informative about the nearby tissue structures for each voxel, they encode the spatial constraints into the classification, thus improving the quality of the estimated tissue probability maps, as also demonstrated in Fig. 2. Then, the tissue probability maps are iteratively updated and fed into the next training iteration. Finally, a sequence of learned classifiers will be obtained.

Similarly, in the testing stage, given a target subject with multi-parametric MRIs, we can obtain the initial tissue probability map by applying the trained classifier in the first iteration using only the multi-parametric MRIs. In the later iterations, along with multi-parametric MRIs, the tissue probability maps resulted from the previous iteration are also fed into the next classifier for refinement. Fig. 3(a)–(c) shows examples by applying a sequence of the learned classifiers for cancer localization in the PZ, CG, and TZ, respectively. As we can see from Fig. 3, the cancer probability maps are updated with iterations and become more and more accurate. In the following, we will detail on how to train the classifiers via random forests.

2.3. Random forests

Random forest [31] is an ensemble learning method for classification or regression. As a machine learning method, random forest has been successfully applied in many fields, such as bioinformatics, computer vision, and medical image analysis [41,42].

In our work, we employ random forest to implement cancer localization for a given testing voxel x , based on its high-dimensional feature representation $f(x, \mathbf{I})$, where \mathbf{I} is a set of multi-parametric MR images as we define in the above. When the random forest is implemented, at each iteration, totally T decision trees, indexed by $t \in [1, T]$, are used as classification trees. In the training stage, in the first iteration, each tree t learns a weak class predictor $p_t(c|f(x, \mathbf{I}))$ using features extracted from a set of multi-parametric MR images \mathbf{I} , where c is the label of cancer or non-cancer. When the training stage starts, according to the feature representation, the training samples are assigned to the left and right child nodes at each parent node through recursive splitting. At each node, to obtain the optimal split for maximizing the information gain [43], a number of random splits on different combinations

of features and thresholds are considered. The tree keeps growing until it reaches the constraint conditions. For example, it reaches a specified depth (d), or the criterion that no tree leaf node should contain less than a certain number of training samples (s_{\min}). Finally, we associate each leaf node (l) with the empirical distribution $p_t^l(c|f(x, I))$ over the two classes by the proportions of the labels of all training samples which reach that leaf node.

In the testing stage, by applying the learned split functions, each testing voxel x is independently tested. When the testing sample arrives at a leaf node l_x in each trained tree t , the empirical distribution of that leaf node is used to determine the class probability of the testing sample x , i.e., $p_t(c|f(x, I)) = p_t^{l_x}(c|f(x, I))$. Finally, we can obtain the class probability for the testing sample x by computing the mean of the class probabilities from all

individual trees, i.e., $p(c|x) = \frac{1}{T} \sum_{t=1}^T p_t(c|f(x, I))$. In our work, what we should solve is a two-class classification problem. Therefore, in the trained random forest, for each i th training subject at the first iteration, only cancer and non-cancer probability maps ($\bar{I}_i = \{I_{PC}^i, I_{nPC}^i, i=1, \dots, N\}$) are generated, as shown in the second column of Fig. 2.

2.4. Appearance features and context features

Random forest can deal with a large number of features efficiently. In other words, for random forest, in our study, any kind of features from multi-parametric MRI and tissue probability maps can be extracted and used for classification, such as the histogram of oriented gradients (HOG) features [44], local binary pattern (LBP) features [45], Gabor features [46], and any kind of gradient features [47]. In fact, in our implementation, considering the efficiency, we use the random Haar-like features [48] for both appearance features and context features. Note that the Haar-like features have been used widely in many studies of pattern recognition or machine learning [49–51]. Specifically, for each voxel x , its Haar-like features are computed as the local mean intensity of any randomly displaced cubical region R_1 or the mean intensity difference over any two randomly displaced, asymmetric cubical regions (R_1 and R_2) within the image patch R [52]:

$$f(x, I) = \frac{1}{|R_1|} \sum_{u \in R_1} I(u) - b \frac{1}{|R_2|} \sum_{v \in R_2} I(v), \quad b \in \{0, 1\} \quad (2)$$

where R is the patch centered at voxel x , I is any kind of source image, and the parameter $b \in \{0, 1\}$ indicates whether one or two cubical regions are used (as shown in Fig. 4, for $b=0$ and $b=1$). In the image patch R , its intensities are normalized to the unit ℓ_2 norm [53,54]. In theory, for each voxel we can determine an infinite number of such features.

3. Experimental results and analysis

In this section, the proposed method will be extensively evaluated on 26 subjects using leave-one-out cross-validation. Results of the proposed method are compared with the manual ground-truth segmentations, as well as other conventional methods.

3.1. Evaluation methods

In the following, we mainly employ the section-based evaluation (SBE) [55], sensitivity, specificity, and Dice ratio, to measure the performances of different methods.

3.1.1. Section-based evaluation (SBE)—Section-based evaluation (SBE) is defined as a ratio of *the number* of the sections that automatic method localizes correctly to *the number* of the total sections that the prostate is divided into:

$$\text{SBE} = \frac{\sum_{s=1}^n L\{\text{label}(A_s), \text{label}(B_s)\}}{n} \quad (3)$$

$$L\{X, Y\} = \begin{cases} 1, & \text{if } X=Y \\ 0, & \text{if } X \neq Y \end{cases} \quad (4)$$

where $\text{label}(\cdot)$ in Eq. (3) represents that the section is cancer or non-cancer; A_s and B_s are the localization results of the section s through automatic method and expert, respectively; and n is the number of the total sections that the prostate is divided into.

The SBE is an effective method widely used for evaluating the performance of the prostate cancer section localization. Generally, in the clinical setting, when a patient is diagnosed as a suspicious prostate cancer sufferer with a high PSA level or an abnormal screening DRE, a biopsy is frequently suggested [56]. Often, when doing a biopsy, the prostate is divided into sections based on the number of biopsy points to use, followed by shooting a needle in each section to extract the respective tissue sample. A popular way is to divide prostate into 6 sections (for 6 point biopsy) [57]: right base, left base, right mid, left mid, right apex, and left apex. Alternatively, some doctors will choose 12 sections by further dividing each of these 6 sections into medial and lateral parts. This method is very popular in prostate cancer diagnosis in clinic, and it is also the standard of the prostate biopsy, as shown in the schematic diagram at the top of Fig. 5. However, 6- or 12-sections partition does not consider the shape of prostate. We take a typical slice of the mid part of the prostate as an example in Fig. 6. The 1st row and 2nd row show the division results on a MRI slice with 6-section division and 12-section division, respectively. It can be seen that the 6- or 12-section division will generate too large section for a biopsy needle, which may result in more shots to different points in each section for accurate extraction of cancer tissue. To reduce the number of the locations for biopsy in each section while still keeping the accuracy of cancerous section localization, we divide the prostate into sections by considering the size of each prostate part (base, mid-gland and apex). In general, as shown in the bottom of Fig. 5, the base and mid parts of the prostate are divided into 9 sections, while the apex is divided into 4 sections. For a typical prostate slice mentioned above, the division result is shown in the 3rd row of Fig. 6. It can be observed that, by dividing the prostate into these fine sections, we can more accurately localize cancer sections, compared with the conventional 6- and 12-section partition strategies.

3.1.2. Sensitivity and specificity—We employ two important statistical measures: sensitivity and specificity. Here, besides the standard voxel-wise definitions of sensitivity and specificity, we also reformulate the sensitivity and specificity based on the divided prostate sections:

$$\text{sensitivity} = \frac{TCS}{TCS + FNS} \quad (5)$$

$$\text{specificity} = \frac{TNS}{TNS + FCS} \quad (6)$$

where TCS is the number of true localized cancerous sections, FNS is the number of cancerous sections localized as normal sections, TNS is the number of true localized normal sections, and FCS is the number of normal sections localized as cancerous sections.

3.1.3. Dice ratio—Dice ratio [58] measures the overlap between our localized cancer region and the ground-truth cancer region. The Dice ratio is formulated as:

$$\text{Dice ratio} = \frac{2|I_{pre} \cap I_{grd}|}{|I_{pre}| + |I_{grd}|} \quad (7)$$

where I_{pre} is our predicted cancer region and I_{grd} is the ground-truth cancer region.

3.2. Parameters setting

In the random forest training, the number of classification trees is 20; the maximum tree depth is 50; the number of randomized Haar-like features is 10,000 for both image appearance and context features; and the minimum sample number for each leaf node is 8. These parameters are determined via cross-validation. Specifically, we first study the impact of the number of trees, T , on cancer section localization, as shown in Fig. 7(a). We set the maximal tree depth as 50 and the minimal number of leaf nodes as 8 according to previous work [59]. In Fig. 7(a), the SBE increases from 0.66 ($T=2$) to 0.87 ($T=15$), the standard deviation decreases from 0.030 ($T=2$) to 0.013 ($T=15$), and the cancer localization performance tends to be stable when $T=20$. Therefore, to produce reasonable and accurate results, we set $T=20$ in our work. Fig. 7(b) shows the impact of the maximum tree depth. We find that the SBE gradually increases as the depth varying from 5 to 20 and then becomes stable as the depth reaches over 40. Fig. 7(c) shows the impact of the minimum sample number for each leaf node. The performance becomes stable when the minimum sample number is less than 8, while the performance starts decreasing when the minimum sample number is larger than 10. The possible reason is that the samples with cancer label and the samples with non-cancer label fall into the same leaf node because of setting too large value for the minimum sample number in each leaf node, which results in fuzzy classification. As for the number of iterations, we find that the satisfactory performance can

be achieved in 3 or 4 iterations, and, after that, the performance becomes stable. In this paper, we conservatively set the iteration number as 10, to obtain both effectiveness and accuracy for our proposed method.

3.3. Feature types

In our work, we choose the random Haar-like features for the appearance features and context features. To demonstrate the advantage of random Haar-like features, we make comparisons with other four types of features: HOG features [44], LBP features [45], Gabor features [46] and gradient features [47]. As shown in Table 2, we can find that the sensitivity with random Haar-like feature is a little bit lower than that with LBP feature, and also the specificity with random Haar-like feature is lower than that with HOG feature. However, the overall performance with random Haar-like feature is much better than the respective performances with other four types of features.

3.4. Importance of context features

In the proposed method, the context features play an important role. They can improve the accuracy of localization iteratively. Fig. 8 and Table 3 show the SBEs of our sections division on 26 patient subjects by sequentially applying the learned classifiers based on the multi-parametric MR images. It can be seen that the SBEs are improved with the iterations and also become stable after a few iterations. Specially, in the second iteration, the SBEs are improved greatly due to the integration of the previously estimated tissue probability maps for guiding the classification. These results demonstrate the importance of using context features for cancer localization.

The sensitivity and specificity of our cancer section localization with iterations are also shown in Table 3. It can be seen that there is a decrease, followed by an increase of both sensitivity and specificity with the increasing number of iterations. The main reason is that, in the first iteration, when no cancer probability maps are available, most of the localized cancer regions are larger than the ground truth, i.e., by the example shown in the second column of Fig. 2 and the example shown in Fig. 3(b); thus the sensitivity is higher and specificity is lower. With the increase of iterations, the localized cancer regions shrink and the final localized cancer regions become smaller than the ground truth, resulting in the decrease of sensitivity while increase of specificity. We can also find that the specificity is improved from 81.3% to 84.1% in the second iteration.

From Table 3, it can be seen that the section-based sensitivity and specificity, respectively, achieve the relative high values, i.e., 91.5% and 85.1%, with a few iterations. This means that almost no any cancer section is localized as non-cancer section and also most of the non-cancer sections are correctly localized. To verify the accuracy of cancer region localization with our method for potentially guiding biopsy, two typical prostate slices are taken as examples. For the first example shown in Fig. 9(a), our method can accurately localize cancer regions. For another example shown in Fig. 9(b), it can be seen that our method can *not only* localize the true cancer region *but also* provide another suspicious cancer region; both of these localized cancer regions are suspicious, which need to be tested by biopsy needle. However, compared with the traditional biopsy, in the second example, the

number of sections to place the biopsy needle is the same, but the accuracy of localizing cancer sections is largely improved because of adoption of fine section division. Therefore, the proposed method with our section division method is very useful to guide prostate biopsy, which can reach a considerable performance gain by reducing the number of suspicious cancer sections as well as guaranteeing the accuracy of cancer section localization.

3.5. Performance of combining multi-parametric MRI

In the past decade, many studies have shown the improvement of cancer diagnosis or detection accuracy with combination of different types of MRI. For example, Wang et al. [60] and Kim et al. [61] validated that the single modality MRI, T2 or ADC map can be used for cancer detection. But the information from the single modality MRI is not rich enough to better differentiate cancer region from normal tissue. To illustrate the superior performance of cancer detection using multi-parametric MRI, we compared the performances using different combinations of image modalities as shown in Fig. 10. The SBE by combining all T2, DWI and dADC is 87.1%, which is better than the results obtained by any of other six combination cases (T2: 76.8%; DWI: 80.4%; dADC: 83.2%; T2 + dADC: 86.8%; T2 + DWI: 80.6%; DWI + dADC: 86.5%). By comparing the results only with single modality MRI, we can see that the performance of using dADC in prostate cancer localization is better than that of using either T2 or DWI, demonstrating that ADC map can provide more useful information for prostate cancer localization. Since ADC map is computed from DWI, these results also provide the evidence as concluded in [27–29]: DWI has outperformance over other existing MRI in prostate cancer localization and detection.

3.6. Comparisons with conventional methods

Due to the unavailability of codes and dataset, we cannot perform a fair comparison with other state-of-the-art prostate cancer localization methods. Therefore, in this section, we compared our method with other classification-based methods on the same dataset. More specifically, we compare the proposed method with four popular classification methods: (1) Naïve Bayes classifier [62], (2) Support Vector Machine (SVM) [63], (3) SVM with radial basis function (RBF) kernel (SVM + RBF) [64], and (4) AdaBoost [65], as detected below. In these methods, to achieve the best performance, we extract *not only* the Haar-like features, *but also* the HOG features [44], LBP features [45], Gabor features [46] and gradient features [47].

- Naïve Bayes classifier technique is based on the so-called Bayesian theorem, and is particularly well-suited when the dimensionality of the inputs is high. Despite its simplicity, Naïve Bayes classifier can often outperform more sophisticated classification methods. In the training stage, using a set of training samples, the method estimates the parameters of a probability distribution (i.e., Gaussian distribution), by assuming that features are conditionally independent given the class. In other words, all model parameters (i.e., class priors and feature probability distributions) can be approximated with the relative frequencies from the training samples. In the testing stage, for any unseen test sample, this method

computes the posterior probability of that sample belonging to each class, and then classifies the test sample according to the largest posterior probability.

- Support Vector Machine (SVM) is a supervised learning model associated with learning algorithms that analyze data and recognize patterns, which can be used for classification and regression. It is often used for solving two-class classification problem. Given a set of training samples, each belonging to one of the two categories, an SVM training algorithm builds a model, i.e., generating two optimal hyperplanes to separate the input samples into two categories and also preventing the samples from falling into the maximum margin (i.e., the gap between the hyperplanes). For the parameter setting, we choose the linear function and RBF as two candidate kernel functions and set both cost functions as 1, and the parameter gamma in RBF kernel is set as $1/k$, where k is the number of training samples. Then, the learning model is generated by training the input training samples. Before training and testing, for each sample, 500 discriminant features among all features are selected using feature selection method (such as Fisher Score [66]). In the application stage, the generated model is used for cancer prediction in the testing image.
- AdaBoost has been a mature and popular machine learning method for classification for decades. When training the input samples, it generates a series of weak learners by updating the weight on each training sample. At each iteration, a weak learner is selected and assigned with a coefficient such that the sum of training error of the boosting classifiers is minimized. These weak learners finally converge to a strong learner. We set the maximum number of iterations as 200 in the weak learner refining process to build a strong classifier. Finally, the generated classifier is used in the testing stage and then the tissue probability map is obtained.

Fig. 11 shows the comparison results using the four methods on the multi-parametric MR images of the same patient as shown in Fig. 6. The results of Naïve Bayes, SVM, SVM + RBF, AdaBoost and our proposed method are shown in the 1st, 2nd, 3rd, 4th and 5th rows, respectively, and further indicated by the orange, brown, cyan, green and blue contours. The 1st, 2nd and 3rd columns represent the localization results with T2, DWI and dADC images, respectively. It can be observed that the result of proposed method is much better than other four methods, and is very similar to the ground truth. In the localization maps of other three methods except SVM + RBF, the localized cancer regions are noisy, which could result in more non-cancer regions localized as the suspicious cancer regions for guiding needle placement. In contrast, from the 5th row, we can see that, by using the proposed method, the localized cancer region is clear. On the other hand, for the result with SVM + RBF, although the localized cancer region is also clear, it is much smaller than the ground truth, and is also not as good as the result obtained by our proposed method.

3.6.1. SBE, sensitivity and specificity—For quantitative evaluation, the SBE, sensitivity and specificity on 26 subjects are shown in Table 4, which further illustrates the outperformance of the proposed method over other four methods. Specifically, the average SBE (87.1%), sensitivity (91.5%) and specificity (85.1%) of the proposed method are much

higher than those by any of other four methods: (1) SBE 68.5%, sensitivity 73.7% and specificity 65.9% by Naïve Bayes; (2) SBE 71.4%, sensitivity 72.6% and specificity 71.0% by SVM; (3) SBE 79.9%, sensitivity 66.1% and specificity 83.8% by SVM + RBF; (4) SBE 68.0%, sensitivity 94.6%, and specificity 60.5% by AdaBoost. It can be seen that the SBE and specificity by our proposed method are significantly higher than those by Naïve Bayes, SVM, SVM + RBF and AdaBoost method. The sensitivity by the proposed method is slightly smaller than that by AdaBoost, but still much better than that by Naïve Bayes, SVM and SVM + RBF, respectively. All these results again demonstrate the performance of our proposed method than other four conventional methods.

3.6.2. Common ROC and ROI-based ROC analysis—To better validate the proposed method, we further employ the receiver operating characteristic (ROC) analysis method [67]. We use two kinds of ROC analysis methods. One is the common ROC analysis method [67,68] and the other is the ROI-based ROC analysis method [69].

The ROC curve is used to analyze the performance of a classifier system in many research fields, such as medicine, radiology, biometrics, machine learning and data mining [67]. The common ROC curve is built by analyzing the results from each testing subject. It illustrates the relationship between the true positive rate and the false positive rate at various threshold settings. On the other hand, the ROI-based ROC analysis was proposed by Obuchowski et al. [69]. They used it to detect and locate breast cancers. They partitioned the images into several ROIs. Every reader assigned a confidence score to each ROI. Finally, they used the ROI-based ROC analysis to access the reader's ability of detecting and locating the abnormalities. In our work, we also use the ROI-based ROC analysis to evaluate the performance of our method. We divide the images into sections as shown in the bottom of Fig. 5. We choose the maximum value from the probabilities of all voxels in a section as the confidence score of each section, which results in localization of all suspicious cancer sections as well as fully convinced non-cancer ones, thus ensuring both accuracy and precision of cancer-section localization. We also calculate the area under the ROC curve (AUC) which can evaluate a classifier system quantitatively.

Fig. 12(a) shows common ROC curves, and Fig. 12(b) shows ROI-based ROC curves. The blue curve illustrates the ROC of our proposed method. As we can see from Fig. 12(a), along with the AUC value (0.832), our proposed method produces slightly better than that obtained by Naïve Bayes method, but much better than other three methods. These results indicate that our method can localize cancer voxels very well. Furthermore, as we can see from Fig. 12(b), along with the area under the ROI-based ROC curve (0.883), our proposed method has obvious superiority over other four methods.

In addition, to evaluate the effectiveness of our method on individual patients, we apply two different ROC analysis methods to the localization results with five different methods on 26 patients. Fig. 13(a) shows the AUC values on 26 patients by the proposed method, Naïve Bayes, SVM, SVM + RBF, and AdaBoost methods. Most of the AUC values by our method are over 0.7, with the highest as 1.0, which indicates that our method works well in cancer voxel localization for the same individual patient. From Fig. 13(b), we can observe that most of the AUC values under the ROI-based ROC curves by our proposed method are over 0.8,

which are better than those achieved by Naïve Bayes, SVM, SVM + RBF and AdaBoost methods on 26 patients. These results further validate the effectiveness of our proposed method in cancer section localization.

3.6.3. Mean Dice ratio—To evaluate the performance of our method for the cancer localization in different zones, we calculate the mean Dice ratios in PZ, CG and TZ, respectively. The average Dice ratios are shown in Table 5. It can be observed that our method achieves higher average Dice ratios 62.0% in PZ, 56.7% in CG and 78.0% in TZ than other four methods.

4. Discussions

In this paper, we have proposed a fully automatic method for prostate cancer localization. In previous works, most methods [8,14] involve a semi-automatic identification process, i.e., pre-outlining the suspicious cancer ROIs by experts and then classifying the ROIs into cancer or non-cancer. Another limitation of most existing methods is that they focused on cancer identification only in the PZ, such as [9,10,15]. In MR images, cancers in the PZ have richer and discriminative textural imaging signatures compared with those in the CG [17] and TZ [18,19], which makes cancer detection in the PZ easier than that in the CG and TZ. In contrast, in our study, we localize the cancer regions in the whole prostate. Other existing methods [17,26–28,60,61,70,71] mainly focused on classification of the manually delineated ROIs as cancer or normal tissue, which is subject to the accuracy of manual experts and also labor-intensive and time-consuming.

Another difference in our study is that we currently aim to guide prostate biopsy through the cancer localization method. Using our localized cancer regions, the clear shooting path to the cancer region could be designed and also the accuracy of biopsy could be improved significantly. Additionally, our method can also be used in other applications, such as focal therapy planning, triage and follow-up of patients with active surveillance, and decision making in treatment selection.

Compared with the works of other researchers, such as [6], there are two limitations in our study: (1) the number of subjects is limited. In [6], totally 165 patients with cancer were included. In our study, only 26 patients were validated. To overcome this limitation, in our experiments, we detected over 100 thousand samples from totally 26 subjects and used voxel level classification method. (2) For each prostate image in our study, cancer ROIs are outlined only on 1–3 discontinuous slices. For some special cancer regions, such as with low Gleason scores, the number of such regions is small and the extracted appearance information is relatively little for localizing or discriminating them from the surrounding normal tissues. This actually causes some cancer ROIs not accurately localized by our method.

In fact, our method can also be extended to other cancer/tumor localization or identification with some adjustment of parameters. For example, we have applied our method to high-grade brain tumor localization. We use 20 subjects from the BRATS 2013 [72] off-site dataset, where each subject has four modality images, including Flair, T2, T1c and T1 MR

images. All of these subjects are annotated by 4 doctors. We set the number of tree as 10; the maximum depth of tree as 100; the number of randomized Haar-like features as 5000; and the patch size as $31 \times 31 \times 31$. We use the Dice ratio to evaluate our method, and obtain the mean Dice ratio of 0.768. The best result is 0.802. It can be seen that our method also achieves a reasonable result on brain tumor localization, which can be further improved in the future.

5. Conclusion

We have proposed a novel framework to localize prostate cancer regions in the in vivo MR images. Our proposed method can directly localize cancer regions from the entire images. Specially, we employ random forests and auto-context model to effectively integrate features from multi-parametric MRI and also the iteratively-estimated probability maps for cancer localization. Experimental results on 26 real patient data show that our method can accurately localize cancer regions, which will be useful for potentially guiding the needle biopsy, targeting the tumor in focal therapy planning, triage and follow-up of patients with active surveillance, and decision making in treatment selection. In our future work, we will further validate our proposed method on more patient subjects.

References

1. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. *CA Cancer J Clin.* 2012; 62:10–29. [PubMed: 22237781]
2. Schröder FH, Hugosson J, Roobol MJ, Tammela TL, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med.* 2009; 360:1320–8. [PubMed: 19297566]
3. Hambrock T, Hoeks C, Hulsbergen-van de Kaa C, Scheenen T, Fütterer J, Bouwense S, et al. Prospective assessment of prostate cancer aggressiveness using 3-T diffusion-weighted magnetic resonance imaging-guided biopsies versus a systematic 10-core transrectal ultrasound prostate biopsy cohort. *Eur Urol.* 2012; 61:177–84. [PubMed: 21924545]
4. Klotz L. Prostate cancer overdiagnosis and overtreatment. *Curr Opin Endocrinol Diabetes Obes.* 2013; 20:204–9. [PubMed: 23609043]
5. Hoeks CM, Barentsz JO, Hambrock T, Yakar D, Somford DM, Heijmink SW, et al. Prostate cancer: multiparametric MR imaging for detection, localization, and staging. *Radiology.* 2011; 261:46–66. [PubMed: 21931141]
6. Litjens G, Debats O, Barentsz J, Karssemeijer N, Huisman H. Computer-aided detection of prostate cancer in MRI. *IEEE Trans Med Imaging.* 2014; 33:1083–92. [PubMed: 24770913]
7. Lim HK, Kim JK, Kim KA, Cho K-S. Prostate cancer: apparent diffusion coefficient map with T2-weighted images for detection—a multireader study 1. *Radiology.* 2009; 250:145–51. [PubMed: 19017927]
8. Niaf E, Rouvière O, Mège-Lechevallier F, Bratan F, Lartizien C. Computer-aided diagnosis of prostate cancer in the peripheral zone using multiparametric MRI. *Phys Med Biol.* 2012; 57:3833. [PubMed: 22640958]
9. Vos PC, Hambrock T, Barentsz JO, Huisman HJ. Computer-assisted analysis of peripheral zone prostate lesions using T2-weighted and dynamic contrast enhanced T1-weighted MRI. *Phys Med Biol.* 2010; 55:1719. [PubMed: 20197602]
10. Chan I, Wells W III, Mulkern RV, Haker S, Zhang J, Zou KH, et al. Detection of prostate cancer by integration of line-scan diffusion, T2-mapping and T2-weighted magnetic resonance imaging; a multichannel statistical classifier. *Med Phys.* 2003; 30:2390–8. [PubMed: 14528961]

11. Liu X, Zhou L, Peng W, Wang C, Wang H. Differentiation of central gland prostate cancer from benign prostatic hyperplasia using monoexponential and biexponential diffusion-weighted imaging. *Magn Reson Imaging*. 2013; 31:1318–24. [PubMed: 23791546]
12. Park SY, Kim CK, Park BK, Ha SY, Kwon GY, Kim B. Diffusion-tensor MRI at 3 T: differentiation of central gland prostate cancer from benign prostatic hyperplasia. *Am J Roentgenol*. 2014; 202:254–62.
13. Firjani A, Elnakib A, Khalifa F, Gimel'farb G, El-Ghar MA, Elmaghraby A, et al. A diffusion-weighted imaging based diagnostic system for early detection of prostate cancer. *J Biomed Sci Eng*. 2013; 6:346.
14. Niaf E, Flamary R, Rouviere O, Lartizien C, Canu S. Kernel-based learning from both qualitative and quantitative labels: application to prostate cancer diagnosis based on multiparametric MR imaging. *IEEE Trans Image Process*. 2014; 23:979–91. [PubMed: 24464613]
15. Langer DL, van der Kwast TH, Evans AJ, Trachtenberg J, Wilson BC, Haider MA. Prostate cancer detection with multi-parametric MRI: logistic regression analysis of quantitative T2, diffusion-weighted imaging, and dynamic contrast-enhanced MRI. *J Magn Reson Imaging*. 2009; 30:327–34. [PubMed: 19629981]
16. Oto A, Kayhan A, Jiang Y, Tretiakova M, Yang C, Antic T, et al. Prostate cancer: differentiation of central gland cancer from benign prostatic hyperplasia by using diffusion-weighted and dynamic contrast-enhanced MR imaging 1. *Radiology*. 2010; 257:715–23. [PubMed: 20843992]
17. Viswanath SE, Bloch NB, Chappelow JC, Toth R, Rofsky NM, Genega EM, et al. Central gland and peripheral zone prostate tumors have significantly different quantitative imaging signatures on 3 Tesla endorectal, in vivo T2-weighted MR imagery. *J Magn Reson Imaging*. 2012; 36:213–24. [PubMed: 22337003]
18. Akin O, Sala E, Moskowitz CS, Kuroiwa K, Ishill NM, Pucar D, et al. Transition zone prostate cancers: features, detection, localization, and staging at endorectal MR imaging 1. *Radiology*. 2006; 239:784–92. [PubMed: 16569788]
19. Chesnais A, Niaf E, Bratan F, Mège-Lechevallier F, Roche S, Rabilloud M, et al. Differentiation of transitional zone prostate cancer from benign hyperplasia nodules: evaluation of discriminant criteria at multiparametric MRI. *Clin Radiol*. 2013; 68:e323–30. [PubMed: 23528164]
20. Zakian KL, Eberhardt S, Hricak H, Shukla-Dave A, Kleinman S, Muruganandham M, et al. Transition zone prostate cancer: metabolic characteristics at 1H MR spectroscopic imaging—initial results 1. *Radiology*. 2003; 229:241–7. [PubMed: 12920178]
21. Noguchi M, Stamey TA, Mc NEAL JE, Yemoto CE. An analysis of 148 consecutive transition zone cancers: clinical and histological characteristics. *J Urol*. 2000; 163:1751–5. [PubMed: 10799175]
22. Ikonen S, Kivisaari L, Tervahartiala P, Vehmas T, Taari K, Rannikko S. Prostatic MR imaging. *Acta Radiol*. 2001; 42:348–54. [PubMed: 11442457]
23. Tiwari P, Kurhanewicz J, Madabhushi A. Multi-kernel graph embedding for detection, Gleason grading of prostate cancer via MRI/MRS. *Med Image Anal*. 2013; 17:219–35. [PubMed: 23294985]
24. Sparks R, Madabhushi A. Statistical shape model for manifold regularization: Gleason grading of prostate histology. *Comput Vis Image Underst*. 2013; 117:1138–46. [PubMed: 23888106]
25. Litjens GJS, Barentsz JO, Karssemeijer N, Huisman HJ. Automated computer-aided detection of prostate cancer in MR images: from a whole-organ to a zone-based approach. *Proc SPIE*. 2012; 8315:83150G–183150G.
26. Miao H, Fukatsu H, Ishigaki T. Prostate cancer detection with 3-T MRI: comparison of diffusion-weighted and T2-weighted imaging. *Eur J Radiol*. 2007; 61:297–302. [PubMed: 17085002]
27. Shimofusa R, Fujimoto H, Akamata H, Motoori K, Yamamoto S, Ueda T, et al. Diffusion-weighted imaging of prostate cancer. *J Comput Assist Tomogr*. 2005; 29:149–53. [PubMed: 15772529]
28. Morgan V, Kyriazi S, Ashley S, DeSouza N. Evaluation of the potential of diffusion-weighted imaging in prostate cancer detection. *Acta Radiol*. 2007; 48:695–703. [PubMed: 17611881]
29. Ozer S, Langer DL, Liu X, Haider MA, van der Kwast TH, Evans AJ, et al. Supervised and unsupervised methods for prostate cancer segmentation with multispectral MRI. *Med Phys*. 2010; 37:1873–83. [PubMed: 20443509]

30. Montillo A, Shotton J, Winn J, Iglesias JE, Metaxas D, Criminisi A. Entangled decision forests and their application for semantic segmentation of CT images. *Inf Process Med Imaging*. 2011;184–96. [PubMed: 21761656]
31. Breiman L. Random forests. *Mach Learn*. 2001; 45:5–32.
32. Tu Z, Bai X. Auto-context and its application to high-level vision tasks and 3D brain image segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2010; 32:1744–57. [PubMed: 20724753]
33. Loog M, Ginneken B. Segmentation of the posterior ribs in chest radiographs using iterated contextual pixel classification. *IEEE Trans Med Imaging*. 2006; 25:602–11. [PubMed: 16689264]
34. Rice, JA. Mathematical statistics and data analysis. 2. Belmont, CA: Duxbury Press; 1995. p. 507-28.
35. Verma S, Rajesh A, Morales H, Lemen L, Bills G, Delworth M, et al. Assessment of aggressiveness of prostate cancer: correlation of apparent diffusion coefficient with histologic grade after radical prostatectomy. *Am J Roentgenol*. 2011; 196:374–81. [PubMed: 21257890]
36. Chenevert TL, Galbán CJ, Ivancevic MK, Rohrer SE, Londy FJ, Kwee TC, et al. Diffusion coefficient measurement using a temperature-controlled fluid for quality control in multicenter studies. *J Magn Reson Imaging*. 2011; 34:983–7. [PubMed: 21928310]
37. Langer DL, van der Kwast TH, Evans AJ, Sun L, Yaffe MJ, Trachtenberg J, et al. Intermixed normal tissue within prostate cancer: effect on MR imaging measurements of apparent diffusion coefficient and T2—sparse versus dense cancers 1. *Radiology*. 2008; 249:900–8. [PubMed: 19011187]
38. Peng Y, Jiang Y, Yang C, Brown JB, Antic T, Sethi I, et al. Quantitative analysis of multiparametric prostate MR images: differentiation between prostate cancer and normal tissue and correlation with Gleason score—a computer-aided diagnosis development study. *Radiology*. 2013; 267:787–96. [PubMed: 23392430]
39. Greve DN, Fischl B. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage*. 2009; 48:63–72. [PubMed: 19573611]
40. Rother, C., Minka, T., Blake, A., Kolmogorov, V. Cosegmentation of image pairs by histogram matching-incorporating a global constraint into MRFS. 2006 IEEE computer society conference on computer vision and pattern recognition; 2006. p. 993-1000.
41. Criminisi A, Shotton J. Decision forests for computer vision and medical image analysis. 2013
42. Qi Y. Random forest for bioinformatics. *Ensemble Mach Learn*. 2012:307–23.
43. Zikic D, Glocker B, Criminisi A. Atlas encoding by randomized forests for efficient label propagation. *Med Image Comput Comput-Assisted Interv MICCAI*. 2013; 2013:66–73.
44. Dalal, N., Triggs, B. Histograms of oriented gradients for human detection. *IEEE computer society conference on computer vision and pattern recognition*, 2005. *CVPR 2005*; 2005. p. 886-93.
45. Timo A, Abdenour H, Matti PI. Face description with local binary patterns: application to face recognition. *IEEE Trans Pattern Anal Mach Intell*. 2006; 28:2037–41. [PubMed: 17108377]
46. Chengjun L, Harry W. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Trans Image Process*. 2002; 11:467–76. [PubMed: 18244647]
47. Huang L-L, Shimizu A, Hagihara Y, Kobatake H. Gradient feature extraction for classification-based face detection. *Pattern Recogn*. 2003; 36:2501–11.
48. Viola P, Jones MJ. Robust real-time face detection. *Int J Comput Vis*. 2004; 57:137–54.
49. Lienhart, R., Maydt, J. An extended set of Haar-like features for rapid object detection. *Proceedings. 2002 international conference on image processing*; 2002; 2002. p. 900-3.
50. Mita, T., Kaneko, T., Hori, O. Joint Haar-like features for face detection. *Tenth IEEE international conference on computer vision*, 2005. *ICCV 2005*; 2005. p. 1619-26.
51. Chen, Q., Georganas, ND., Petriu, EM. Real-time vision-based hand gesture recognition using Haar-like features. *Instrumentation and measurement technology conference proceedings*, 2007. *IMTC 2007*; IEEE; 2007. p. 1-6.
52. Han X. Learning-boosted label fusion for multi-atlas auto-segmentation. *Mach Learn Med Imaging*. 2013:17–24.

53. Wright J, Ma Y, Mairal J, Sapiro G, Huang TS, Yan S. Sparse representation for computer vision and pattern recognition. *Proc IEEE*. 2010; 98:1031–44.
54. Cheng, H., Liu, Z., Yang, L. Sparsity induced similarity measure for label propagation. 2009 IEEE 12th international conference on computer vision; 2009. p. 317-24.
55. Farsad M, Schiavina R, Castellucci P, Nanni C, Corti B, Martorana G, et al. Detection and localization of prostate cancer: correlation of 11C-choline PET/CT with histopathologic step-section analysis. *J Nucl Med*. 2005; 46:1642–9. [PubMed: 16204714]
56. Slongo LE, Sugisawa MC, Ioshii SO, Tâmbara Filho R, Rocha LC. Efficiency of 6- and 12-punctures biopsies to detect prostate cancer in patients with PSA < 10 ng/mL and normal digital rectal examination. *Int Braz J Urol*. 2003; 29:24–9. [PubMed: 15745463]
57. Wang Z, Ramin SA, Tsai C, Lui P, Ruckle HC, Beltz RE, et al. Telomerase activity in prostate sextant needle cores from radical prostatectomy specimens. *Urol Oncol*. 2001; 6:57–62. [PubMed: 11166622]
58. Dice LR. Measures of the amount of ecologic association between species. *Ecology*. 1945; 26:297–302.
59. Wang L, Gao Y, Shi F, Li G, Gilmore JH, Lin W, et al. LINKS: Learning-based multi-source IntegrationN frameworkK for Segmentation of infant brain images. *Neuroimage*. 2015; 108:160–72. [PubMed: 25541188]
60. Wang, H., Viswanath, S., Madabuhshi, A. Discriminatively weighted multi-scale Local Binary Patterns: applications in prostate cancer diagnosis on T2W MRI. 2013 IEEE 10th international symposium on biomedical imaging (ISBI); 2013. p. 398-401.
61. Kim JH, Kim JK, Park B-W, Kim N, Cho KS. Apparent diffusion coefficient: prostate cancer versus noncancerous tissue according to anatomical region. *J Magn Reson Imaging*. 2008; 28:1173–9. [PubMed: 18837000]
62. Mitchell, TM. Machine learning. Vol. 45. Burr Ridge, IL: McGraw Hill; 1997.
63. Hsu C-W, Lin C-J. A comparison of methods for multiclass support vector machines. *IEEE Trans Neural Netw*. 2002; 13:415–25. [PubMed: 18244442]
64. Gilad-Bachrach, R., Navot, A., Tishby, N. Margin based feature selection-theory and algorithms. Proceedings of the twenty-first international conference on Machine learning; 2004. p. 43
65. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. *Comput Learn Theory*. 1995:23–37.
66. Zhang D, Chen S, Zhou Z-H. Constraint Score: a new filter method for feature selection with pairwise constraints. *Pattern Recognit*. 2008; 41:1440–51.
67. Fawcett T. An introduction to ROC analysis. *Pattern Recogn Lett*. 2006; 27:861–74.
68. Obuchowski NA. Fundamentals of clinical research for radiologists. *Am J Roentgenol*. 2005; 184:364–72. [PubMed: 15671347]
69. Obuchowski NA, Lieber ML, Powell KA. Data analysis for detection and localization of multiple abnormalities with application to mammography. *Acad Radiol*. 2000; 7:516–25. [PubMed: 10902960]
70. Juntu J, Sijbers J, De Backer S, Rajan J, Van Dyck D. Machine learning study of several classifiers trained with texture analysis features to differentiate benign from malignant soft-tissue tumors in T1-MRI images. *J Magn Reson Imaging*. 2010; 31:680–9. [PubMed: 20187212]
71. Tanimoto A, Nakashima J, Kohno H, Shinmoto H, Kuribayashi S. Prostate cancer screening: the clinical value of diffusion-weighted imaging and dynamic MR imaging in combination with T2-weighted imaging. *J Magn Reson Imaging*. 2007; 25:146–52. [PubMed: 17139633]
72. <http://martinos.org/rtim/miccai2013/data.html>.

Biographies

Chunjun Qian is currently pursuing the Ph.D. degree at the department of mathematics in NJUST. His advisor is Prof. Xiaoping Yang. He earned his bachelor degrees at the

department of mathematics in Nanjing University of Science & Technology. His research interests focus on medical image processing, machine learning.

Li Wang is a research instructor from the University of North Carolina at Chapel Hill, USA. He works with Professor Dinggang Shen in Medical Image Analysis field. His research interests focus on image segmentation, image registration, cortical surface analysis, machine learning, etc. Contact: li.wang8401@gmail.com; li_wang@med.unc.edu.

Yaozong Gao is currently pursuing the Ph.D. degree at the computer science department in UNC-Chapel Hill. His advisor is Prof. Dinggang Shen. He earned his bachelor and master degrees at CS department in Zhejiang University. His research interests focus on machine learning, especially on its application to medical image analysis.

Ambereen Yousuf is currently involved in several research projects at the University of Chicago including the Philips HIFU trial. She received her medical degree in 1999. Her interest in clinical research subsequently led her to pursue a certificate in Clinical Trials Management at the University of Chicago.

Xiaoping Yang is a Professor in the Department of Mathematics at NJUST. He has served as a reviewer for many international journals and conferences. He has published many articles in journals and proceedings of international conferences. His research interests include Partial Differential Equations and Applications, Geometric Analysis, Image Processing.

Aytekin Oto is an expert in abdominal radiology. His research interests include imaging of the prostate, MR guided biopsies and MR guided focal therapy. He has published over 120 articles. He also serves as a reviewer for the NIH and on the editorial board of the World Journal of Radiology.

Dinggang Shen is a Professor in the Department of Radiology and BRIC at UNC-CH. He has served as a reviewer for numerous international journals and conferences. He has published 600 articles in journals and proceedings of international conferences. His research interests include medical image analysis, pattern recognition and computer vision.

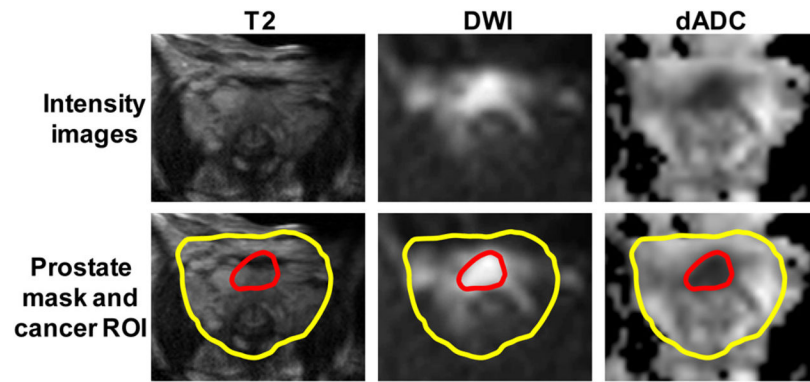


Fig. 1.

T2, DWI, and dADC images are used for cancer localization. In the 2nd row, the prostate boundary and cancer region are indicated by the yellow contour and red contour, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

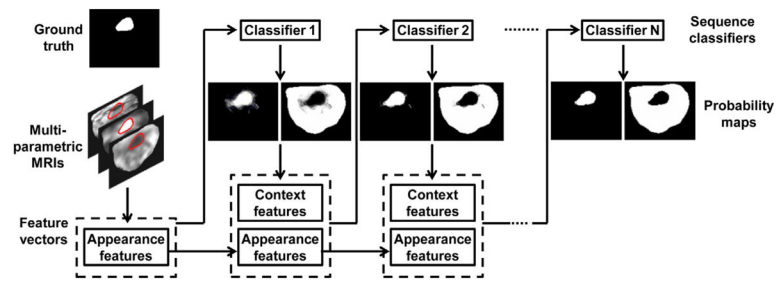


Fig. 2.

Flowchart of the training procedure for our proposed method with multi-source images, i.e., T2, DWI and dADC MR images, along with cancer probability maps. The appearance features from multi-parametric MR images are used for training the first classifier, and then both appearance features and the context features from probability maps are employed for training the subsequent classifiers.

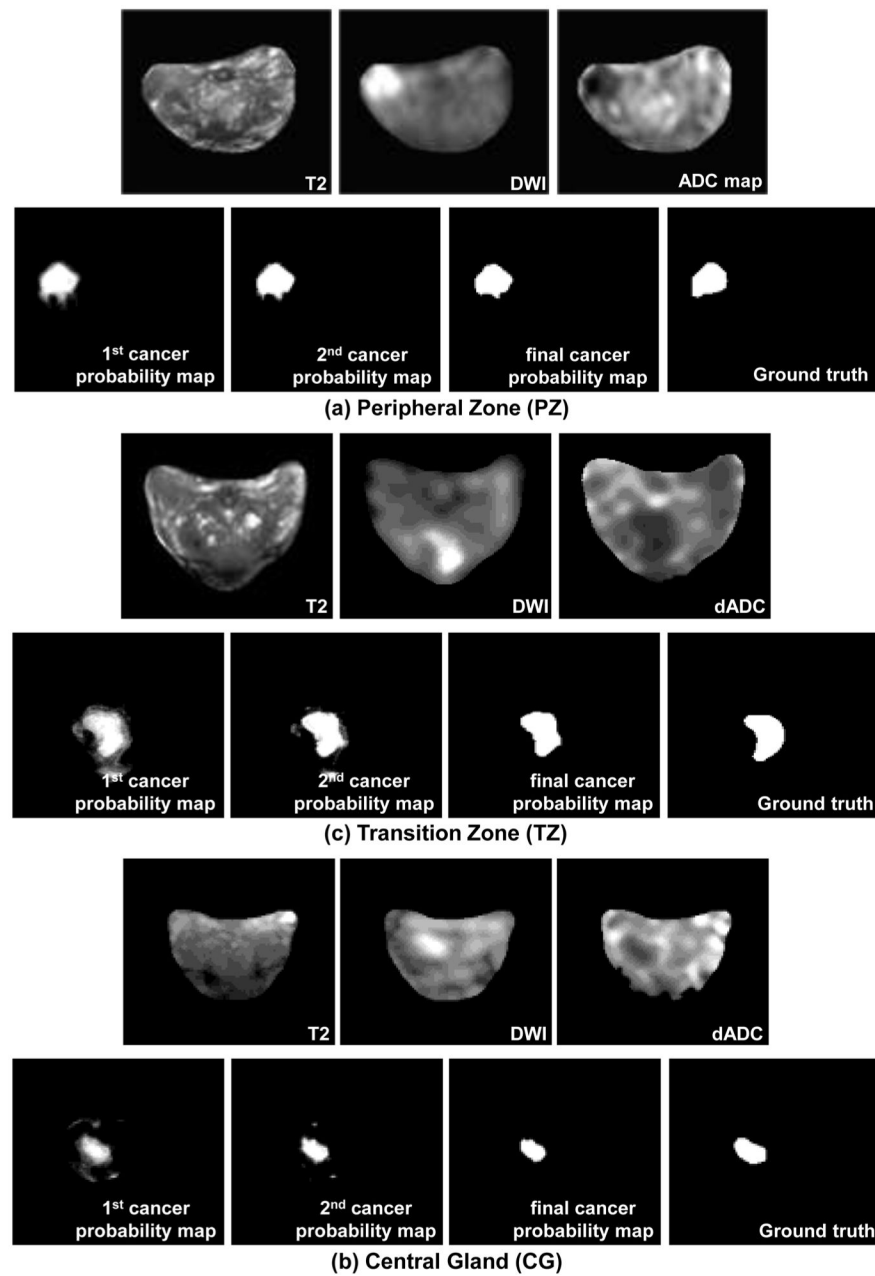


Fig. 3.

Three typical examples, showing the performances of applying the proposed method to cancer localization in (a) PZ, (b) CG, and (c) TZ. We use the trained classifier from multi-parametric MRIs in the first iteration to obtain the initial tissue probability maps. Then, in the later iterations, the tissue probability maps resulted from previous iteration are also fed into the next classifier for refinement.

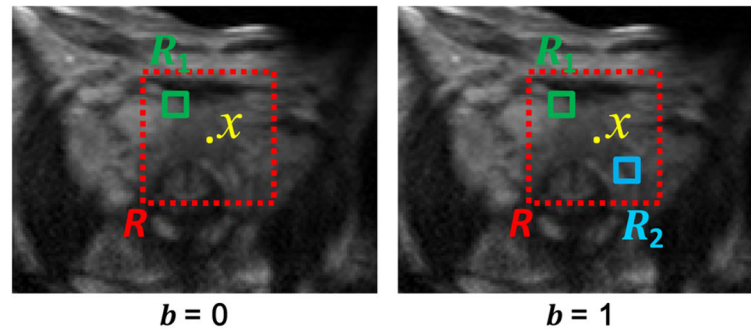


Fig. 4.

Computation of Haar-like features. The red rectangle indicates the patch centered at voxel x . Haar-like features are computed as the local mean intensity of any randomly displaced cubical region R_1 , or the mean intensity difference over any two randomly displaced, asymmetric cubical regions (R_1 and R_2) within the image patch R . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

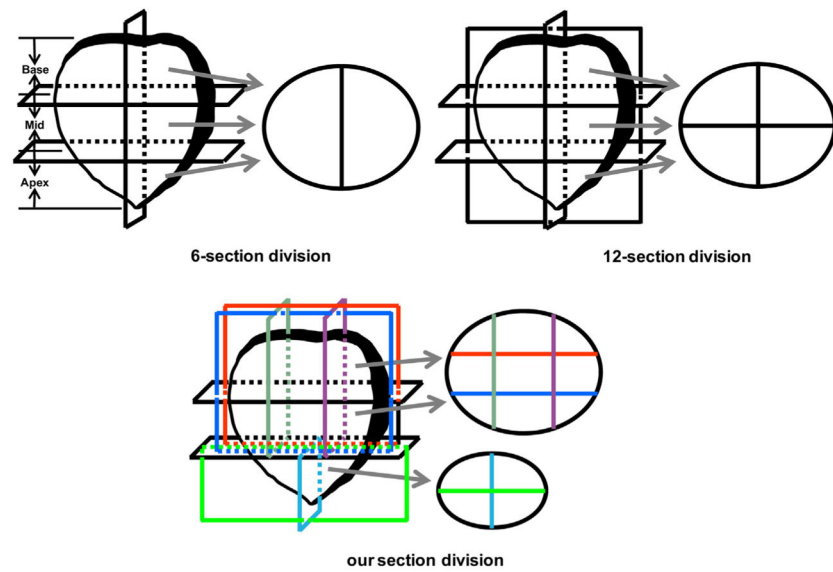


Fig. 5. Schematic diagram of the conventional prostate division (i.e., 6-section division and 12-section division) and our section division.

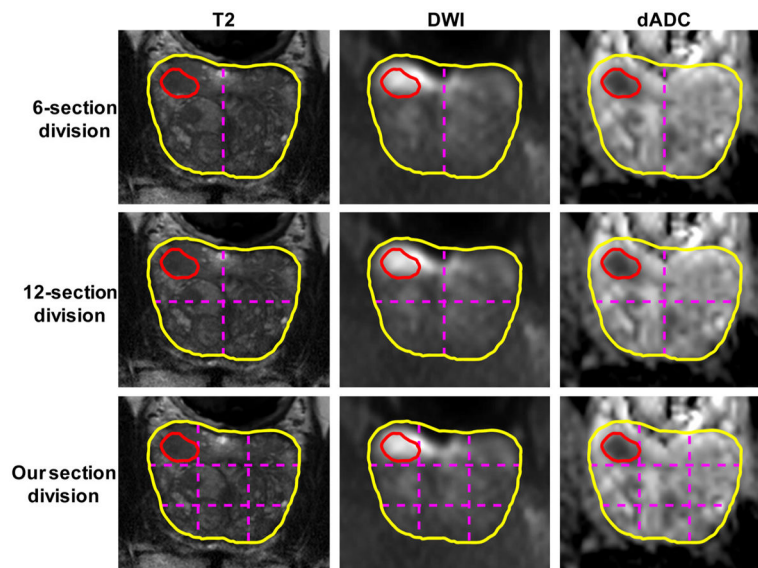


Fig. 6.

Different section division methods for a typical prostate slice. Here, the prostate is divided by pink dashed lines. The red contour indicates the manual segmentation of the cancer ROI, and yellow contour represents the prostate boundary. The slice is divided into 2, 4 and 9 parts, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

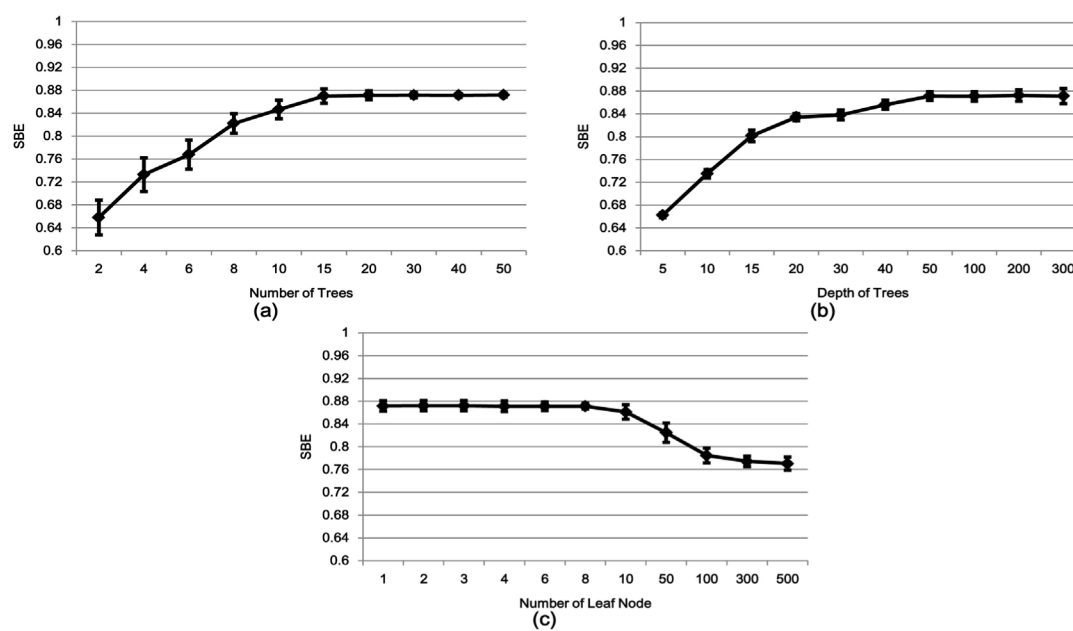


Fig. 7. Influences of (a) the number of classification trees, (b) the maximal depth of trees and (c) the minimum sample number in each leaf node.

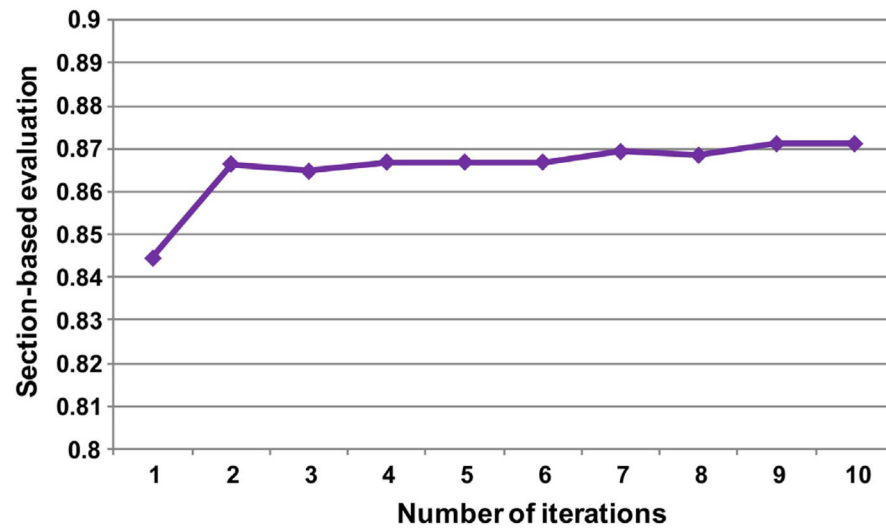


Fig. 8.

Improvement of SBE with classification iterations. Specifically, in the second iteration, the SBE is improved greatly, and, after a few iterations, it tends to be stable at around 87%.

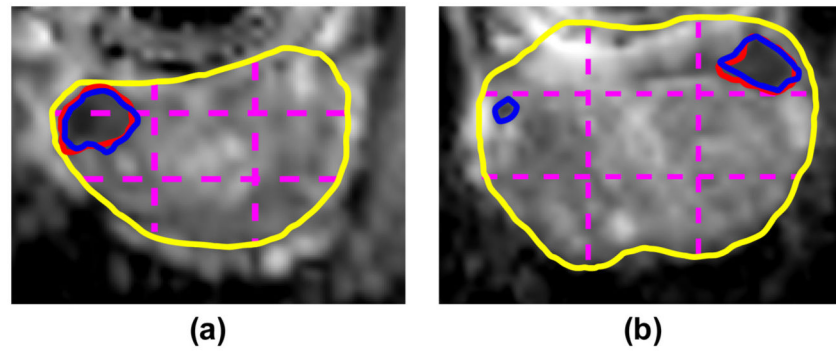


Fig. 9.

Two examples of cancer section localization by our method. Prostate is divided by pink dashed lines. Yellow contour indicates the prostate boundary. Red contours indicate the manual delineations, and blue contours indicate the localization results with our method. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

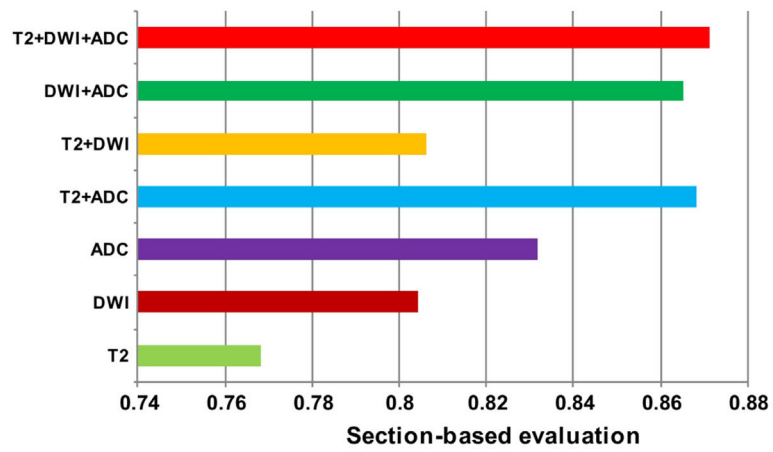


Fig. 10.

The SBE values obtained by using different combinations of multi-parametric MRI. The result by combining all T2, DWI and dADC images is the best, compared to other six combinations.

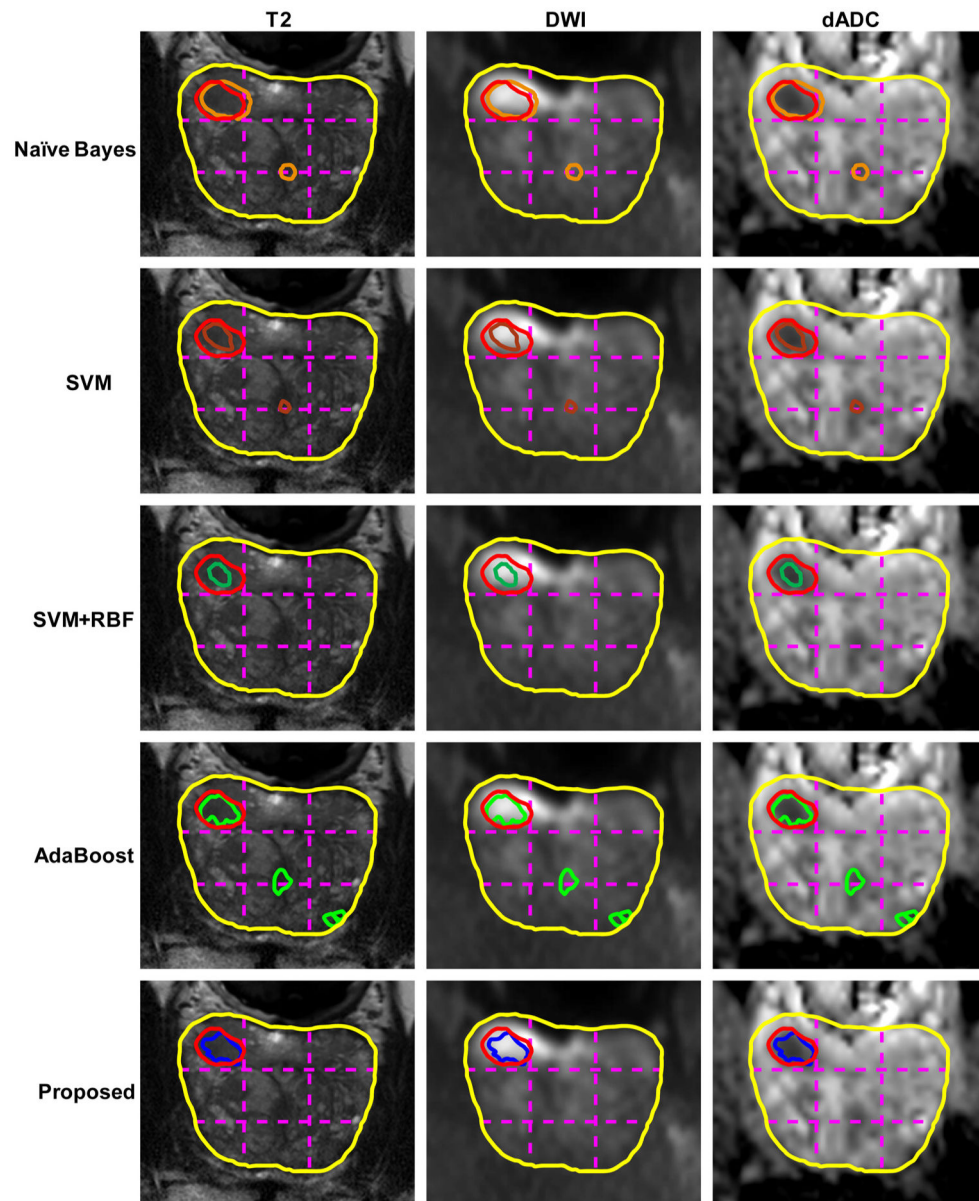


Fig. 11.

Comparison of cancer localization results by Naïve Bayes, SVM, SVM + RBF, AdaBoost, and the proposed method, on a typical prostate slice also as shown in Fig. 6. The yellow contour indicates the prostate region. The red contour indicates manual segmentation, along with the automated segmentation results by Naïve Bayes, SVM, SVM + RBF, AdaBoost, and our proposed method indicated by the orange, brown, cyan, green, and blue contours, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

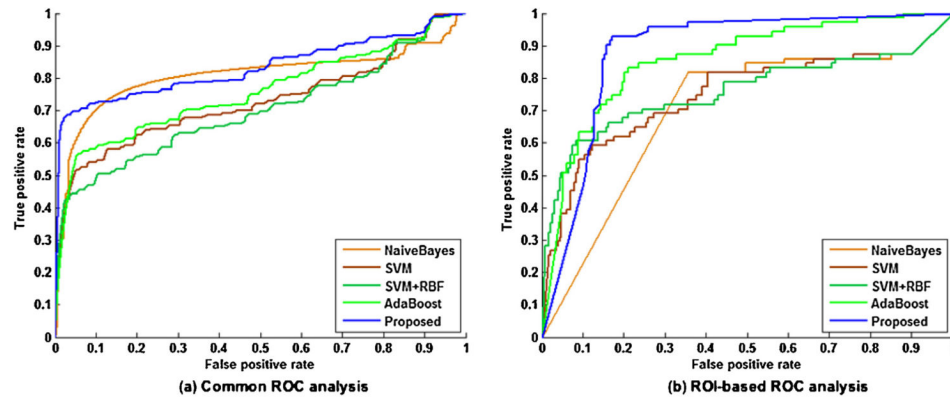


Fig. 12.

(a) Common ROC curves; (b) ROI-based ROC curves. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

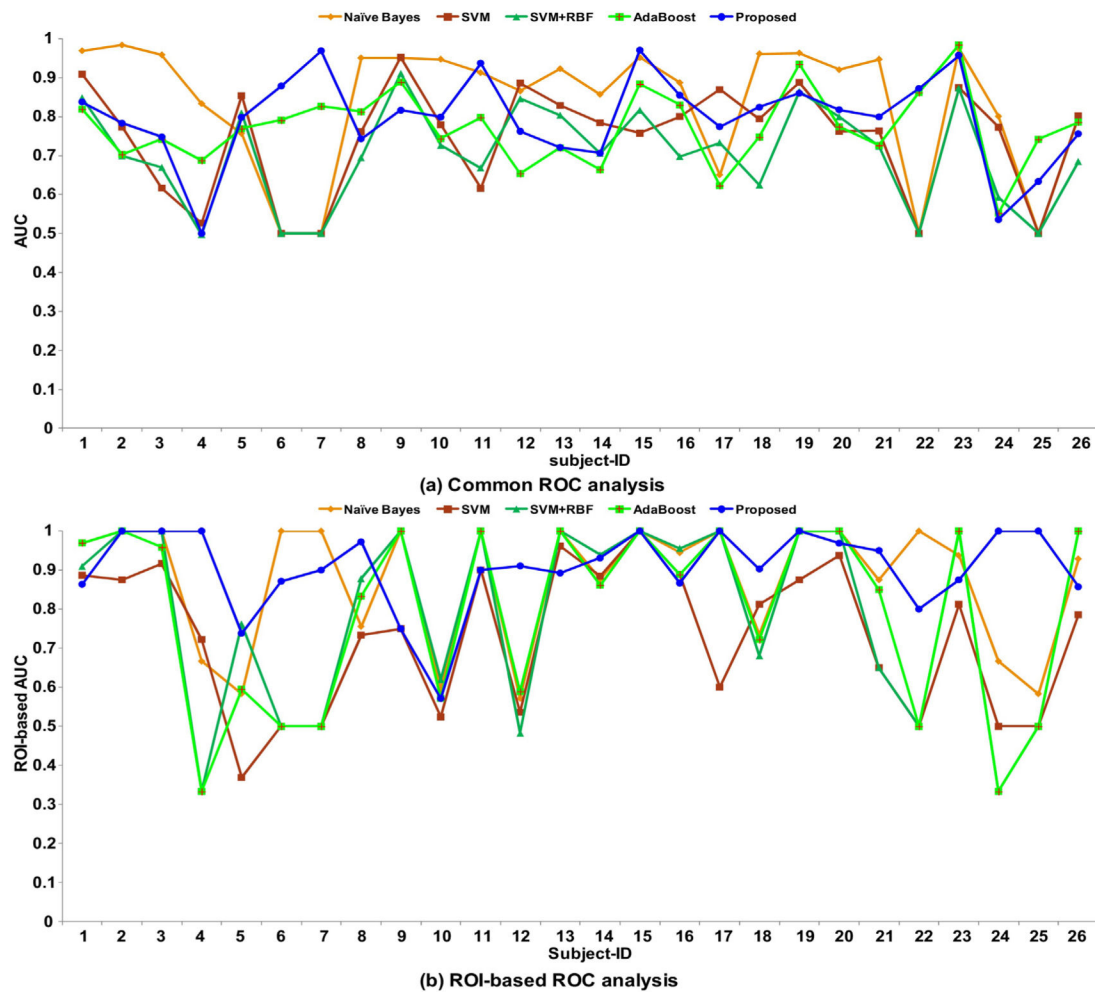


Fig. 13.

(a) Comparison of AUC values among the proposed method, Naïve Bayes, SVM, SVM + RBF, and AdaBoost methods on 26 patients. Most of AUC values are over 0.7 and the highest is 1.0 by our proposed method. (b) Comparison of AUC values under the ROI-based ROC curves by the proposed method, Naïve Bayes, SVM, SVM + RBF, and AdaBoost on 26 patients. Most AUC values by our proposed method are over 0.8. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

Table 1

MR image acquisition parameters.

Multi-parametric MRI	Sequence type	Repetition time (ms)	Echo time (ms)	Resolution (mm ³)	Flip angle (degree)
T2-weighted	Fast spin echo	3166–6581	90–120	$0.3125 \times 0.3125 \times 3$	90
DWI	Fast spin echo, echo planar imaging	2948–8616	71–85	$1.406 \times 1.406 \times 3$	90

Table 2

SBE, sensitivity and specificity by our proposed method using each type of features: HOG, LBP, Gabor, Gradient, and random Haar-like.

Features	HOG	LBP	Gabor	Gradient	Random Haar-like
SBE	0.726	0.583	0.756	0.813	0.871
Sensitivity	0.217	0.917	0.721	0.660	0.915
Specificity	0.895	0.462	0.745	0.835	0.851

The number in “bold” font indicates the best result achieved by different feature types in terms of SBE, Sensitivities, and Specificity.

Table 3

SBE, sensitivity and specificity of our cancer section localization *with respect to* different number of iterations.

	Iteration	SBE	Sensitivity	Specificity
1 (first cancer probability map)		0.844	0.941	0.813
	2	0.866	0.928	0.841
	3	0.865	0.915	0.841
	4	0.867	0.915	0.846
	5	0.867	0.915	0.846
	6	0.867	0.915	0.846
	7	0.869	0.909	0.851
	8	0.868	0.896	0.851
	9	0.871	0.915	0.851
	10	0.871	0.915	0.851

Table 4

SBE, sensitivity and specificity of cancer section localization with five different methods on 26 subjects.

Method	SBE	Sensitivity	Specificity
Naïve Bayes	0.685	0.737	0.659
SVM	0.714	0.726	0.710
SVM + RBF	0.799	0.661	0.838
AdaBoost	0.680	0.946	0.605
Proposed	0.871	0.915	0.851

The number in “bold” font indicates the best result achieved by different methods in terms of SBE, Sensitivities, and Specificity.

Table 5

Mean Dice ratios in different zones by five methods.

Method	Naive Bayes	SVM	SVM + RBF	AdaBoost	Proposed
PZ	0.443	0.364	0.367	0.456	0.620
CG	0.400	0.468	0.503	0.493	0.567
TZ	0.439	0.568	0.536	0.580	0.780

The number in “bold” font indicates the best result achieved by different methods in the zones of PZ, CG and TZ.