

Atlas registration and ensemble deep convolutional neural network-based prostate segmentation using magnetic resonance imaging

Haozhe Jia ^{a,b,c}, Yong Xia ^{a,b,*}, Yang Song ^c, Weidong Cai ^c, Michael Fulham ^{b,c,d,e}, David Dagan Feng ^c

^a Shaanxi Key Lab of Speech & Image Information Processing (SAIIP), School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, PR China

^b Centre for Multidisciplinary Convergence Computing (CMCC), School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, PR China

^c Biomedical and Multimedia Information Technology (BMIT) Research Group, School of Information Technologies, University of Sydney, NSW 2006, Australia

^d Department of PET and Nuclear Medicine, Royal Prince Alfred Hospital, NSW 2050, Australia

^e Sydney Medical School, University of Sydney, NSW 2006, Australia

ARTICLE INFO

Article history:

Received 14 June 2017

Revised 22 September 2017

Accepted 25 September 2017

Available online xxx

Communicated by Jun Yu

Keywords:

MRI prostate segmentation
Deep neural network
Pixel classification
Ensemble learning

ABSTRACT

Automatic segmentation of prostate in magnetic resonance (MR) images has been more and more applied to the diagnosis of prostate disease and various clinical applications. However, due to the inhomogeneous and varying anatomical appearance around prostate boundary, the segmentation of prostate MR images faces great challenges. Since deep learning shows superior performance in computer vision, we propose a coarse-to-fine segmentation strategy by using deep neural networks to tackle the segmentation problem of the endorectal coil prostate images and non-endorectal coil prostate images separately. First, we present a registration-based coarse segmentation to the pre-processed prostate MR images to get the potential boundary region. Second, we train deep neural networks as pixel-based classifier to predict whether the pixel in the potential boundary region is prostate pixel or not. To improve the discriminability of the algorithm, we further introduce ensemble learning for fine segmentation. Finally, a boundary refinement is used to eliminate the outlier and smooth the boundary. The proposed method has been extensively evaluated on the PROMIS12 challenge dataset and PROSTATEEx17 challenge dataset. Experimental results show superior segmentation performance (0.910 ± 0.036 in dice ratio, 1.583 ± 0.441 in average boundary distance and 4.579 ± 1.791 in Hausdorff distance), which demonstrates the effectiveness of the proposed algorithm.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

National Cancer Institute statistics indicate that more than 220,800 men were diagnosed with prostate cancer in the United States in 2015 and prostate cancer is identified as a major health threat [1]. Magnetic resonance (MR) imaging, due to its superior spatial resolution and tissue contrast, is the main imaging modality used to evaluate the prostate gland [2,3]. Assessment of the tumor volume and extension through the gland capsule are important for diagnosis, management and prognosis. Currently, MR

prostate segmentation is usually undertaken by radiologists and it is based on visual inspection, slice-by-slice, which is time consuming, requires a high degree of skill/concentration, is prone to intra- and inter-operator bias and is unsuitable for the evaluation of large-scale samples [4]. A number of semi or fully automated methods [5–7] have been proposed for the segmentation of various organs/tissues in medical images, however, the automated segmentation of prostate MR images remains a challenging task. The challenges mainly relate to the variability in size/shape/contours of the gland, heterogeneity in signal intensity around endorectal coils (ERCs), imaging artifacts and low contrast between the gland and adjacent structures [6,8,9].

In this paper, we propose a coarse-to-fine prostate segmentation approach based on a probabilistic atlas-based coarse segmentation and ensemble deep convolutional neural networks (DCNNs)-

* Corresponding author at: Shaanxi Key Lab of Speech & Image Information Processing (SAIIP), School of Computer Science and Engineering, Northwestern Polytechnical University, Xi'an 710072, PR China

E-mail address: yxia@nwpu.edu.cn (Y. Xia).

based fine segmentation. The motivation of the proposed method is to introduce deep learning, which shows superior performance in a wide variety of vision problems, to prostate segmentation. In our approach, we first perform atlas registration based coarse segmentation to determine the potential boundary region. As a result, the ensemble DCNNs-based fine segmentation only needs to focus on this small boundary region, which can significantly reduce the time consumption of the proposed method. In fine segmentation, we fine-tune a set of pre-trained DCNNs using image patches and employ the image ensemble to predict the class label of voxels in the boundary region indicated by the coarse segmentation.

There are two main contributions of this paper. First, we show that the use of pre-trained VGG-19 can alleviate overfitting and transfer the knowledge about image representation learned on the ImageNet dataset to characterizing prostate images. Second, the experimental results demonstrate the use of ensemble learning can substantially improve the performance of prostate segmentation.

Pilot data from this research was reported at the 2017 International Symposium of Biomedical Imaging (ISBI 2017, Melbourne, Vic, Australia, 18–21 April 2017). When compared to the pilot study, our current algorithm replaces the LeNet-5 network with a much deeper model VGG-19 network, which has shown superior performance in the ImageNet Challenge. The VGG-19 network has been pre-trained on the ImageNet Challenge dataset, and thus can largely overcome the difficulties caused by a small training dataset via transferring the knowledge about image representation learned on large scale natural images to characterize prostate MR images. We evaluated our approach against several state-of-the-art approaches on the Prostate MR Image Segmentation Challenge 2012 (PROMIS12) [8] and SPIE-AAPM-NCI PROSTATEx Classification Challenge 2017 (PROSTATEx17) datasets [10].

2. Related work

Current approaches to prostate MR image segmentation include those based on anatomical atlases, deformable models and machine learning techniques. The traditional approach using MR images is the atlas-based joint registration-comparison that registers an MR image to an anatomical atlas and maps anatomical structures from the atlas to the image [11,12]. Klein et al. [13] proposed a three-stage automated atlas-based segmentation using non-rigid registration. Stojanov and Koceski [5] suggested a hybrid topological prostate segmentation method based on a set of pre-labeled MR atlas images. Atlas-based methods have low accuracy and are not robust due to the anatomical variation across subjects.

A variety of deformable models have been used. Skalski et al. [14] adapted the active contour with gradient vector flow method and incorporated the prior shape knowledge into the energy function. Artan et al. [15] used a shape prior in a graph-based active contour model. Yang et al. [16] reported on a novel hierarchical level set clustering approach that made full use of statistical information from the manual segmentation result together with the shape prior. Guo et al. [17] suggested an integrated deformable model with distributed discriminative dictionary learning in a non-parametric and discriminative fashion. Ghose et al. [18] combined a probabilistic atlas with the active appearance model (AAM). Meanwhile, Yang et al. [6] applied discriminant analysis to a statistical shape model to measure the distinctiveness of the leaned scale invariant feature transformation (SIFT) descriptors for each landmark of the statistical shape model directly. All these deformable model-based approaches can fail when the margins of the gland are diffuse/ill-defined or in close proximity to structures with similar appearances such as the bladder.

Machine learning-based methods also performed well in prostate MR segmentation. Gao et al. [7] used multi-channel MR images to train a decision forest and fuse the global and local

template-based classifiers. Cheng et al. [19] developed a supervised learning approach that combined the atlas-based AAM with a support vector machine (SVM).

Deep learning is a subfield of machine learning, which learns deep-level abstractions in data by utilizing hierarchical architectures [20]. More recently, it has been applied to a wide variety of problems, such as transfer learning [21,22], natural language processing [23] and most prominently in computer vision [24–27], where DCNNs showed excellent results in the ImageNet 2012 Challenge [28]. It is now commonly accepted that DCNNs can improve the accuracy of image classification [29,30] and perform well on local tasks with a structured output [31].

The availability of a large amount of annotated medical image data has made it feasible to use DCNNs for medical image segmentation and classification. Holger et al. [32] used a multi-level DCNNs based scheme to segment the pancreas; it used progressive pruning and in a bottom-up fashion: from dense labeling of image patches, to regions and the entire organ. Havaei et al. [33] reported on fully automated brain tumor segmentation that was based on a cascade DCNN architecture, in which the output of a basic DCNN is treated as an additional source of information for the subsequent DCNN, and local and global contextual features were exploited. Kainz et al. [34] chose DCNNs as pixel classifiers to achieve semantic segmentation of the colon and trained two DCNNs to identify gland objects and gland-separating structures. Maji et al. [35] used a computational imaging framework with deep and ensemble learning for the detection of blood vessel and non-blood vessel areas on colored images of the optic fundus. Dvorak and Menze [36] used DCNNs as learning algorithms to characterize the local structures of brain tumors. Xu et al. [37] applied a DCNN-based feature learning to automatically segment and classify epithelial and stromal regions from digitized tumor tissue microarrays.

3. Methods

As previously mentioned, the common challenges of automated segmentation of prostate MR images, such as variable gland size/shape/contours, heterogeneous signal intensity and low image contrast can have severe impacts on the segmentation performance. Therefore, in this paper, we proposed a coarse-to-fine prostate segmentation approach, which combines the atlas registration and ensemble deep learning to achieve superior segmentation results. The pipeline of this approach is shown in Fig. 1.

3.1. Voxel value normalization

The MR scans in the PROMIS12 dataset show marked variations in the dynamic range, voxel size and appearance. Hence, we used the re-slicing procedure in the statistical parametric mapping (SPM) [38] toolbox as a pre-processing step to achieve a uniform voxel size of $0.65 \times 0.65 \times 1.5 \text{ mm}^3$ for each scan. Then the scans were separated into those with ERCs and those without to reduce the impact of voxel values on the segmentation. The steps of segmenting both groups of scans then include voxel value normalization, atlas-based coarse segmentation, ensemble DCNN-based fine segmentation, and boundary refinement.

The following truncated linear map was used to normalize the voxel values of each non-ERCs study

$$I'_x = \begin{cases} 255 * (I_x - I_{\min}) / (I_{\max} - I_{\min}) & , I_x \leq \tau \\ 255 & , I_x > \tau \end{cases} \quad (1)$$

where I'_x is the normalized value of voxel x , I_{\max} is the maximum and I_{\min} the minimum voxel values in the volume and τ is the truncation threshold, which is empirically set to 4096 if

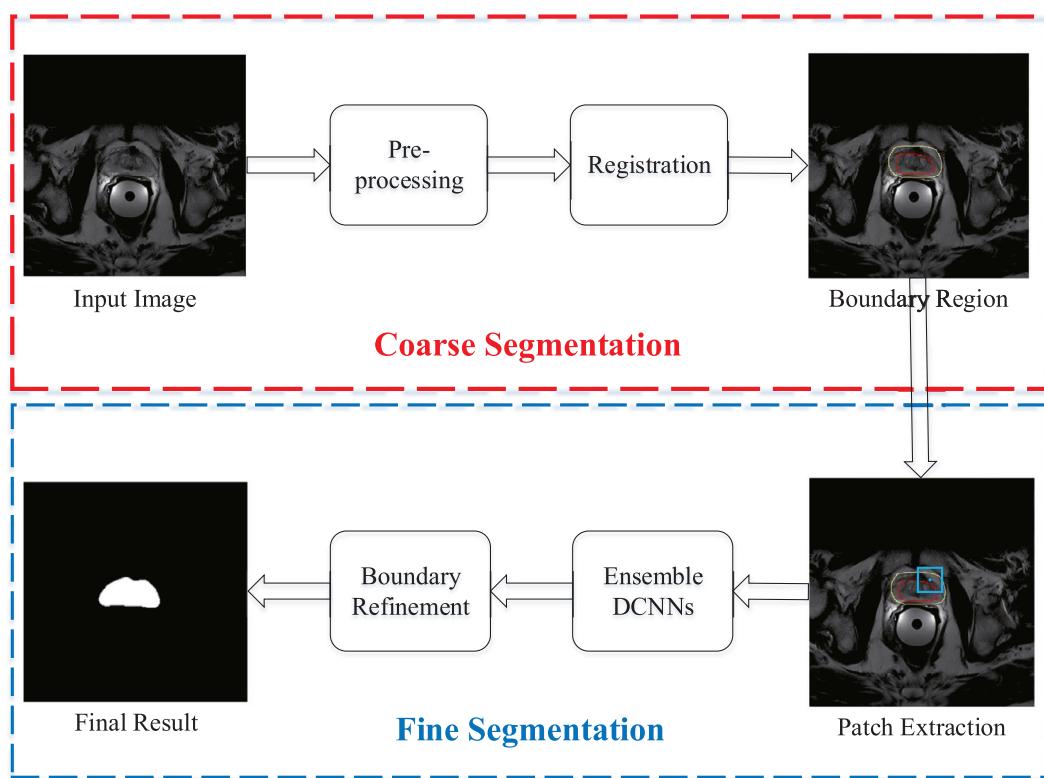


Fig. 1. Pipeline of the proposed method. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

$I_{\max} > 4096$ and 1024 otherwise. Scans where the ERC was used usually have intense spikes near the coil that may result in segmentation artifacts. Therefore, we applied the Poisson image editing [39,40] to each scan to alleviate the impact of those intense spikes in following steps:

Step 1: The region near the ERC that contains spikes was extracted by a threshold equal to the truncation threshold τ defined for non-ERC scans;

Step 2: The voxel value normalization problem was converted into seeking an adjusted image $f: \Omega \rightarrow R$, so that the boundary of Ω matches the spike region and the gradient within Ω is similar to the high pass version of the image

$$E(f) = \min \int_{\Omega} |\nabla f - \nabla g|^2 dx, \quad (2)$$

where $\Omega \subset R^2$ is the spike region, $f: \Omega \rightarrow R$ is the adjusted image intensity, $f = I$ on the boundary of Ω , R is the set of real number, R^2 is two-dimensional real number vector space and $g(x) = (I - G_\sigma * I)(x)$ is the high pass filtered image. The minimizer of Eq. (2) is a solution to the Poisson equation

$$\nabla^2 f = \nabla^2 g \quad (3)$$

Step 3: Voxel values in the spike region were replaced by the corresponding values on the adjusted image f ;

Step 4: The spike suppressed image is applied to Eq. (1) to further normalize the voxel values. An example of an ERC study scan showing the detected spike region and the result of voxel value normalization is shown in Fig. 2, which illustrates that the heterogeneity in signal intensity has been substantially suppressed after the normalization.

3.2. Atlas-based coarse segmentation

Now a probabilistic atlas is constructed for each study, and the coarse segmentation of the gland was achieved via an atlas-based

joint registration comparison analysis. The diagram of this atlas-based coarse segmentation procedure is shown in Fig. 3. For each group of studies, let the scan to-be-segmented (target image) be denoted by S . Each training case consists of a co-aligned MR scan I_i and its binary segmentation ground truth L_i . The deformable registration via attribute matching and mutual-saliency weighting (DRAMMS) [41], algorithm is adopted to modulate registration and to estimate a nonlinear transformation $T_i^{(S)}$ that maps the training scan I_i to the target scan S . Then, the estimated transformation $T_i^{(S)}$ is applied to the ground truth L_i , and thus generates a prostate atlas $A_i^{(S)}$ for S , in which each element $a_{is}^{(S)} \in [0, 1]$ gives the probability of the voxel s in S belonging to the prostate estimated according to the training scan I_i . The probabilistic atlas $A^{(S)}$ is constructed by averaging the atlas produced by using all training cases to deal with the anatomical variation of the human gland

$$A^{(S)} = \frac{1}{N} \sum_{i=1}^N A_i^{(S)}, \quad (4)$$

where N is the number of training studies in the group. This probabilistic atlas represents the heuristic knowledge provided by training data on the prior probability of each voxel belonging to the prostate gland. The procedure of constructing the probabilistic atlases $A^{(S)}$ for study S is summarized in Fig. 3. Next, the target scan was partitioned into positive, boundary, and negative volumes by applying a low threshold 0.25 and a high threshold 0.75 to the probabilistic atlas, as shown in the coarse segmentation result of Fig. 3. The positive region is inside the lower boundary (highlighted in red) and gives the core volume of the prostate. The negative region is outside the higher boundary (highlighted in yellow) and gives the background. The edges of prostate gland are located inside the annular potential boundary region (between the red and yellow lines). In the fine segmentation stage, all the pixels inside the potential region are classified to determine the accurate prostate boundary.

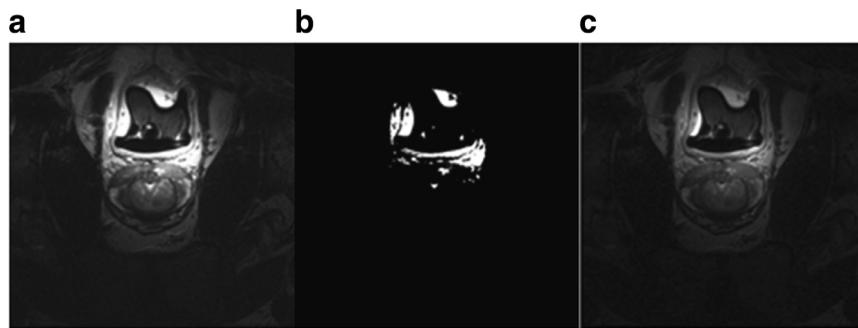


Fig. 2. MR images, using the ERC, showing: (a) the transaxial image, (b) detected spike region and (c) result from voxel value normalization.

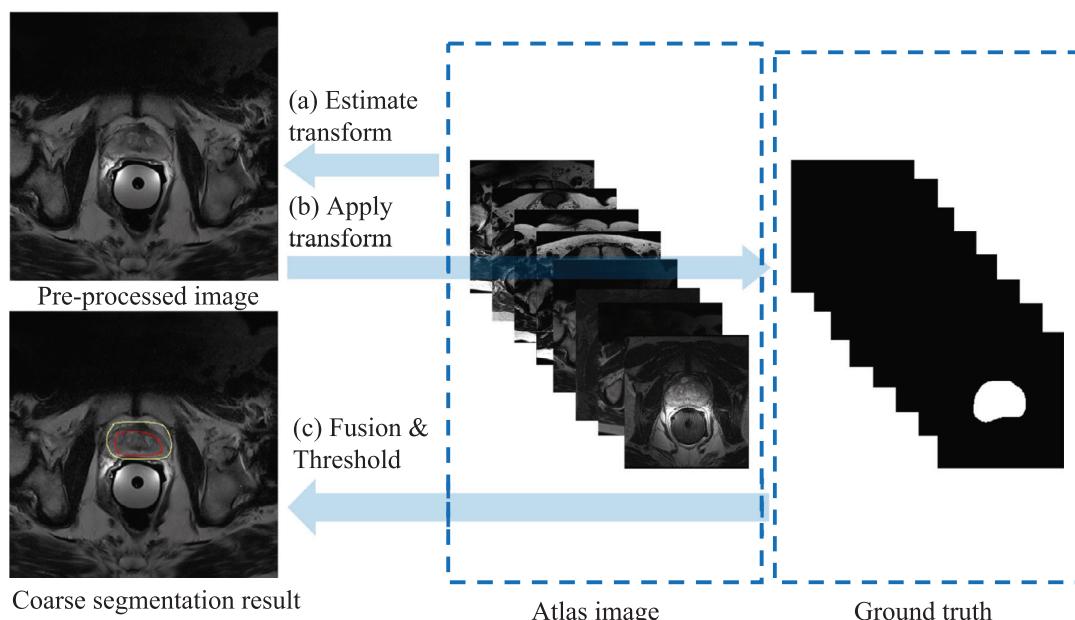


Fig. 3. Outline of atlas-based coarse prostate segmentation. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

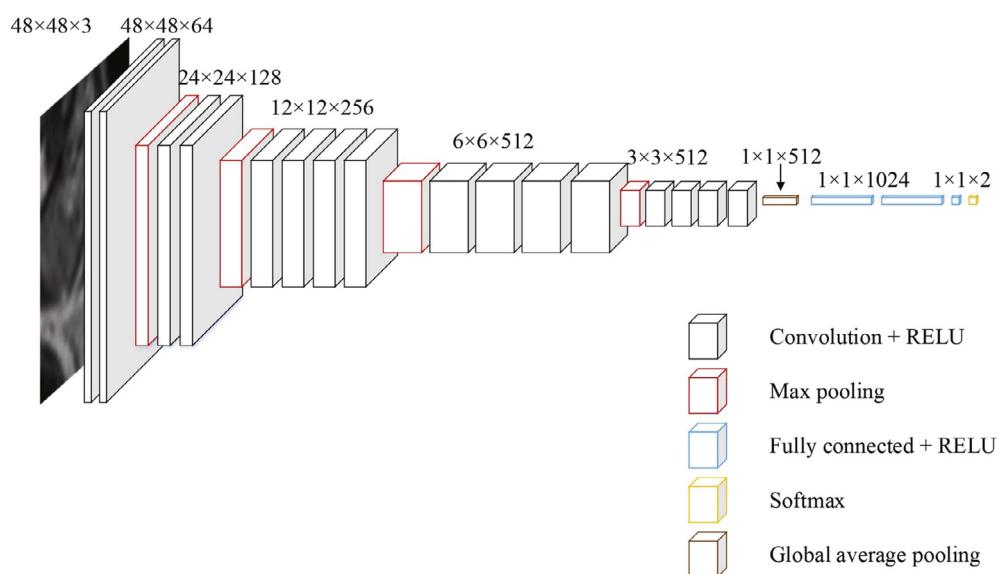


Fig. 4. Architecture of the VGG-19 model used for this study.

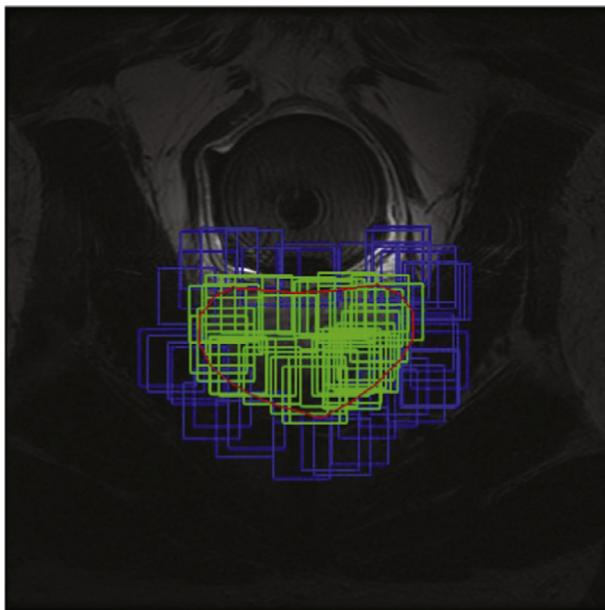


Fig. 5. Prostate patches (green) and non-prostate patches (blue) on a training slice. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Ensemble DCNN-based fine segmentation

After the atlas-based coarse segmentation determines the class label of voxels in the positive and negative volumes, the fine segmentation step further classifies each voxel in the boundary volume into prostate or non-prostate using the ensemble DCNN classifier. Since the distance between transverse slices is much larger than the distance between two voxels within the slice, fine segmentation is performed on a slice-by-slice basis from the axial view. Since the image representation ability learned by a DCNN model on large-scale natural images can be transferred to solving other computer vision tasks [26,29], we used the widely used VGG-19 network [29] to construct the ensemble DCNNs classifier. The VGG-19 consists of 19 learnable layers, including 16 convolutional layers and three fully connected layers, and five pooling layers (see Fig. 4). In each convolutional layer, the input convolves with multiple 3×3 kernels and results in a group of feature maps, which will be the input of the subsequent layer. The VGG-19 network is equipped with rectified linear units (ReLUs) as activation functions. ReLU was first proposed in AlexNet by Krizhevsky et al.

[26] in 2012. Compared to traditional activation functions, such as the sigmoid function, hyperbolic tangent function and radial basis function, ReLU has a much simpler form and significantly reduces the cost of training a deep neural network. Interestingly, ReLU enables deep models to have even better performance, partly due to its desirable property that it does not require input normalization to prevent it from saturation. The numbers of kernels in different convolutional layers range from 64 to 512. A pooling layer down-samples the feature maps obtained in the previous convolutional layer to introduce invariance to local translations to reduce the number of parameters and computation in the network. The max pooling with 2×2 receptive fields is used in the VGG-19 model. Three fully connection layers that play the role of feature transformation, contain 4096, 4096 and 1000 neurons, respectively. We used the dropout technique [42] in the fully connection layer with the disabling rate of 0.5 to ensure the network had a better generalization ability. Before producing the prediction of the class label of the center pixel, we implemented a softmax-loss layer at the end of the network, which is conceptually identical to a softmax layer followed by a binary cross-entropy loss layer, but it provides a more numerically stable gradient [43]. The VGG-19 model was previously trained on the ImageNet training set. The ImageNet training set is a 1000-category natural image database [44]. To adapt this models to our prostate segmentation problem, we randomly selected two neurons in the last fully connected layer and removed other output neurons and the weights attached to them.

The VGG-19 model used for this study was designed by the VGG team [29] for the ILSVRC-2014 competition [28] and re-implemented using Keras and re-trained using the ImageNet dataset by the Keras Group [45]. The pre-trained VGG-19 model was fine-tuned by using image patches extracted from the training studies. On each slice of a training study, a boundary region was defined as the difference between the dilation and erosion of the ground truth slice using a disk whose radius was 20 pixels. Within this boundary region, seed pixels were sampled with a 5×5 sliding window with a stride of 5 and only one non-zero element. For each seed pixel, a 48×48 image patch, centered on it, was extracted and duplicated three times as a training sample, whose class label was the label of the seed pixel. Thus, both prostate patches and non-prostate patches were extracted on both sides of the true boundary, as shown in Fig. 5, where prostate patches are highlighted in green and non-prostate patches in blue. Then, all independently sampled patches are used to fine-tune the parameters of each VGG-19 model. We set the learning rate to 0.00001 and batch size to 100. Our ensemble DCNNs classifiers comprised 7 individual VGG-19 models which were trained in the same fash-

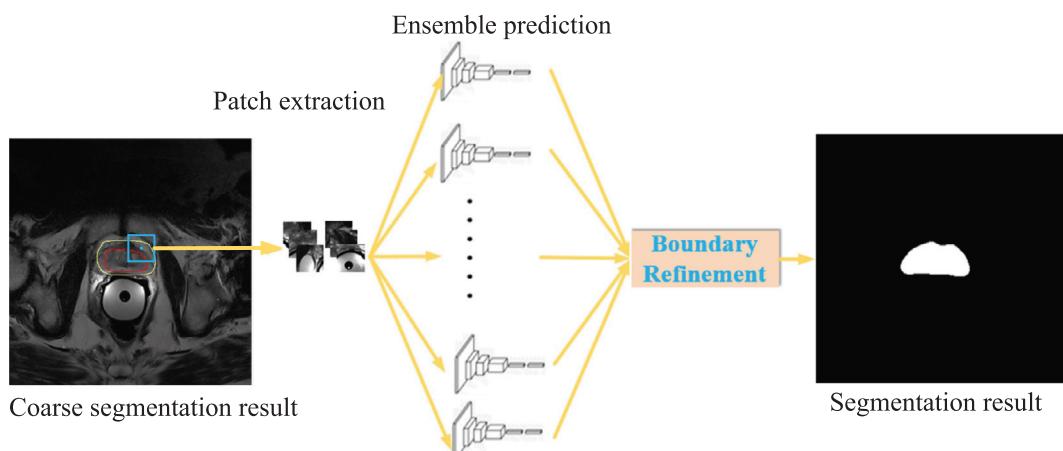


Fig. 6. Diagram for ensemble DCNN-based fine segmentation. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Table 1
Details of the dataset acquisition protocols.

Center	Field strength	ERC	Resolution	Manufacturer
Hk	1.5	Y	0.625/3.6	Siemens
BIDMC	3	Y	0.25/2.2-3	GE
UCL	1.5/3	N	0.325–0.625/3.3-6	Siemens
RUNMC	3	N	0.5–0.75/3.6–4.0	Siemens

ion. In fine segmentation, the image patch centered on each pixel in the boundary volume was extracted and applied to each individual fine-tuned VGG-19 model, and the outputs of these models were averaged to predict the class label of the corresponding pixel [46]. Finally, we applied a simple boundary refinement procedure to the raw output, since the temporary segmented images may have irregular regions on the boundary. The operations in this process included 3×3 median filtering, which removes isolated spurs and holes, and a boundary fitting procedure, which cleans up the boundary for an accurate and smooth segmentation result. In boundary fitting, we first calculated the distances between consecutive boundary points and the centroid, then removed 10% boundary points whose distance was most different from the mean distance, and finally fitted a cubic B-spline to the remaining boundary points to obtain the refined segmentation. The ensemble DCNN-based fine segmentation is shown in Fig. 6.

4. Experiments and results

4.1. Dataset

We used the PROMIS12 and PROSTATEx17 datasets in this study. Generally, the voxel intensities and appearance of prostate MR images can vary greatly between different acquisition protocols, field strength and scanners [47]. Common causes for the differences in appearance include the signal-to-noise ratio [48], the bias field [49], resolution [50] and lack of standardization of intensity values [51]. The PROMIS12 dataset comprises 100 MR scans from four medical centers – the Haukeland University Hospital (HK) in Norway, the Beth Israel Deaconess Medical Center (BIDMC) in the US, University College London (UCL) in the UK and the Radboud University Nijmegen Medical Centre (RUNMC) in the Netherlands. Data were acquired using an ERC at HK and BIDMC; UCL and RUNMC acquired data without an ERC. Each center provided 25 transverse T2-weighted MR images, together with the corresponding segmentation ground truth annotated by an experienced reader. The details of the acquisition protocols are listed in Table 1, and one scan from each center is shown in Fig. 7. The PROSTATEx17 database has 204 training MR studies from each distinct patient at a single exam. Each study has four image series: T2-weighted (transaxial and sagittal, DICOM format), Ktrans (computed from dynamic contrast-enhanced images, MDH format) and apparent diffusion

coefficient images (computed from diffusion-weighted imaging, DICOM format). This dataset does not have a corresponding segmentation ground truth, since aim of the Challenge is to provide quantitative image analysis methods for the diagnosis of clinically important prostate lesions. Representative images are shown in Fig. 8.

4.2. Experiment setting and evaluations

The first experiment was performed on the PROMIS12 dataset with a four-fold cross-validation, in which ERC and non-ERC prostate images were distributed in each fold equally. The segmentation accuracy was assessed by five metrics – the Dice similarity coefficient (DSC) [52], relative volume difference (RVD) [52], average boundary distance (ABD) [52], 95% Hausdorff distance (95%HD) [53], and the Hausdorff distance (HD) [53]. DSC is one of the most commonly-used overlap-based metrics and was calculated as the ratio between the intersection and union of the obtained volume X and ground truth volume Y

$$D(X, Y) = \frac{2|X \cap Y|}{|X| + |Y|} \quad (5)$$

where $|\cdot|$ is the cardinality of a set. The value of DSC ranges from 0 to 1, with a higher value representing a more accurate segmentation result. RVD was calculated as the percentage of the difference between the obtained volume X and ground truth volume Y

$$RVD(X, Y) = 100 \times \left(\frac{|X|}{|Y|} - 1 \right) \quad (6)$$

A positive RVD reflects under-segmentation of the prostate, whereas a negative RVD reflects over-segmentation. ABD and HD are classical shape distance-based evaluation metrics. Hence, let the surfaces of the obtained prostate volume be X_s and ground truth be Y_s . Both metrics can be calculated as

$$HD_{asym}(X_s, Y_s) = \max_{x \in X_s} (\min_{y \in Y_s} \|x - y\|) \quad (7)$$

$$ABD(X_s + Y_s) = \frac{1}{|X_s| + |Y_s|} \left(\sum_{x \in X_s} \min_{y \in Y_s} \|x - y\| + \sum_{y \in Y_s} \min_{x \in X_s} \|x - y\| \right), \quad (8)$$

where $\|x - y\|$ is the Euclidean distance between voxel x and y . Since HD is sensitive to outliers, we also adopted 95%HD, which is the 95th percentile of the asymmetric HD.

4.3. Result

Figs. 9 and 10 show the segmentation results obtained by applying our algorithm to ERC and non-ERC MR scans. For both figures, the obtained prostate boundary is highlighted in green and

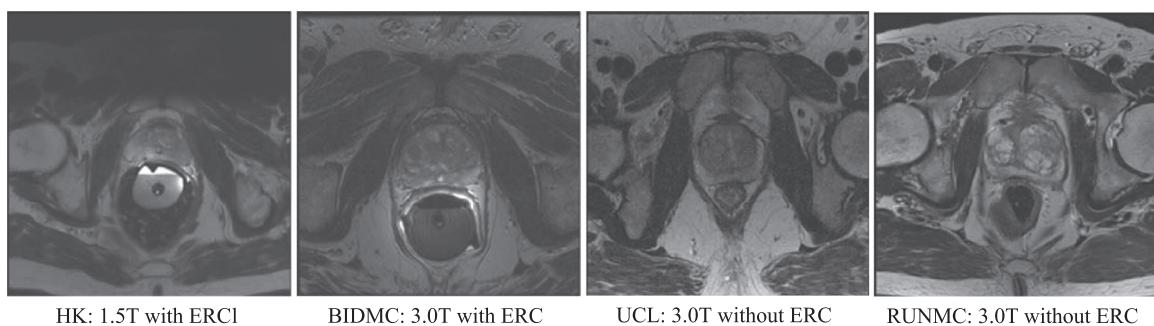


Fig. 7. Example image slices from PROMIS12 dataset (left two scans with ERC; right: two scans without ERC).

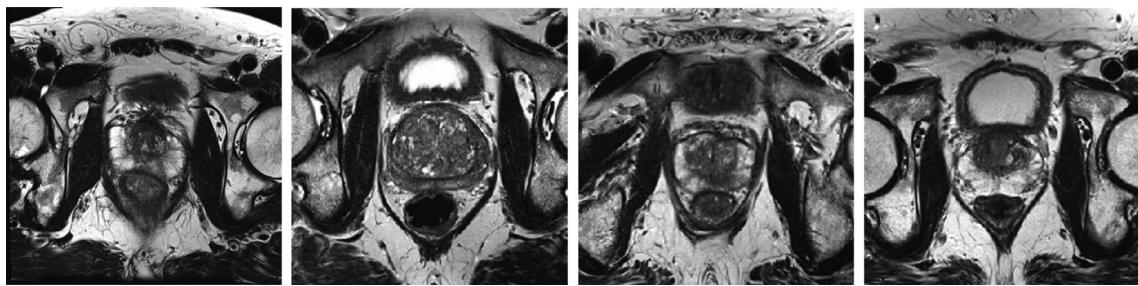


Fig. 8. Example image slices from the PROSTATEx dataset.



Fig. 9. ERC segmentation results with our algorithm; obtained boundaries are highlighted in green; the ground truth is outlined in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

the ground truth is marked in red. Although a mismatch was found in some locations, the obtained boundary was very near to the ground truth in most studies. Any inaccuracies can be largely ascribed to the complexity of soft tissues adjacent to the prostate, which may result in an over- or under-segmented prostate boundary. **Table 2** provides the mean and standard deviation of DSC, ABD,

95%HD and HD obtained when comparing our approach to the other 5 methods on the PROMIS12 dataset. Since the experimental details from the other approaches were not available, we took the results directly from the papers rather than reproducing them in the test. In the comparison, our algorithm achieved the highest segmentation accuracy regardless of the metric. Given that there

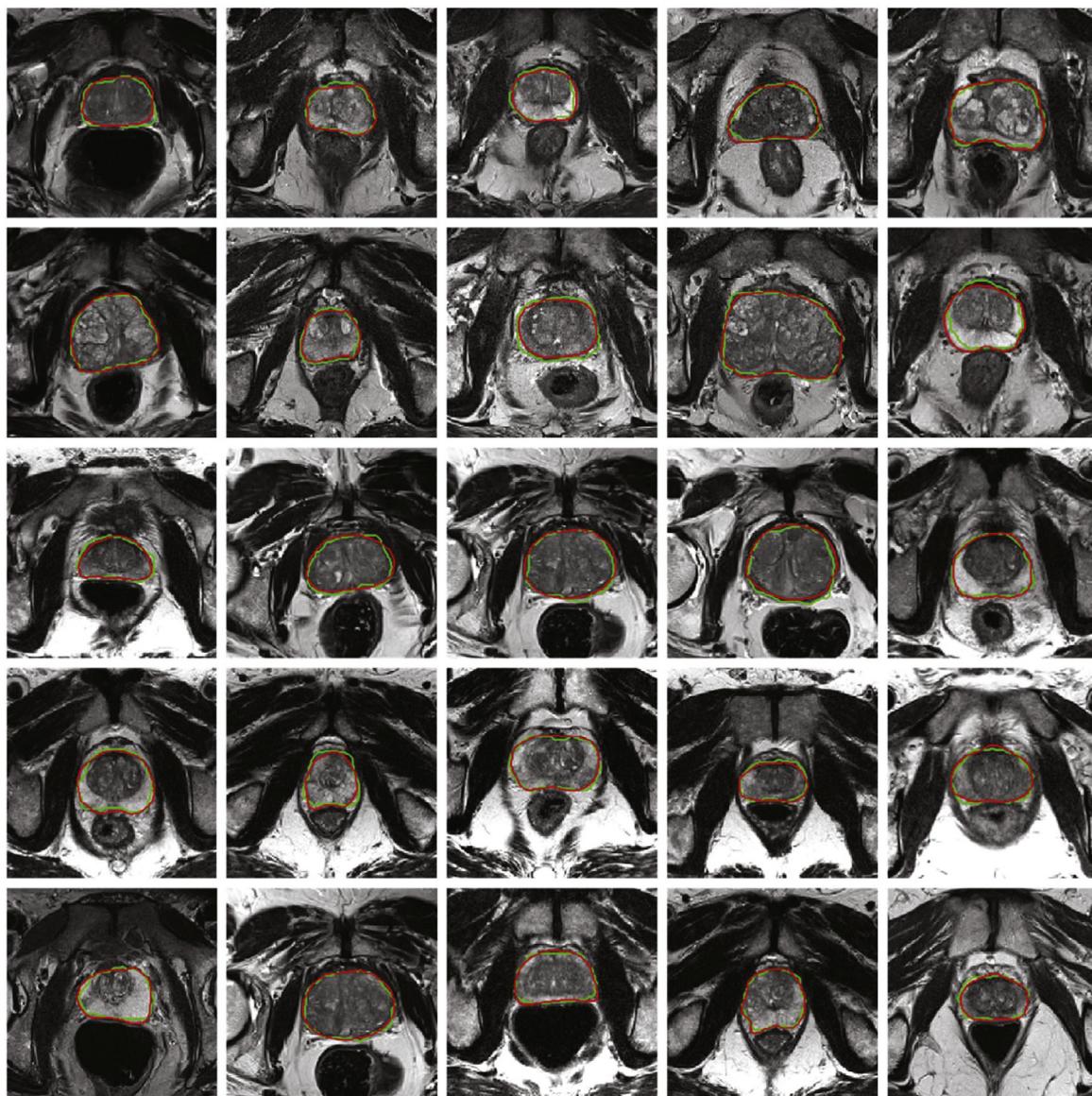


Fig. 10. Non-ERC segmentation results of our algorithm; obtained boundaries are highlighted in green; the ground truth is outlined in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2

Mean \pm standard deviation of quantitative results for segmentations obtained by different algorithms.

Algorithm	DSC	RVD(%)	ABD(mm)	HD(mm)	95%HD(mm)
3D AAM (one shape model) [54]	0.784 ± 0.120	/	/	7.320 ± 4.910	/
3D AAM (two shape model) [54]	0.810 ± 0.120	/	/	6.430 ± 4.630	7.300 ± 4.900
Atlas fusion (local atlases and patch weighting) [55]	0.847 ± 0.044	/	/	/	/
Probabilistic ASM [56]	0.860 ± 0.006	/	1.600 ± 0.630	/	9.510 ± 2.730
Automated AAM [37]	0.880 ± 0.030	/	/	4.170 ± 1.350	/
Our algorithm	0.910 ± 0.036	4.674 ± 9.401	1.583 ± 0.441	2.813 ± 1.292	4.579 ± 1.791

was no ground truth for the prostate boundaries in the PROSTATE17 dataset, we directly tested our approach on the PROSTATE17 dataset using the network model trained with the PROMIS12 dataset. The qualitative segmentation results are shown in Fig. 11. It shows that although some prostate glands were over- and under-segmented, our segmentation results largely matched the outline of the gland in each study. These results suggest that our approach can be applied to a different dataset.

5. Discussion

5.1. Parameter settings

Our algorithm has a number of parameters but using the pre-trained VGG-19 model, most of the parameters take the default values. We specified the remaining parameters as follows. When fine-tuning those pre-trained models, we set the learning rate to 0.00001, since the size of our training dataset is relatively small and a large learning rate may lead to over-fitting or an early stop.

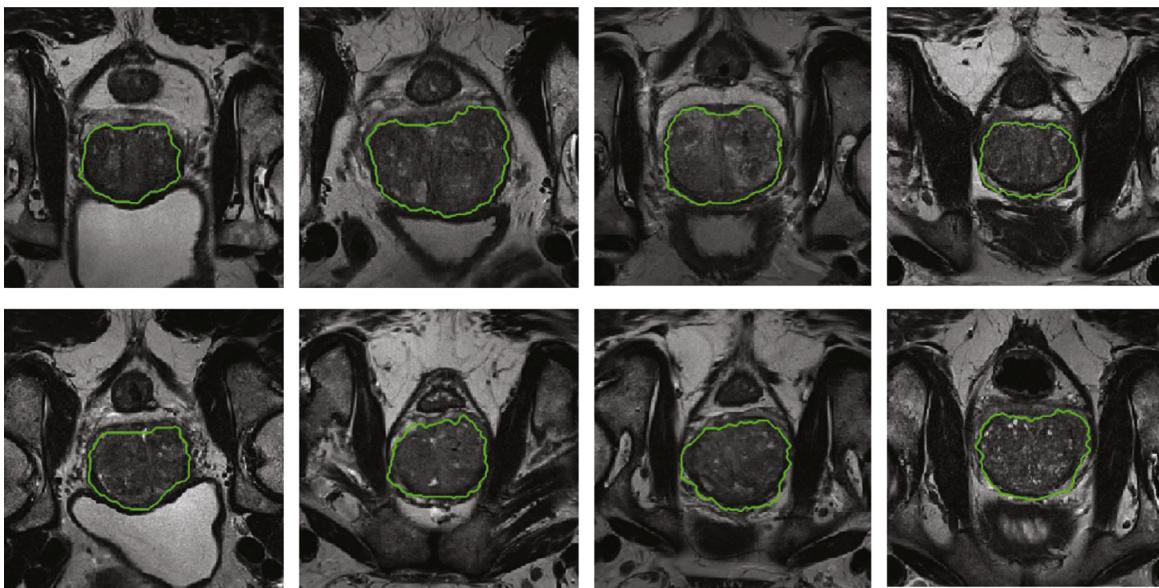


Fig. 11. Segmentation results using our approach on examples from the PROSTATEx dataset. (For interpretation of the references to color in this figure, the reader is referred to the web version of this article.)

Table 3
Comparison of segmentation performance with different patch sizes.

Patch size	DSC	RVD(%)	ABD(mm)	HD(mm)	95%HD(mm)
48	0.887 ± 0.041	6.367 ± 13.745	1.715 ± 0.356	3.788 ± 1.153	5.736 ± 1.574
56	0.882 ± 0.042	5.927 ± 15.658	1.883 ± 0.400	5.043 ± 1.260	7.579 ± 1.791
64	0.867 ± 0.054	9.174 ± 21.419	2.056 ± 0.520	6.352 ± 1.450	8.991 ± 1.626

Table 4
Comparison of segmentation performance with different ensemble numbers.

Ensemble number	DSC	RVD(%)	ABD(mm)	HD(mm)	95%HD(mm)	Total training time(h)
3	0.891 ± 0.040	6.300 ± 12.666	1.700 ± 0.354	3.715 ± 1.232	4.890 ± 1.733	24
5	0.903 ± 0.034	5.410 ± 12.700	1.667 ± 0.335	3.295 ± 1.122	4.612 ± 1.593	40
7	0.910 ± 0.036	4.674 ± 9.407	1.583 ± 0.441	2.813 ± 1.292	4.579 ± 1.791	56

We set the size of image patches to 48×48 , which is the smallest input size accepted by the VGG-19 model, since using large patches increases the number of mixed patches where there are prostate and non-prostate regions. We also tested our approach with a single pre-trained VGG-19 model of different patch sizes. The results, in [Table 3](#), show that our approach has a superior performance when the patch size is 48×48 . The number of the VGG-19 networks in the ensemble mode may also impact our approach. We set this parameter to different values and tested the performance of the proposed algorithm on the PROMIS12 dataset (see [Table 4](#)). [Table 4](#) shows that our algorithm achieved the highest accuracy when seven pre-trained VGG-19 models were used. The time cost for model training, however, surges with the increase of the ensemble number. Therefore, we empirically set this parameter to seven in our experiments as a trade-off between accuracy and complexity.

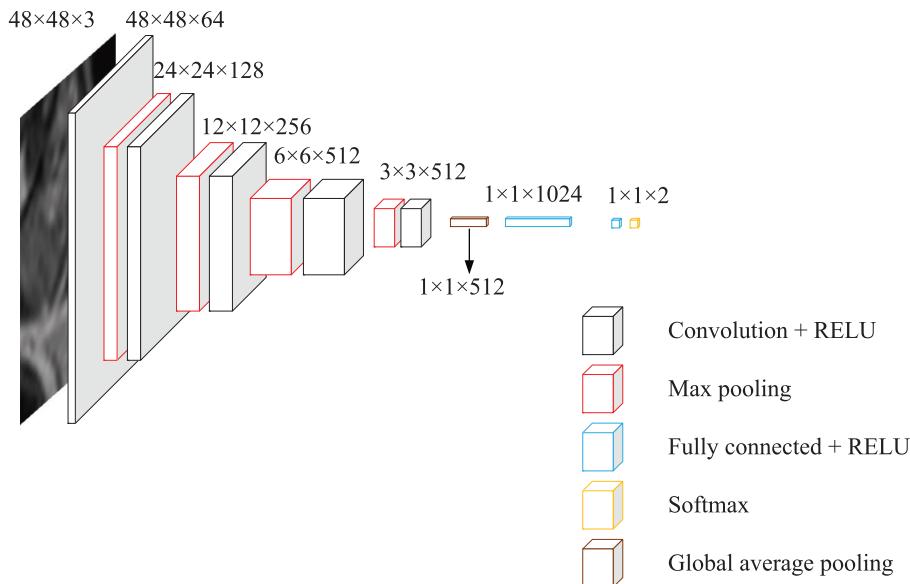
5.2. Pre-trained versus fully-trained DCNN

To show the advantage of using pre-trained DCNN models, we replaced the pre-trained VGG-19 model with the LeNet-5 model, which was randomly initialized and fully-trained by using extracted image patches, and re-performed the segmentation experiments on the PROMIS12 dataset. As shown in [Fig. 12](#), this DCNN is based on the LeNet-5 model and comprises four convolution layers, four pooling layers and two fully connected layers. The segmentation accuracy using both DCNN models is shown in

[Table 5](#), where with or without ensemble learning, the fine-tuned pre-trained VGG-19 model always outperformed the fully trained LeNet-5 model. This finding relates to the small size of the training dataset which is not able to effectively train a deep model without suffering from over-fitting, and that the knowledge about image representation learned on the ImageNet dataset can be transferred to prostate MR images.

5.3. Computational complexity

Our algorithm has four main steps: voxel value normalization, probabilistic atlas-based coarse segmentation, ensemble DCNN-based fine segmentation and boundary refinement. Of these steps, the construction of the probabilistic atlas for each study and training the ensemble DCNN model are the most time consuming. We constructed the probabilistic atlas, using the DRAMMS algorithm, on a Mac OS X 10.10 system (Intel Core i5 CPU 1.4GHz and 8 GB memory) and DCNN models were trained on an Ubuntu 14.04 64bit system (Intel Core i7-4790 CPU 3.2GHz, NVidia GTX Titan X GPU and 32GB memory). [Table 6](#) outlines the details and times to perform the segmentations using our approach. During the construction of the probabilistic atlas, the time taken to spatially normalize one training case took 4–5 min. During fine segmentation, the off-line training time of each randomly weighted initialized DCNN model took about 120 min and the pre-trained weights initialized DCNN model took 60 min. The online test of ensemble DCNN, including patches extraction and voxel classification took

**Fig. 12.** Architecture of the LeNet-5 model.**Table 5**Mean \pm standard deviation of quantitative results for segmentations obtained by different algorithms.

Algorithm	DSC	RVD(%)	ABD(mm)	HD(mm)	95%HD(mm)
Single model	LeNet-5	0.859 \pm 0.065	10.079 \pm 14.695	1.858 \pm 0.626	5.320 \pm 1.612
	VGG-19	0.887 \pm 0.041	6.367 \pm 13.745	1.715 \pm 0.356	3.788 \pm 1.153
Ensemble model	LeNet-5	0.877 \pm 0.072	7.320 \pm 10.268	1.753 \pm 0.687	5.071 \pm 1.698
	VGG-19	0.910 \pm 0.036	4.674 \pm 9.401	1.583 \pm 0.441	2.813 \pm 1.292
					4.579 \pm 1.791

Table 6

Details and times for our approach.

Parameter	Value	
	Registration	Other work
Algorithm	Platform:	Mac OS X 10.10 system
	Language:	C++
	Libraries:	DRAMMS
	Multi-threaded:	None
	User interaction:	None
	Machine	Intel core i5
Machine	CPU:	CPU 1.4 GHz
	CPU clock speed:	8 GB
	Machine memory:	None
	GPU:	Nvidia GTX titan X
Time	Normalization and enhancement:	≤ 1 s (per study)
	Registration:	30–40 min (per study registration by 10 atlas)
	Label fusion:	≈ 1 s (per study)
	Off-line DCNNs training:	3.5 h (randomly initialization)
	DCNNs prediction:	≈ 2 min (per study)
	Boundary refinement:	≤ 1 s (per study)
	Total segmentation:	≈ 40 min (per study)

about 1 min. The time for boundary refinement was less than 1 s. To shorten the time required for our approach we will need to use a more efficient spatial normalization method to replace the DRAMMS algorithm or employ a more powerful deep learning model to avoid the construction of the probabilistic atlas.

6. Conclusions

We present an automated coarse-to-fine segmentation approach for prostate MR studies obtained with and without ERC. The coarse segmentation was achieved by using a probabilistic atlas constructed for each scan to determine the potential boundary region; the fine segmentation was done using a cohort of trained DCNNs. Our results suggest that ensemble DCNNs initialized with

pre-trained weights substantially improve segmentation accuracy and our approach produces more accurate segmentation than current state-of-the-art probabilistic ASM, 3D AAM-based algorithms on the PROMIS12 dataset and shows good qualitative segmentation results on the PROSTATEx17 dataset. Our approach, however, is time consuming and in future work we will explore approaches to minimize the time required.

Conflict of interests

The authors declare that there is no conflict of interests regarding the publication of this article.

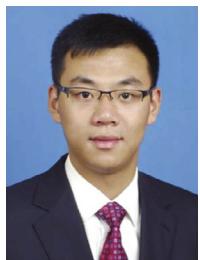
Acknowledgment

This work was supported in part by the National Natural Science Foundation of China under Grants 61471297 and 61771397, and in part by the Australian Research Council under Grants DP160103675, LP160101162 and DP170104304. We appreciate the efforts devoted by the organizers of the 2012 Prostate MR Image Segmentation Challenge to collect and share the data for comparing interactive and (semi)-automatic segmentation algorithms for MRI of the prostate.

References

- [1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2016, CA: A Cancer J. Clin. 66 (1) (2016) 7–30.
- [2] A. Gubern-Merida, M. Kallenberg, R.M. Mann, R. Marti, N. Karssemeijer, Breast segmentation and density estimation in breast MRI: a fully automatic framework, IEEE J. Biomed. Health Inform. 19 (1) (2015) 349–357.
- [3] J.L. Leake, R. Hardman, V. Ojili, I. Thompson, A. Shanbhogue, J. Hernandez, J. Barentsz, Prostate MRI: access to and current practice of prostate MRI in the united states, J. Am. Coll. Radiol. 11 (2) (2014) 156–160.
- [4] Y. Yuan, B. Li, M.Q.-H. Meng, Bleeding frame and region detection in the wireless capsule endoscopy video, IEEE J. Biomed. Health Inform. 20 (2) (2016) 624–630.
- [5] D. Stojanov, S. Koceski, Topological MRI prostate segmentation method, in: Proceeding of the Federated Conference on Computer Science and Information Systems (FedCSIS), IEEE, 2014, pp. 219–225.
- [6] M. Yang, X. Li, B. Turkbey, P.L. Choyke, P. Yan, Prostate segmentation in MR images using discriminant boundary features, IEEE Trans. Biomed. Eng. 60 (2) (2013) 479–488.
- [7] Q. Gao, A. Asthana, T. Tong, Y. Hu, D. Rueckert, P. Edwards, Hybrid decision forests for prostate segmentation in multi-channel MR images, in: Proceedings of the Twenty Second International Conference on Pattern Recognition (ICPR), IEEE, 2014, pp. 3298–3303.
- [8] G. Litjens, R. Toth, W. van de Ven, C. Hoeks, S. Kerkstra, B. van Ginneken, G. Vincent, G. Guillard, N. Birbeck, J. Zhang, et al., Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge, Med. Image Anal. 18 (2) (2014) 359–373.
- [9] S. Ghose, A. Oliver, R. Martí, X. Lladó, J.C. Vilanova, J. Freixenet, J. Mitra, D. Sidibé, F. Meriaudeau, A survey of prostate segmentation methodologies in ultrasound, magnetic resonance and computed tomography images, Comput. Methods Prog. Biomed. 108 (1) (2012) 262–287.
- [10] SPIE-AAPM-NCI prostatex challenges, (<https://wiki.cancerimagingarchive.net/display/Public/SPIE-AAPM-NCI+PROSTATEx+Challenges>).
- [11] G. Topping, K. Dinelle, S. McCormick, R. Kornelsen, V. Sossi, Methods for parkinsons rat model PET image analysis with regions of interest, in: Proceedings of the IEEE Nuclear Science Symposium Conference Record, vol. 4, IEEE, 2007, pp. 2917–2921.
- [12] H.D. Protas, P.M. Thompson, K.M. Hayashi, C.-L. Yu, M. Bergsneider, S.-C. Huang, Regional quantitative analysis of cortical surface maps of FDG PET images, in: Proceedings of the IEEE Nuclear Science Symposium Conference Record, vol. 3, IEEE, 2005, pp. 4–pp.
- [13] S. Klein, U.A. van der Heide, B.W. Raaymakers, A.N. Kotte, M. Staring, J.P. Pluim, Segmentation of the prostate in MR images by atlas matching, in: Proceeding of the Fourth IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE, 2007, pp. 1300–1303.
- [14] A. Skalski, J. Lagwa, P. Kedzierawski, T. Zielinski, T. Kuszewski, Automatic prostate segmentation in MR images based on 3D active contours with shape constraints, in: Proceeding of the Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA), IEEE, 2013, pp. 246–249.
- [15] Y. Artan, M.A. Haider, I.S. Yetik, Graph-based active contours using shape priors for prostate segmentation with MRI, in: Proceeding of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro, IEEE, 2011, pp. 1459–1462.
- [16] X. Yang, S. Zhan, D. Xie, H. Zhao, T. Kurihara, Hierarchical prostate MRI segmentation via level set clustering with shape prior, Neurocomputing 257 (2017) 154–163.
- [17] Y. Guo, Y. Zhan, Y. Gao, J. Jiang, D. Shen, MR prostate segmentation via distributed discriminative dictionary (DDD) learning, in: Proceedings of the IEEE Tenth International Symposium on Biomedical Imaging (ISBI), IEEE, 2013, pp. 868–871.
- [18] S. Ghose, J. Mitra, A. Oliver, R. Martí, X. Lladó, J. Freixenet, J.C. Vilanova, D. Sidibé, F. Meriaudeau, A coupled schema of probabilistic atlas and statistical shape and appearance model for 3D prostate segmentation in MR images, in: Proceedings of the Nineteenth IEEE International Conference on Image Processing (ICIP), IEEE, 2012, pp. 541–544.
- [19] R. Cheng, B. Turkbey, W. Gandler, H.K. Agarwal, V.P. Shah, A. Bokinsky, E. McCreedy, S. Wang, S. Sankineni, M. Bernardo, et al., Atlas based AAM and SVM model for fully automatic MRI prostate segmentation, in: Proceedings of the IEEE Thirty-Sixth Annual International Conference on Engineering in Medicine and Biology Society (EMBC), IEEE, 2014, 2881–2858.
- [20] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, M.S. Lew, Deep learning for visual understanding: a review, Neurocomputing 187 (2016) 27–48.
- [21] Z. Huang, S.M. Siniscalchi, C.-H. Lee, A unified approach to transfer learning of deep neural networks with applications to speaker adaptation in automatic speech recognition, Neurocomputing 218 (2016) 448–459.
- [22] D.C. Cireşan, U. Meier, J. Schmidhuber, Transfer learning for Latin and Chinese characters with deep neural networks, in: Proceedings of the International Joint Conference on Neural Networks (IJCNN), IEEE, 2012, pp. 1–6.
- [23] R. Collobert, J. Weston, A unified architecture for natural language processing: deep neural networks with multitask learning, in: Proceedings of the International Joint Conference on Neural Networks (IJCNN), ACM, 2008, pp. 160–167.
- [24] J. Yu, X. Yang, F. Gao, D. Tao, Deep multimodal distance metric learning using click constraints for image ranking, IEEE Trans. Cybern. PP (99) (2016) 1–11.
- [25] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
- [26] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet Classification with Deep Convolutional Neural Networks, in: Proceedings of the Advances in Neural Information Processing Systems, Curran Associates, Inc., 2012, pp. 1097–1105.
- [27] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.
- [28] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, et al., Imagenet large scale visual recognition challenge, Int. J. Comput. Vis. 115 (3) (2015) 211–252.
- [29] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, (2014). ArXiv e-print arxiv:[1409.1556](https://arxiv.org/abs/1409.1556).
- [30] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
- [31] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
- [32] H.R. Roth, L. Lu, A. Farag, H.-C. Shin, J. Liu, E.B. Turkbey, R.M. Summers, Deeporgan: multi-level deep convolutional networks for automated pancreas segmentation, in: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 556–564.
- [33] M. Haweai, A. Davy, D. Warde-Farley, A. Biard, A. Courville, Y. Bengio, C. Pal, P.-M. Jodoin, H. Larochelle, Brain tumor segmentation with deep neural networks, Med. Image Anal. 35 (2017) 18–31.
- [34] P. Kainz, M. Pfeiffer, M. Urschler, Semantic segmentation of colon glands with deep convolutional neural networks and total variation segmentation, (2015). ArXiv e-print arxiv:[1511.06919](https://arxiv.org/abs/1511.06919).
- [35] D. Maji, A. Santara, P. Mitra, D. Sheet, Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images, (2016). ArXiv e-print arxiv:[1603.04833](https://arxiv.org/abs/1603.04833).
- [36] P. Dvorkin, B. Menze, Structured prediction with convolutional neural networks for multimodal brain tumor segmentation, in: Proceedings of the Multimodal Brain Tumor Image Segmentation Challenge, 2015, pp. 13–24.
- [37] J. Xu, X. Luo, G. Wang, H. Gilmore, A. Madabhushi, A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images, Neurocomputing 191 (2016) 214–223.
- [38] K.J. Friston, A.P. Holmes, K.J. Worsley, J.-P. Poline, C.D. Frith, R.S. Frackowiak, Statistical parametric maps in functional imaging: a general linear approach, Hum. Brain Mapp. 2 (4) (1994) 189–210.
- [39] P. Pérez, M. Gangnet, A. Blake, Poisson image editing, ACM Trans. Gr. (TOG) 22 (3) (2003) 313–318.
- [40] N. Birkbeck, J. Zhang, M. Requardt, B. Kiefer, P. Gall, S. Kevin Zhou, Region-specific hierarchical segmentation of MR prostate using discriminative learning, in: Proceedings of the Conference on MICCAI Grand Challenge: Prostate MR Image Segmentation, 2012.
- [41] Y. Ou, A. Sotiras, N. Paragios, C. Davatzikos, DRAMMS: deformable registration via attribute matching and mutual-saliency weighting, Med. Image Anal. 15 (4) (2011) 622–639.
- [42] N. Srivastava, G.E. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, Dropout: a simple way to prevent neural networks from overfitting, J. Mach. Learn. Res. 15 (1) (2014) 1929–1958.
- [43] C. Farabet, C. Couprie, L. Najman, Y. LeCun, Learning hierarchical features for scene labeling, IEEE Trans. Pattern Anal. Mach. Intell. 35 (8) (2013) 1915–1929.
- [44] ImageNet, (<http://www.image-net.org/>).
- [45] Keras VGG-19 model, (<https://keras.io/applications/#vgg19>).
- [46] T.G. Dietterich, et al., Ensemble methods in machine learning, Mult. Classif. Syst. 1857 (2000) 1–15.
- [47] J.O. Barentsz, J. Rickenberg, R. Clements, P. Choyke, S. Verma, G. Villeirs, O. Rouviere, V. Logager, J.J. Fütterer, ESUR prostate MR guidelines 2012, Eur. Radiol. 22 (4) (2012) 746–757.
- [48] X. Li, W. Huang, W.D. Rooney, Signal-to-noise ratio, contrast-to-noise ratio and pharmacokinetic modeling considerations in dynamic contrast-enhanced magnetic resonance imaging, Magn. Resonance Imaging 30 (9) (2012) 1313–1322.
- [49] K. Sung, B.L. Daniel, B.A. Hargreaves, Transmit B1+ field inhomogeneity and T1 estimation errors in breast DCE-MRI at 3 tesla, J. Magn. Resonance Imaging 38 (2) (2013) 454–459.
- [50] A. Tanimoto, J. Nakashima, H. Kohno, H. Shimoto, S. Kurabayashi, Prostate cancer screening: the clinical value of diffusion-weighted imaging and dynamic MR imaging in combination with T2-weighted imaging, J. Magn. Resonance Imaging 25 (1) (2007) 146–152.
- [51] L.G. Nyúl, J.K. Üdupa, et al., On standardizing the MR image intensity scale, Magn. Reson. Med. 42 (1999) 1072–1081.

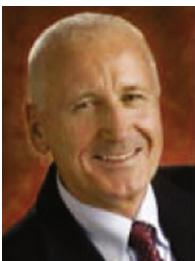
- [52] T. Heimann, B. Van Ginneken, M.A. Styner, Y. Arzhaeva, V. Aurich, C. Bauer, A. Beck, C. Becker, R. Beichel, G. Bekes, et al., Comparison and evaluation of methods for liver segmentation from CT datasets, *IEEE Trans. Medi. Imaging* 28 (8) (2009) 1251–1265.
- [53] S.S. Chandra, J.A. Dowling, K.-K. Shen, P. Raniga, J.P. Pluim, P.B. Greer, O. Salvado, J. Fripp, Patient specific prostate segmentation in 3-D magnetic resonance images, *IEEE Trans. Medi. Imaging* 31 (10) (2012) 1955–1964.
- [54] B. Maan, F. van der Heijden, Prostate MR image segmentation using 3D active appearance models, in: Proceedings of the Grand Challenges in Medical Image Analysis, PROMISE12, 2012, pp. 44–51.
- [55] G. Qinquan, et al., An automatic multi-atlas based prostate segmentation using local appearance-specific atlases and patch-based voxel weighting, in: Proceedings of the MICCAI-PROMISE, 2012.
- [56] M. Kirschner, F. Jung, S. Wesarg, Automatic prostate segmentation in MR images with a probabilistic active shape model, in: Proceedings of the MICCAI Grand Challenge: Prostate MR Image Segmentation, 2012.



Haozhe Jia received his B.E. in Computer Science from Northwestern Polytechnical University (NPU), Xi'an, China, in 2015. He is currently a Master student at the School of Computer Science and Engineering, NPU. His research areas include neuroimaging, computer-aided diagnosis and machine learning.



Weidong Cai received his Ph.D. Degree in Computer Science from the Basser Department of Computer Science, The University of Sydney, in 2001. He is currently an Associate Professor in the School of Information Technologies, Director of the Multimedia Laboratory in The University of Sydney. He has been a Lead Investigator/Visiting Professor on medical image analysis and medical computer vision at Surgical Planning Laboratory (SPL), Harvard Medical School during his 2014 SSP. His research interests include medical image analysis, image/video processing and retrieval, bioimaging informatics, computational neuroscience, computer vision & pattern recognition, and multimedia computing.



Michael Fulham graduated in 1979 (MBBS Hons) from The University of New South Wales. He trained in Internal Medicine and Neurology in Australia and then spent 5 years at the National Institutes of Health from 1988 to 1993 in the Neuroimaging Branch of NINDS. He returned to Australia in 1993. He is currently a full-time senior clinician and Director of the Department of PET and Nuclear Medicine at Royal Prince Alfred (RPA) Hospital; he is also the Clinical Director for Medical Imaging throughout the Sydney Health District. He has directed the first Australian PET program at RPA since 1993 and Medical Imaging Stream in the various iterations of Area Health Services (AHSs) and Local Health Districts (LHDs) since 2000. He is a Clinical Professor in the Sydney Medical School and Adjunct Professor in the Faculty of Engineering and Information Technologies at the University of Sydney. Throughout his career the major theme of his work has been to improve patient care through leadership, translational research and technological solutions in cancer and the neurosciences.



David Dagan Feng received the M.E. Degree in Electrical Engineering & Computer Science from Shanghai Jiao Tong University, Shanghai, China, in 1982, the M.Sc. Degree in Biocybernetics and the Ph.D. Degree in Computer Science from the University of California, Los Angeles, CA, USA, in 1985 and 1988, respectively, where he received the Crump Prize for Excellence in Medical Engineering. He is currently the Head of School of Information Technologies and the Director of the Institute of Biomedical Engineering and Technology, University of Sydney, Sydney, Australia, a Guest Professor of a number of Universities, and a Chair Professor of Hong Kong Polytechnic University, Hong Kong. He is a fellow of ACS, HKIE, IET, and the Australian Academy of Technological Sciences and Engineering.



Yong Xia received his B.E., M.E. and Ph.D. in Computer Science and Technology from Northwestern Polytechnical University (NPU), Xi'an, China, in 2001, 2004 and 2007, respectively. He is currently a Professor at the School of Computer Science and Engineering, NPU. He is also with the Shaanxi Key Lab of Speech & Image Information Processing (SAIIP) and Centre for Multidisciplinary Convergence Computing (CMCC) at the School. His research interests include medical image analysis, computer-aided diagnosis, pattern recognition, machine learning, and data mining.



Yang Song is currently an Australian Research Council (ARC) Discovery Early Career Researcher Award (DECRA) Fellow at the School of Information Technologies, the University of Sydney. She received her Ph.D. Degree in Computer Science from the University of Sydney, in 2013. Her research interests include biomedical imaging informatics, computer vision, and machine learning.