Fully automated segmentation of prostate whole gland and transition zone in diffusion-weighted MRI using convolutional neural networks

Tyler Clark
Junjie Zhang
Sameer Baig
Alexander Wong
Masoom A. Haider
Farzad Khalvati

Fully automated segmentation of prostate whole gland and transition zone in diffusion-weighted MRI using convolutional neural networks

Tyler Clark,^a Junjie Zhang,^a Sameer Baig,^a Alexander Wong,^b Masoom A. Haider,^a and Farzad Khalvati^{a,*}

^aUniversity of Toronto, Department of Medical Imaging, Sunnybrook Research Institute, Toronto, Canada

^bUniversity of Waterloo, Department of Systems Design Engineering, Waterloo, Canada

Abstract. Prostate cancer is a leading cause of cancer-related death among men. Multiparametric magnetic resonance imaging has become an essential part of the diagnostic evaluation of prostate cancer. The internationally accepted interpretation scheme (Pi-Rads v2) has different algorithms for scoring of the transition zone (TZ) and peripheral zone (PZ) of the prostate as tumors can appear different in these zones. Computer-aided detection tools have shown different performances in TZ and PZ and separating these zones for training and detection is essential. The TZ-PZ segmentation which requires the segmentation of prostate whole gland and TZ is typically done manually. We present a fully automatic algorithm for delineation of the prostate gland and TZ in diffusion-weighted imaging (DWI) via a stack of fully convolutional neural networks. The proposed algorithm first detects the slices that contain a portion of prostate gland within the three-dimensional DWI volume and then it segments the prostate gland and TZ automatically. The segmentation stage of the algorithm was applied to DWI images of 104 patients and median Dice similarity coefficients of 0.93 and 0.88 were achieved for the prostate gland and TZ, respectively. The detection of image slices with and without prostate gland had an average accuracy of 0.97. © 2017 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.4.4.041307]

Keywords: prostate cancer; diffusion-weighted MRI; prostate segmentation; convolutional networks.

Paper 17130SSR received May 8, 2017; accepted for publication Sep. 27, 2017; published online Oct. 17, 2017.

1 Introduction

Prostate diseases [e.g., prostate cancer, prostatitis, and benign prostate hyperplasia (BPH)] are common afflictions in men. In particular, prostate cancer is the most diagnosed cancer and the third leading cause of cancer death in Canadian men according to the Canadian Cancer Society.¹ Multiparametric MRI (mpMRI) has been widely adopted as a way of detecting and localizing prostate cancer.^{2,3} An internationally accepted scheme of scoring and interpretation has been adopted by radiologists for prostate mpMRI (Prostate Imaging - Reporting and Data System: Pi-Rads v2).⁴ In order to perform this interpretation, the radiologist must first divide the prostate into the anatomic regions of peripheral zone (PZ) and transition zone (TZ). Diffusion-weighted imaging (DWI) is one of the components of the mpMRI and is dominant in scoring of lesions in the PZ. In the TZ, T2-weighted imaging (T2w) is of more importance with DWI images having a secondary weighting. Despite this standardization, interobserver variability in scoring remains an issue.⁵ As a result and given the time-consuming and tedious nature of manual contouring of prostate (whole gland and TZ-PZ), there has been interest in developing computer-aided diagnostic approaches to Pi-Rads interpretation. An important step in this process would be the segmentation of prostate into PZ and TZ. Furthermore, accurate segmentation of prostate and related anatomic structures is an essential task for a number of clinical workflows including radiation treatment planning.

Computer-aided detection (CAD) algorithms proposed for automated detection of prostate cancer rely on segmentation of prostate gland as a preprocessing step.^{6–11} It has also been shown that the transition and PZs of prostate have different imaging features and hence, to improve the detection accuracy, CAD tools train separate classifiers for these two zones.¹² Consequently, there is a real need for developing automatic methods that offer robust and accurate segmentation of prostate and TZ-PZ in DWI.

Automatic segmentation of prostate in DWI images, however, is a challenging task.¹³ First, different MR images have global interscan variability and intrascan intensity variations due to different MR scan protocols and prostate signal inhomogeneity. Second, prostates exhibit large contrast variation in image response, due to heterogeneous anatomic structures of prostate and its surrounding tissue (e.g., blood vessels, bladder, rectum, and seminal vessels), which adds to the difficulty in the determination of prostate boundaries. Third, prostates have a wide range of morphological variation (e.g., shape, size, and volume) among different subjects. Moreover, segmentation of the TZ is more challenging as there is wide variation in TZ morphology related to benign prostatic hyperplasia (BPH), which is a part of the normal aging process in most men. We hypothesize that convolutional neural networks may be well suited to the task of TZ-PZ segmentation given that the TZ-PZ boundary is readily observable by radiologists with minimal training. To obtain TZ-PZ segmentation, one solution is to segment the whole gland and the TZ. This will give both the TZ and PZ boundaries. As will be discussed in this paper, our proposed

*Address all correspondence to: Farzad Khalvati, E-mail: farzad.khalvati@utoronto.ca

method automatically segments both the prostate whole gland and TZ.

1.1 Related Work

With the growing importance of prostate MRI segmentation, over the past few years, several segmentation methods have been proposed to meet the challenges, including registration-based methods,^{14–17} deformable models,^{18–21} and machine learning-based methods.^{10,11,22,23}

Registration-based methods involve registration of a reference image to the target image, and applying the corresponding transformation to the manually segmented mask of the reference to produce the final segmentation. Atlas-based models, belonging to registration-based methods, represent a class of popular approaches in the literature for prostate segmentation. For example, Klein et al.¹⁴ and Langerak et al.²⁴ reported similar multi-atlas-based methods, in which several template images with corresponding segmentation results (or labels) are registered to the target image using a nonrigid registration method, and then the deformed labels are fused using various voting techniques to generate a consensus label of prostate segmentation. Deformable models, such as levelsets¹⁹ and active appearance models (AAMs),²¹ use prior geometric and regional statistics information, such as smoothness and curvature of the boundary, mean, and standard deviation of regions, in an energy minimization or correlation maximization framework to achieve segmentation. Another category of approaches proposed for prostate segmentation uses feature-based machine learning methods, such as Gaussian mixture model,¹¹ random forest classifier,²³ and marginal spacing learning,¹⁰ to cluster or classify the image into prostate and background regions. Hybrid methods are also common by taking advantage of different models to obtain improved performance. For instance, Martin et al.²⁵ proposed a two-stage procedure for prostate segmentation combining a probabilistic atlas model and a deformable model. Ghose et al.²² and Zhang et al.¹¹ used different machine learning approaches to first produce a probabilistic segmentation of the prostate that was then further refined with a levelset to obtain final segmentation.

Recently, deep convolutional networks (ConvNets) have led to a series of breakthroughs in computer vision and have achieved promising results in different vision tasks such as classification, segmentation, and object detection.^{26–28} With powerful feature representation capability, ConvNets learn a hierarchical level of abstraction of an image through the network layers, in which the shallow layer grasps low-level local features while deeper layers whose receptive fields are much boarder capture high-level global information of the image.²⁹

Early application of ConvNets has been mostly on classification tasks where the output of an image is a single class label.^{26,30,31} A fully connected layer is placed at the end of the networks to convert the convoluted features to the probabilities of certain labels. However, in some visual tasks such as segmentation, the label of each pixel is most relevant. Long et al.²⁷ proposed fully convolutional networks (FCN) trained end-to-end to implement pixelwise segmentation, in which the networks learn to combine coarse, high layer information with fine, low layer information to make dense predictions at the size of the input image. The architecture of FCN has been modified and extended to meet the challenges in biomedical image analysis. Ronneberger et al.³² proposed U-Net for cell segmentation in microscopy images, which features a symmetric

architecture by supplementing an expansive path to the usual contracting network in FCNs. High-resolution features in the contracting path are combined with the corresponding upsampling phase to propagate context information for better prediction. With the impressive performance achieved by deep learning, some studies started to employ ConvNets for automatic prostate segmentation in MRIs.^{33–36} Milletari et al.³⁴ extended U-Net to three-dimensions (3-D) to make volumetric segmentation of prostate. Yu et al.³⁵ proposed a volumetric ConvNets with mixed residual connections for automatic prostate segmentation from 3-D MR images. This method ranked first in the open MICCAI PROMISE12 challenge,³⁷ outperforming other methods by a large margin. Clark et al.³⁶ applied a tailored FCN to segment prostate gland using DWI data.

1.2 Proposed Algorithm

Most of the related work has been on segmenting the prostate in T2-weighted MRI (T2w). However, it is of great relevance to perform segmentation on other MR modalities such as DWI. In this work, we significantly extend upon the work in Clark et al.³⁶ (which was for prostate gland only) and propose an architecture with several ConvNets in its core to segment the prostate gland, as well as the TZ, in DWI data. This adds to the difficulty of the task, considering the already reduced spatial resolution and contrast of tissues on DWI compared to T2w. Given its anterior location, the TZ is known to have a heterogeneous appearance but accurate interpretation of the TZ holds great value in the diagnosis of BPH as well as TZ cancer.³⁸ To our best knowledge, only few preliminary studies have reported on automatic segmentation of TZ using T2w images,^{39–41} and one on semi-automatic TZ segmentation using DWI.¹⁷ Another extension of our work compared to the one in Ref. 36 is that we implement the architecture on mpMRI (several MRI modalities as opposed to single DWI modality), which can be regarded as a two-dimensional (2-D) image with multiple channels comprised of different MR modalities and their derivatives. The testing data used for experiments in this paper are much larger compared to that in Ref. 36 (104 versus 30 patients). Moreover, the architecture proposed in this paper is different than that in Ref. 36 in the following aspects.

Although max pooling has been shown to be effective for classification convolutional neural networks,³⁰ when it is used to reduce the dimensions of the feature map, it creates a loss of information. Specifically, the position of the neuron in the pools receptive field is lost. This creates a small amount of translation invariance that is compounded through the layers. When pooling is used, it also creates an asymmetry between the downsampling (pooling) and upsampling methods. To increase the symmetry as well as preserve features detail information, strided convolutions were used on the downsampling path; and fractionally strided convolutions (also known as deconvolutions) were used for upsampling. In particular, this increased the Dice similarity coefficient (DSC) of the segmentation of the smaller regions of prostate.

Another important factor is the extent to which the spatial resolution of the feature map must be reduced. Given the lower resolution of DWI images, the input images were only reduced three times compared to four from the original paper.³² The number of filters, number of convolutions, and kernel size were also modified to address the requirements of the problem at hand and achieve better results. Residual connections were added to the connections between up- and downsampling

paths. This was especially important in the first layer skip connection which had only undergone one convolution operation. This helped to reduce areas of false positive.

The task of segmenting prostate (and TZ) in 2-D MRI slice can actually be divided into two consecutive cognitive procedures: detection (or classification) and segmentation. In our proposed architecture, we incorporate two components to reflect the two procedures: the detection component determines whether there is prostate (or TZ) in the image, which consists of a standard VGG³⁰ ConvNet designed for classification; if detected, it proceeds to the segmentation component, in which a deep ConvNet based on U-Net was adopted, modified, and configured to perform the segmentation task. The detection of the presence of the prostate gland is another novelty of our proposed algorithm; in almost all previous work on prostate segmentation, it is assumed that the given 3-D image starts and ends with portions of prostate gland. Nevertheless, in real-world scenarios, the 3-D images contain several slices before (after) prostate gland starts (ends). The combination of detection-segmentation makes the proposed algorithm a fully automated method with potentially high impact in the clinical workflow. This is another difference between this work and the algorithm proposed in Ref. 36 where in the latter, the user has to first determine the presence of the prostate gland in a given slice before applying the segmentation algorithm.

The proposed ConvNet segmentation method copes with limited training data, commonly encountered in medical image analysis, by harnessing data augmentation methods as proposed in Krizhevsky et al.²⁶ to expand the size of samples. This alleviates overfitting and limited representation caused by small data sets and encourages the learning of invariance to spatial deformations and intensity shifting of images. Spatial dropout was also implemented to reduce overfitting. We further replace the regular convolution blocks in U-Net with inception reduction blocks to reduce computational cost and improve performance.⁴² Finally, we define the objective function as DSC calculated by comparing the predicted segmentation result with the ground truth.

2 Methods

In this section, the data used for the experiments, and the proposed deep convolutional neural networks used for classification and segmentation are presented.

2.1 Imaging Data and Preprocessing

In this retrospective study, mpMRI images of 104 patients were acquired using a Philips Achieva 3.0T MRI at Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada. Institutional research ethics board approval was granted and written informed consent was waived. Each patient had a 3-D volume divided into 12 to 34 slices for which the prostate was manually segmented by a radiologist who delineated the whole gland and TZ-PZ boundary for every slice on b_0 images of DWI ($b = 0 \text{ s/mm}^2$). On average, 10 slices or 41% of the MRI volume did not contain the prostate. For each slice location through the prostate, the following pulse sequence images or their derivatives were used: b_0 , b_{100} , b_{400} (or b_{600}), apparent diffusion coefficient (ADC) map, computational b_{2000} , T2w, and b_0 gradient. The b_0 gradient was created using a Sobel edge detection filter from the b_0 image.

2.1.1 Data augmentation

To increase data size, we used different data augmentation methods which include:

- The image is flipped on both X and Y axes.
- Random rotation: The image is rotated randomly between 0 and 7 deg.
- Random channel shift: Channel shifting is the process of taking pixels in an image and applying those values to pixels in different positions on the image.
- Image is randomly shifted vertically and horizontally between $\pm[0, 4]$ pixels.
- Elastic deformation as described in Ref. 43 is performed by applying random affine displacement fields to images.

There were total of 1498 slices out of which 1248 were used for training in each cross-validation. Data augmentation increased the number of training slices to 16,224 (data size increase of 13 $\times$) while the test set remained unchanged.

2.2 Architecture Overview

The architecture of the proposed approach can be decomposed into four stages, which are shown in Fig. 1. The first stage contains a classification ConvNet (Fig. 2) that detects the presence of the prostate gland. The second stage of the approach performs segmentation of the entire prostate with another convolutional network (Fig. 3). The third stage classifies the image (same architecture as Fig. 2) to determine whether there is TZ contained within the prostate. Then, if it is detected, the TZ is also segmented using the network of the same structure as in Fig. 3.

2.3 Classification Architecture

The structure of the classification ConvNet (Fig. 2) was modeled based on a recent network design known as a VGG³⁰ network. The hyperparameters of A and B blocks can be found in Tables 1 and 2, respectively; where N is the number of filters in the layer shown in Fig. 2. The stacked A blocks learn filters then use zero padding to maintain feature tensor dimensionality while increasing feature complexity with 3×3 convolutions. The B blocks contain three fully connected layers that have connections to all neurons in the previous layer. The last fully connected layer has a binary output that produces the classification results (whether or not prostate is present in the 2-D image slice).

Fig. 1 Algorithm overview.

Fig. 2 Classification architecture.

Fig. 3 Segmentation architecture.

2.4 Segmentation Architecture

2.4.1 U-Net

The architecture for the segmentation of both prostate gland and TZ is based on the U-Net architecture³² with modifications tailored for the problem at hand as described below. A diagram of the structure is shown in Fig. 3.

2.4.2 Downblocks

The segmentation architecture is composed of a series of four “downblocks.” Downblocks have an inception block (Fig. 4),⁴² followed by a convolution of stride 2 that reduces the dimensionality. Inception blocks apply four convolution and pooling operations in parallel then concatenate the feature tensors at the end of the block. Merging of signals after parallel operations has been shown theoretically and experimentally to increase segmentation and classification accuracy.⁴² Table 3 contains all of the hyperparameters. N is the layer number shown in Fig. 3.

2.4.3 Upblocks

The “upblocks” concatenate the convolutional features of their reciprocal “downblock” counterpart in a similar fashion to the U-Net.³² To increase feature map dimensions, deconvolution blocks were tested and found to be more effective than uppooling and were therefore used. A deconvolution (transposed convolution or fractionally strided convolution) is a convolution operation that increases the spatial feature dimensions.⁴⁴ Table 4 contains all of the hyperparameters.

2.4.4 Binarization layer

The binarization layer uses a $1 \times 1 \times 1$ convolution with a sigmoid activation.

The optimizer used for gradient descent was Adam.⁴⁵ The loss function used was the DSC. The DSC essentially measures the overlap between the predicted mask and ground truth.

$$\text{DSC} = 2 \frac{|P \cap G|}{|P| + |G|}, \quad (1)$$

which can be expressed for the k 'th pixel.

Table 1 A-block.

Layer	Filter	Stride	Dropout	Activation	Batch normalization
Zero padding	1×1	n/a	None	n/a	n/a
Convolution block	$3 \times 3 \times N$	1	None	relu	No
Zero padding	1×1	n/a	None	n/a	n/a
Convolution block	$3 \times 3 \times N$	1	None	relu	No
Zero padding	1×1	n/a	None	n/a	n/a
Convolution block	$3 \times 3 \times N$	1	None	relu	No
Convolution	$2 \times 2 \times N$	2×2	None	relu	No

Table 2 B-block.

Layer	Filter	Stride	Dropout	Activation	Batch normalization
Dense	4096	N/A	0.6	relu	None
Dense	4096	N/A	0.6	relu	None
Dense	2	N/A	None	Softmax	None

$$DSC = 2 \frac{\sum P_k G_k}{\sum P_k + \sum G_k}, \quad (2)$$

where P_k is the k 'th pixel's predicted mask, G_k is the ground truth mask, and both masks are binary.

2.5 Computer System

The processor used for the proposed algorithm was an Nvidia GEFORCE GTX-980m GPU with Cuda parallel processing. The language used was Python with a Keras wrapper with a Theano back-end.⁴⁶ In total, there were 80 iterations for combined algorithms, which translate to 600 training hours.

3 Results

Analysis of the algorithm was done using fourfold cross-validation.

3.1 Classification

The results of the classification algorithm that delineated the prostate from nonprostate slices for the prostate whole gland and TZ are presented in Tables 5 and 6, respectively. The average accuracy for the prostate gland detection was 97.1% with an average specificity and sensitivity of 94.1% and 99.4%,

Fig. 4 Residual and inception blocks.**Table 3** Downblock.

Layer	Filter	Stride	Dropout	Activation	Batch normalization
Inception convolution block	$32 \times 2^{N-1}$		none	relu	Yes
Convolution	$3 \times 3 \times (32 \times 2^{N-1})$	2×2	0.6	relu	Yes

Table 4 Upblock.

Layer	Filter	Stride	Dropout	Activation	Batch normalization
Deconvolution block	$3 \times 3 \times (32 \times 2^{N-1})$	2×2	None	relu	Yes
Convolution block	$1 \times 1 \times (32 \times 2^{N-1})$	1×1	None	relu	No
Residual convolution block	$32 \times 2^{N-1}$		None	relu	Yes
Merge block	N/A	N/A	0.6	None	No
Inception convolution block	$32 \times 2^{N-1}$		None	relu	Yes

Table 5 Prostate gland classification results.

Fold	Sensitivity	Specificity	Accuracy
1	1	0.953	0.979
2	0.987	0.943	0.966
3	0.992	0.938	0.970
4	0.996	0.931	0.970

Table 6 TZ classification results.

Fold	Sensitivity	Specificity	Accuracy
1	1	0.859	0.938
2	1	0.862	0.938
3	0.995	0.826	0.901
4	1	0.854	0.928

respectively. The average accuracy for the TZ detection was 92.6% with an average specificity and sensitivity of 85.0% and 99.9%, respectively.

3.2 Segmentation

3.2.1 Prostate whole gland

The results of the cross-validation are presented in Table 7. The average mean and median DSC over the fourfolds are 0.886 and

Table 7 Prostate gland segmentation accuracy (DSC).

Fold	Mean	Median
1	0.872	0.920
2	0.894	0.928
3	0.889	0.934
4	0.889	0.937

Fig. 5 Fourth fold cross-validation results (DSC) for prostate whole gland segmentation.

0.930, respectively. The histogram of the fourth fold of DSC results is shown in Fig. 5. Two example segmentation results of the prostate gland are shown in Fig. 6 where the green line is the ground-truth mask outline and blue line is the automated segmentation mask outline.

3.2.2 Transition zone

The results of the cross-validation for TZ segmentation can be found in Table 8. The average mean and median DSC over the fourfolds are 0.847 and 0.885, respectively. The histogram of the fourth fold of cross-validation of DSC results is shown in Fig. 7. Two example segmentation results for TZ are shown in Fig. 8 where the green line is the ground-truth mask outline and blue line is the automated segmentation mask outline.

3.3 Validation on External Data

Although the proposed algorithm in this paper was designed for segmentation of the prostate whole gland and TZ in DWI, it was validated by application to an external dataset of T2w images of the whole prostate gland for 50 patients from the 2012 MICCAI PROMISE competition.³⁷ The data were picked to be from multiple centers and multiple MRI device vendors as well as patients with and without an endorectal coil. The T2w data were of much greater resolution in comparison to our DWI data, and due to the sensitivity of our proposed segmentation algorithm to input image sizes, all T2w volumes were down-sampled to $144 \times 144 \times 26$. We divided the 50 patients' data

Fig. 6 Prostate whole gland segmentation example. Blue: segmentation result; Green: ground truth.

Table 8 TZ segmentation accuracy (DSC).

Fold	Mean	Median
1	0.842	0.881
2	0.847	0.876
3	0.846	0.889
4	0.854	0.894

Fig. 7 Fourth fold cross-validation results for TZ segmentation.

into a training and validation set of 75% and 25%, respectively, which amounted to an average of 37 patients with an average of 479 slices containing the prostate in the training set, and 12 patients with 152 slices for the test set for each fold of validation. We also included all of our data of 104 patients into the training set, which totaled 141 patients per training set. The segmentation accuracy for this external T2w dataset for prostate gland was mean of 0.862 ± 0.008 and a median of 0.893 ± 0.018 .

4 Discussion

In this study, we have proposed a generic architecture based on convolutional neural networks to tackle the prostate

segmentation problem in MRI, which is similar to human’s cognitive procedure: first, it determines the presence of the object of interest in the image, next it proceeds to delineation of the object. The former is global perception, in which relative coarse and broad receptive fields are required; the latter is more visual intensive, involving interpretation of details in local context. Two ConvNets, as high-level abstraction of the cognitive procedure, are incorporated in the architecture in a consecutive manner to perform the segmentation task.

This architecture brings several advantages to our method: automatically detecting the presence of prostate in MRI images increases the approach’s efficiency and clinical applicability. This is due to the fact that prostate MRI volumes have a significant number of slices that do not contain the gland. The detection component in the first stage eliminates the need for the radiologist’s intervention to locate slices where the prostate resides, which makes the algorithm fully automatic.

As was discussed in Sec. 1.2, the proposed algorithm was based on a significant modification of U-Net algorithm.³² While our proposed algorithm achieved a mean DSC of 0.886 for segmentation of prostate gland in DWI, the conventional/ original U-Net proposed in Ref. 32 applied to the same data only yielded a mean DSC of 0.705 for the prostate whole gland. This is an indication of the significant difference between the original algorithm and the one proposed in this paper (mean DSC of 0.705 versus 0.886) and the fact that the latter has been designed to address the shortcomings of the previous one.

One limitation of using the DSC as a loss function and metric is that it heavily penalizes small errors on the smaller regions of the prostate (e.g., a segmentation that is 30 pixels and is 3 pixels off will receive a DSC of 0.90). The majority of slices with a DSC below 0.75 were below the mean prostate surface area. In total, only 27 slices were below this DSC threshold. In Fig. 9, the blue plot is a histogram of the prostate surface area per slice and the green is the histogram of segmentation results below a DSC cutoff of 0.75. As discussed, the DSC metric strongly penalizes small errors for small segmentations; and expert variability also increases in these cases where the prostate cross section is small in 2-D slices. The lack of reliability analysis with respect to interobserver variability is another limitation of this work. In a future work, we will use the manual segmentation of three readers to test the reliability of the algorithm with respect to interobserver variability.

Fig. 8 TZ segmentation example. Blue: segmentation results; Green: ground truth.

Fig. 9 Distribution of prostate surface size (# of pixels) (blue) and the number of cases with segmentation accuracy (DSC) below 0.75 threshold (green).

Several works in the literature on prostate segmentation are based on the implication that the location of the prostate has been determined without specifying the detection methods used, whether manually or not. Kirschner et al.¹⁸ extended Viola-Jones object detected algorithm, which was originally developed for face detection, to identify a prostate region on the slice per segmentation; however, the detection accuracy was not reported explicitly. The ConvNet in our algorithm for prostate detection achieved an accuracy of 0.97. It was also found that performing prostate detection in the first step resulted in roughly an average 3% increase in DSC score, compared with only applying segmentation to the MR images where in some cases images with no prostate would be segmented and compared to ground truth resulting in DSC of 0.

For the prostate segmentation task, conventional approaches often rely on handcrafted features, empirically designed image descriptors or functions, suffering from limited representation and learning abilities.⁴⁷ Atlas-based approaches, as the most reported registration-based method, may be affected by serious errors when the processed prostate instances are dissimilar to the atlas, despite the nonrigid registration. Some other registration-based methods may require intensive user interaction and image registration algorithms. Several prostate contours¹⁵ or bounding boxes¹⁷ need to be drawn on key slices for the algorithm to work effectively.

Deformable models, such as levelsets,¹⁹ AAMs,^{20,21} and active shape models,¹⁸ usually require good initialization of contours to produce plausible results.²⁰ To solve the optimal contour in an energy minimization framework, these models often involve incorporating various energy terms. There is no solid standard in designing the energy terms and different studies propose their own sets of energy terms, not to mention the empirical setting of weights for combining those energy terms and parameters within the energy terms.^{18–21}

Machine learning-based methods perform segmentation by clustering or classifying voxels belonging to the prostate based on feature vectors extracted from mpMRI data.^{22,23} Thus, they can be sensitive to noise and intensity variations within the gland and neighboring tissues. This is why postprocessing is usually necessary to refine the results produced in the first step, which could be a deformable model²² or a probabilistic model²³ that enable inclusion of contextual information or a shape prior such as Markov random fields to improve the segmentation. Nevertheless, the hybrid approaches still could not eliminate some of the inherent drawbacks of individual methods.

Conventional neural networks, on the other hand, hold great potential of being the next breakthrough in medical image analysis. Two pure convolutional neural networks have been reported for volumetric prostate segmentation considering the 3-D nature of prostate MR images. Milletari et al.³⁴ extended the U-Net architecture from Ronneberger et al.³² by replacing 2-D operations with their 3-D counterparts and used a residual function after convolutional blocks to improve the learning. Yu et al.³⁵ proposed a learning architecture with mixed long and short residual connections for automated prostate segmentation from 3-D MR images. This model now leads the PROMISE12 challenge³⁷ and outperforms other methods by a large margin.

Highlighted by the success of convolutional neural networks in many visual recognition tasks including prostate segmentation, a 2-D ConvNet extended from U-Net is adopted in the second stage of our architecture, considering segmentation of prostate on 2-D MR image slices is still of great relevance. First, 2-D ConvNet can be applied to arbitrary stand-alone images or directly to slices of interest, whereas 3-D ConvNet needs to be applied to the full volumetric data. In scenarios where only key images are available because of limited bandwidth or other reasons, 2-D ConvNet is more flexible and easier to be deployed to provide the segmentation function. Second, due to high-volume data in computation and constraint of

limited GPU memory, 3-D ConvNets need to resort to tiling strategy and run on overlapped 3-D patches to perform the segmentation on volumetric MRI data,³⁵ which raises efficiency issues. In our proposed architecture, the detection component in the first stage removes irrelevant slices, and the segmentation then focuses only on slices where prostate is present, greatly reducing the computation cost.

Although a direct comparison with state-of-the-art methods would be unfeasible given the variability of data quality, size, and types of MRI modalities used, it could be certainly valuable to attempt a rough comparison. As the top entry in PROMISE12 challenge, the volumetric ConvNet proposed by Yu et al.³⁵ was trained on 50 monoparametric T2w image series and tested on a separate test set of 30 cases. We had a relatively larger data set of 104 mpMRI series and evaluated our model with fourfold cross-validation. Our mean DSC of the prostate segmentation for DWI is 0.886, comparable and very close to 0.894 achieved in Yu et al.³⁵ For PROMISE12 T2w data, our algorithm produced mean DSC of 0.862 (median of 0.893). The drop in the DSC compared to that of Yu et al.³⁵ is due to the fact that we had to downsample T2w data to be compatible with our segmentation algorithm. Further research is needed to extend our 2-D ConvNet to 3-D and investigate how and to what extent 3-D spatial information encoded in 3-D ConvNet can possibly improve the segmentation accuracy. The 3-D U-Net by Milletari et al.,³⁴ which is another entry in PROMISE12, performs similar to our 2-D ConvNet (DSC of 0.869 versus 0.862) for PROMISE12 T2w data.

The utilization of mpMRI data in our model relates to another aim of our study, which is to segment the TZ of the prostate; a more complex structure to segment compared to the prostate whole gland. More modalities, such as ADC map, contain additional information on zonal distribution within the prostate to better differentiate the transition and PZs.⁴⁷ To our best knowledge, this is the first study using a ConvNet for the segmentation of prostate TZ in MRI data. The mean DSC of the TZ segmentation of our model is 0.847 (median of 0.885), which is much higher than the most recent report of 0.79 by Toth et al.⁴¹ using a multiple-levelset AAM method. The algorithm proposed in Ref. 41 was applied to T2w images, in which as discussed before, the boundary between TZ-PZ is much clearer than that in DWI. Nevertheless, our proposed algorithm outperformed the one in Ref. 41 with a large margin (0.847 versus 0.79).

Fully automatic segmentation of prostate gland and TZ with high accuracy is an important step in reducing interuser variability in the localization of prostate cancer. It can also significantly increase the reliability of CAD algorithms. We will use the proposed segmentation algorithm as part of a CAD system to create the Pi-Rads zonemap automatically and then validate whether this reduces the interuser variability among clinicians when interpreting prostate mpMRI images. We will also increase the training data size (e.g., 500) to investigate the upperbound of segmentation accuracy with respect to training data size. Finally, we will further extend our 2-D ConvNet to 3-D to investigate whether segmentation accuracy for the whole prostate gland as well as TZ improves.

5 Conclusions

We presented an approach based on multiple convolutional neural networks that perform automated classification and segmentation of MRI prostate images in DWI. Our proposed algorithm

is fully automated, meaning that slices containing prostate gland are automatically detected, with no need for user intervention to find the prostate gland. In addition to prostate gland segmentation, we applied our segmentation algorithm to the prostate's TZ, which is a more challenging and a necessary task for classification congruent with Pi-Rads v2 interpretation used in clinical practice. Future works will aim at using this algorithm as a preprocessing step for CAD algorithms for prostate cancer. We will also validate whether automatic generation of Pi-Rads zone-map can reduce interuser variability in localization of prostate cancer.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

References

1. Canadian Cancer Society's Advisory Committee on Cancer Statistics, *Canadian Cancer Statistics 2016*, <http://www.cancer.ca/> (10 October 2017).
2. C. M. Hoeks et al., "Prostate cancer: multiparametric MR imaging for detection, localization, and staging," *Radiology* **261**(1), 46–66 (2011).
3. D. J. Lee et al., "Multiparametric magnetic resonance imaging in the management and diagnosis of prostate cancer: current applications and strategies," *Curr. Urol. Rep.* **15**(3), 390 (2014).
4. J. O. Barentsz et al., "ESUR prostate MR guidelines 2012," *Eur. Radiol.* **22**(4), 746–757 (2012).
5. B. Muller et al., "Prostate cancer: interobserver agreement and accuracy with the revised prostate imaging reporting and data system at multiparametric MR imaging," *Radiology* **277**(3), 741–750 (2015).
6. S. Wang et al., "Computer aided-diagnosis of prostate cancer on multiparametric MRI: a technical review of current research," *BioMed Res. Int.* **2014**, 789561 (2014).
7. G. Litjens et al., "Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge," *Med. Image Anal.* **18**, 359–373 (2014).
8. F. Khalvati et al., "Automated prostate cancer detection via comprehensive multi-parametric magnetic resonance imaging texture feature models," *BMC Med. Imaging* **15**(1), 27 (2015).
9. A. Cameron et al., "MAPS: a quantitative radiomics approach for prostate cancer detection," *IEEE Trans. Biomed. Eng.* **63**(6), 1145–1156 (2016).
10. Y. Zheng and D. Comaniciu, *Marginal Space Learning for Medical Image Analysis*, Springer, New York (2014).
11. J. Zhang et al., "Segmentation of prostate in diffusion MR images via clustering," in *Machine Learning for Medical Image Computing, 14th Int. Conf. on Image Analysis and Recognition (ICIAR)*, Montreal, Québec, Canada, pp. 1–8 (2017).
12. F. Khalvati et al., "Flipping the computer aided diagnosis (CAD) training paradigm for prostate cancer: using pirads reporting of multiparametric MRI (mpMRI) to train a cad system and then validating with pathology," in *Imaging Network Ontario (ImNO) Symp.* (2017).
13. D. Mahapatra and J. M. Buhmann, "Prostate MRI segmentation using learned semantic knowledge and graph cuts," *IEEE Trans. Biomed. Eng.* **61**(3), 756–764 (2014).
14. S. Klein et al., "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information," *Med. Phys.* **35**(4), 1407–1417 (2008).
15. F. Khalvati et al., "Inter-slice bidirectional registration-based segmentation of the prostate gland in MR and CT image sequences," *Med. Phys.* **40**, 123503 (2013).
16. F. Khalvati et al., "Sequential registration-based segmentation of the prostate gland in MR image volumes," *J. Digital Imaging* **29**(2), 254–263 (2016).
17. J. Zhang et al., "A local ROI-specific atlas-based segmentation of prostate gland and transitional zone in diffusion MRI," *J. Comput. Vision Imaging Syst.* **2**(1) (2016).

18. M. Kirschner, F. Jung, and S. Wesarg, "Automatic prostate segmentation in MR images with a probabilistic active shape model," in *PROMISE12—MICCAI 2012 Grand Challenge on Prostate MR Image Segmentation* (2012).
19. X. Liu et al., "Unsupervised segmentation of the prostate using MR images based on level set with a shape prior," in *Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pp. 3613–3616, IEEE (2009).
20. G. Vincent, G. Guillard, and M. Bowes, "Fully automatic segmentation of the prostate using active appearance models," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI '12)* (2012).
21. R. Toth and A. Madabhushi, "Multifeature landmark-free active appearance models: application to prostate MRI segmentation," *IEEE Trans. Med. Imaging* **31**, 1638–1650 (2012).
22. S. Ghose et al., "A random forest based classification approach to prostate segmentation in MRI," in *MICCAI Grand Challenge: Prostate MR Image Segmentation*, pp. 20–27 (2012).
23. E. Moschidis and J. Graham, "Automatic differential segmentation of the prostate in 3-D MRI using random forest classification and graph-cuts optimization," in *9th IEEE Int. Symp. on Biomedical Imaging (ISBI '12)*, pp. 1727–1730, IEEE (2012).
24. T. R. Langerak et al., "Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (SIMPLE)," *IEEE Trans. Med. Imaging* **29**(12), 2000–2008 (2010).
25. S. Martin, J. Troccaz, and V. Daanen, "Automated segmentation of the prostate in 3D MR images using a probabilistic atlas and a spatially constrained deformable model," *Med. Phys.* **37**(4), 1579–1590 (2010).
26. A. Krizhevsky, G. E. Hinton, and I. Sutskever, "ImageNet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, pp. 1–9 (2012).
27. J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '14)*, pp. 3431–3440 (2014).
28. K. He et al., "Deep residual learning for image recognition," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778 (2016).
29. M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European Conf. on Computer Vision*, Vol. **8689**, pp. 818–833 (2014).
30. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Int. Con. on Learning Representations (ICLR)* (2014).
31. C. Szegedy et al., "Going deeper with convolutions," in *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR '15)*, pp. 1–9, IEEE (2015).
32. O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI '15)*, pp. 234–241 (2015).
33. R. Cheng et al., "Active appearance model and deep learning for more accurate prostate segmentation on MRI," *Proc. SPIE* **9784**, 97842I (2016).
34. F. Milletari, N. Navab, and S.-A. Ahmadi, "V-net: fully convolutional neural networks for volumetric medical image segmentation," in *Fourth Int. Conf. on 3D Vision (3DV)*, pp. 565–571, IEEE (2016).
35. L. Yu et al., "Volumetric convnets with mixed residual connections for automated prostate segmentation from 3D MR images," in *Proc. of the Thirty-First AAAI Conf. on Artificial Intelligence*, 4–9 February 2017, San Francisco, California, pp. 66–72 (2017).
36. T. Clark et al., "Fully deep convolutional neural networks for segmentation of the prostate gland in diffusion-weighted MR images," in *Machine Learning for Medical Image Computing, 14th Int. Conf. on Image Analysis and Recognition (ICIAR)*, pp. 1–8, Montreal, Québec, Canada (2017).
37. MICCAI Grand Challenge: Prostate MR Image Segmentation 2012, "PROMISE12," <https://promise12.grand-challenge.org/> (2012).
38. S. I. Jung et al., "Transition zone prostate cancer: incremental value of diffusion-weighted endorectal MR imaging in tumor detection and assessment of aggressiveness," *Radiology* **269**(2), 493–503 (2013).
39. N. Makni et al., "Zonal segmentation of prostate using multispectral magnetic resonance images," *Med. Phys.* **38**, 6093–6105 (2011).
40. G. Litjens et al., "A pattern recognition approach to zonal segmentation of the prostate on MRI," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI '12)*, Vol. **15**, No. 2, pp. 413–420, Springer, Berlin, Heidelberg (2012).
41. R. Toth et al., "Simultaneous segmentation of prostatic zones using active appearance models with multiple coupled levelsets," *Comput. Vision Image Understanding* **117**, 1051–1060 (2013).
42. C. Szegedy, S. Ioffe, and V. Vanhoucke, "Inception-v4, Inception-ResNet and the impact of residual connections on learning," in *Advancement of Artificial Intelligence (AAI)*, pp. 4278–4284 (2017).
43. P. Simard, D. Steinkraus, and J. Platt, "Best practices for convolutional neural networks applied to visual document analysis," in *Proc. Seventh Int. Conf. on Document Analysis and Recognition* (2003).
44. V. Dumoulin and F. Visin, "Convolution arithmetic," https://github.com/vdumoulin/conv_arithmetic (2016).
45. D. P. Kingma and J. L. Ba, "Adam: a method for stochastic optimization," in *Int. Conf. on Learning Representations* (2015).
46. F. Chollet, "Keras," 2015, <https://github.com/fchollet/keras> (10 October 2017).
47. L. Rundo et al., "Automated prostate gland segmentation based on an unsupervised fuzzy C-means clustering technique using multispectral T1w and T2w MR imaging," *Information* **8**(2), 49 (2017).

Biographies for the authors are not available.