A novel registration-based algorithm for prostate segmentation via the combination of SSM and CNN

Chunxia Qin, Puxun Tu, Xiaojun Chen, Jocelyne Troccaz

A novel registration-based algorithm for prostate segmentation via the combination of SSM and CNN

Chunxia Qin[a,b]; Puxun Tu[a]; Xiaojun Chen[a]; Jocelyne Troccaz[c]
a. School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China;
b. School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China;
c. Univ. Grenoble Alpes, CNRS, Grenoble INP, TIMC, Grenoble, France

April 16, 2022

Chunxia Qin and Puxun Tu contributed equally to this work.

Corresponding author: Xiaojun Chen. email: xiaojunchen@sjtu.edu.cn.

Abstract

Purpose: Precise determination of target is an essential procedure in prostate interventions, such as prostate biopsy, lesion detection, and targeted therapy. However, the prostate delineation may be tough in some cases due to tissue ambiguity or lack of partial anatomical boundary. In this study, we proposed a novel supervised registration-based algorithm for precise prostate segmentation, which combine the convolutional neural network (CNN) with a statistical shape model (SSM).
Methods: The proposed network mainly consists of two branches. One called SSM-Net branch was exploited to predict the shape transform matrix, shape control parameters, and shape fine-tuning vector, for the generation of the prostate boundary. Furtherly, according to the inferred boundary, a normalized distance map was calculated as the output of SSM-Net. Another branch named ResU-Net was employed to predict a probability label map from the input images at the same time. Integrating the output of these two branches, the optimal weighted sum of the distance map and the probability map was regarded as the prostate segmentation.
Results: Two public datasets PROMISE12 and NCI-ISBI 2013 were utilized to evaluate the performance of the proposed algorithm. The results demonstrate that the segmentation algorithm achieved the best performance with an SSM of 9500 nodes, which obtained a dice of 0.907 and an average surface distance of 1.85 mm. Compared with other methods, our algorithm delineates the prostate region more accurately and efficiently. In addition, we verified the impact of model elasticity augmentation and the fine-tuning item on the network segmentation capability. As a result, both factors have improved the delineation accuracy, with dice increased by 10% and 7% respectively.
Conclusions: Our segmentation method has the potential to be an effective and robust approach for prostate segmentation.

Keywords: registration-based segmentation, statistical shape mode, probability map, boundary distance map

Contents

I. Introduction

troduction

With Magnetic Resonance (MR) imaging becoming an increasingly important non-invasive imaging modality[b1], prostate MR segmentation has been paid close attention in recent years, as it is crucial for the clinical diagnosis, therapeutic procedure, and treatment planning of various prostate disorders (e.g., prostate cancer, prostatitis or prostatic hypertrophy)[b2]. For example, prostate delineation is widely applied for the precise localization of prostate boundary in radiotherapy for dose planning. Besides, in the image-guided computer-assisted surgery, the segmentation of the prostate on preoperative MRI is an essential reference for the inter-operative low-quality image, like ultrasound image[b3]. However, until now prostate on MR images is still mostly segmented manually by radiologists. The handcrafted delineation of prostate boundary is a time-consuming and labor-intensive operation with a low reproducibility because of its high dependence on medical experience. Moreover, those problems are further aggravated when the borderline is indistinct. As pointed out by Yu et al.[b4], automatic prostate segmentation is also a challenging task due to the issue of intensity inhomogeneity, variation of anatomical appearance, and lack of boundary discriminability.

In order to address this challenging task, different automatic or semi-automatic segmentation approaches have been reported in recent years. Martin et al.[b5,b6] proposed a semi-automatic prostate segmentation method, in which a rigid intensity-based registration algorithm and a non-rigid hybrid registration framework were employed successively to align an atlas to the patient image. In their work, 18 MRI series were involved to construct the atlas. Two accuracy metrics respectively based on volume and surface distance were used to investigate the segmentation performance. Results showed that the segmentation accuracy of the apex region and the central region is higher than the base part. In another publication, to add additional knowledge into the segmentation procedure, Korsager et al.[b7] combined the spatial information of a prostate atlas with the intensity information in a graph cut segmentation framework to achieve automatic prostate delineation. Their validation experiment was investigated on 76 axial MR images. As a result, a mean Dice similarity coefficient (DSC) of 0.88 and a mean surface distance of 1.45 mm were reported. Besides, Tian et al.[b8] utilized a superpixel-based graph cut framework to acquire the prostate surface on MRI. A superpixel is a group of pixels that have similar characteristics such as intensity or location. Due to the capacity to carry wider information, serving as a more

convenient and compact representation of the original image, the superpixel image has been widely used in image segmentation algorithms. In Tian's study, a graph cuts algorithm and an active contour model were integrated for cross-promotion. According to their experiment results, the verification on 43 MRI examples obtained a mean dice of 0.893.

Recently, the performance of deep learning approaches has outperformed to the trandi-tional state-of-the-art methods in many fields, especially in computer vision[b9,b10] and medical image processing[b11,b12]. In those researches, neural networks often work as information extractors to eliminate the tedious procedure of traditional feature choice and collection. For precise segmentation of MRI prostate, Guo et al.[b13] used more concise and effective hierarchical features from MRI prostate image by utilizing a stacked sparse autoencoder. Based on the extracted features, a sparse patch matching method was employed to deduce the corresponding prostate likelihood map, which was further combined with a sparse shape model for the final segmentation. Besides, Mun et al.[b14] integrated encoding, bridge, decoding, and classification modules to develop a baseline convolutional neural network (CNN) to extract volumetric information. In the meantime, Jia et al.[b15] researched a coarse-to-fine algorithm for MRI prostate segmentation through a deep learning method. In their algorithm, a registration-based segmentation was firstly used to obtain a rough prostate region. Then a pixel-wise recognizer based on a neural network was further adopted to classify the prostate boundary from the image background. Finally, a refinement algorithm was applied to smooth the contour. Similar to Jia's work, He et al.[b16] exploited another coarse-to-fine prostate segmentation system via different algorithms. They firstly proposed an adaptive feature learning probability boosting tree for prostate pre-segmentation. Next, a CNN method was developed to obtain the prostate profile model by the judgment of the inner, external, and boundary points. For the last step, an active shape model was employed for the final surface refinement. Their results showed the method is accurate and robust for prostate segmentation, as the neural network was utilized for the extraction of latent image features and the prediction of the prostate boundary. In addition, Wang et al.[b17] introduced a 3D deep-supervised full CNN with group dilated convolution, aiming to preserve extra image information for prostate delineation. Their method achieved a dice of 0.86. Generally, the aforementioned methods demonstrated that compared with the traditional segmentation algorithms, the approaches based on the deep learning method can delineate the target more accurately in less time. However, the aforementioned ways only involved the information of

the specific input image while without any prior knowledge constraints that are potentially helpful to improve the prostate segmentation accuracy.

Statistical shape model (SSM) is a geometric model containing a mean shape and multiple compressed primarily shape variations of a collection of similar shapes. Due to the ability to represent prior geometric information, SSM has been widely applied in different medical modalities for the segmentation and registration of various anatomical structures, including brain, bone, liver, heart, prostate and so on[b18,b21]. In this paper, based on a registration approach, we proposed a novel segmentation algorithm to tackle the prostate extraction problem by combining a boundary predictor and a label classifier. Specifically, a GoogLeNet-based branch (SSM-Net) was involved as the boundary predictor to deduce the prostate contour to obtain a boundary distance map, serving as the target contour constraints. Meanwhile, a 3D residual U-net branch (ResU-Net) was employed as the label classifier to predict a probability label map from the input images, to judge the class possibility of each pixel. In the inference step, the optimal weighted sum of the distance map and the probability map was regarded as the final prostate segmentation. In our validation experiment, six different SSMs with various nodes were built to investigate the accuracy and efficiency of the whole algorithm.

II. Materials and methods

::section2

Fig. 1 is an overview of the proposed prostate segmentation algorithm. The segmentation network comprises two branches, serving as a boundary predictor and a label classifier respectively. As shown by Fig. 1A, the first SSM-Net branch is used to deduce from the input, three variables: the SSM global transform, shape control parameters, and the point-wise fine-tuning vector. Then, a patient-specific deformation field is calculated based on the predicted variables, and a surface of target is further generated via a coordinate sampler. The second ResU-Net branch is employed to predict a probability label map from the input images, instead of directly using its binarized result as the target region. The parameters of the whole network can be optimized by minimizing the loss function value, which is related to two parts (Sec. II.D.): 1) the dice coefficient between the predicted probability map and the input binary label, defined as Loss 1. 2) the similarity between the SSM deformed surface and the input prostate, defined as Loss 2. In the inference phase(Fig. 1B), according to the

generated contour obtained by deforming the SSM by using the deformation field predicted by the SSM-Net branch, a distance map is directly calculated to serve as the target boundary constraints. Then, the weighted sum of the distance map and the probability map is regarded as the final prostate segmentation.

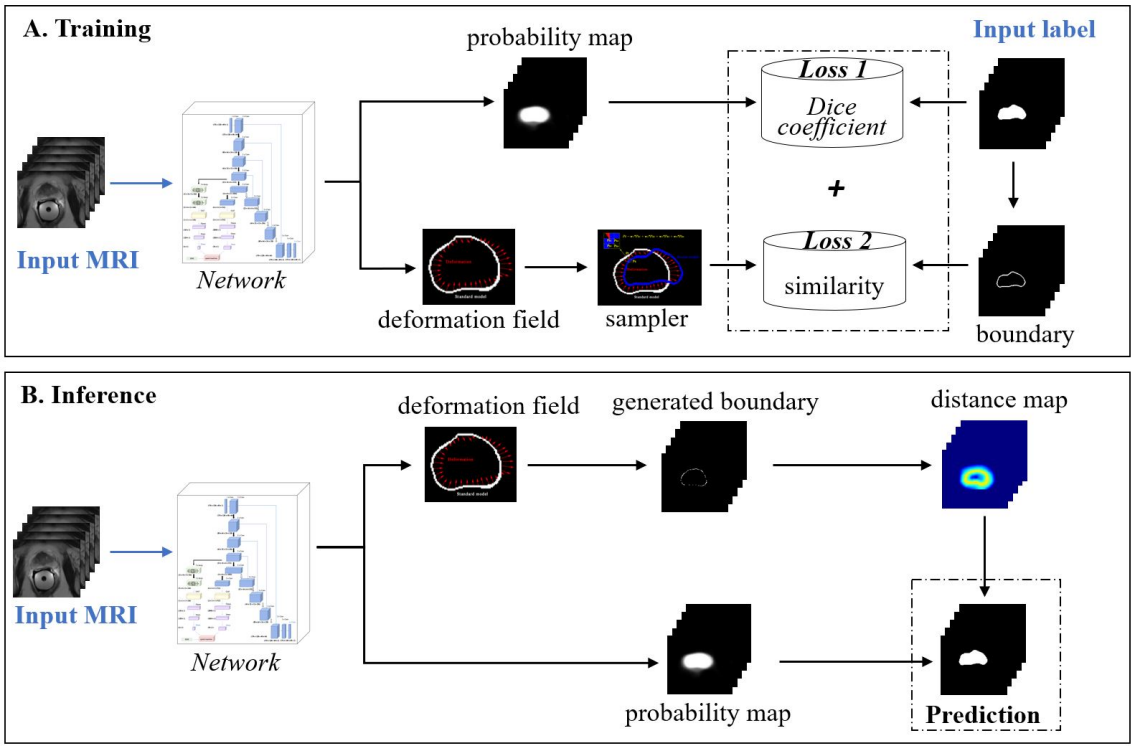

Figure 1: The overview of the proposed prostate segmentation framework.

figure1

II.A. Statistical shape model

section2.1

Due to the capacity to carry prior geometric information of numerous examples in different medical modalities, SSM has been widely applied in object recognition, image process, surgery implant design[22]. Generally, it involves two parts to describe the statistical spatial information of a collection of objects: a geometric model for the representation of the mean shape, and a series of variation vectors to depict the principal components of divergences between the objects and the mean shape. As the most prevalent SSM type, Principal Component Analysis (PCA) based shape model can model the variability of various types of objects such as images, displacement fields, surface meshes, and volumetric meshes. In our

b22

work, we built PCA SSMs based on triangulated surface meshes, to represent the prostate surface on the patient image. For the prostate, the procedures of building the PCA model can be summarized as following steps: 1) Segment the target from the original image volume. 2) Construct and refine 3D surface triangle mesh based on the binary target label. 3) Subdivide and decimate the meshes to a specific number of points. 4) Transform the meshes to the same posture and align them correspondingly. 5) Establish PCA-based SSMs based on the corresponding objects. 6) Augment the SSM flexibility by employing a Gaussian process.

Table1 Table 1: Mathematic notations of statistical shape model

A. Variables: scalars, vectors, matrices	
N	Number of SSM nodes
M	Number of SSM variations
$\bar{u}_{3N\times1}(\bar{u}_{3N})$	SSM mean shape
$\psi_{3N\times M}$	SSM variation / PCA basis
$\kappa_{M\times1}(\kappa_M)$	Variance of SSM variations
$\theta_{M\times1}(\theta_M)$	SSM shape control parameters
$t_3 : T, R$	Transform parameters: translation and rotation
$I_{3\times3}(I)$	Identity matrix
B. Operators and functions	
$diag(\vec{v})$	Diagonal matrix of vector $\vec{v}$
$V(\psi_{3N\times M}, \kappa_M, \theta_M, \bar{u}_{3N})$	Model surface deformation

According to the theory of the PCA-based model, an arbitrary shape can be represented by superimposing a deformation field to the mean shape. In our work, a deformation field is described as the sum of global transform, weighted variations, and the mean shape. That means, according to the notations defined in Table 1, shape can be written as follows:

$$\begin{aligned} u' &= \bar{u}_{3N} + V(\psi_{3N\times M}, \kappa_M, \theta_M, t_3) \\ &= \bar{u}_{3N} + \psi_{3N\times M} \cdot diag(\kappa_M) \cdot \theta_M \cdot R + T \end{aligned} \tag{1}$$

equation1

The deformation ability of SSM depends on its node number (N) and the variation matrix $\psi_{3N\times M}$. As only very limited datasets were used to build the SSM, the model is insufficient to explain all possible shape variations. Two approaches were employed to solve this problem. Firstly, according to the principle of statistic shape, augmenting the example

shapes by involving small and very smoothly varying deformations, can make the variation matrix (noted as $\psi_{3N\times M}\cdot diag(\kappa_M)$ in Eq. (1)) more representative. $\psi_{3N\times M}\cdot diag(\kappa_M)$ is PCA dimensionality reduction form of variation matrix $\psi_{3N\times M}$.

$$\psi_{3N\times 3N}=\begin{bmatrix} k_{SSM}\left(x_1,x_1\right) & \cdots & k_{SSM}\left(x_1,x_N\right) \\ \vdots & \ddots & \vdots \\ k_{SSM}\left(x_N,x_1\right) & \cdots & k_{SSM}\left(x_N,x_N\right) \end{bmatrix} \tag{2}$$

$$k_{SSM}\left(x,x'\right)=\frac{1}{n-1}\sum_{i=1}^{n}\left(u_i\left(x\right)-\mu_{SSM}\left(x\right)\right)\left(u_i\left(x'\right)-\mu_{SSM}\left(x'\right)\right)^T \tag{3}$$

Combining sample covariance kernel ($k_{SSM}\left(x,x'\right)$) calculated from sample data, and Gaussian kernel $k_g^{(s,\sigma)}$ can enlarge the flexibility of the model. For the Gaussian kernel in this paper, the smoothness $\sigma=10mm$ and the scale of the deformation $s=2mm$.

$$k_{aug}\left(x,x'\right)=k_{SSM}\left(x,x'\right)+k_g^{(2,10)}(x,x') \tag{4}$$

$$k_g^{(s,\sigma)}(x,x')=s\cdot diag(\exp(-\frac{\left\|x-x'\right\|^2}{\sigma^2})) \tag{5}$$

Secondly, supplementing a point-by-point item ξ_{3N} to the deformed target (u') was another feasible approach to represent more possible targets, as shown as Eq. (6). Item ξ_{3N} was predicted by the SSM-net branch automatically. As shown in Figure 2, the prediction of the offset vector shared the same residual CNN structure as the prediction of shape control parameters (noted as θ_M in Eq. (6)).

$$\begin{aligned} u^{'} &= \bar{u}_{3N}+V(\psi_{3N\times M},\kappa_M,\theta_M,t_3)+\xi_{3N} \\ &= \bar{u}_{3N}+\psi_{3N\times M}\cdot diag(\kappa_M)\cdot\theta_M\cdot R+T+\xi_{3N} \end{aligned} \tag{6}$$

The surface model is commonly defined in the physical spatial coordinate system to maintain the visualization invariance on different platforms. The transformation between model space and image space in this work is shown as follows:

$$P_i=\lfloor(P_m-P_0)/s+0.5\rfloor \tag{7}$$

where s is the image resolution, and P_0 is the position of image origin in the physical spatial coordinate system. P_m indicates the coordinate of model node m in spatial space and P_i is the corresponding coordinate of P_m in image space.

II.B. CNN architecture of the segmentation framework

section2.2

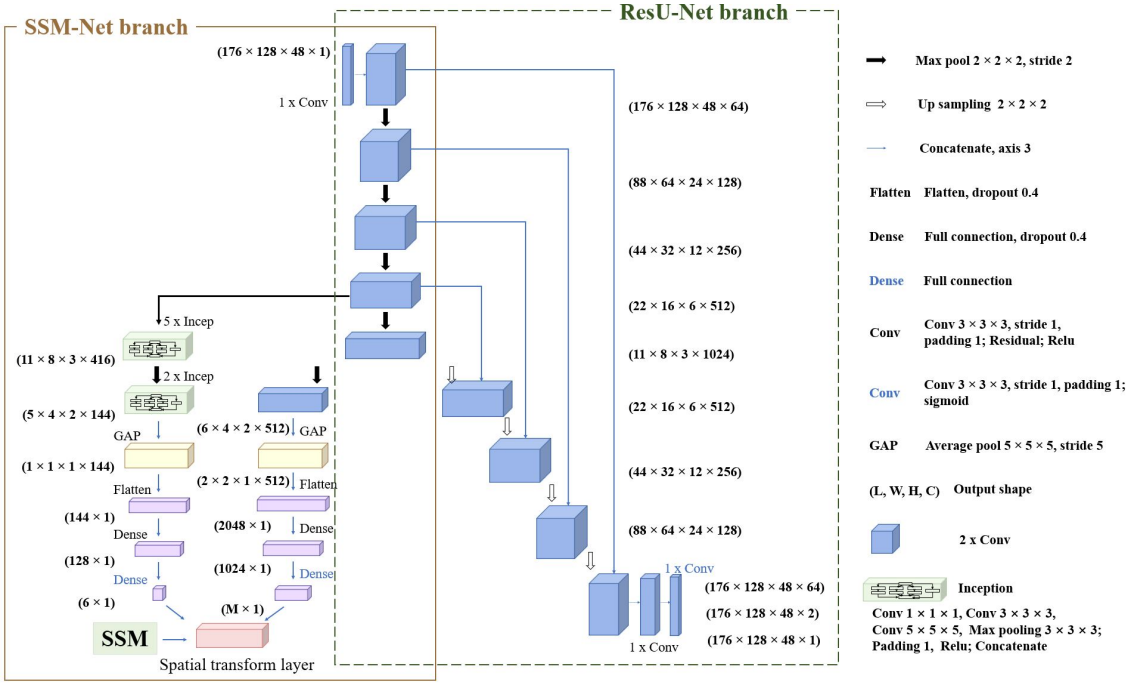

figure2

Figure 2: The structure of the whole network.

This subsection describes the architecture of the network involved in our segmentation framework. As illustrated in Fig. 2, the input image was size of (176, 128, 48, 1). For the SSM-Net branch, the inception block of GoogLeNet was utilized for the extraction of the prostate position. There, GoogLeNet[23] is a CNN originally designed for the ILSVRC (ImageNet Large Scale Visual Recognition Competition). Its adopted inception blocks are composed of several convolutional filters of various sizes, making easier the exploration of image details at different scales. On the basis of the original GoogLeNet, a dense layer with a size of 128 is connected to its flatten layer, for the prediction of the global transform (Dx, Dy, Dz, Rx, Ry, Rz), including three parameters for translation and three parameters for rotation, as shown by the left column of SSM-Net branch in Fig. 2. The right column of the SSM-Net branch in Fig. 2A shows the network structure for the prediction of shape

control parameters with the size of M × 1 (M is the number of shape variations). The involved convolution layers following the "contracting path" with the kernel size 3 × 3 × 3 used 1 pixel stride, and the employed max-pooling layers used pooling size of 2. Similarly, the prediction of fine-tuning vector shared the same residual CNN structure, expect that the average pooling layer is replaced by a max-pooling layer with the pooling size of 2, the stride of 2. In the last step of the SSM-net branch, the variables including the transform matrix, the shape control parameters, and the fine-tuning vector, are input to the last spatial transformation layer, to yield the prostate contour.

As illustrated by the ResU-net branch in Fig. 2, a residual U-Net is employed to infer the probability label map. U-net[24] is a widely used network with high accuracy for object segmentation. In our segmentation framework, each residual block consists of two convolution layers with a kernel size of 3 × 3 × 3, pixel stride of 1. And the max-pooling layers use a pooling size equal to the stride and the size of the up-sample layers is set to 2. The sigmoid function is utilized as the activation function of the last layer to limit output values to [0,1]. For the whole network, the structure of each layer is shown in Fig. 2.

II.C. Grid mapping

As shown in the overview of the segmentation framework (Fig. 1), the patient-specific prostate shape can be obtained by superimposing the predicted deformation field to the standard SSM surface in the inference procedure. While in the training phase, in order to calculate the loss, a predicted binary surface image for SSM (annotated as "generated model surface" in Fig. 1.) is generated by interpolating the input boundary based on the deformation field.

The output binary surface of SSM $g^t \in \mathbb{R}^{L^t W^t H^t}$ is defined on a regular grid $G^t = \{G_i^t\} = \{(x_i^t, y_i^t, z_i^t)\}, i \in [L^t W^t H^t]$, where L^t,W^t,H^t represent the length, width and height of the output. Similarly, let $G_s = \{(x_i^s, y_i^s, z_i^s)\}, x_i^s \in [0, L^s], y_i^s \in [0, W^s], z_i^s \in [0, H^s]$ be the input grid, where, L^s,W^s,H^s are the length, width and height of the input binary mask respectively. The relationship between the output grid G^t and the input grid G^s can be written as follows.

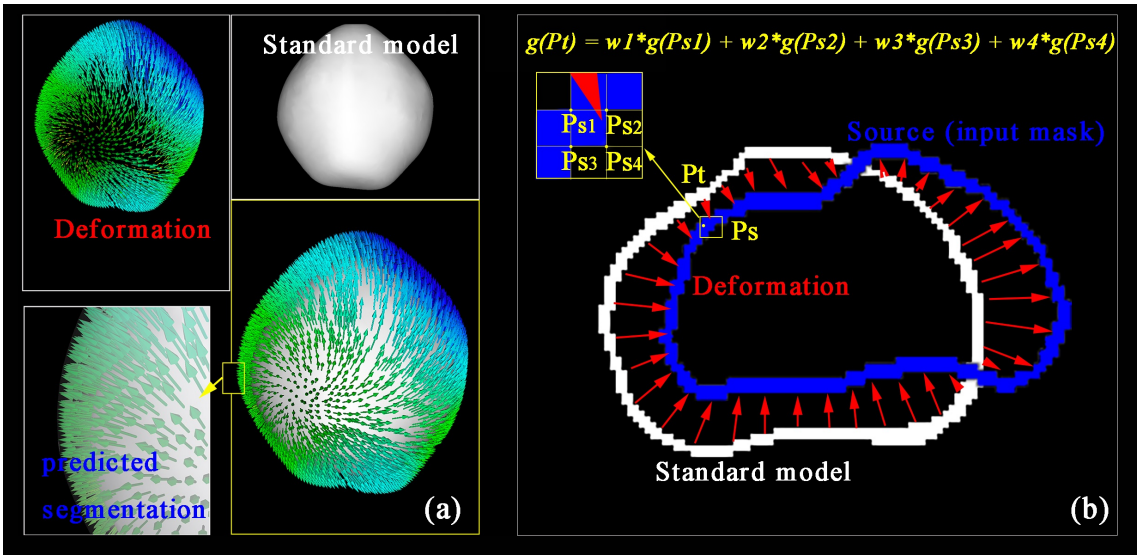

Figure 3: The schematic diagram of segmentation generation and image interpolation. (a) shows that the spatial structure of a predicted segmentation is equal to the sum of the SSM and a deformation field. (b) illustrates the 2D calculation strategy of the gray value on the model boundary according to the binary input mask and the deformation field. indicated with red arrows. g(P) represents the gray value of point P.

figure3

For $\forall i \in [1...L^t W^t H^t]$,

$$G_i^s = D_i(G_i^t) = \begin{pmatrix} x_i^t + d_{i,x} \\ y_i^t + d_{i,y} \\ z_i^t + d_{i,z} \end{pmatrix} = \begin{pmatrix} x_i^s \\ y_i^s \\ z_i^s \end{pmatrix} \quad (8)$$

equation8

$D_i(G_i^t)$ is the deformation field predicted by SSM-Net branch, with size of $L^t \times W^t \times H^t \times 3$. Based on Eq. (8), for each node in the output grid, we can trace its corresponding position on the input mask according to the deformation field. It means that the gray value of output point $G_i^t = (x_i^t, y_i^t, z_i^t)$ depends on the gray value of the relevant position $G_i^s = (x_i^s + d_{i,x}, y_i^s + d_{i,y}, z_i^s + d_{i,z})$ in the input mask.

II.D. Loss function

section2.4

The loss function of the proposed network consists of two parts: the part for SSM-Net branch and the part for ResU-Net branch.

For SSM-Net branch, according to the mean shape of SSM and the predicted deforma-

tion field, we can calculate the final segmentation directly. To evaluate the accuracy of the prediction, the specific energy function of SSM-Net branch was employed to optimize the deformation field, which is described as follows:

$$\begin{aligned} E &= L(g^t \diamond G^s) + \lambda \|\varphi\| \\ &= L(g^t \diamond G^s) + \lambda_1 \|\theta\| + \lambda_2 \|t\| + \lambda_3 \|\nabla\xi\| \end{aligned} \tag{9}$$

equation9

where g represents the gray label in image space. Specifically, g^t denotes the boundary of the input mask, and $(\theta,\ t,\ \xi)$ is the predicted deformation field. In this paper, we defined L as:

$$L(g^t \diamond G^s) = 1 - (g^t \diamond G^s)/N \tag{10}$$

equation10

Then, the optimization function can be written as follows:

$$\varphi = (\theta, t, \xi) = \arg\min(1 - (g^s \diamond G^t)/N) + \lambda_1 \|\theta\| + \lambda_2 \|t\| + \lambda_3 \|\nabla\xi\| \tag{11}$$

equation11

where N represents the node number of SSM.

To obtain $g^t \diamond G^s$, an interpolation is required to calculate the gray value of arbitrary position $G_i^s = (x_i^s, y_i^s, z_i^s)$. In this work, 3D bilinear interpolation method[b25,b26] is employed. Let g_i^t represents its gray level at point $G_i^t = (x_i^t, y_i^t, z_i^t)$. $g_i^t \diamond G_i^s$ mean the gray level at point $G_i^s = (x_i^s, y_i^s, z_i^s)$, and $\{(m,n,p)\}^s$ represents the point (m,n,p) of input mask respectively. According to Eq. (4), For $\forall i \in [1...L^tW^tH^t]$,

equation4

$$\begin{aligned} g_i^t \diamond G_i^s = g^s_{(x_i^s, y_i^s, z_i^s)} = \sum_m^{L^s} \sum_n^{W^s} \sum_p^{H^s} g^t_{(m,n,p)} \cdot max(0, 1 - |x_i^s \\ - n|) \cdot max(0, 1 - |y_i^s - m|) \cdot max(0, 1 - |z_i^s - p|) \end{aligned} \tag{12}$$

equation12

The partial derivatives with respect to gray $g^t_{(m,n,p)}$ and coordinate position (x_i^t, y_i^t, z_i^t) for the backpropagation of loss can be written as follows ($\partial g_i^t \diamond G_i^s / \partial y_i^s /$, $\partial g_i^t \diamond G_i^s / \partial z_i^s$ are similar with $\partial g_i^t \diamond G_i^s / \partial x_i^s$):

$$\begin{aligned} \frac{\partial g_i^t \diamond G_i^s}{\partial g^t_{(m,n,p)}} = \sum_m^{L^s} \sum_n^{W^s} \sum_p^{H^s} max(0, 1 - |x_i^s - n|) \cdot \\ max(0, 1 - |y_i^s - m|) \cdot max(0, 1 - |z_i^s - p|) \end{aligned} \tag{13}$$

equation13

$$\frac{\partial g_i^t \diamond G_i^s}{\partial x_i^s} = \sum_m^{L^s} \sum_n^{W^s} \sum_p^{H^s} g_{(m,n,p)}^t \cdot max(0, 1 - |y_i^s - m|)$$
$$\cdot max(0, 1 - |z_i^s - p|) \cdot \begin{cases} 0, & |x_i^s - n| > 1, \\ -1, & x_i^s > n, \\ 1, & x_i^s < n. \end{cases} \quad (14)$$ equation14

For ResU-Net branch, the following loss function is adopted, where, S_{mask} represent the ground truth. ϑ is the network parameters and P_{ResU} is the output probability map. The values of range from 0 to 1.

$$L(\vartheta, S_{mask}) = 1 - \frac{2 \times \|P_{ResU}(\vartheta) \times S_{mask}\|}{\|P_{ResU}\| + \|S_{mask}\|} \quad (15)$$ equation15

Thus, the parameterized ResU-Net branch can be optimized during the training procedure:

$$\hat{\vartheta} = \arg\min_{\vartheta} L(\vartheta, S_{mask}) \quad (16)$$ equation16

Based on Eq. (equation10) and Eq. (equation15), the loss function of the whole network is defined as Eq. (equation17), to optimize the network weights. Where α and β are trainable parameters.

$$L_{combined} = \alpha L(l, M, \phi) + \beta L(\vartheta, S_{mask}) \quad (17)$$ equation17

II.E. Inference of prostate region

section2.5

In the prostate inference step, the binarized weighted sum of the deduced probability map and distance map is regarded as the final prostate segmentation. In terms of probability label map, a bigger value means a higher probability for a pixel to belong to the prostate region. The distance map obtained from the boundary image predicted by the SSM-branch, is such that only the pixel on the boundary has a value of 1, while others are equal to 0. It is calculated according to the following equation:

$$D_{ssm}(P_i) = 1 - \|P_i - \hat{P}_i\|/10 \quad (18)$$ equation18

where, P_i represents pixel i, $D_{ssm}(P_i)$ is the value of P_i on the distance map. $\hat{P}_i$ is the closest boundary point of P_i on the input boundary image, and $\|P_i - \hat{P}_i\|$ is the Euclidean

distance between P_i and $\hat{P}_i$. 10 is the calculation range, which should be changed according to the image resolution because it decides the prostate voxel range in images. As the image volume was resampled to the same resolution, the calculation range is a constant in our work. Furtherly, as the interior prostate gland being segmented should be evaluated with a large value to reflect its high probability, for the pixel P_i in the interior region with $D_{ssm}(P_i)$ less than 0.5, its D_{ssm} value is reassigned to 1.

The probability map generated by the Res-Unet branch is such that the central region of the prostate has a probability of 1 whilst the marginal prostate region is valued with a lower value, which contributes to most of the prediction deviation, especially when the contour of the prostate is indistinct. In contrast, in this case, the SSM-Net branch can deduce a relatively reasonable boundary due to the representation of prostate shape prior and provided complementary information for the result of the ResU-Net branch. However, the distance map calculated from the SSM-Net branch might not match the segmentation results from the ResU-Net branch, therefore, we have investigated the segmentation performance of the proposed method under different combinations of SSM-Net and ResU-Net branches, as shown as follows:

$$P_{combined} = w_1 P_{ResU} + (1 - w_1) D_{ssm} \tag{19}$$ equation19

Two metrics were introduced to evaluate the performance of the proposed segmentation framework, including the dice similarity coefficient (DSC) and the average over the shortest distance between the boundary points of the volumes(ABD). The DSC is formulated as follows:

$$Dice = 2 * \|S_{pred} \cap S_{mask}\| / (\|S_{pred}\| + \|S_{mask}\|) \tag{20}$$ equation20

where S_{pred} and S_{mask} respectively present the predicted segmentation and the input ground truth.

III. Results

:section3

III.A. Data acquisition and experiment set up

section3.1

We validated the proposed network on two public datasets: the MICCAI PROMISE12 challenge dataset[b27] and NCI-ISBI 2013 challenge dataset[b28]. The first database contains 50

prostate transversal T2 MRIs for training and 30 prostate images for testing. And the second database respectively involves 60, 10, and 10 cases for training, leaderboard, and testing. The two datasets share 11 common volumes, and the ground truths of the PROMISE12 testing set are unavailable. Thus, a total of 119 T2 MRI image cases are collected for our experiment. As the gathered data have different voxel spacing and image size, the input images were resized to the shape of $176 \times 128 \times 48$ beforehand.

From the 119 volumes, 40 randomly selected cases contributed to the SSM establishment, and the rest data were used for the network optimization. Specifically, 63 out of the 79 image volumes (3/4) were applied for the network training and the remaining (1/4) for the validation. The modeling procedure and the network segmentation performance are illustrated in Section III.B. and Section III.C. respectively. And Section III.D. analyzes and discusses the influence of SSM flexibility on the framework segmentation accuracy. The network training and all experiments were conducted on a computer with Intel® Core (TM) I7-8700K with a 3.70 GHz CPU, 8 GB memory, and two graphic cards of 8GB NVIDIA GeForce GTX 1080 Ti.

III.B. SSM establishment and analysis

section3.2

The node number and the principal component number are two dominant factors for the flexibility of SSM. The former is decided according to the node number of the counted meshes, and the latter is determined by the compactness of SSM. The compactness is measured via the accumulation of SSM variations which are arranged according to their eigenvalues. In this work, we selected the first M principle components to keep 95% of the total eigenvalues. To investigate the influence of the two factors on the segmentation accuracy of the segmentation method, six SSMs with various nodes and principal components were built. 3D slicer (https://www.slicer.org) was utilized for the refinement of 3D surfaces. In addition, the Gaussian process model building, model fitting, and the PCA-based model building are completed via the Statismo library [22]. The details of SSMs are presented in Table 2. The third column shows the models after flexibility augmentation.

Table 2: The details of statistical shape models with various number of nodes

Table2

Number of nodes	Original SSMs	Augmented SSMs
	Variation(3N x M)	Variation(3N x M)
1625	4875 × 49	4875 × 50
3250	9750 × 49	9750 × 50
6500	19500 × 49	19500 × 50
9750	29250 × 49	29250 × 50
13000	39000 × 48	39000 × 50
16250	48750 × 47	48750 × 50

III.C. Accuracy evaluation and analysis

section3.3

figure4

Figure 4 illustrates the flexibilities of the primary SSMs. The green models in the middle column depicts the mean shapes with increasing node numbers from top to button. From the middle to the rightmost, the three columns respectively exhibit the deformed shapes generated from the mean shape of SSM by weighting the first three principles variations with three times of their corresponding deviations. Accordingly, the three-column shapes on the left describe the generated models deformed by negative triple deviations. Generally, each row horizontally reveals the influence of the first three principal variations on the deformation of each SSM, and each column vertically compares the different performances of SSMs. In conclusion, the last three SSMs with node numbers of 9750, 13000, 15250 had similar interior and exterior characteristics, while the first three SSMs behaved quite differently.

equation17

Based on Eq. (17), the network weights was optimized via Adam optimizer with the learning rate ranged from 10^{-5} to 1, and the three regularization coefficients λ_1, λ_2 and λ_3 ranged from 0 to 1 respectively. During the training period, we compared the performance of the proposed segmentation framework applying different values of regularization coefficients λ_1, λ_2 and λ_3 in Eq. (11). The experiments showed that the best result in terms of the Dice coefficient is achieved when λ_1, λ_2, and λ_3 are equal to 0, 0.01, and 0.01. In the following work, λ_1, λ_2 and λ_3 are set to 0, 0.01, and 0.01. In addition, we figured out the proposed segmentation framework performs best when the "hyperparameter" w_1 in Eq. (19) equals 0.6 by trial and error.

equation11

equation19

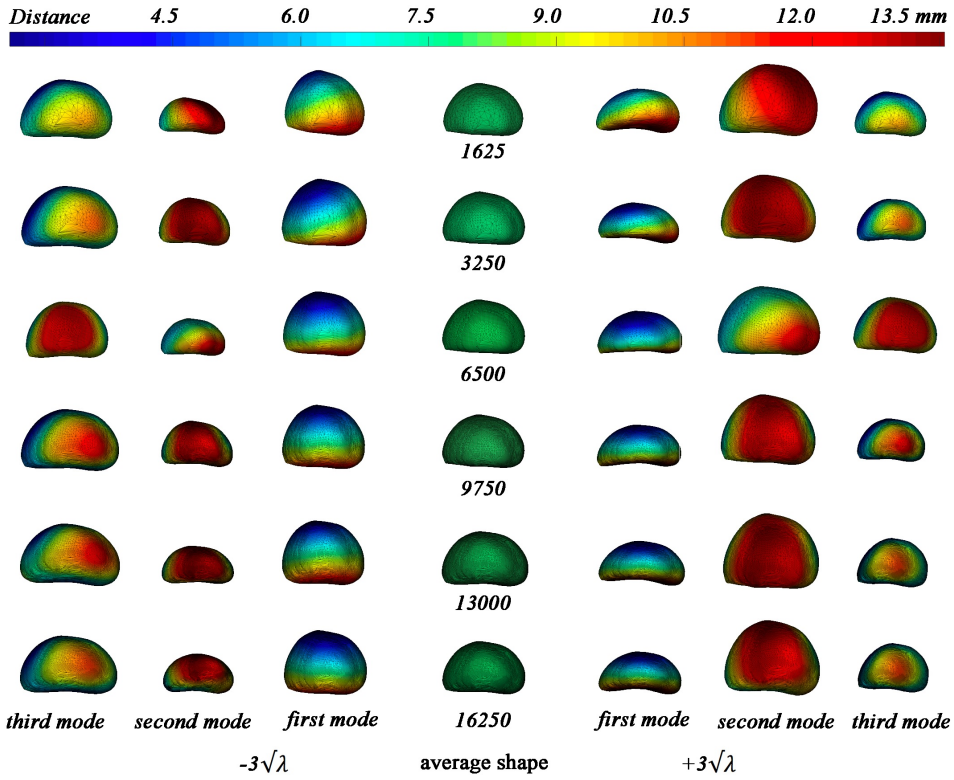

Figure 4: The flexibility of the statistical shape models with different node numbers. The middle column with green color exhibits the mean models, and the left three columns and the right three columns respectively show the deformed models drove by $3\sqrt{\lambda}$ times of the first three principal components of variations. λ is the corresponding deviation of each component.

figure4

figure5

Fig. 5 illustrates the DSC and ABD of the proposed prostate segmentation framework when adopted different SSMs. 4-fold cross-validation was conducted for each group. The DSC and ABD of the ResU-Net branch are constant over the different groups. In terms of the SSM-Net branch, when the network adopts SSMs with 1625 to 9750 nodes (referred as network nodes in the following text), the dice result has significant improved from 0.69 to 0.90 (paired t-test, p <0.001). At the same time, the ABD value has an opposite steep trend, decreasing from 2.63 mm to 2.39 mm. Both dice and ABD have the best result when network node number is 9750, with an average dice of 0.862 and an ABD of 2.04 mm. As regards the performance of the whole framework, it has a similar trend with the SSM-Net branch, with the dice reached a peak of 0.907 and the ABD declined to the lowest of 1.85 mm. For the two groups with more than 9750 nodes, they performed slightly inferior with a

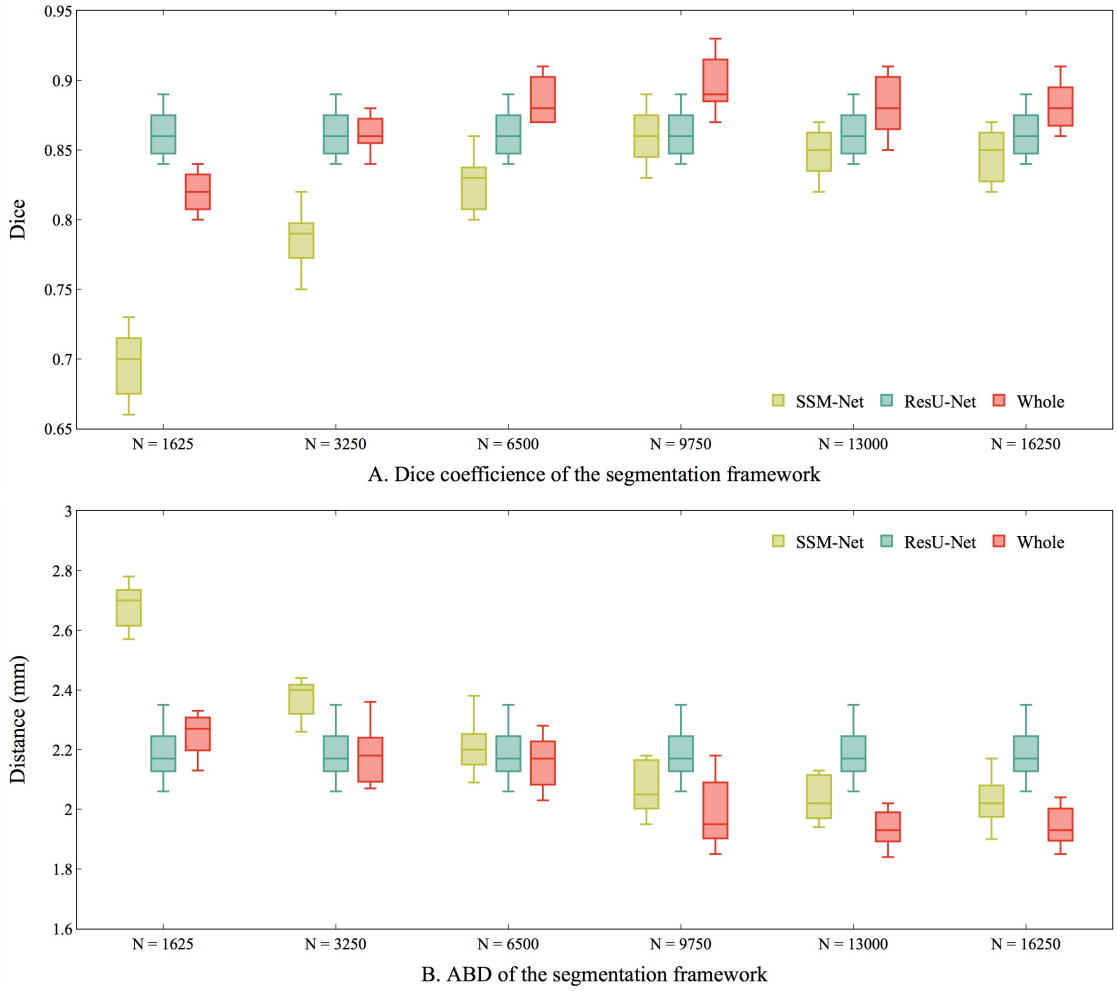

Figure 5: The DSC and ABD results of the whole framework, SSM-Net branch, and ResU-Net branch respectively.

figure5

dice of 0.89. According to the record data, we concluded that the SSM with 9750 nodes is optimal to employ in our framework for the representation of the prostate spatial boundary.

The segmentation procedure of the proposed framework is shown in Fig. 6. The purple model in (c) represents the SSM, whose center is initially positioned at the origin (0,0,0) of the anatomical coordinate system of LPS (Left, Posterior, Superior). As shown in (c), according to the input image, the SSM-Net branch respectively predicts the global transform, weight parameters, and an offset vector for the calculation of the deformation field. The 3D white surface in (d) represents the generated prostate boundary by applying the deformation field to the SSM. Fig. 7 illustrates the delineation results of four images series. Four slices range from number 16 to number 46 with an interval of 10, are selected to display the recognition

figure6

figure7

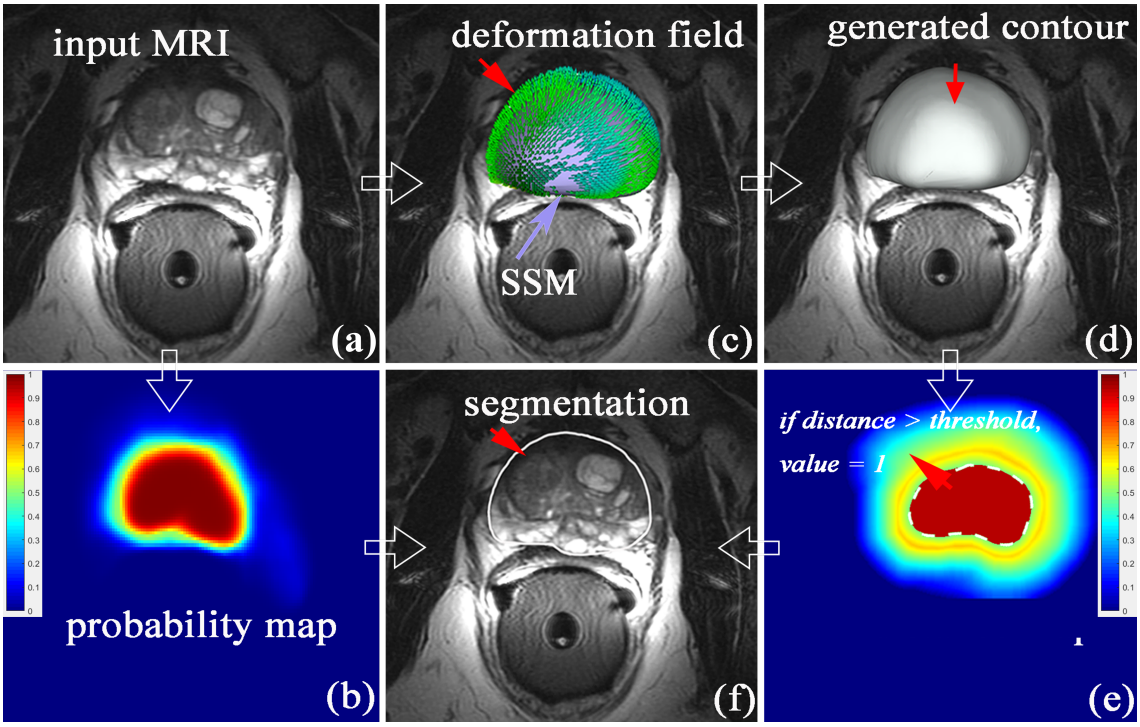

Figure 6: The whole segmentation procedure. (a) The input MR image for segmentation. (b) The probability map. (c) The purple model represents the SSM, whose center initially positioned at the origin (0,0,0). (d) The generated contour of the prostate from the SSM. (e) The distance map. (f) The segmentation result.

figure6

performances on different prostate zones. Each row stands one example, and accordingly, each column shows the delineation results of the same layers of different examples. From the exhibition, we deduced that the segmentation on the prostate central zone has higher accuracy than the base of the prostate (more complex to delineate). For the case of the severe hyperplastic prostate gland which is larger than the mean shape of SSM, the segmentation framework has relatively poor performance with the maximum ABD was 2.7 mm and the dice coefficient was 0.83. Compared to the segmentation approach employing only the SSM method, our framework can achieve more accurate delineation when the target is beyond the SSM deformation range.

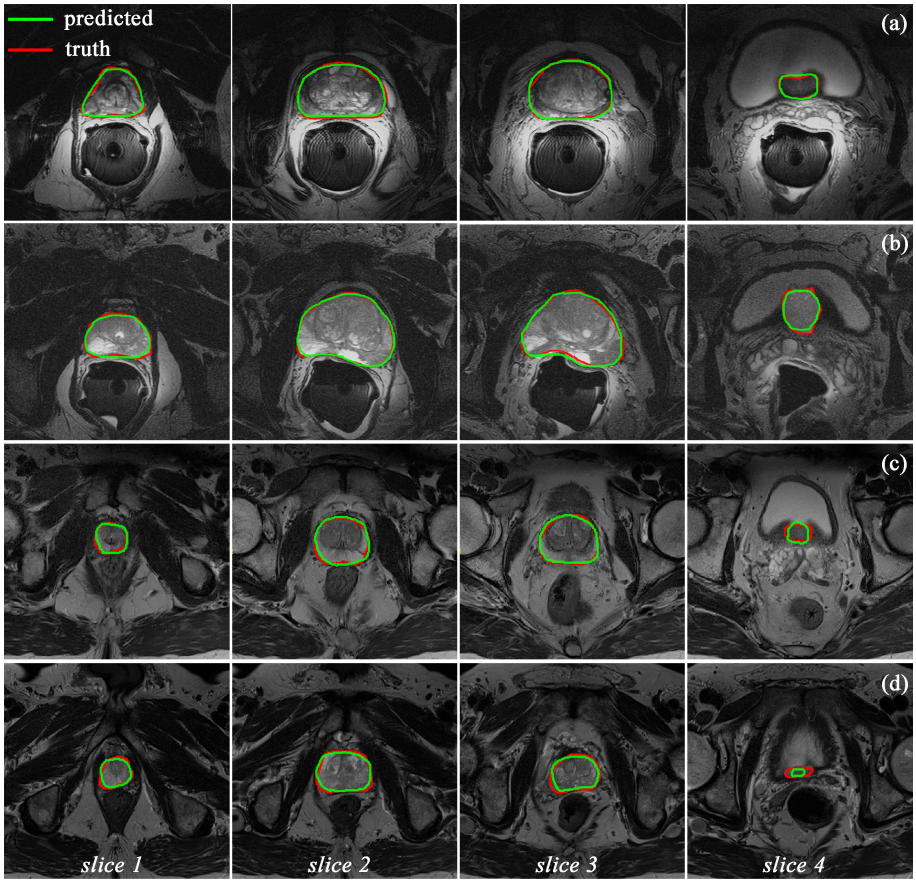

Figure 7: The segmentation results of four image cases via the proposed method. Each row stands one example, and each column shows their segmentation results of the same layers of different examples.

figure7

III.D. The influence of network flexibility on segmentation accuracy

section3.4

According to the analysis in Section II.B., network node number, flexibility augmentation of SSM, and fine-tuning item (offset) are the three dominant factors to affect the network deformation ability. To investigate their impacts on the network delineation accuracy, we verified the individual segmentation performance of the SSM-Net branch under different combinations of the three elastic determinants. Fig. 8 presents the DSC and ABD values of the SSM-Net branch with various nodes and different utilization situations of augmentation and offset item. As shown by those statistical trend lines, regardless of the model augmentation or fine-tuning item, the SSM-Net branch with 9750 nodes or more outperformed other situations. In terms of the model flexibility augmentation, its application has improved the

sec::section2.2

figure8

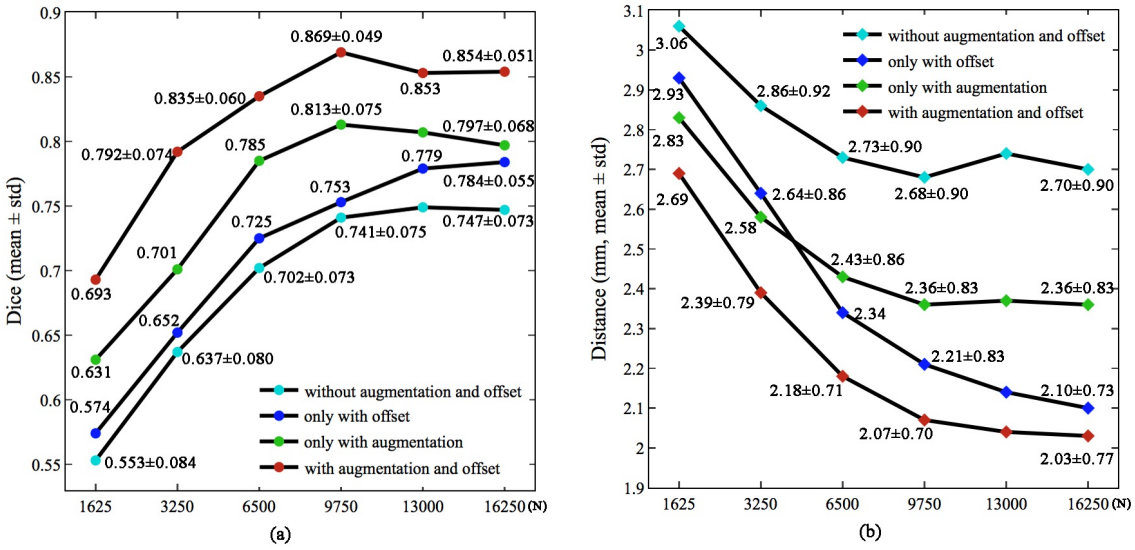

Figure 8: The influence of the deformation ability of SSM-Net branch on its segmentation performances, including dice scores and ABD values). The network elastic ability is mainly determined by three factors: node numbers, flexibility augmentation and the offset item.

figure8

network accuracy. As shown by the lines with cyan and blue dots in Fig. 8, compared with the results of the network without augmentation and offset, the best average dice and distance of the network adopting model elastic augmentation were improved to 0.81 and 2.36 mm respectively. And the network employing both augmentation and offset had higher average dice and smaller average distance than the network utilizing only offset item, as informed by the lines with green and red dots in the figure. Similarly, we figure out that the employment of offset item also contributed to the improvement of the SSM-Net branch. Specifically, compared with the segmentation result of the network without augmentation (the line with cyan dots in subgraph (a)), the involvement of fine-tuning item (the line with blue dots in subgraph (a)) increased the highest average dice of the six groups from 0.74 to 0.79. Based on the network which applies only flexibility augmentation (the lines with blue dots), the utilization of offset items (the lines with red dots) improved the average dice coefficient by 0.06 and the distance by 0.25 mm, making the SSM-Net branch achieve the best result.

Table 3: Comparison between our method and other automatic segmentation methods. All the methods were trained and tested on the same dataset.

Work	Method	DSC + std	ABD [mm])	Time
Maan et al.[32]	3D AAM	0.81 ± 0.13	3.09 ± 0.96	4 min
Vincent et al.[33]	AAM	0.86 ± 0.07	2.17 ± 0.63	8 min
Toth et al.[36]	Deformation landmark AAM	0.77 ± 0.18	3.64 ± 1.39	3 min
Ou et al.[30]	Multi-atlas	0.84 ± 0.06	2.85 ± 0.72	40 min
Gao et al.[31]	Multi-atlas + patch-based voxel	0.82 ± 0.02	2.86 ± 0.82	30 min
Milletari et al.[35]	V-Net	0.86 ± 0.11	2.13 ± 0.86	<1 min
Yu et al.[4]	Volumetric ConvNet	0.87 ± 0.24	2.05 ± 0.69	<1 min
Karimi et al.[34]	CNN + SSM	0.88 ± 0.09	2.16 ± 0.77	<1 min
Tian et al.[37]	PSNet	0.86 ± 0.40	2.72 ± 0.90	<1 min
Jia et al.[38]	Multi-atlas +VGG-19	0.92 ± 0.05	1.63 ± 0.38	40 min
Ours	SSM-Net + ResU-Net	0.90 ± 0.08	1.85 ± 0.75	<1min

III.E. Comparison with other methods

We compared the performance of our method with existing automatic segmentation methods (including atlas-based method, deformable model-based method, and deep learning-based method). Non-open-source methods were excluded for comparison. All deep-learning-based methods are trained on the same dataset (consisting of 63 volumes), and tested on the same dataset (consisting of 16 volumes). For methods that are not based on deep learning, their performances are evaluated on the same test dataset (consisting of 16 volumes).. The comparison results are summarizes in Table 3. Results illustrated that our method can achieve the second best accuracy than other methods. The inference time consumption of our method is much less than atlas or deformable model based methods, and is comparable with deep learning network methods.

IV. Discussion

Prostate segmentation facilitates the diagnosis and treatment of prostate diseases. For example, the determination and location of the prostate region are essential information for radiotherapy or high-intensity focused ultrasound operations. However, its clinical application is still limited, because of the segmentation challenges like inhomogeneous intensity,

various anatomical appearances, and indiscernible boundaries. Therefore, in this work, we proposed a novel prostate segmentation framework, based on CNN and SSM, which has been widely used in prostate segmentation[b29]. The results demonstrate that the network has the optimal dice of 0.907 and DSC of 1.85 mm under the network nodes of 9750. And both model elastic augmentation and offset applications have positive effects. The performance on the collected clinical data demonstrates that our prostate segmentation framework is feasible, and it has the potential to be a useful clinical tool for the diagnosis, treatment design, and therapeutic procedure of variable prostate disease.

As shown in Figure 4, driven by their first three principal components of variations, the first four statistical shape models with nodes number from 1625 to 9750 perform quite different from each other, while the two statistic shape models with nodes number of 13000 and 15250 have almost similar performance with the model with 9750. Thus, the mesh shape with nodes number of 9750 is accurate enough to represent the anatomical structure of a regular prostate in the physical coordinate system. However, when nodes number is more than 9750, the accuracy (including DSC and ABD) does not improve with the increase of nodes number. That's because when the model owns a substantial number of nodes, one pixel in the image space may correspond to more than one node in the physical space, resulting in decreased performance.

In the past three decades, three major categories of automatic prostate segmentation have been introduced, including atlas-based algorithms, deformable model-based approaches, and deep learning-based methods. For the first category, the atlas-based strategy has been widely utilized in medical image segmentation and registration[b30–b31]. Its major principle is to align an atlas that contained spatial prior knowledge to the target images by registration approaches. Then apply the alignment information to deform the atlas label for the final target segmentation. Secondly, in terms of the model-based method, a deformable model is firstly constructed for the representation of standard prostate contour. Then the information extracted from the target image was further applied to drive the model to generate the specific shape. Several groups in the list[b32–b34] employed deformable model (AAM, ASM, SSM) for the prediction of prostate. For the third group, deep learning-based approaches, especially CNNs are widely introduced in medical image processing because of the powerful feature extraction and non-linear learning capability[b4,b34,b35].

Several teams have combined deep learning algorithms with atlas or deformable models in their researches. Similarly, our segmentation combined neural network with SSM for high precise MRI prostate segmentation. In this way, prior knowledge is introduced by SSM to serve as the boundary constraint and a rough reference, and furtherly combined with a neural network (ResU-Net in our work) which can obtain details from the target image, for the precise recognization of the target region. Besides, as SSM is built based on lots of medical image data, the generated model can reasonably represent the deformation. Therefore, our method required less time (approximately 3 s) while performed a satisfactory segmentation accuracy with high robustness.

In medical image processing, the collection of the training dataset limits the application of various learning-based approaches. Fortunately, SSM utilizes geometric information rather than intensity information. Therefore, all images in different modalities such as computed tomography, ultrasound and MRI can contribute to the construction of SSM. It is worth mentioning that, if the dataset is sufficient, active shape model(ASM) and statistical deformation models (SDM) could become superior training supervisors than SSM, as the former can carry intensity information of images and the latter is capable to represent the statistical information of deformation field of a collection of examples. In addition, the introduction of the finite element model (FEM) is worth considering for our further development, because the biomechanical information in FEM can contribute to the delineation of a specific target.

The proposed method has a very good potential for clinical application. After establishing the SSM model and training the network, the MR images containing the prostate can be segmented automatically. The 3D model of the prostate can be reconstructed from the segmented results, which is useful for morphological analysis, volume calculation, etc. Moreover, such workflow can be integrated into our previously developed image-guided surgical system [39,40] to improve its efficiency and automation.

V. Conclusions

:section5

In this study, we introduced a novel registration-based algorithm that combines CNN and SSM and applied it to the task of precise prostate segmentation. A two branches structure was designed, through which the prior knowledge introduced by SSM and boundary features

extracted by the CNN were fully used for prostate segmentation. Extensive experimental results conducted on two public datasets demonstrated that the proposed network can achieve better performance than several state-of-the-art algorithms for prostate segmentation.

Acknowledgments

This work was supported by grants from the National Natural Science Foundation of China (81971709; M-0019; 82011530141), PHC CAI YUANPEI Program(41366SA), Agence Nationale de la Recherche (Investissements d'Avenir Program – grants ANR-11-LABX-0004 and ANR-19-P3IA-0003), Région Auvergne Rhône-Alpes (ProNavIA project), the Foundation of Science and Technology Commission of Shanghai Municipality (19510712200; 20490740700), Shanghai Jiao Tong University Foundation on Medical and Technological Joint Science Research (YG2019ZDA06; YG2021ZD21; YG2021QN72; YG2022QN056), and Hospital Funded Clinical Research(21XJMR02), Xinhua Hospital Affiliated to Shanghai Jiao Tong University, School of Medicine.

Conflict of Interest

The authors have no relevant conflicts of interest to disclose.

References

b1 [1] Futterer JJ, Briganti A, De Visschere P, et al. Can Clinically Significant Prostate Cancer Be Detected with Multiparametric Magnetic Resonance Imaging? A Systematic Review of the Literature. *European Urology*. 2015;68(6):1045-1053.

b2 [2] Litjens G, Debats O, Barentsz J, Karssemeijer N, Huisman H. Computer-Aided Detection of Prostate Cancer inMRI. *IEEE Transactions on Medical Imaging*. 2014;33(5):1083-1092.

b3 [3] Khallaghi S, Sanchez CA, Rasoulian A, et al. Statistical Biomechanical Surface Registration: Application to MR-TRUS Fusion for Prostate Interventions. *IEEE Transactions on Medical Imaging*. 2015;34(12):2535-2549.

b4 [4] Yu L, Yang X, Chen H, Qin J, Heng P-A, Aaai. Volumetric ConvNets with Mixed Residual Connections for Automated Prostate Segmentation from 3D MR Images. Paper presented at: 31st AAAI Conference on Artificial Intelligence; 2017 Feb 04-09, 2017; San Francisco, CA.

b5 [5] Martin S, Daanen V, Troccaz J. Atlas-based prostate segmentation using an hybrid registration. *International Journal of Computer Assisted Radiology and Surgery*. 2008;3(6):485-492.

b6 [6] Martin S, Troccaz J, Daanen V. Automated segmentation of the prostate in 3D MR images using a probabilistic atlas and a spatially constrained deformable model. *Medical Physics*. 2010;37(4):1579-1590.

b7 [7] Korsager AS, Fortunati V, van der Lijn F, et al. The use of atlas registration and graph cuts for prostate segmentation in magnetic resonance images. *Medical Physics*. 2015;42(4):1614-1624.

b8 [8] Tian Z, Liu L, Zhang Z, Fei B. Superpixel-Based Segmentation for 3D Prostate MR Images. *IEEE Transactions on Medical Imaging*. 2016;35(3):791-801.

b9 [9] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Networks. *Communications of the Acm*. 2020;63(11):139-144.

b10 [10] Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE transactions on pattern analysis and machine intelligence*. 2017;39(6):1137-1149.

b11 [11] Pereira S, Pinto A, Alves V, Silva CA. Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images. *IEEE Transactions on Medical Imaging*. 2016;35(5):1240-1251.

b12 [12] Tajbakhsh N, Shin JY, Gurudu SR, et al. Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Transactions on Medical Imaging*. 2016;35(5):1299-1312.

b13 [13] Guo Y, Gao Y, Shen D. Deformable MR Prostate Segmentation via Deep Feature Learning and Sparse Patch Matching. *IEEE Transactions on Medical Imaging*. 2016;35(4):1077-1089.

b14 [14] Mun J, Jang W-D, Sung DJ, Kim C-S, Ieee. Comparison of objective functions in cnn-based prostate magnetic resonance image segmentation. Paper presented at: 24th IEEE International Conference on Image Processing (ICIP); 2017 Sep 17-20, 2017; Beijing, PEOPLES R CHINA.

b15 [15] Jia H, Xia Y, Cai W, Fulham M, Feng DD, Ieee. Prostate segmentation in mr images using ensemble deep convolutional neural networks. Paper presented at: IEEE 14th International Symposium on Biomedical Imaging (ISBI) - From Nano to Macro; 2017 Apr 18-21, 2017; Melbourne, AUSTRALIA.

b16 [16] He B, Xiao D, Hu Q, Jia F. Automatic Magnetic Resonance Image Prostate Segmentation Based on Adaptive Feature Learning Probability Boosting Tree Initialization and CNN-ASM Refinement. *IEEE Access*. 2018;6:2005-2015.

b17 [17] Wang B, Lei Y, Tian S, et al. Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation. *Medical Physics*. 2019;46(4):1707-1718.

b18 [18] Shen DG, Herskovits EH, Davatzikos C. An adaptive-focus statistical shape model for segmentation and shape modeling of 3-D brain structures. *IEEE Transactions on Medical Imaging*. 2001;20(4):257-270.

b19 [19] Seim H, Kainmueller D, Heller M, Lamecker H, Zachow S, Hege H-C. Automatic Segmentation of the Pelvic Bones from CT Data Based on a Statistical Shape Model. *VCBM*. 2008;8:93-100.

b20 [20] Zhang X, Tian J, Deng K, Wu Y, Li X. Automatic Liver Segmentation Using a Statistical Shape Model With Optimal Surface Detection. *IEEE Transactions on Biomedical Engineering*. 2010;57(10):2622-2626.

b21 [21] Alba X, Pereanez M, Hoogendoorn C, et al. An Algorithm for the Segmentation of Highly Abnormal Hearts Using a Generic Statistical Shape Model. *IEEE Transactions on Medical Imaging*. 2016;35(3):845-859.

b22 [22] Lüthi M, Blanc R, Albrecht T, et al. Statismo-A framework for PCA based statistical models. *The Insight Journal*. 2012;2012:1-18.

b23 [23] Szegedy C, Liu W, Jia Y, et al. Going Deeper with Convolutions. Paper presented at: IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2015 Jun 07-12, 2015; Boston, MA.

b24 [24] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Paper presented at: 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI); 2015 Oct 05-09, 2015; Munich, GERMANY.

b25 [25] Jaderberg M, Simonyan K, Zisserman A. Spatial transformer networks. *Advances in neural information processing systems*. 2015;28:2017-2025.

b26 [26] Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca AV. Voxelmorph: a learning framework for deformable medical image registration. *IEEE transactions on medical imaging*. 2019;38(8):1788-1800.

b27 [27] Litjens G, Toth R, van de Ven W, et al. Evaluation of prostate segmentation algorithms for MRI: The PROMISE12 challenge. *Medical Image Analysis*. 2014;18(2):359-373.

b28 [28] Bloch N, Madabhushi A, Huisman H, et al. NCI-ISBI 2013 challenge: automated segmentation of prostate structures. *The Cancer Imaging Archive*. 2015;370:6.

b29 [29] Heimann T, Meinzer H-P. Statistical shape models for 3D medical image segmentation: A review. *Medical Image Analysis*. 2009;13(4):543-563.

b30 [30] Ou Y, Doshi J, Erus G, Davatzikos C. Multi-atlas segmentation of the prostate: A zooming process with robust registration and atlas selection. *Medical Image Computing and Computer Assisted Intervention (MICCAI) Grand Challenge: Prostate MR Image Segmentation*. 2012;7:1-7.

b31 [31] Gao Q, Rueckert D, Edwards P. An automatic multi-atlas based prostate segmentation using local appearance-specific atlases and patch-based voxel weighting. *MICCAI Grand Challenge: Prostate MR Image Segmentation*. 2012;2012.

b32 [32] Maan B, van der Heijden F. Prostate MR image segmentation using 3D active appearance models. *MICCAI Grand Challenge: Prostate MR Image Segmentation*. 2012;2012.

b33 [33] Vincent G, Guillard G, Bowes M. Fully automatic segmentation of the prostate using active appearance models. *MICCAI Grand Challenge: Prostate MR Image Segmentation*. 2012;2012:2.

b34 [34] Karimi D, Samei G, Kesch C, Nir G, Salcudean SE. Prostate segmentation in MRI using a convolutional neural network architecture and training strategy based on statistical shape models. *International Journal of Computer Assisted Radiology and Surgery*. 2018;13(8):1211-1219.

b35 [35] Milletari F, Navab N, Ahmadi S-A, Ieee. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation. Paper presented at: 4th IEEE International Conference on 3D Vision (3DV); 2016 Oct 25-28, 2016; Stanford Univ, Stanford, CA.

b36 [36] Toth R, Madabhushi A. Deformable landmark-free active appearance models: application to segmentation of multi-institutional prostate MRI data. *MICCAI Grand Challenge: Prostate MR Image Segmentation*. 2012;2012.

b37 [37] Tian Z, Liu L, Zhang Z, Fei B. PSNet: prostate segmentation on MRI based on a convolutional neural network. *Journal of Medical Imaging*. 2018;5(2).

b38 [38] Jia H, Xia Y, Song Y, Cai W, Fulham M, Feng DD. Atlas registration and ensemble deep convolutional neural network-based prostate segmentation using magnetic resonance imaging. *Neurocomputing*. 2018;275:1358-1369.

b39 [39] Qin C, Cao Z, Fan S, et al. An oral and maxillofacial navigation system for implant placement with automatic identification of fiducial points. *International journal of computer assisted radiology and surgery*. 2019;14(2):281-289.

b40 [40] Tu P, Qin C, Guo Y, et al. Ultrasound image guided and mixed reality-based surgical system with real-time soft tissue deformation computing for robotic cervical pedicle screw placement. *IEEE Transactions on Biomedical Engineering*. 2022. doi:10.1109/TBME.2022.3150952.