IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 27, 2021, accepted July 8, 2021, date of publication July 12, 2021, date of current version August 4, 2021.

*Digital Object Identifier 10.1109/ACCESS.2021.3096665*

# ProSegNet: A New Network of Prostate Segmentation Based on MR Images

**YUEJING QIAN**
Zhejiang Industry and Trade Vocational College, Wenzhou, Zhejiang 325000, China
e-mail: yz_qian@163.com

This work was supported by the Natural Science Foundation of Zhejiang Province under Grant LGF19F020004.

**ABSTRACT** Prostate cancer is the most common cancer in men after lung cancer. Generally, the segmentation of the prostate is the preprocessing work for the diagnosis of prostate cancer. Aiming at the variety of prostate and the similarity of visual characteristics between prostates and their surroundings, this paper proposes a new prostate segmentation network based on MR images, denoted as ProSegNet. ProSegNet consists of two parts: encoder and decoder. To improve the feature extraction capability of the encoder, we use dense blocks as the feature extraction unit, and at the same time introduce a cross-stage partial (CSP) structure to reduce the amount of calculation. In the design of the decoder structure, we integrated the spatial attention mechanism and the channel attention mechanism to enable it to focus on the important features while ignoring the invalid features. In addition, to segment the prostate more accurately, we add a prostate contour segmentation branch to the output of the segmentation network to learn the contour features of the prostate. Finally, to alleviate the problem of small intensity difference between the prostate and surrounding tissues, we designed a truncated intensity stretching image enhancement method. The performance of ProSegNet has been experimentally verified in the Promise12 and ProstateX datasets. On the Promise12 dataset, the dice similarity coefficient (DSC) and hausdorff (Haus) distance are 0.908 and 9.87 respectively. On the ProstateX dataset, the DSC and Haus reach the results of 0.892 and 10.45, respectively. Experimental results show that the ProSegNet can obtain a competitive performance.

**INDEX TERMS** Prostate segmentation, magnetic resonance image, convolutional neural network, attention mechanism, multi-task learning.

## I. INTRODUCTION

The prostate is the largest substantial organ in the accessory glands of the male genitalia, which has the physiological function of secreting and storing prostate fluid. It is located at the bottom of the pelvic cavity, with a bladder on the top, a pubic bone in the front, and a rectum in the back [1]. It is in a relatively secret location in the human body and is difficult to directly observe and touch. The size and appearance of the prostate resemble an inverted and slightly flat chestnut, with a wide upper end, a narrow lower end and a flat back. Most adult men have a prostate that weighs 7 to 16 grams and is about 20 cm in size. In clinical diagnosis, if the volume of the prostate exceeds 30 cm, it can be diagnosed as the benign prostatic hyperplasia [2].

Due to the bad living habits of modern people, such as prolonged sitting for a long time, excessive drinking, irregular work and rest, etc., the incidence of prostate disease is increasing year by year. Prostate disease is common in middle-aged and elderly people, but in recent years, the affected population has gradually shown a younger trend [3]. High-incidence diseases of the prostate include prostatitis, benign prostatic hyperplasia, and prostate cancer. In particular, prostate cancer has become the second leading cause of death among American men [4]. According to statistics, prostate cancer is the most common male cancer in 84 countries, with a high incidence in developed countries, and the number of patients in developing countries is also increasing year by year [5].

Common diagnostic methods for prostate cancer include serum prostate specific antigen (PSA) examination, transrectal ultrasound (TRUS), computer tomography (CT) image examination, magnetic resonance (MR) image inspection, biopsy inspection, etc [6]. Serum PSA is one of the serum markers of prostate cancer, but prostate enlargement caused by diseases such as prostatitis and benign prostatic

The associate editor coordinating the review of this manuscript and approving it for publication was Larbi Boubchir.

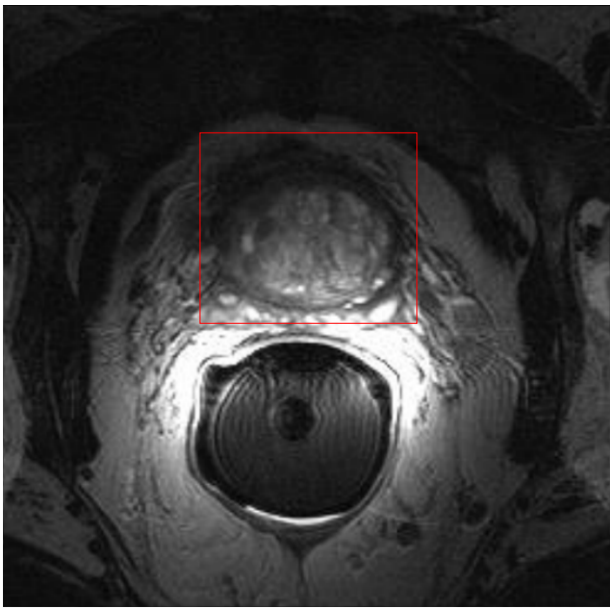

**FIGURE 1.** The imaging characteristics of prostate in MR images.

hyperplasia can also cause serum PSA to rise. Therefore, serum PSA is highly sensitive to prostate cancer but has low specificity [7]. TRUS, CT images, and MR image inspections are all imaging detection technologies. TRUS is fast and low cost, but the imaging quality is poor [8]; CT images also have the problem of low imaging quality of human soft tissues; compared to TRUS and CT, MR images have a clearer imaging effect on human soft tissues [9]. In vivo puncture, usually after the aforementioned multiple detection methods determine the presence of suspicious lesions, as the final diagnosis method, the puncture process will cause greater pain to the patient. In summary, MR images as an effective detection method have been widely used in the detection of prostate cancer [10].

Due to the morphological and positional characteristics of the prostate, it is difficult to accurately identify the prostate even in MR images. Fig. 1 shows the imaging characteristics of the prostate in the MR image. It can be observed that the prostate has the characteristics of small size, closeness to surrounding tissues, and unclear borders, which brings certain difficulties to doctors' clinical observation. Manual labeling of the prostate not only requires certain clinical knowledge, but also consumes a lot of time and energy for doctors. In recent years, thanks to the development of hardware, the deep learning method represented by Convolutional Neural Network (CNN) [11] has made breakthrough progress. Based on the good performance of CNN in the field of image processing, it has been applied to solve computer vision problems on a large scale. As a typical medical image data, MR image can also be segmented using CNN method to obtain more accurate segmentation results [12]–[15].

This paper proposes a new method of prostate segmentation based on MR images, denoted as ProSegNet. ProSegNet can automatically segment the prostate from the MR image and accurately determine the position, boundary, and volume of the prostate, so as to help clinicians quickly locate the prostate, identify abnormal shapes, and calculate the PSA concentration based on the volume. Besides, prostate segmentation is also the pre-processing work for prostate lesion detection and benign and malignant judgment. In general, our technical contributions in this work mainly include the following four points.

(1) A truncated intensity stretching image enhancement method is designed to alleviate the problem of small intensity differences between the prostate and surrounding tissues.

(2) The dense connection structure and the cross-stage partial (CSP) network structure are applied to the encoder to improve the feature extraction ability of the encoder without increasing the amount of calculation.

(3) Combine the spatial attention mechanism and channel attention mechanism, and apply it to the decoder, so that the decoder can better focus on the effective features.

(4) To obtain more accurate segmentation results, we add a prostate contour segmentation branch to the output of the segmentation network to learn the contour features of the prostate and further refine the segmentation results.

## II. RELATED WORK

Regarding the existing prostate segmentation methods, this paper divides them into two categories to explain, namely, the prostate segmentation method based on traditional image processing and the prostate segmentation method based on deep learning.

### A. PROSTATE SEGMENTATION METHOD BASED ON TRADITIONAL IMAGE PROCESSING

Zwiggelaar *et al.*, proposed a semi-automatic prostate MR image segmentation method, which uses prostate anatomical features to convert the original image into a polar coordinate transformation space for representation, and then use methods such as boundary tracking and non-maximum suppression to finally determine the boundary of the prostate [16]. Similarly, Samiee *et al.*, divided the prostate into four quadrants and utilized the prior knowledge of the prostate boundary in different quadrants to track the prostate boundary [17]. Tapia *et al.*, applied the different manifestations of signal singularities and noise in the wavelet domain, and tracked the prostate boundary based on the prior knowledge of spatial variation rules [18]. Moreover, Cootes *et al.*, published a prostate segmentation method based on Active Shape Model (ASM), which designed a shape template based on statistical data, and then used ASM for prostate contour segmentation [19]. Klein *et al.*, designed a prostate segmentation method based on atlas non-rigid registration. The method first non-rigid registration of the image in the atlas with the target patient image, and then fused the tags of the atlas to obtain the segmentation result [20]. In a similar manner, Dowling *et al.*, also designed a prostate segmentation method based on multi-atlas. This method first performs non-rigid

registration, and then dynamically merges the multiple atlases to further improve the accuracy of prostate segmentation [21].

To make better use of the shape information of the prostate, Vikal *et al.*, constructed a semi-automatic prostate segmentation method. This method first uses template matching in one slice to find the outline of the prostate, and then spreads to other slices to obtain a three-dimensional prostate structure [22]. Zhang *et al.*, also designed a semi-automatic prostate segmentation method, which specifies the foreground and background in a human-computer interaction manner, and performs prostate segmentation based on the active contour model [23]. Further, Toth *et al.*, presented an ASM-based segmentation method, which initializes the ASM by automatically identifying the prostate spectrum in the MR image, which improves the problem of manually initializing the ASM. And by clustering the segmentation results of multiple slices, the expansion from 2D segmentation to 3D segmentation is realized [24]. Singh *et al.*, used Chan-Vese active contour model and morphological operations to segment the prostate[25]. In addition, Martin *et al.*, proposed a prostate segmentation method based on atlas and key point matching, which segment the prostate by registering images of multiple patients with images of the target patient [26]. Firjany *et al.*, adopted a graph cut method to segment the prostate. This method optimizes the energy function of the first-order image operator, the second-order spatial variant homogeneity operator, and the prostate shape operator to segment the prostate in the MR image [27]. Richard *et al.*, applied the clustering method to the prostate segmentation method. They clustered each pixel based on the features extracted by the four texture energy measurement methods, thereby marking each pixel in the image as prostate tissue or non-prostatic tissue [28].

### B. PROSTATE SEGMENTATION METHOD BASED ON DEEP LEARNING

Allen *et al.*, published an automatic prostate segmentation method, which classifies each voxel point in the peripheral zone and the central zone, and uses a 3D statistical shape modeling method to segment the prostate [29]. Makni *et al.*, used deformable models and probabilistic frames to segment the prostate. They first use the statistical shape model as prior knowledge, then use the Gaussian mixture model for histogram matching, and finally perform label optimization based on Bayesian posterior classification [30]. Further, Mohamed *et al.*, proposed a supervised learning method of prostate segmentation [31]. This method extracts and analyzes the spectral features of the prostate in the TRUS image, then uses the Gabor filter [32] and the image's frequency and spatial features to locate the prostate, and finally uses the support vector machine [33] to classify the voxel points in the prostate region. Liu *et al.*, proposed an unsupervised prostate cancer segmentation method [34], which uses fuzzy Markov random field [35] to segment the prostate in MR images. Khurd *et al.*, first used the classifier trained by the GMM-EM algorithm to locate the center of the prostate, and then combined the shape model and MLRW to segment the prostate [36]. Moreover, Mahapatra *et al.*, reported an automatic prostate segmentation method based on random forest (RF) and graph cuts. They use supervoxel segmentation to automatically select regions of interest, and then use image features and RF classifiers to classify them [37].

In view of the wide application of the CNN method on natural images, many CNN-based prostate segmentation methods have been published in recent years. For example, Zhu *et al.*, designed a prostate segmentation method based on a deep convolutional neural network. This method deepens the network depth of CNN to extract advanced semantic features, and uses Dicecoefficient as the loss function to improve the uneven distribution of foreground and background in the image [38]. Karimi *et al.*, proposed a prostate segmentation method based on CNN and statistical shape model, which uses CNN to predict the center position of the prostate and the parameters of the shape model to determine the location of key points on the surface of the prostate [39]. Similarly, Moradi *et al.*, first utilized CNN to locate the region of the prostate, then adopted the probability atlas to obtain the initial segmentation results, and finally applied the statistical shape model to limit the contour to the allowed shape [40]. Clark *et al.*, used a fully convolutional neural network (FCNN) to segment the prostate in diffusion-weighted imaging (DWI). They first locate the prostate in the three-dimensional DWI image and then perform segmentation [41]. Besides, Tian *et al.*, utilized a deep FCNN to automatically segment the prostate, which is essentially a pixel-level classification network [42]. Jia *et al.*, first used image registration to perform coarse segmentation of the prostate; then trained a pixel-level CNN classification model to predict whether the pixels in the candidate region are prostate pixels; finally introduced ensemble learning for fine segmentation [43]. Zhang *et al.*, designed a new CNN structure, denoted as Z network, to segment the prostate from MR images. The Z network can capture more effective features by using cascading and dense connections [44]. Yaniv *et al.*, segmented the prostate based on 3D V-net [45] and optimized the speed and memory limitations of this network [46]. Additionally, Jin *et al.*, applied bicubic interpolation to preprocess the low frequency part of the prostate MR image, and then used an improved 3D V-Net (3D PBV-Net) to segment the prostate [47]. Silva *et al.*, proposed a two-stage prostate segmentation method. They first classified the prostate and non-prostatic tissues based on clustering and probability atlas in the CNN model combined with particle swarm optimization algorithm, and then adopted the 3D Chan-Vese active contour model to obtain the prostate mask [48].

## III. THE PROPOSED METHOD

The prostate segmentation method proposed in this paper will be described in detail below. The method is mainly composed of three parts, namely: data preprocessing, segmentation network design and network model training.

### A. TRUNCATED INTENSITY STRETCHING IMAGE ENHANCEMENT

The prostate MR image contains a variety of image sequences, such as T2 weighted sequence, DWI image, ADC image, and KTrans image. Because the resolution of T2-weighted images is high, and the tissue structure can be clearly observed in T2-weighted images, we will perform prostate segmentation in T2-weighted images. To make the prostate region clearer in the T2-weighted image, we first use Eqs.(1)-(3) to stretch the contrast of the T2-weighted image. Then use contrast limited adaptive histogram equalization (CLAHE) [49] for image enhancement, where the clipping limit is set to 0.05.

$$\min_{value} = Sort(gray_{list})\,[len(gray_{list})/100] \quad (1)$$

$$max_{value} = Sort(gray_{list})\,[len(gray_{list})/100 \times 99] \quad (2)$$

$$\text{stretch}_{img} = \frac{gray_{img} - \min_{value}}{\max_{value} - \min_{value}} \times 255 \quad (3)$$

where "Sort" means ascending sorting; "$\text{gray}_{\text{img}}$" means the original intensity image; "$\text{gray}_{\text{list}}$" means the one-dimensional intensity list corresponding to "$\text{gray}_{\text{img}}$"; "$\text{len}(\text{gray}_{\text{list}})$" indicates the length of the list "$\text{gray}_{\text{list}}$"; "/" indicates division operation; "L[i]" indicates the i-th element of list L, and "$\text{stretch}_{\text{img}}$" indicates the image after intensity stretch.

### B. ProSegNet: PROSTATE SEGMENTATION NETWORK

The structure diagram of the prostate segmentation network (ProSegNet) designed in this paper is shown in Fig. 2. The main structure of the network is similar to U-net [50]. The left side of ProSegNet is the encoder, which is a CSP structure based on dense blocks; the right side is the decoder, which combines the spatial attention mechanism and the channel attention mechanism; the middle is the skip connection structure. The output of ProSegNet has two prediction branches, one is the mask prediction branch and the other is the contour prediction branch.

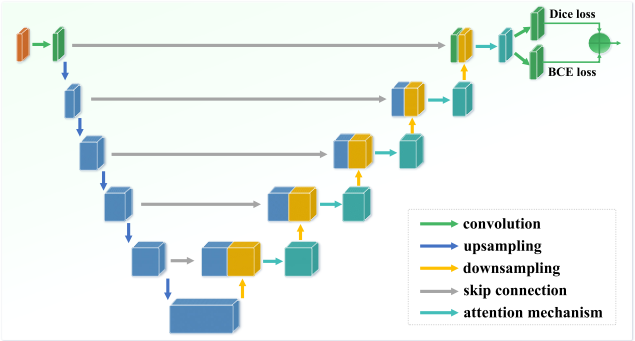


**FIGURE 2.** The structure diagram of ProSegNet.

#### 1) ENCODER BASED ON CROSS-STAGE PARTIAL

To improve the feature extraction capability of the encoder, we introduced the feature extraction unit DenseBlock [51] proposed by Huang *et al.* DenseBlock repeatedly concatenates the feature map of the current layer to each subsequent layer to better extract features. However, this complex network structure will lead to an increase in the amount of calculation. In response to this problem, literature [52] reported an efficient cross-stage partial network structure. The feature map is first divided into two parts, so that the gradient can be propagated back from different network paths, and then the two parts are merged. This can not only achieve richer gradient combinations, but also reduce the amount of calculation and improve the speed and accuracy of inference. We combine the CSP structure with DenseBlock, and the specific structure diagram is shown in Fig. 3. Among them, each letter in CBR represents a network layer, C represents the convolutional layer, B represents the batch normalization layer, R represents the Relu activation function, and CBR represents a module composed of three network layers of C, B and R connected in turn. The DenseBlock structure includes 6 interconnected BottleNeck structures, and the number of channels k of the feature map output by each BottleNeck is a fixed value of 32. The Transition Layer in DenseBlock consists of a $1 \times 1$ Conv and a $2 \times 2$ average pooling layer with a stride size of 2.

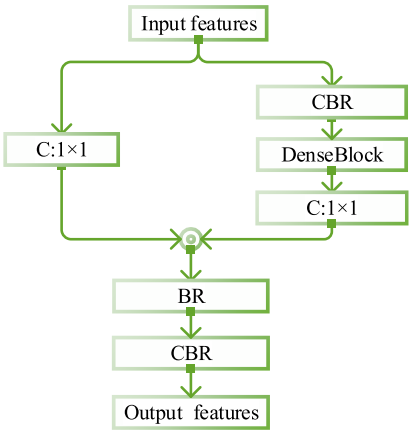


**FIGURE 3.** CSP structure based on DenseBlock.

The input feature is divided into left and right paths. The right path first passes through the CBR module, the channel dimension is reduced to half of the original, and then passes through the DenseBlock and the $1 \times 1$ convolution layer. The left path reduces the channel dimension to half of the original through a $1 \times 1$ convolutional layer, and is spliced with the right path, and then passes through a BR module and a CBR module.

#### 2) DECODER BASED ON ATTENTION MECHANISM

In the decoder, a deconvolution layer with a stride size of 2 is used for upsampling to gradually restore the size of the feature map. Then, the feature map of the decoder is spliced with the feature map of the same size from the encoder through the skip connection structure. After that, through

two convolutional layers for feature extraction, in order to restore the feature information lost due to downsampling. Finally, they are sent to the attention module to obtain more representative features.

In recent years, the channel attention mechanism has been one of the research hotspots in the field of semantic segmentation. This structure assigns weights to each feature channel to increase the proportion of important feature channels, so as to obtain the features that require the most attention. In 2017, Hu *et al.*, proposed SENet [53], which through explicit modeling of the interdependence between channels, adaptively calibrate the channel feature response, and enhance the expressive ability of the network. In 2018, Yu *et al.*, proposed a Smooth Network (SN), and applied the channel attention mechanism in the network, which alleviated the problem of intra-class inconsistency in semantic segmentation to a certain extent [54]. We know that different regions in the image contribute differently to the task. The spatial attention mechanism is designed to find and highlight important regions in the image. Among them, the representative paper is the Spatial Transformer Network (STN) proposed by Google DeepMind. STN completes the preprocessing operation of the task by learning the deformation of the input image [55].

The spatial channel attention module designed in this paper is divided into two parts, namely the spatial attention module and the channel attention module. The corresponding structure diagram is shown in Fig. 4. The spatial channel attention module is different from the SCA-CNN [56] designed by Chen *et al.* The SCA in SCA-CNN is a cascaded structure, that is, the features output by the feature extractor pass through the spatial attention module, and then its output is sent to the channel attention module. The spatial channel attention module is a dual-branch structure, that is, the features output by the feature extractor will be sent to the channel and the spatial attention module respectively, and then their output results will be merged.

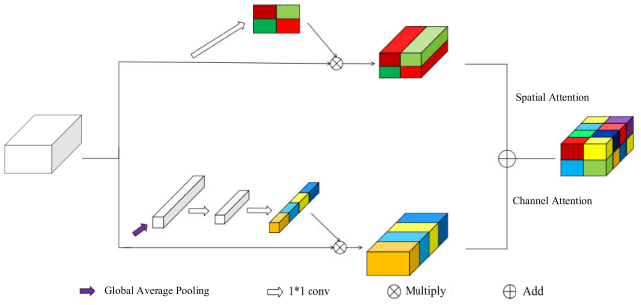


**FIGURE 4.** Spatial channel attention mechanism.

Specifically, when the input feature map enters the spatial attention branch, the feature map is first reduced to a one-dimensional feature map through $1 \times 1$ Conv, which represents the attention weight of pixels at different locations. Then the weight and the original feature map are multiplied pixel by pixel to obtain the spatial attention feature map. When the input feature map enters the channel attention branch, it is reduced to a feature map with only one pixel per channel through global average pooling, and each pixel value represents the attention weight of the channel. Then multiply it with the original feature map to obtain the channel attention feature map. Finally, the spatial attention feature map and the channel attention feature map are added pixel by pixel to obtain the spatial channel attention feature map.

### C. MULTI-TASK JOINT TRAINING

Image segmentation is generally for complete mask segmentation, and the same is true for prostate segmentation, that is, the network only learns complete prostate mask features. However, the internal pixels of the prostate occupy most of the area of the entire mask, which will cause the neural network to pay too much attention to learning the internal features of the prostate and ignore the learning of the contour features of the prostate. The contour of the prostate is blurred and difficult to observe and segment. If the network can learn the contour characteristics of the prostate, it can complete the segmentation of the prostate more accurately. Meanwhile, the contours of most prostates are similar. By learning the contour features of the prostate, a shape constraint can be formed to alleviate the problem of over-segmentation.

To learn the contour features of the prostate, a contour prediction branch is added to the output of the network. During training, the neural network predicts the complete mask and contour mask of the prostate at the same time, thereby simultaneously learning the internal features and contour features of the prostate.

We use the binary cross entropy loss (BCELoss) to calculate the loss of the contour branch, and use the dice loss (DiceLoss) to calculate the loss of the mask branch. The general form of BCEloss is defined as shown in Eq.(4), where $y$ is the true label value and o is the predicted value.

$$\mathrm{B}CELoss = -\frac{1}{2}\sum [y \ln o + (1-y)\ln(1-o)] \tag{4}$$

The DiceLoss is defined as shown in Eq.(5), where G is the real mask image and P is the predicted mask image.

$$\mathrm{DiceLoss} = 1 - \frac{2 \times |G \cap P|}{|G| \cup |P|} \tag{5}$$

The loss function (SegLoss) used in training is the weighted sum of BCELoss and DiceLoss, as shown in Eq.(6).

$$\mathrm{Seg}Loss = \alpha \times BCELoss + \beta \times DiceLoss \tag{6}$$

According to the experimental results of adjusting the values of $\alpha$ and $\beta$ five times, it is recommended to set the values of $\alpha$ and $\beta$ to 1.0 and 1.2, respectively.

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

Due to the small scale of the prostate dataset, overfitting is inevitable. To alleviate this problem, we adopted a training strategy of transfer learning. The model trained on the ImageNet dataset will be used as a pre-training model, and then the model will be retrained using prostate MR image data.

Meanwhile, in order to further alleviate the over-fitting problem, the training strategy of early stopping is used [57]. In this strategy, the segmentation performance is not improved, and the number of epochs to continue training is 10. In addition, the optimizer used during training is Adam, the initial learning rate is 0.001, the decay rate is 0.5, the batch size is 16, and the total number of training generations is 30. Finally, the image size is scaled to $256 \times 256$, and the intensity distribution interval is mapped to [0,1] for model training.

### A. DATA

The experimental data used in this paper comes from two public datasets: PROMISE12 [58], [59] and ProstateX [60], [61].

PROMISE12 (Prostate MR Image Segmentation) is a prostate segmentation competition organized by the International Medical Image Processing Organizing Committee in 2012. The dataset provided by the competition contains 50 training samples and 30 testing samples. Each training sample contains the T2 weighted sequence of the prostate MR image and its corresponding prostate mask. Each testing sample only provides the T2-weighted sequence of the MR image, and the prostate mask is annotated by a professional radiologist who cooperates with us.

ProstateX is a medical image recognition competition jointly initiated by the American Physicists Association, the International Society of Optics and Photonics, and the National Cancer Institute. The dataset provided by the competition contains multiple sequences of MR images, such as T2-weighted images and proton density-weighted images. The training set contains 204 patients with a total of 330 lesions, and the testing set contains 142 patients with a total of 208 lesions. Furthermore, the label file provided by ProstateX only contains the location of the lesion and the benign and malignant information, and no mask information of the prostate. To verify the generalization performance of the prostate segmentation algorithm, we asked a professional radiologist to mark the prostate mask on the MR images of 40 patients randomly selected from the ProstateX dataset.

### B. EVALUATION CRITERIA

There are many indicators to evaluate the performance of image segmentation algorithms. We use the four most commonly used indicators, Dice Similarity Coefficient (DSC)[62], Hausdorff (Haus) distance[63], recall (REC) rate and precision (PRE) rate to evaluate the performance of the prostate segmentation algorithm.

DSC is a variation of Intersection of Union (IOU). The larger the value, the better the segmentation accuracy. The calculation formulas for DSC, SEN and PRE are as shown in Eqs.(7)-(9) respectively.

$$\mathrm{DSC} = \frac{2 \times S(P \cap G)}{S(P) + S(G)} \tag{7}$$

$$\mathrm{REC} = \frac{S(P \cap G)}{S(G)} \tag{8}$$

$$\mathrm{PRE} = \frac{S(P \cap G)}{S(P)} \tag{9}$$

where P represents the predicted mask region, G represents the region corresponding to the ground truth, and S(*) represents the area counted in pixels.

To quantify the degree of agreement between the boundaries of the two regions, Haus is introduced as another evaluation indicator. The calculation formula of Haus is as shown in Eq.(10).

$$\mathrm{Haus(P,G)} = \max\left\{\sup_{\mathrm{p}\in\mathrm{P}} \inf_{g\in G} d(p,g), \sup_{\mathrm{g}\in\mathrm{G}} \inf_{p\in P} d(p,g)\right\} \tag{10}$$

The meanings of P and G are the same as in Eq. (7), d(i, j) represents the Euclidean distance of pixels i and j, sup(supremum) and inf(infimum) represent the supremum and infimum respectively.

### C. OVERALL PERFORMANCE

To evaluate the performance of the proposed prostate segmentation method ProSegNet reasonably and accurately, this paper adopts a 5-fold cross-validation method for evaluation. Besides, we train the model on the training set of the PROMISE12 dataset. Table 1 shows the results of the 5-fold cross-validation of ProSegNet on the two testing sets corresponding to the ProstateX and PROMISE12 datasets.

**TABLE 1. ProSegNet's 5-fold cross-validation results on the testing sets corresponding to the ProstateX and PROMISE12 datasets.**

| 5-fold cross validation | Dataset | DSC | Haus | REC | PRE |
|---|---|---|---|---|---|
| 1 | ProstateX | 0.899 | 10.28 | 0.931 | 0.888 |
| | PROMISE12 | 0.915 | 9.30 | 0.936 | 0.897 |
| 2 | ProstateX | 0.884 | 10.71 | 0.951 | 0.859 |
| | PROMISE12 | 0.902 | 10.10 | 0.960 | 0.865 |
| 3 | ProstateX | 0.898 | 10.31 | 0.943 | 0.874 |
| | PROMISE12 | 0.912 | 9.36 | 0.952 | 0.882 |
| 4 | ProstateX | 0.896 | 10.29 | 0.936 | 0.882 |
| | PROMISE12 | 0.913 | 9.25 | 0.941 | 0.890 |
| 5 | ProstateX | 0.884 | 10.68 | 0.944 | 0.867 |
| | PROMISE12 | 0.899 | 10.29 | 0.957 | 0.881 |
| Average | ProstateX | 0.892 | 10.45 | 0.941 | 0.874 |
| | PROMISE12 | 0.908 | 9.87 | 0.949 | 0.883 |

Through the analysis of the experimental data in Table 1, it can be seen that the DSC of ProSegNet on the testing set corresponding to the ProstateX dataset has been as high as 89%, and the DSC on the testing set of PROMISE12 has exceeded 90%. Additionally, it can be seen that the performance of ProSegNet on the testing set corresponding to the PROMISE12 dataset is better than that of the ProstateX dataset. This is because the ProSegNet model is trained on the PROMISE12 training set, so the performance of ProSegNet on the PROMISE12 testing set is higher than ProstateX. Fig. 5 shows the visualization results of 4 samples randomly selected from the testing set of the PROMISE12 and ProstateX datasets. Among them, the red curve represents the boundary of ground truth, and the yellow curve represents the boundary of the prediction result. It can be seen from

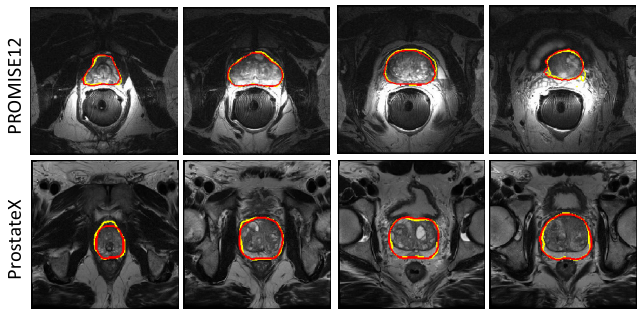


**FIGURE 5.** Visualization results of ProSegNet prostate segmentation method.

Fig. 5 that the ProSegNet segmentation method can accurately segment the prostate.

### D. ABLATION STUDY

To verify the effectiveness of each component in the ProSegNet method, we designed an ablation experiment as shown in Table 2. Table 2 shows the performance after gradually adding each component on the basis of UNet.

**TABLE 2.** Ablation experiment. Note that, "PreWeight" indicates the pre-trained model trained on ImageNet; "DataPrepro" indicates the data preprocessing method of truncated intensity stretching image enhancement; "DenseCSP" indicates the designed CSP structure based on DenseBlock; "Attention" indicates the designed spatial channel attention mechanism; "BSELoss" indicates the multi-task joint training method based on BCE loss and dice loss.

| Methods | UNet | PW-UNet | DP-UNet | DCSP-UNet | Atten-UNet | ours |
|---|---|---|---|---|---|---|
| PreWeight | | ✓ | ✓ | ✓ | ✓ | ✓ |
| DataPrepro | | | ✓ | ✓ | ✓ | ✓ |
| DenseCSP | | | | ✓ | ✓ | ✓ |
| Attention | | | | | ✓ | ✓ |
| BCEDice | | | | | | ✓ |
| REC | 0.873 | 0.881 | 0.895 | 0.914 | 0.922 | 0.941 |
| PRE | 0.814 | 0.820 | 0.829 | 0.847 | 0.856 | 0.874 |
| DSC | 0.825 | 0.832 | 0.846 | 0.865 | 0.874 | 0.892 |
| Haus | 15.76 | 14.89 | 13.19 | 12.75 | 11.81 | 10.45 |

After analyzing the ablation experiment results in Table 2, the following conclusions can be drawn:

#### 1) EFFECTIVENESS OF THE PRE-TRAINED MODEL

By comparing the second and third rows of experimental data in Table 2, the following conclusions can be found. That is, using the model trained on ImageNet as the pre-training model of the segmentation network encoder, and then training on the PROMISE12 dataset is better than training directly on the PROMISE12 training set. This experimentally verifies that for prostate segmentation tasks with less training data, it is necessary to use a pre-trained model based on ImageNet.

#### 2) EFFECTIVENESS OF THE DATA PREPROCESSING

By comparing the third and fourth rows of experimental data in Table 2, it can be found that after the introduction of the data preprocessing method of truncated intensity stretching image enhancement, both DSC and Haus have been greatly improved, especially DSC increased by 1.4 %. This shows that this data preprocessing method is effective for prostate segmentation.

#### 3) EFFECTIVENESS OF DenseBlock-BASED CSP STRUCTURE

By comparing the experimental data in the 4th and 5th rows of Table 2, we can find that the introduction of the DenseBlock-based CSP structure on the basis of UNet, although Haus did not improve much, but the DSC increased by 1.9%. Overall, the designed CSP structure based on DenseBlock helps to improve the performance of prostate segmentation.

#### 4) EFFECTIVENESS OF SPATIAL CHANNEL ATTENTION MECHANISM

Comparing the experimental data in rows 5 and 6 of Table 2, it can be found that DSC has increased by 0.9% and Haus has shrunk by 0.95 pixels. This shows that introducing the designed spatial channel attention mechanism into the decoder is effective for prostate segmentation.

#### 5) EFFECTIVENESS OF MULTI-TASK JOINT TRAINING METHOD

By comparing the last two rows of experimental data in Table 2, it can be found that after combining Dice loss and BCE loss for multi-task joint training, the performance of DSC and Haus has been steadily improved. Specifically, DSC increased by 1.8%, Haus reduced by 1.36 pixels, which proved the effectiveness of multi-task joint training.

### E. EXPERIMENTAL COMPARISON

To verify the superiority of ProSegNet segmentation method, we compare it with the prostate segmentation method published in 2018-2020. Table 3 shows the experimental comparison results of our method and other methods on the testing sets of the ProstateX and PROMISE12 datasets.

**TABLE 3.** Comparison experiments between our method and other prostate segmentation methods.

| Methods | Dataset | DSC | Haus | REC | PRE |
|---|---|---|---|---|---|
| Moradi et al. [40] | ProstateX | 0.853 | 12.76 | 0.896 | 0.828 |
| | PROMISE12 | 0.859 | 12.18 | 0.899 | 0.833 |
| Yaniv et al. [46] | ProstateX | 0.861 | 12.07 | 0.902 | 0.835 |
| | PROMISE12 | 0.874 | 11.12 | 0.914 | 0.846 |
| Tian et al. [42] | ProstateX | 0.860 | 12.23 | 0.904 | 0.832 |
| | PROMISE12 | 0.870 | 12.06 | 0.912 | 0.849 |
| Jia et al. [43] | ProstateX | 0.869 | 12.18 | 0.916 | 0.852 |
| | PROMISE12 | 0.881 | 11.26 | 0.926 | 0.857 |
| Jin et al. [47] | ProstateX | 0.880 | 11.17 | 0.924 | 0.854 |
| | PROMISE12 | 0.893 | 10.34 | 0.931 | 0.859 |
| Zhang et al. [44] | ProstateX | 0.878 | 11.79 | 0.921 | 0.849 |
| | PROMISE12 | 0.889 | 10.55 | 0.932 | 0.861 |
| Silva et al. [48] | ProstateX | 0.883 | 11.02 | 0.930 | 0.863 |
| | PROMISE12 | 0.897 | 10.13 | 0.937 | 0.872 |
| Ours | ProstateX | 0.892 | 10.45 | 0.941 | 0.874 |
| | PROMISE12 | 0.908 | 9.87 | 0.949 | 0.883 |

It is not difficult to see from Table 3 that on the two datasets of ProstateX and PROMISE12, the prostate segmentation method proposed in this paper shows excellent performance whether it is compared with the 2D CNN or 3D CNN-based prostate segmentation methods. It should be noted that, to ensure the fairness of comparison, the methods based on 2D CNN all use the pre-training model on ImageNet and the data preprocessing method of truncated intensity stretching image enhancement. Since 3D CNN cannot be pre-trained on ImageNet, they only use the data preprocessing method designed in this paper. In addition, for the training and testing of the model, we use an image resolution of 256 × 256. Including the methods listed in the experimental comparison, the image resolution of 256 × 256 size is also used. To further verify the superiority of the proposed method, we randomly selected 4 samples from the testing set for visual comparison. Fig. 6 shows the results of visual comparison between our method and other methods. It can be seen from Fig. 6 that our method in this paper is better than other methods.

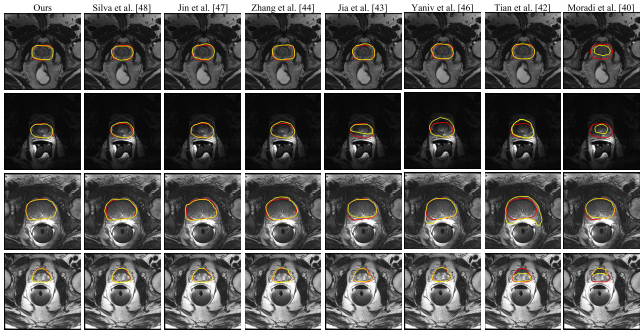


**FIGURE 6.** Visual comparison results of various prostate segmentation methods.

## V. CONCLUSION

In this study, we propose a new MR image-based prostate segmentation method, which can not only reduce the pressure of doctors to read the film, but also improve the diagnostic efficiency of prostate cancer. Specifically, because there is an irrelevant background in the MR image, and the boundary of the prostate is also relatively blurred, this brings certain difficulties to the training of the model. To alleviate this problem, we use the proposed truncated intensity stretching image enhancement method to improve the contrast of the image. In addition, to improve the feature extraction capability of the encoder without increasing the amount of calculation, we integrated the DenseBlock-based CSP structure into the encoder. And, in order to enable the decoder to extract effective features well, we combine the spatial attention mechanism and the channel attention mechanism to integrate into the decoder. Next, to segment the prostate more accurately, we added a prostate contour prediction branch to the output of the segmentation network. In other words, if the network can learn the contour features of the prostate, it can segment the prostate more accurately. Finally, we verify the effectiveness of the various components of the proposed ProSegNet segmentation method according to the ablation study, and show the superiority of the ProSegNet method according to the comparative experiment.

## REFERENCES

[1] J. E. McNeal, "The zonal anatomy of the prostate," *Prostate*, vol. 2, no. 1, pp. 35–49, 1981, doi: 10.1002/pros.2990020105.
[2] C. G. Roehrborn, "Pathology of benign prostatic hyperplasia," *Int. J. Impotence Res.*, vol. 20, no. S3, pp. S11–S18, Dec. 2008, doi: 10.1038/ijir.2008.55.
[3] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2020," *CA A, Cancer J. Clinicians*, vol. 70, no. 4, pp. 7–30, 2020, doi: 10.3322/caac.21590.
[4] Y. Shao, J. Wang, B. Wodlinger, and S. E. Salcudean, "Improving prostate cancer (PCa) classification performance by using three-player minimax game to reduce data source heterogeneity," *IEEE Trans. Med. Imag.*, vol. 39, no. 10, pp. 3148–3158, Oct. 2020, doi: 10.1109/TMI.2020.2988198.
[5] P. D. Baade, D. R. Youlden, and L. J. Krnjacki, "International epidemiology of prostate cancer: Geographical distribution and secular trends," *Mol. Nutrition Food Res.*, vol. 53, no. 2, pp. 171–184, Feb. 2009, doi: 10.1002/mnfr.200700511.
[6] G. Litjens, O. Debats, J. Barentsz, N. Karssemeijer, and H. Huisman, "Computer-aided detection of prostate cancer in MRI," *IEEE Trans. Med. Imag.*, vol. 33, no. 5, pp. 1083–1092, May 2014, doi: 10.1109/TMI.2014.2303821.
[7] G. L. Andriole, E. D. Crawford, R. L. Grubb III, S. S. Buys, D. Chia, T. R. Church, M. N. Fouad, C. Isaacs, P. A. Kvale, D. J. Reding, and J. L. Weissfeld, "Prostate cancer screening in the randomized prostate, lung, colorectal, and ovarian cancer screening trial: Mortality results after 13 years of follow-up," *JNCI J. Nat. Cancer Inst.*, vol. 104, no. 2, pp. 125–132, Jan. 2012, doi: 10.1093/jnci/djr500.
[8] L. Marks, S. Young, and S. Natarajan, "MRI-ultrasound fusion for guidance of targeted prostate biopsy," *Current Opinion Urol.*, vol. 23, no. 1, pp. 43–50, 2013, doi: 10.1097/mou.0b013e32835ad3ee.
[9] K. Kitajima, Y. Kaji, Y. Fukabori, K.-I. Yoshida, N. Suganuma, and K. Sugimura, "Prostate cancer detection with 3 t MRI: Comparison of diffusion-weighted imaging and dynamic contrast-enhanced MRI in combination with T2-weighted imaging," *J. Magn. Reson. Imag.*, vol. 31, no. 3, pp. 625–631, Mar. 2010, doi: 10.1002/jmri.22075.
[10] W. Yang, Y. Shi, S. H. Park, M. Yang, Y. Gao, and D. Shen, "An effective MR-guided CT network training for segmenting prostate in CT images," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 8, pp. 2278–2291, Aug. 2020, doi: 10.1109/JBHI.2019.2960153.
[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst., 26th Annu. Conf. Neural Inf. Process. Syst.*, Dec. 2012, pp. 1106–1114.
[12] H. Yang, W. Huang, K. Qi, C. Li, X. Liu, M. Wang, H. Zheng, and S. Wang, "CLCI-Net: Cross-level fusion and context inference networks for lesion segmentation of chronic stroke," in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2019*, D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, and A. Khan, Eds. Cham, Switzerland: Springer, 2019, pp. 266–274, doi: 10.1007/978-3-030-32248-9_30.
[13] Y. Zhou, W. Huang, P. Dong, Y. Xia, and S. Wang, "D-UNet: A dimension-fusion u shape network for chronic stroke lesion segmentation," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 3, pp. 940–950, May 2021, doi: 10.1109/TCBB.2019.2939522.
[14] X. Liu, H. Yang, K. Qi, P. Dong, Q. Liu, X. Liu, R. Wang, and S. Wang, "MSDF-Net: Multi-scale deep fusion network for stroke lesion segmentation," *IEEE Access*, vol. 7, pp. 178486–178495, 2019, doi: 10.1109/ACCESS.2019.2958384.
[15] K. Qi, H. Yang, C. Li, Z. Liu, M. Wang, Q. Liu, and S. Wang, "X-Net: Brain stroke lesion segmentation based on depthwise separable convolution and long-range dependencies," in *Medical Image Computing and Computer Assisted Intervention—MICCAI 2019*, D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P.-T. Yap, and A. Khan, Eds. Cham, Switzerland: Springer, 2019, pp. 247–255, doi: 10.1007/978-3-030-32248-9_28.

[16] R. Zwiggelaar, Y. Zhu, and S. Williams, "Semi-automatic segmentation of the prostate," in *Proc. 1st Iberian Conf. Pattern Recognit. Image Anal., (IbPRIA)*, Puerto de Andratx, Mallorca, Spain, Jun. 2003, pp. 1108–1116, doi: 10.1007/978-3-540-44871-6_128.

[17] M. Samiee, G. Thomas, and R. Fazel-Rezai, "Semi-automatic prostate segmentation of MR images based on flow orientation," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, Aug. 2006, pp. 203–207, doi: 10.1109/ISSPIT.2006.270797.

[18] D. Flores-Tapia, G. Thomas, N. Venugopal, B. McCurdy, and S. Pistorius, "Semi automatic MRI prostate segmentation based on wavelet multiscale products," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, pp. 3020–3023, doi: 10.1109/IEMBS.2008.4649839.

[19] T. F. Cootes, A. Hill, C. J. Taylor, and J. Haslam, "The use of active shape models for locating structures in medical images," in *Information Processing in Medical Imaging*, H. H. Barrett and A. F. Gmitro, Eds. Berlin, Germany: Springer, 1993, pp. 33–47, doi: 10.1007/BFb0013779.

[20] S. Klein, U. A. van der Heide, I. M. Lips, M. van Vulpen, M. Staring, and J. P. Pluim, "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information," *Med. Phys.*, vol. 35, no. 4, pp. 1407–1417, 2008, doi: 10.1118/1.2842076.

[21] J. A. Dowling, J. Fripp, S. Chandra, J. P. W. Pluim, J. Lambert, J. Parker, J. Denham, P. B. Greer, and O. Salvado, "Fast automatic multi-atlas segmentation of the prostate from 3D MR images," in *Prostate Cancer Imaging. Image Analysis and Image-Guided Interventions*, A. Madabhushi, J. Dowling, H. Huisman, and D. Barratt, Eds. Berlin, Germany: Springer, 2011, pp. 10–21, doi: 10.1007/978-3-642-23944-1_2.

[22] S. Vikal, S. Haker, C. Tempany, and G. Fichtinger, "Prostate contouring in MRI-guided biopsy," *Proc. SPIE*, vol. 7259, Mar. 2009, Art. no. 72594A, doi: 10.1117/12.812433.

[23] Y. Zhang, B. J. Matuszewski, A. Histace, F. Precioso, J. Kilgallon, and C. Moore, "Boundary delineation in prostate imaging using active contour segmentation method with interactively defined object regions," in *Prostate Cancer Imaging. Computer-Aided Diagnosis, Prognosis, and Intervention*, A. Madabhushi, J. Dowling, P. Yan, A. Fenster, P. Abolmaesumi, and N. Hata, Eds. Berlin, Germany: Springer, 2010, pp. 131–142, doi: 10.1007/978-3-642-15989-3_15.

[24] R. Toth, B. N. Bloch, E. M. Genega, N. M. Rofsky, R. E. Lenkinski, M. A. Rosen, A. Kalyanpur, S. Pungavkar, and A. Madabhushi, "Accurate prostate volume estimation using multifeature active shape models on T2-weighted MRI," *Academic Radiol.*, vol. 18, no. 6, pp. 745–754, Jun. 2011, doi: 10.1016/j.acra.2011.01.016.

[25] D. Singh, V. Kumar, C. J. Das, A. Singh, and A. Mehndiratta, "Segmentation of prostate zones using probabilistic atlas-based method with diffusion-weighted MR images," *Comput. Methods Programs Biomed.*, vol. 196, Nov. 2020, Art. no. 105572, doi: 10.1016/j.cmpb.2020.105572.

[26] S. Martin, V. Daanen, and J. Troccaz, "Atlas-based prostate segmentation using an hybrid registration," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 3, no. 6, pp. 485–492, Dec. 2008, doi: 10.1007/s11548-008-0247-0.

[27] A. Firjany, A. Elnakib, A. El-Baz, G. Gimel'farb, M. A. El-Ghar, and A. Elmagharby, "Novel stochastic framework for accurate segmentation of prostate in dynamic contrast enhanced MRI," in *Prostate Cancer Imaging. Computer-Aided Diagnosis, Prognosis, and Intervention*, A. Madabhushi, J. Dowling, P. Yan, A. Fenster, P. Abolmaesumi, and N. Hata, Eds. Berlin, Germany: Springer, 2010, pp. 121–130, doi: 10.1007/978-3-642-15989-3_14.

[28] W. D. Richard and C. G. Keen, "Automated texture-based segmentation of ultrasound images of the prostate," *Comput. Med. Imag. Graph.*, vol. 20, no. 3, pp. 131–140, 1996, doi: 10.1016/0895-6111(96)00048-1.

[29] P. D. Allen, J. Graham, D. C. Williamson, and C. E. Hutchinson, "Differential segmentation of the prostate in MR images using combined 3D shape modelling and voxel classification," in *Proc. 3rd IEEE Int. Symp. Biomed. Imag., Macro Nano*, Apr. 2006, pp. 410–413, doi: 10.1109/ISBI.2006.1624940.

[30] N. Makni, P. Puech, R. Lopes, A. S. Dewalle, O. Colot, and N. Betrouni, "Combining a deformable model and a probabilistic framework for an automatic 3D segmentation of prostate on MRI," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 4, no. 2, p. 181, 2008, doi: 10.1007/s11548-008-0281-y.

[31] S. S. Mohamed, A. M. Youssef, E. F. El-Saadany, and M. M. A. Salama, "Prostate tissue characterization using TRUS image spectral features," in *Image Analysis and Recognition*, A. Campilho and M. Kamel, Eds. Berlin, Germany: Springer, 2006, pp. 589–601, doi: 10.1007/11867661_53.

[32] T. P. Weldon, W. E. Higgins, and D. F. Dunn, "Efficient Gabor filter design for texture segmentation," *Pattern Recognit.*, vol. 29, no. 12, pp. 2005–2015, 1996, doi: 10.1016/S0031-3203(96)00047-7.

[33] M. Pal and G. M. Foody, "Feature selection for classification of hyperspectral data by SVM," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2297–2307, May 2010, doi: 10.1109/TGRS.2009.2039484.

[34] X. Liu, D. L. Langer, M. A. Haider, Y. Yang, M. N. Wernick, and I. S. Yetik, "Prostate cancer segmentation with simultaneous estimation of Markov random field parameters and class," *IEEE Trans. Med. Imag.*, vol. 28, no. 6, pp. 906–915, Jun. 2009, doi: 10.1109/TMI.2009.2012888.

[35] Z. Kato and T.-C. Pong, "A Markov random field image segmentation model for color textured images," *Image Vis. Comput.*, vol. 24, no. 10, pp. 1103–1114, 2006, doi: 10.1016/j.imavis.2006.03.005.

[36] P. Khurd, L. Grady, K. Gajera, M. Diallo, P. Gall, M. Requardt, B. Kiefer, C. Weiss, and A. Kamen, "Facilitating 3D spectroscopic imaging through automatic prostate localization in MR images using random Walker segmentation initialized via boosted classifiers," in *Prostate Cancer Imaging. Image Analysis and Image-Guided Interventions*, 2011, pp. 47–56.

[37] D. Mahapatra and J. M. Buhmann, "Prostate MRI segmentation using learned semantic knowledge and graph cuts," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 3, pp. 756–764, Mar. 2014, doi: 10.1109/TBME.2013.2289306.

[38] Q. Zhu, B. Du, B. Turkbey, P. L. Choyke, and P. Yan, "Deeply-supervised CNN for prostate segmentation," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 178–184, doi: 10.1109/IJCNN.2017.7965852.

[39] D. Karimi, G. Samei, C. Kesch, G. Nir, and S. E. Salcudean, "Prostate segmentation in MRI using a convolutional neural network architecture and training strategy based on statistical shape models," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 13, no. 8, pp. 1211–1219, Aug. 2018, doi: 10.1007/s11548-018-1785-8.

[40] H. Moradi, A. H. Foruzan, and Y.-W. Chen, "Automatic segmentation of prostate in MR images using deep learning and multi-atlas techniques," in *Proc. 25th Nat. 3rd Int. Iranian Conf. Biomed. Eng. (ICBME)*, Nov. 2018, pp. 1–4, doi: 10.1109/ICBME.2018.8703532.

[41] T. Clark, J. Zhang, S. Baig, A. Wong, M. A. Haider, and F. Khalvati, "Fully automated segmentation of prostate whole gland and transition zone in diffusion-weighted MRI using convolutional neural networks," *J. Med. Imag.*, vol. 4, no. 4, 2017, Art. no. 41307, doi: 10.1117/1.jmi.4.4.041307.

[42] Z. Tian, L. Liu, Z. Zhang, and B. Fei, "PSNet: Prostate segmentation on MRI based on a convolutional neural network," *J. Med. Imag.*, vol. 5, no. 2, pp. 1–6, Jan. 2018, doi: 10.1117/1.JMI.5.2.021208.

[43] H. Jia, Y. Xia, Y. Song, W. Cai, M. Fulham, and D. D. Feng, "Atlas registration and ensemble deep convolutional neural network-based prostate segmentation using magnetic resonance imaging," *Neurocomputing*, vol. 275, pp. 1358–1369, Jan. 2018, doi: 10.1016/j.neucom.2017.09.084.

[44] Y. Zhang, J. Wu, W. Chen, Y. Chen, and X. Tang, "Prostate segmentation using Z-Net," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 11–14, doi: 10.1109/ISBI.2019.8759554.

[45] F. Milletari, N. Navab, and S.-A. Ahmadi, "V-Net: Fully convolutional neural networks for volumetric medical image segmentation," in *Proc. 4th Int. Conf. 3D Vis. (3DV)*, Oct. 2016, pp. 565–571, doi: 10.1109/3DV.2016.79.

[46] O. Yaniv, O. Portnoy, A. Talmon, N. Kiryati, E. Konen, and A. Mayer, "V-Net light–parameter-efficient 3-D convolutional neural network for prostate MRI segmentation," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 442–445, doi: 10.1109/ISBI45749.2020.9098643.

[47] Y. Jin, G. Yang, Y. Fang, R. Li, X. Xu, Y. Liu, and X. Lai, "3D PBV-net: An automated prostate MRI data segmentation method," *Comput. Biol. Med.*, vol. 128, Jan. 2021, Art. no. 104160, doi: 10.1016/j.compbiomed.2020.104160.

[48] G. L. F. da Silva, P. S. Diniz, J. L. Ferreira, J. V. F. França, A. C. Silva, A. C. de Paiva, and E. A. A. de Cavalcanti, "Superpixel-based deep convolutional neural networks and active contour model for automatic prostate segmentation on 3D MRI scans," *Med. Biol. Eng. Comput.*, vol. 58, no. 9, pp. 1947–1964, Sep. 2020, doi: 10.1007/s11517-020-02199-5.

[49] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *J. VLSI signal Process. Syst. Signal, Image Video Technol.*, vol. 38, no. 1, pp. 35–44, 2004, doi: 10.1023/B:VLSI.0000028532.53893.82.

[50] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015*, 2015, pp. 234–241.

[51] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jul. 2017, pp. 2261–2269, doi: 10.1109/CVPR.2017.243.
[52] C.-Y. Wang, H.-Y. Mark Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, and I.-H. Yeh, "CSPNet: A new backbone that can enhance learning capability of CNN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2020, pp. 1571–1580, doi: 10.1109/CVPRW50498.2020.00203.
[53] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 7132–7141, doi: 10.1109/CVPR.2018.00745.
[54] C. Yu, J. Wang, C. Peng, C. Gao, G. Yu, and N. Sang, "Learning a discriminative feature network for semantic segmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Jun. 2018, pp. 1857–1866, doi: 10.1109/CVPR.2018.00199.
[55] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, "Spatial transformer networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 28, 2015, pp. 2017–2025.
[56] L. Chen, H. Zhang, J. Xiao, L. Nie, J. Shao, W. Liu, and T.-S. Chua, "SCA-CNN: Spatial and channel-wise attention in convolutional networks for image captioning," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6298–6306, doi: 10.1109/CVPR.2017.667.
[57] R. Caruana, S. Lawrence, and C. L. Giles, "Overfitting in neural nets: Backpropagation, conjugate gradient, and early stopping," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 13. Denver, CO, USA, 2000, pp. 402–408.
[58] G. Litjens, R. Toth, W. van de Ven, C. Hoeks, S. Kerkstra, B. van Ginneken, G. Vincent, G. Guillard, N. Birbeck, J. Zhang, and R. Strand, "Evaluation of prostate segmentation algorithms for MRI: The PROMISE12 challenge," *Med. Image Anal.*, vol. 18, no. 2, pp. 359–373, 2014, doi: 10.1016/j.media.2013.12.002.
[59] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587, doi: 10.1109/CVPR.2014.81.
[60] D. L. Pham and J. L. Prince, "An adaptive fuzzy segmentation algorithm for three-dimensional magnetic resonance images," in *Information Processing in Medical Imaging*, A. Kuba, M. Šáamal, and A. Todd-Pokropek, Eds. Berlin, Germany: Springer, 1999, pp. 140–153, doi: 10.1007/3-540-48714-X_11.
[61] J. K. Udupa and S. Samarasekera, "Fuzzy connectedness and object definition: Theory, algorithms, and applications in image segmentation," *Graph. Models Image Process.*, vol. 58, no. 3, pp. 246–261, May 1996, doi: 10.1006/gmip.1996.0021.
[62] S. Valverde, A. Oliver, E. Roura, S. González-Villà, D. Pareto, J. C. Vilanova, L. Ramió-Torrentà, À. Rovira, and X. Lladó, "Automated tissue segmentation of MR brain images in the presence of white matter lesions," *Med. Image Anal.*, vol. 35, pp. 446–457, Jan. 2017, doi: 10.1016/j.media.2016.08.014.
[63] D. Karimi and S. E. Salcudean, "Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks," *IEEE Trans. Med. Imag.*, vol. 39, no. 2, pp. 499–513, Feb. 2020, doi: 10.1109/TMI.2019.2930068.

**YUEJING QIAN** received his Master degree in Computer Technology from Hangzhou Dianzi University of Zhejiang. He is now an Associate Professor of Zhejiang Industry and Trade Vocational College. His current research interests involve medical image processing and deep learning.