

# Mutually communicated model based on multi-parametric MRI for automated segmentation and classification of prostate cancer

Kewen Liu<sup>1,2</sup> | Piqiang Li<sup>1,2</sup> | Martins Otikovs<sup>3</sup> | Xinzhou Ning<sup>1</sup> | Liyang Xia<sup>1,2</sup> | Xiangyu Wang<sup>4</sup> | Lian Yang<sup>5</sup> | Feng Pan<sup>5</sup> | Zhi Zhang<sup>1</sup> | Guangyao Wu<sup>6</sup> | Han Xie<sup>1</sup> | Qingjia Bao<sup>1</sup> | Xin Zhou<sup>1,7,8</sup> | Chaoyang Liu<sup>1,7,8</sup>

<sup>1</sup>State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan, P.R. China

<sup>2</sup>School of Information Engineering, Wuhan University of Technology, Wuhan, P.R. China

<sup>3</sup>Weizmann Institute of Science, Department of Chemical and Biological Physics, Rehovot, Israel

<sup>4</sup>First Affiliated Hospital of Shenzhen University, Shenzhen, P.R. China

<sup>5</sup>Department of Radiology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, P.R. China

<sup>6</sup>Shenzhen University General Hospital, Shenzhen, P.R. China

<sup>7</sup>University of Chinese Academy of Sciences, Beijing, P.R. China

<sup>8</sup>Wuhan National Laboratory for Optoelectronics, Huazhong University of Science and Technology-Optics Valley Laboratory, Wuhan, P.R. China

## Correspondence

Qingjia Bao and Chaoyang Liu, State Key Laboratory of Magnetic Resonance and Atomic and Molecular Physics, Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan 430071, P.R. China.  
Email: [baqingjiahaoba@gmail.com](mailto:baqingjiahaoba@gmail.com) and [chylu@apm.ac.cn](mailto:chylu@apm.ac.cn)

## Funding information

National Major Scientific Research Equipment Development Project of China, Grant/Award Number: 81627901; National key of R&D Program of China, Grant/Award Numbers: 2018YFC0115000, 2016YFC1304702; National Natural Science Foundation of China, Grant/Award Numbers: 11575287, 11705274; Chinese Academy of Sciences, Grant/Award Number: YZ201677

## Abstract

**Background:** Multiparametric magnetic resonance imaging (mp-MRI) is introduced and established as a noninvasive alternative for prostate cancer (PCa) detection and characterization.

**Purpose:** To develop and evaluate a mutually communicated deep learning segmentation and classification network (MC-DSCN) based on mp-MRI for prostate segmentation and PCa diagnosis.

**Methods:** The proposed MC-DSCN can transfer mutual information between segmentation and classification components and facilitate each other in a bootstrapping way. For classification task, the MC-DSCN can transfer the masks produced by the coarse segmentation component to the classification component to exclude irrelevant regions and facilitate classification. For segmentation task, this model can transfer the high-quality localization information learned by the classification component to the fine segmentation component to mitigate the impact of inaccurate localization on segmentation results. Consecutive MRI exams of patients were retrospectively collected from two medical centers (referred to as center A and B). Two experienced radiologists segmented the prostate regions, and the ground truth of the classification refers to the prostate biopsy results. MC-DSCN was designed, trained, and validated using different combinations of distinct MRI sequences as input (e.g., T2-weighted and apparent diffusion coefficient) and the effect of different architectures on the network's performance was tested and discussed. Data from center A were used for training, validation, and internal testing, while another center's data were used for external testing. The statistical analysis is performed to evaluate the performance of the MC-DSCN. The DeLong test and paired *t*-test were used to assess the performance of classification and segmentation, respectively.

**Results:** In total, 134 patients were included. The proposed MC-DSCN outperforms the networks that were designed solely for segmentation or classification. Regarding the segmentation task, the classification localization information helped to improve the IOU in center A: from 84.5% to 87.8% ( $p < 0.01$ ) and in center B: from 83.8% to 87.1% ( $p < 0.01$ ), while the area under curve (AUC) of PCa classification was improved in center A: from 0.946 to 0.991 ( $p < 0.02$ ) and in center B: from 0.926 to 0.955 ( $p < 0.01$ ) as a result of the additional information provided by the prostate segmentation.

**Conclusion:** The proposed architecture could effectively transfer mutual information between segmentation and classification components and facilitate each other in a bootstrapping way, thus outperforming the networks designed to perform only one task.

#### KEYWORDS

deep learning, multi-parametric magnetic resonance imaging, prostatic cancer classification, prostate segmentation

## 1 | INTRODUCTION

Prostate cancer (PCa) is currently the most common cancer in the male urinary system worldwide.<sup>1</sup> Early detection and risk assessment of PCa helps distinguish clinically significant PCa, which risks progress while avoiding overdiagnosis. Therefore, studying PCa's characteristics (classification) is an important research direction in PCa diagnosis. For the diagnosis of PCa, systematic biopsy has remained the standard diagnostic route despite its associated risks.<sup>2–4</sup> Recent studies have demonstrated that multi-parametric magnetic resonance imaging (mp-MRI) can provide a noninvasive way to study characteristics and assist in diagnosing PCa.<sup>5,6</sup> Moreover, MRI-based prostate segmentation is also important in guiding radiotherapy, biopsy, and focal therapy.

For the diagnosis of PCa, systematic biopsy has remained the standard diagnostic route despite its associated risks.<sup>2</sup> Due to a limited number of biopsy samples and low resolution of transrectal ultrasound (TRUS), lesions may be missed, or the Gleason score determined from biopsy samples is likely to be biased.<sup>2,3</sup> Recent studies have demonstrated that multi-parametric MR imaging (mp-MRI) can provide a noninvasive way to study characteristics and assist in diagnosis of PCa.<sup>5,7</sup> On the other hand, MRI-based prostate segmentation is also important in guiding radiotherapy, biopsy, and focal therapy and its application in diagnosis.

Many previous studies have demonstrated that no single MRI sequence is definitive for PCa detection. Mp-MRI with both anatomical and functional sequences is necessary for PCa detection and diagnosis. The Prostate Imaging Reporting and Data System (PI-RADS) scoring system defines that mp-MRI for PCa detection and diagnosis typically includes T2w imaging, diffusion-weighted imaging (DWI), MR spectroscopic imaging (MRSI), and dynamic contrast-enhanced

(DCE) MRI.<sup>8</sup> However, the accuracy of PI-RADS largely depends on the experience of the radiologist. Moreover, manually interpreting mp-MRI images requires substantial expertise of a radiologist. Therefore, automated and accurate PCa detection from mp-MRI contributes to alleviating radiologists' burden, reducing the risk of over-/under-treatment, and enabling large-scale PCa screening.<sup>9</sup>

Computer-aided diagnosis (CAD) has been a rapidly growing and active research area over the past two decades. Several CAD systems for the radiological assessment of thalamic nuclei, left-ventricular, and prostate malignancies have been developed,<sup>10–12</sup> and encouraging results have been reported. CAD systems for PCa comprise two main tasks: prostate segmentation and classification.<sup>7–13</sup> The segmentation task involves detecting the prostate's location and boundaries, whereas the classification's goal is to determine PCa type (e.g., benign, malignant). However, both tasks are still challenging. The heterogeneous anatomical structure of the prostate and its wide variation in size and shape among different subjects make it difficult to determine prostate boundaries. Likewise, the PCa classification is challenging as both PCa, and benign prostatic hyperplasia tissues have similarly abnormal and lower signal responses.

In recent years, growing evidence indicates that deep learning-based methods can effectively learn multi-parametric information. So, it might be a potential alternative to conventional hand-crafted methodologies for solving image classification problems.<sup>14,15</sup> And deep learning methods have become mainstream CAD systems for prostate segmentation and PCa classification due to their accuracy and efficiency.<sup>16</sup> Several computer-aided systems based on deep learning have been developed in the past decade for accurate and automated PCa diagnosis. Song et al. presented deep convolution neural network (CNN)-based enhanced prediction to

diagnose PCa based on mp-MRI, and achieved an area under curve (AUC) of 0.944 for the PCa detection.<sup>17</sup> Yang et al. developed a co-trained CNN with a single-stage SVM classifier for PCa detection and lesion localization from mp-MRI, and achieved a sensitivity of 0.97 on a dataset containing 160 patients.<sup>18</sup> Sanford et al. developed an artificial intelligence solution for PI-RADS classification and compared its performance with an expert radiologist using targeted biopsy results.<sup>19</sup> With the growing importance of prostate segmentation, several segmentation methods have been proposed over the past few years to meet the challenges. Zhu et al. developed a fully automatic approach to segmenting the outer prostate contour and the peripheral zone (PZ) contour with high efficacy.<sup>20</sup> Alkadi et al. proposed an approach based on deep learning to detect and localize PCa in T2w MRI by extracting 3D prostate information using the 3D sliding window method.<sup>21</sup> Ghavami et al. investigated the impact of network architecture on the accuracy of volume measurement based on the 3D prostate T2w MRI data set.<sup>22</sup> Wang et al. employed deeply supervised 3D fully convolutional networks with group dilated convolution for automatic segmentation of prostate T2w MR images.<sup>23</sup>

However, the networks mentioned above are generally designed for either segmentation or classification, ignoring the potential benefits of jointly performing both tasks, while several studies have shown that segmentation and classification are two highly related tasks.<sup>24–26</sup> For example, segmentation can help eliminate the interference of irrelevant regions in the image and help improve the accuracy of PCa diagnosis. In addition, categorized diagnostic information can help to highlight lesion areas and hence contribute to segmentation. In this work, we develop a new mutually communicated deep learning segmentation and classification network (MC-DSCN) for prostate segmentation and PCa diagnosis. The proposed MC-DSCN can transfer mutual information between segmentation and classification components and facilitate each other in a bootstrapping way. The MC-DSCN is trained and validated on mp-MRI data from two medical centers.

## 2 | MATERIAL AND METHOD

### 2.1 | Dataset

The mp-MRI datasets were acquired from two medical centers (center A: First Affiliated Hospital of Shenzhen University, center B: Union Hospital of Huazhong University of Science and Technology). This retrospective study was approved by the Institutional Review Board of both centers, and written informed consent was waived. Between November 2019 and November 2021, 63 and 71 patients were retrieved from medical centers A and B, respectively. The inclusion criteria were: (1) The patients

were scanned by routine prostate MRI examination, and the scanning parameters of each sequence were consistent with the research requirements; (2) Prostate biopsy was performed within 1 month after MR examination; and (3) The patient underwent mp-MRI before radical prostatectomy.

The body coil and 18-channel abdominal phased array coil were used as transmitter and receiver, respectively. The slice package center was located 2.0 cm above the pubic symphysis. During the scan axial T1 weighted, axial, sagittal, and coronal T2 weighted, and axial diffusion weighted (DW) images were acquired. T2W parameters: (1) center A: TE (Echo Time) = 67 ms, TR (Repetition Time) = 4000 ms, matrix = 256 × 256, FOV (Field of View) = 200 × 200 mm, slice thickness = 3 mm; (2) center B: TE = 93 ms, TR = 4500 ms, matrix = 136 × 160, FOV = 170 × 200 mm<sup>2</sup>, slice thickness = 4 mm. DWI parameters: (1) center A: TE = 95 ms, TR = 4500 ms, matrix = 102 × 160, FOV = 128 × 200 mm<sup>2</sup>, slice thickness = 3 mm; (2) center B: TE = 93 ms, TR = 4500 ms, matrix = 136 × 160, FOV = 170 × 200 mm<sup>2</sup>, slice thickness = 4 mm. The Apparent Diffusion Coefficient (ADC) maps were calculated by using linear least squares curve fitting of pixels (in log scale) in the four diffusion-weighted images against their corresponding *b* values (0/100/400/800 s/mm<sup>2</sup>).

All 134 patients from centers A and B were scanned on a Siemens Prisma 3.0 T scanner. After applying inclusion and exclusion criteria, 134 patients (mean age: 65 years; range: 53–77 years; nonmalignant/malignant = 82/52) were included in the study from two medical centers. Herein, “Classification” indicated a binary category of “benign” or “malignant” (clinically significant lesions), which was pathologically confirmed in the following surgery or needle biopsy. Patients from medical center A were divided randomly into training and internal testing cohorts with a ratio of 4:1 using stratified sampling. Patients from medical center B were utilized as the external testing cohort.

The image registration was performed using coordinate information stored in the DICOM headers to match the position and resolution of ADC and T2w images. And, the co-registered ADC-T2w images were augmented using a nonrigid image deformation method.<sup>27</sup> 1270 slices of co-registered ADC-T2w images (530 slices for 63 patients in center A, 740 slices for 71 patients in center B) were augmented to 4364 slices using a nonrigid image deformation method. The details of the data processing can be found in the supplementary material (Figure S1). Firstly, we choose 9 points that are distributed evenly around the prostate, denoted as  $P_n = \{p_1, p_2 \dots p_9\}$ . Then every point was shifted randomly and by amount independent from shift of other points within  $[-6, 6]$  pixels to obtain the new position of these 9 point  $TP_n = \{Tp_1, Tp_2 \dots, Tp_9\}$ . Finally, the deformed T2 and ADC images are derived based on

linear moving least square with these 9 points from  $P_n$  to  $TP_n$ , respectively.

The transformation was derived based on moving least squares.<sup>28</sup> Let's assume the 9 points as control points  $X(x, y)$ . Then for every control point, we randomly shift the  $x$ - $y$  coordinates of the point by  $[-6, 6]$  pixels, that is,  $x' = x \pm 0, 1, 2, \text{ or } 3$ , and  $y' = y \pm 0, 1, 2, \text{ or } 3$ , yielding a new transformed 9 points  $X'(x', y')$ . Then we compute the deformation function for transforming  $X(x, y)$  to  $X'(x', y')$ ,  $f\{X(x, y)\} = X'(x', y')$ . The function  $f$  is calculated based on these properties: (1)  $f$  can handle 9 points and (2)  $f$  should produce smooth deformations. Lastly, we can apply the calculated deformation function  $f$  to the original image to generate augmented data.

As the prostate region is only a small part of the whole image, a square-shaped region ( $128 \times 128$ ) was selected and the pairs of ADC-T2w images and images were cropped to this selection. Then, the intensities of the images within the selected region were normalized before serving as an input to the network.

## 2.2 | Developing a mutually communicated segmentation and classification architecture

This work aims to design a MC-DSCN, which jointly performs segmentation based on pixel-level information and classification based on image-level information. The architecture of this new network is shown in Figure 1. It contains three components: the coarse segmentation component, the classification component, and the fine segmentation component. The proposed MC-DSCN model can transfer the coarse prostate masks produced by the coarse segmentation component to the classification component to exclude irrelevant regions and facilitate classification. Furthermore, this model can transfer the high-quality localization information (cancer response map, CRM) learned by the classification component to the fine segmentation component to mitigate the impact of inaccurate localization on segmentation results.

The coarse segmentation component's main function is to generate prostate masks that provide preliminary information about the prostate location for the classification component. The architecture of this component is based on the residual Unet with attention blocks. Unet is commonly used for medical image segmentation tasks because of its multi-scale feature fusion capability.<sup>29</sup> The residual Unet takes advantage of the Unet and the residual mechanism to solve degradation while training the network. The attention block is placed right before the concatenation operation to highlight salient features passed through the skip connections. The details of the segmentation component are described in Figure 2.

The input of the classification component contains cropped T2w images, ADC maps, and coarse prostate masks. As mentioned above, coarse prostate masks boost the localization and discrimination ability of the classification component as the surrounding tissue negatively affects the PCa diagnosis performance. This component takes advantage of both the bottleneck blocks and the co-training block. The bottleneck block is introduced to extract features and solve the degradation problem while training the network.<sup>30</sup> In the bottleneck, we use a  $1 \times 1$  convolutional layer (shown in Figure 1(b)) to fuse the multi-channel feature map, yielding a single-channel feature map. After bilinear interpolation, this fused feature map can be potentially directly used as the CRM. A greater value at location  $(x, y)$  indicates a greater likelihood of this position being cancerous.

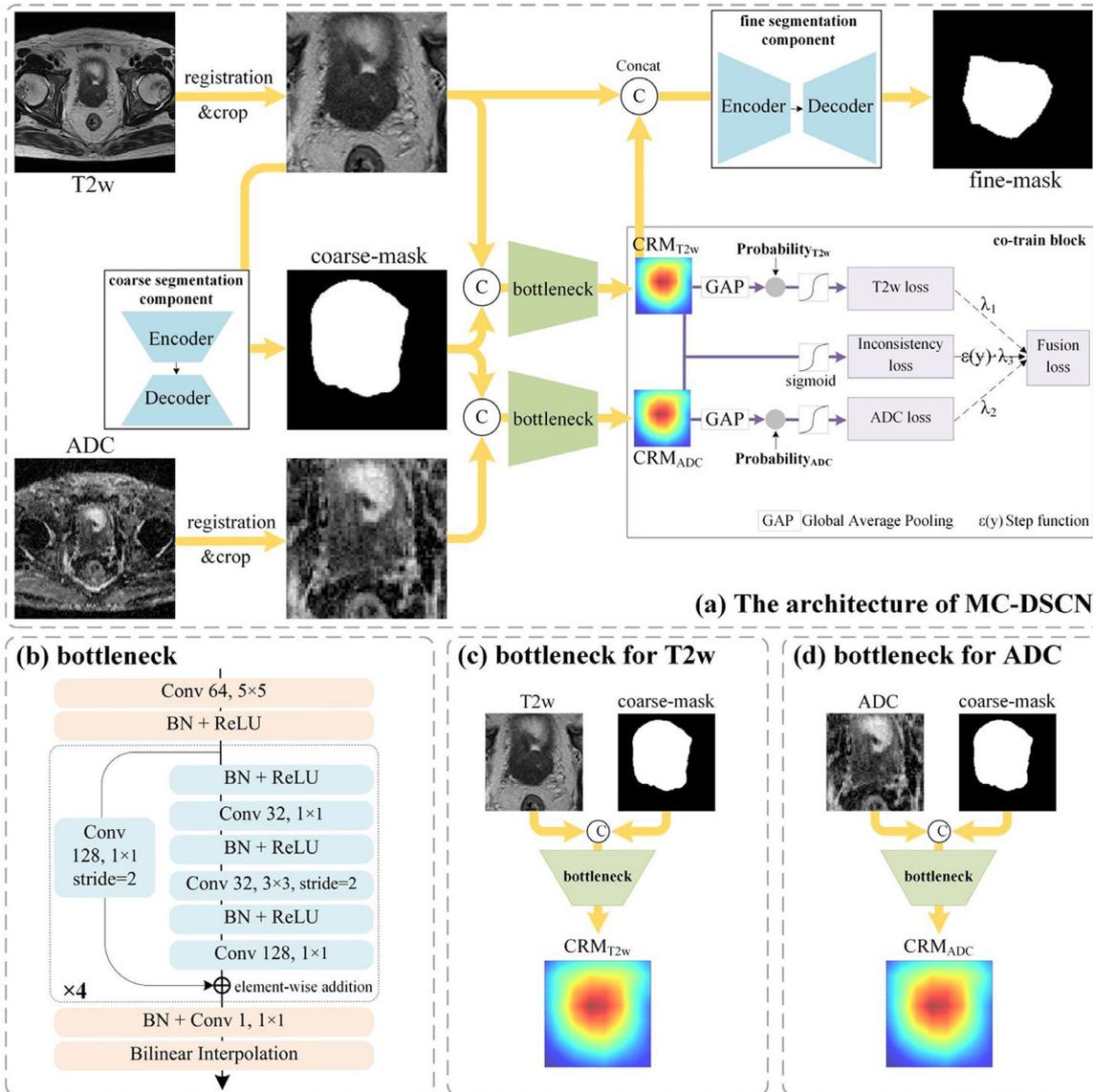
As the ADC and T2w images from the same patient are co-registered, the prostate would be expected to appear in the same location, and the cancer response map of an ADC slice should be consistent with one of the corresponding T2w slices. However, training two independent classification components for ADC and T2w images separately cannot guarantee consistency of CRM derived from the two modalities. Therefore, we introduce a co-training block to fuse the multimodal feature information learned from T2W and ADC images. The co-training block utilizes the classification hybrid loss that includes two classification losses for T2w and ADC images, and the loss to ensure the consistency between the ADC and T2w derived CRMs. The total classification loss is calculated following these steps: First, the prediction probabilities ( $\text{Probability}_{T2w}$  and  $\text{Probability}_{ADC}$ ) are obtained by Global Average Pooling (GAP) and sigmoid function. Then the cross-entropy losses (T2w loss and ADC loss, respectively) are calculated. Thirdly, the consistency loss between the  $\text{CRM}_{T2w}$  and  $\text{CRM}_{ADC}$  is calculated; Finally, the T2w loss, ADC loss, and consistency loss are multiplied by different weighting factors and added together as the total loss. Besides, to ensure the meaningfulness of the consistency of CRMs, we calculate the consistency of CRMs only for the malignant images.

The architecture of the fine segmentation component is the same as for the coarse segmentation component. The difference between the coarse and fine segmentation is that we concatenate the T2w images with CRM as an input to the fine segmentation block to combine the multiparametric location information (shown in Figure 2(b) and (c)).

## 2.3 | The loss of MC-DSCN

### 2.3.1 | Hybrid loss of segmentation

To optimize the performance of the segmentation tasks, we propose a hybrid loss function that not only can



**FIGURE 1** (a) Architecture of the MC-DSCN for segmentation and classification. The coarse segmentation component is constructed to generate coarse masks. Next, the classification component is used to produce CRM. Finally, the fine segmentation component generates fine masks. (b) Architecture of the bottleneck for the classification component. (c) The process of using a bottleneck to produce a CRM for a T2w image. (d) The process of using a bottleneck to produce a CRM for an ADC image. ADC, apparent diffusion coefficient; CRM, cancer response map; MC-DSCN, mutually communicated deep learning segmentation and classification network.

quantify segmentation results similarity by generalized Dice loss at the image level (distribution-based loss) but also include cross-entropy loss at the pixel level (region-based loss).<sup>31</sup> Moreover, we introduce rank loss (boundary-based loss) that pays more attention to the boundary pixels.

The hybrid loss function can be denoted as:

$$Loss_{hybrid} = \lambda_1 Loss_{GDL} + \lambda_2 Loss_{WCE} + \lambda_3 Loss_{rank} \quad (1)$$

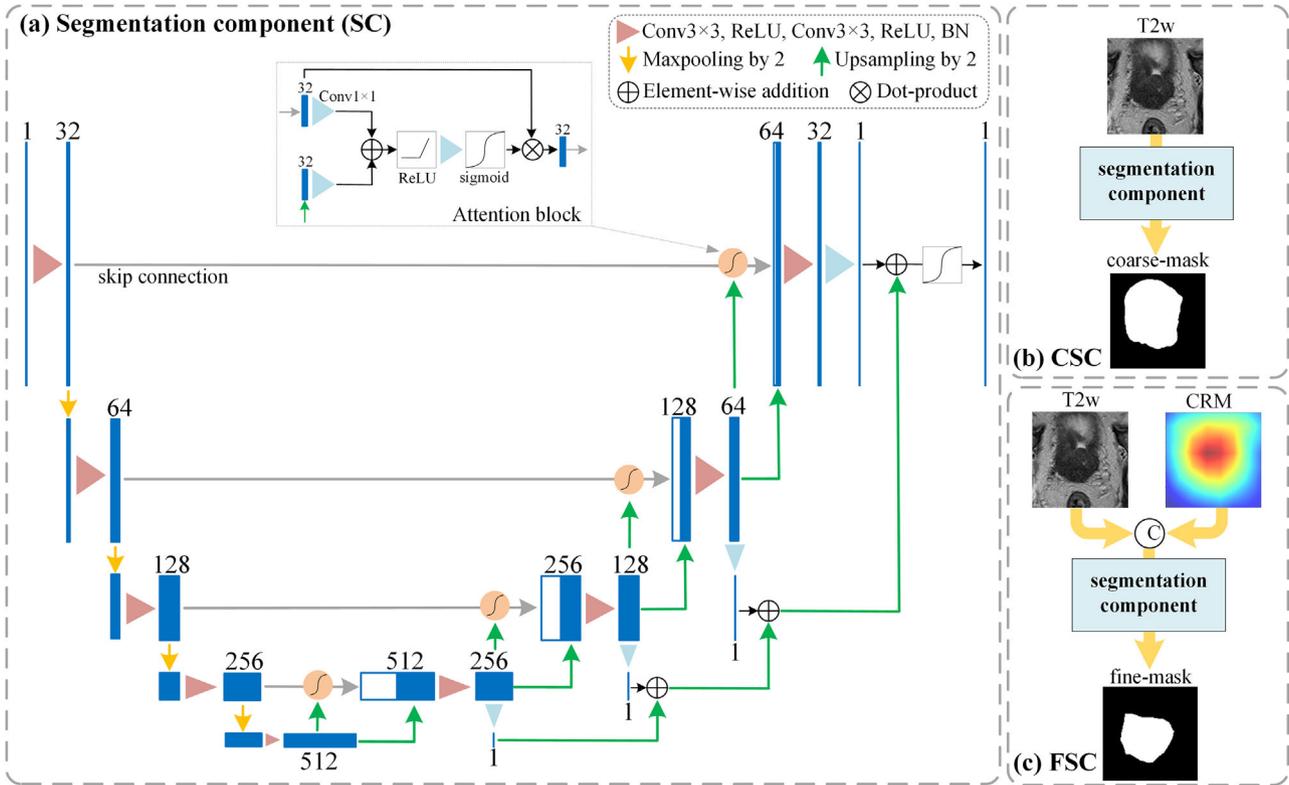
where  $Loss_{GDL}$  is generalized Dice loss,  $Loss_{WCE}$  is the weighted cross-entropy loss,  $Loss_{rank}$  is a rank loss.  $\lambda_1, \lambda_2, \lambda_3$  are weighting factors and set to 1, 1, and 0.1, respectively.

Generalized Dice loss quantifies the degree of agreement between the prediction and the ground truth. It weights each of the categories to overcome gradient fluctuation caused by pixels of small targets being predicted incorrectly and is calculated as follows:

$$Loss_{GDL} = 1 - \frac{2 \cdot \sum_{l=1}^2 w_l \sum_{n=1}^N y_{ln} p_{ln}}{\sum_{l=1}^2 w_l \sum_{n=1}^N (y_{ln} + p_{ln}) + \epsilon} \quad (2)$$

in which

$$w_l = \frac{1}{(\sum_{n=1}^N y_{ln})^2} \quad (3)$$



**FIGURE 2** (a) Overall segmentation component. The architecture is based on the residual Unet with attention blocks. (b) The process of coarse segmentation. (c) The process of fine segmentation.

where  $l$  denotes the number of categories (e.g., background, prostate),  $N$  denotes the number of pixels,  $n$  denotes the index,  $w_l$  represents the weight of each category,  $p_{ln}$  represents the predicted probability of the  $n$ th pixel belonging to the prostate,  $y_{ln}$  represents the ground-truth label of the  $n$ th pixel, and  $\varepsilon$  is set to 0.01 as a smoothing factor.

The second term is weighted cross-entropy loss. Compared to the generalized Dice loss, which measures the profile similarity at the image level, the weighted cross-entropy loss constrains the prediction results at the pixel level, and it is calculated as follows:

$$\text{Loss}_{WCE} = -\frac{1}{N} \left( \sum_{l=1}^2 \sum_{n=1}^N w_l y_{ln} \log p_{ln} + (1 - y_{ln}) \log (1 - p_{ln}) \right) \quad (4)$$

in which

$$w_l = \frac{N - \sum_{n=1}^N p_{ln}}{\sum_{n=1}^N p_{ln}} \quad (5)$$

where  $w_l$  represents the weight of each category.

The third term is rank loss which pays more attention to the boundary pixels. The pixels inside the prostate and in the background usually can be easily recognized and contribute little to the optimization. In contrast, pixels near the boundary are harder to segment and provide more information for the learning process. The rank loss pays more attention to these “hard” pixels near the boundary, which is prone to a higher probability of suffering from prediction errors. Specifically, after the forward propagation of each batch, we will rank the pixels by their prediction errors in the prostate and background, respectively. The top  $K$  pixels with the largest error in prostate or background are selected as “hard” pixels. The rank loss function is calculated as follows:

$$\text{Loss}_{rank} = \frac{1}{K^2} \sum_{i=1}^K \sum_{j=1}^K \max \{0, p_{0i} - p_{1j} + \text{margin}\} \quad (6)$$

where the  $p_{0i}$  and  $p_{1j}$  are the prediction values of the  $i$ th hard pixel of background and the  $j$ th hard pixel of the prostate for the input image, respectively. We enforce  $p_{1j} > p_{0i} + \text{margin}$  in the training stage to ensure that the network pays more attention to the hard pixels. The value of the margin is set to 0.2 in this study.

### 2.3.2 | Hybrid loss of classification

To optimize the performance of the classification components, we propose the following hybrid loss function:

$$\begin{aligned} \text{Los } s_E = & \lambda_1 I(p_{T2w}, y) + \lambda_2 I(p_{ADC}, y) \\ & + \varepsilon(y) \lambda_3 I_{ni}(CRM_{T2w}, CRM_{ADC}) \end{aligned} \quad (7)$$

The  $I(p_{T2w}, y)$  and  $I(p_{ADC}, y)$  indicate the classification loss functions (cross-entropy loss) of T2w and ADC, respectively. And the cross-entropy loss function is calculated as follows:

$$I(p, y) = -[y \log(\sigma(p)) + (1 - y) \log(1 - \sigma(p))] \quad (8)$$

where  $\sigma(\cdot)$  is the sigmoid function, and  $y \in \{0, 1\}$  is the slice-level label.

The  $I_{ni}(CRM_{T2w}, CRM_{ADC})$  indicates the normalized consistency loss function representing the consistency of  $CRM_{T2w}$  and  $CRM_{ADC}$ .  $CRM_{T2w}$  and  $CRM_{ADC}$  indicate the CRM, which are calculated from T2w and ADC images, respectively. The normalized consistency loss function is calculated as follows:

$$\begin{aligned} I_{ni}(CRM_{T2w}, CRM_{ADC}) \\ = \frac{1}{N} \|\sigma(CRM_{T2w}) - \sigma(CRM_{ADC})\| \end{aligned} \quad (9)$$

where  $N$  is total number of pixels in an image, and  $\|\sigma(CRM_{T2w}) - \sigma(CRM_{ADC})\|$  calculates the differences between the  $CRM_{T2w}$  and  $CRM_{ADC}$  as follows:

$$\begin{aligned} \sigma(CRM_{T2w}) - \sigma(CRM_{ADC}) \\ = \sqrt{\sum_{x,y} (\sigma(CRM_{T2w}(x, y)) - \sigma(CRM_{ADC}(x, y)))^2} \end{aligned} \quad (10)$$

The weighting factors  $\lambda_1, \lambda_2, \lambda_3$  are set to 1, 1, 0.25 in the network, and  $\varepsilon(y)$  is a step function. The step function constrains the classification hybrid loss and calculates the consistency only with the malignant images.

### 2.4 | Implementation

The Adam algorithm with a batch size of 16,  $\beta_1 = 0.9$ ,  $\beta_2 = 0.999$  ( $\beta_1$  and  $\beta_2$  are exponential decay rates for the moment estimates), and  $\text{decay} = 1e-6$  was adopted to optimize the segmentation components.<sup>32</sup> And, the SGD algorithm with momentum = 0.9 and batch size = 16 was adopted to optimize the classification components. The initial learning rate for both the segmentation and classification component was set to

0.001. The weights of the networks were initialized using the default initialization mechanism of the Keras framework. We also compared our segmentation network with Unet and Unet++,<sup>33</sup> and our classification network with ResNet50 and VGG16. All the experiments were performed in the Keras framework. The training strategies were optimized in the same computer system with Intel i5-8300H CPU, 16 GB RAM, and a GeForce GTX 1080 graphics processing unit (GPU). The details of the training process and the learning curves can be found in the supplementary materials (Figures S2–S5).

### 2.5 | Statistical analysis

Four quantitative metrics were used to evaluate the quality of the outputs of the segmentation components: intersection over union (IOU), Dice similarity coefficient (DSC), recall, and precision. Classification results were evaluated using four performance metrics: AUC, precision, sensitivity, and specificity.

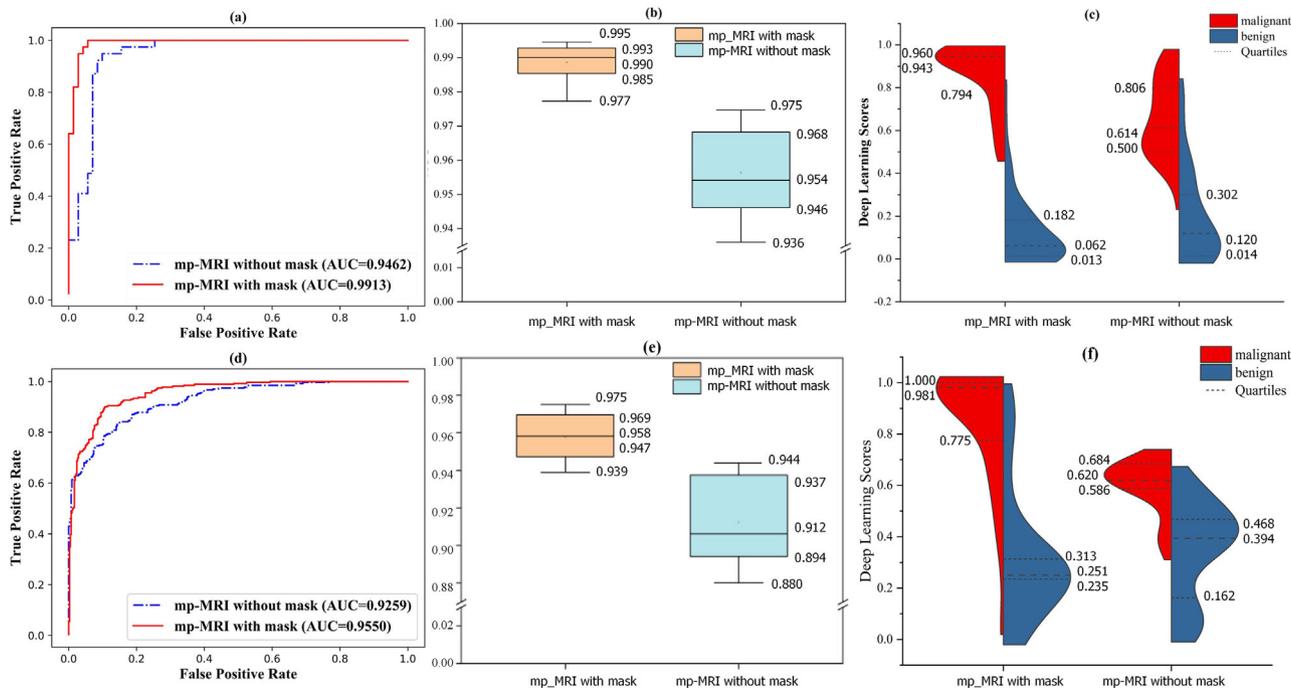
Paired *t*-test was applied to evaluate the contribution of the CRM, the hybrid segmentation loss, and the architecture of the segmentation network. The DeLong test<sup>34</sup> was used to obtain statistical significance (*p*-value) by comparing AUC between the baseline methods (Unet) and the proposed classification network. Statistical significance was defined as  $p < 0.05$ . The model was selected based on the Youden index to ensure a fair comparison.

## 3 | RESULTS

### 3.1 | Evaluation of the contribution of segmentation to classification

To verify the contribution of prostate segmentation to PCa classification, we compared the PCa classification performance with and without using coarse prostate masks generated by the coarse segmentation component.

Figure 3(a) compares ROC curves of the classification component obtained with/without using the coarse masks for center A data. The AUC of the results with coarse masks (0.991, 95% CI: 0.981–0.998,  $p < 0.02$ ) is substantially higher than that of the results without masks (0.946, 95% confidence interval [CI]: 0.914–0.981). Figure 3(b) shows the statistical results of the AUC from the top 10 optimized trained-models with weights at different epochs (ranked by the Youden index). The average AUC calculated from these top 10 optimized models also shows that the accuracy of PCa classification is significantly improved with the coarse segmentation component (from 0.954 to 0.991,  $p < 0.01$ ). Moreover, Figure 3(c) shows the deep learning score (the probability of the malignancy) by violin



**FIGURE 3** Each of the columns from left to right correspond to the area under curve (ROC) curve of the classification component, the AUC boxplot of the classification component, and violin plots of classification scores (representing cases when the coarse mask is or is not used), respectively. In the top row data from center A is presented, while in bottom row data from center B is presented.

plots. With the bootstrapping of the coarse segmentation component, the median value of the deep learning score for malignancy was improved from 0.613 to 0.943, for the nonmalignancy score was decreased from 0.120 to 0.062. Similar statistical results are shown in panels (d), (e), and (f) for center B data.

To further validate the improvement of the network's performance when coarse segmentation masks are included, we also compare the heatmaps obtained by the classification component with or without using prostate masks in the supplementary materials (Figure S6).

### 3.2 | Evaluation of the contribution of classification to segmentation

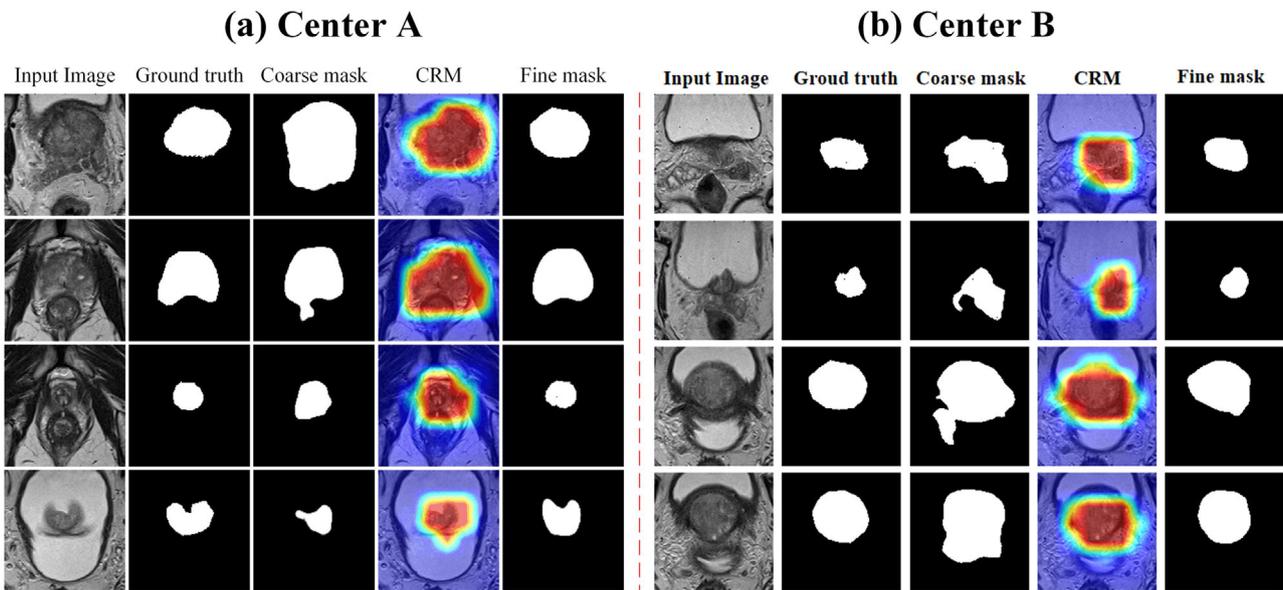
To evaluate the contribution of the CRM generated by the classification component to the segmentation results, we compared the segmentation results from the coarse segmentation component and the fine segmentation component for center A and center B (shown in Figure 4). The center B dataset is used for external testing. As mentioned above, the only difference between the coarse and fine segmentation is that the input to the fine segmentation is T2w images concatenated with CRM, while the input to the coarse segmentation component is only T2w images. In Figure 4, the columns from left to right correspond to input images, ground

truth masks, prostate masks after coarse segmentation, the CRM generated by the classification, and the fine prostate masks output from the fine segmentation component. Note that the segmentation results are more accurate with the additional input of the localization maps learned by the classification component.

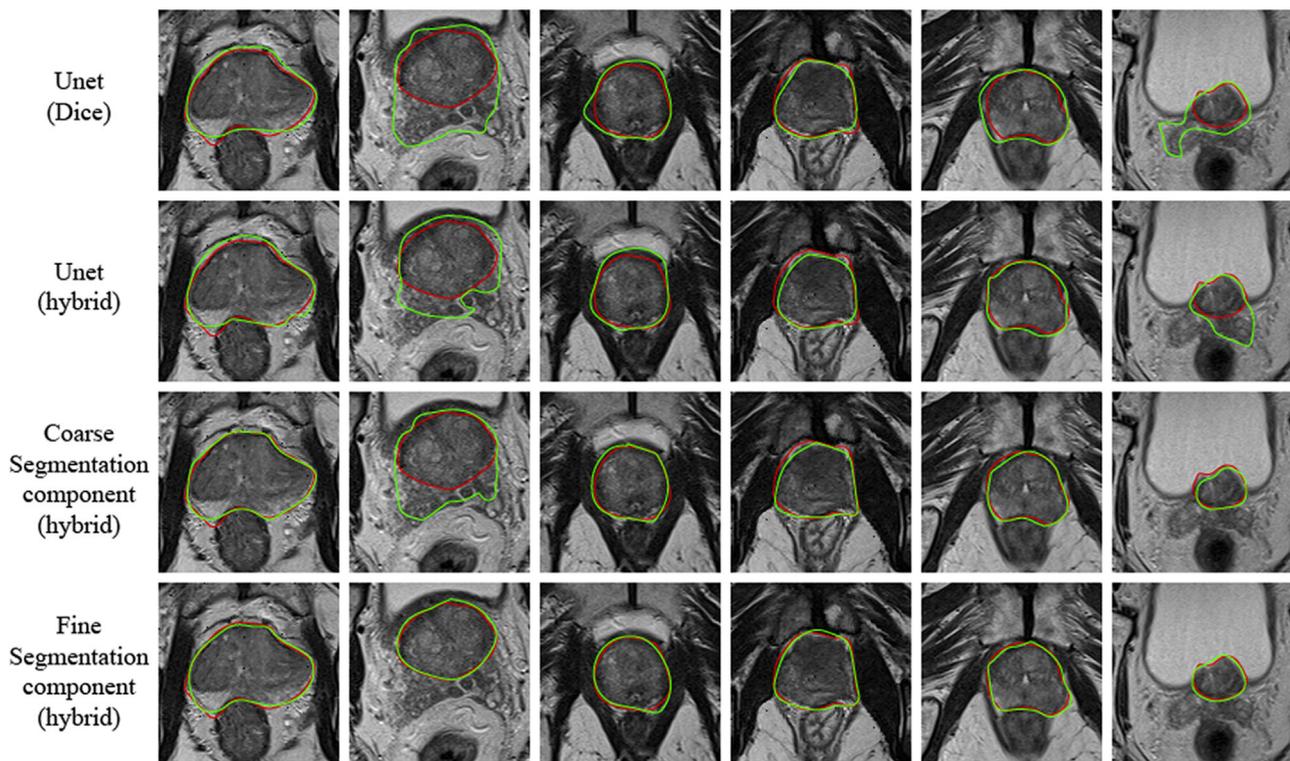
### 3.3 | Segmentation results of MC-DSCN

To demonstrate the effectiveness of the segmentation component in MC-DSCN, we compared our segmentation network with the widely used Unet. Unet is one of the most widely used networks for the segmentation of medical images due to its robustness and generally good performance. The visual results and statistical metrics are shown in Figures 5 and 6, respectively.

In Figure 5, as mentioned above, the main advantage of our model arises from the incorporation of the prostate location information in the segmentation. Note the significant improvement based on this modification by comparing the last two rows (coarse segmentation and fine segmentation). Comparison of Unet when using only Dice loss to Unet with our hybrid loss demonstrates that better segmentation results can be obtained, indicating the advantage of the hybrid loss as it can constrain the segmentation results both at the image and the pixel level. Also note that the segmentation results of our model are closer to the ground truth, as our model



**FIGURE 4** Comparison of the segmentation results of the coarse and fine segmentation components for center A and center B.

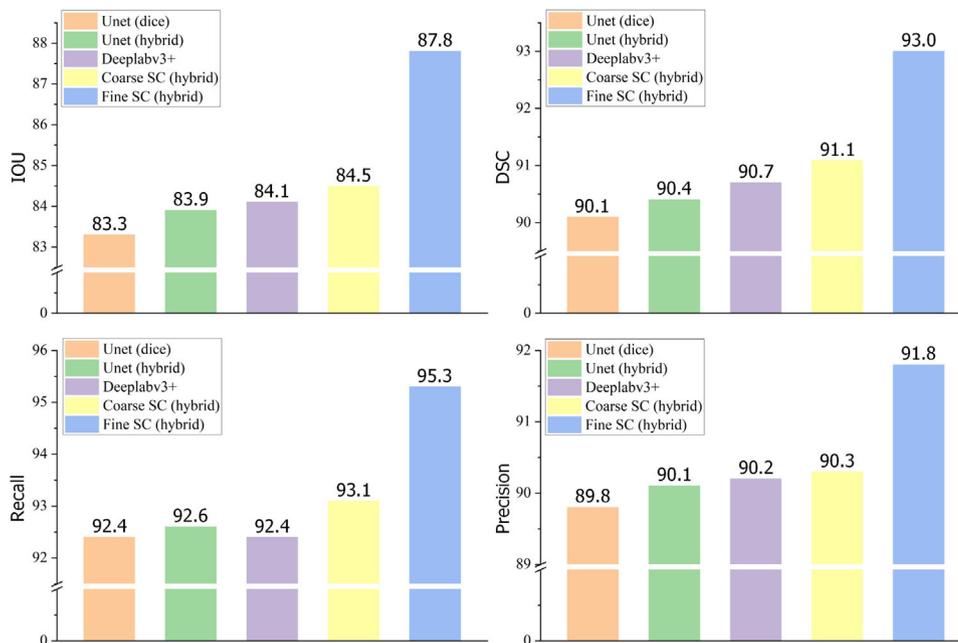


**FIGURE 5** Representative results of prostate segmentation. Red and green contours correspond to the ground truth and prediction of the prostate, respectively.

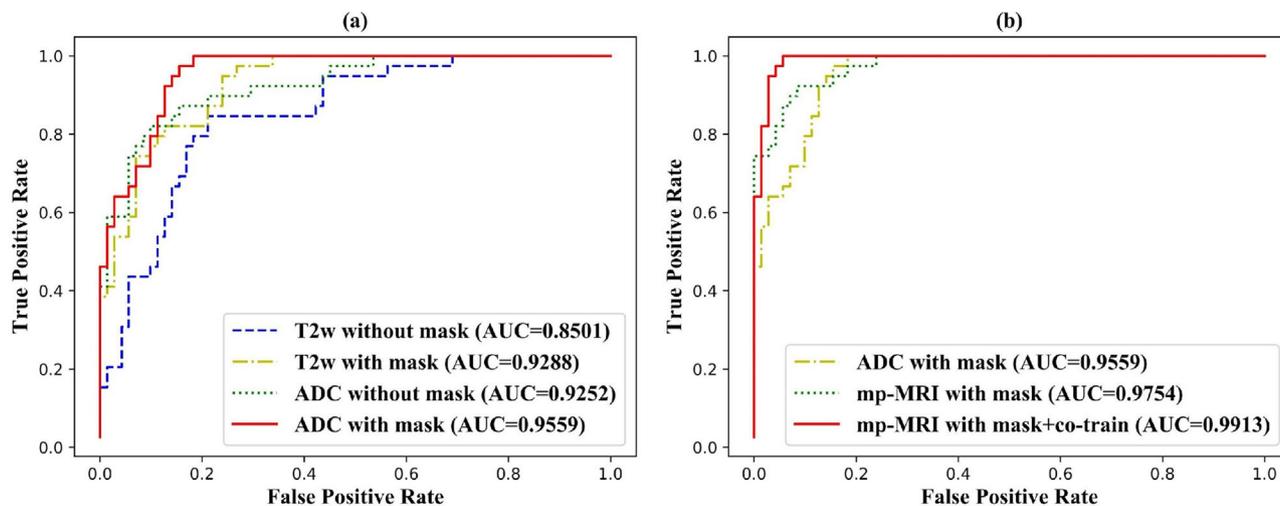
takes advantage of both the attention blocks and the residual mechanism.

In Figure 6 are presented statistical metrics corresponding to different models for center A. All quantitative metrics of segmentation show that our method outperforms Unet. For example, the IOU of the fine

segmentation component is better than the coarse segmentation component (improved from 84.5% to 87.8%,  $p < 0.01$ ). Both coarse and fine segmentation results are better than for Unet (83.9%,  $p < 0.05$ ). And the statistical metrics corresponding to different models for center B can be found in the supplementary materials (Figure



**FIGURE 6** Comparison of quantitative metrics for different segmentation networks, including Unet with Dice and hybrid loss, coarse segmentation component (Coarse SC), and fine segmentation component (Fine SC) for center A. The loss is indicated in brackets, including Dice loss (Dice) and segmentation hybrid loss (hybrid). Both Coarse SC and Fine SC are based on a residual Unet with attention blocks.



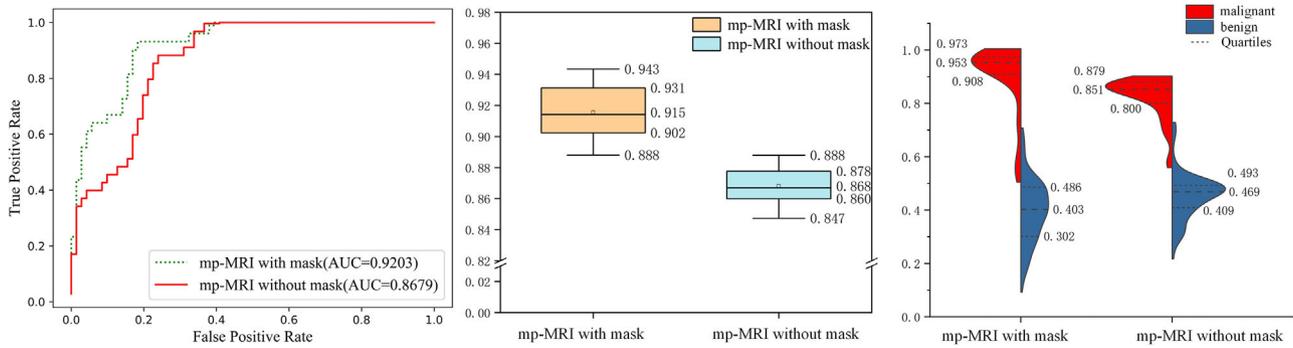
**FIGURE 7** (a) ROC curves of the classification component for different data sets used as an input. Dashed, dotted, dash-dotted, and solid lines correspond to the ROC curves of the model with input data being T2w or ADC images without masks, and T2w or ADC images along with masks, respectively. (b) Dash-dotted, dotted, and solid lines correspond to the ROC curve of the model with imported data that included ADC images and masks, mp-MRI input (T2w and ADC along with masks), and mp-MRI input (T2w and ADC, with masks and co-trained), respectively.

S7). Moreover, we also compared our segmentation network with Unet++, the visual results and statistical metrics can be found in the supplementary materials (Figures S8 and S9).

### 3.4 | Classification results of MC-DSCN

Figure 7 compared the ROC curves of the classification component for different input modes (T2w with

mask, ADC with mask, mp-MRI with mask, et al.). From Figure 7(a), it can be noted that for both T2w and ADC images, the AUC with coarse masks is always higher than the AUC without coarse masks. For T2w images, the AUC is improved from 0.850 to 0.929 ( $p < 0.05$ ). And for ADC images, the AUC is improved from 0.925 to 0.956 ( $p < 0.05$ ). In PI-RADS 2.0, it has been pointed out that T2w images reflect the prostate's anatomical information, and it is useful for delineating suspicious lesions because of their shorter "T2 relaxation time" compared



**FIGURE 8** The external testing results for center A dataset with the network trained with only center B dataset. Each of the columns from left to right corresponds to the ROC curve of the classification component, the area under curve (AUC) boxplot of the classification component, and violin plots of classification scores obtained with/without using the coarse masks, respectively.

to normal glandular tissue.<sup>35,36</sup> ADC maps, which reflect the degree of water diffusivity in the prostate, show low values in cancerous regions due to higher cellularity. In addition, ADC values reflect the aggressiveness of the PCa, with lower ADCs associated with higher-grade tumors.<sup>36</sup>

Figure 7(b) shows a comparison of the AUC for ADC and mp-MRI. Several studies have demonstrated that mp-MRI will significantly increase the sensitivity and specificity of the classification because mp-MRI includes anatomical and functional information. Also, in this study, the AUC of mp-MRI data (AUC: 0.975) was higher than for ADC (AUC: 0.956,  $p < 0.05$ ). Additionally, we evaluated the contribution of the co-training block in our model. As mentioned above, the function of the co-training block is to ensure the consistency of CRM between the different modalities. Compared to the model without the co-training block, AUC increased from 0.975 to 0.991 ( $p < 0.05$ ). In addition, we compared our classification network with ResNet50 and VGG16 for classification, and the classification results can be found in the supplementary materials (Figure S10).

### 3.5 | Evaluation of the generalization ability of the network

After the deep learning network training is completed, it is essential to test the generalization capability of the network. Thus, two medical centers (center A: First Affiliated Hospital of Shenzhen University, center B: Union Hospital) datasets were collected, one for training and the other one for external testing. As was mentioned in the Methods section, all the results presented above are based on the model trained with the center A dataset, and center B used for external testing. In the following test, we try to use center B data for training and data from center A for external testing, and the results are shown in Figure 8. Figure 8(a) are compared ROC curves of the classification component obtained

with/without using the coarse masks for center A data, and the network is trained with center B dataset. The AUC of the results with coarse masks (0.920, 95% CI: 0.902–0.931,  $p < 0.02$ ) is substantially higher than that of the results without masks (0.868, 95% confidence interval [CI]: 0.860–0.878). Similar to Figures 3 and 8(b) shows the statistical results of the AUC from the top 10 optimized trained-models with weights at different epochs. The average AUC calculated from these top 10 optimized trained-models also shows that the accuracy of PCa classification is significantly improved with the coarse segmentation component (from 0.868 to 0.915,  $p < 0.01$ ). Figure 8(c) shows the deep learning score (the probability of the malignancy) using violin plots.

## 4 | DISCUSSION

### 4.1 | Advantages of mutual bootstrapping

This study has developed a new method (MC-DSCN) for prostate segmentation and PCa classification. The proposed model can not only transfer the pixel-level information of prostate location produced by the segmentation network (coarse segmentation) to the classification network, but also transfer the image-level prostate localization information learned by classification network to the fine-segmentation network to alleviate the impact of inaccurate localization on segmentation results.

Firstly, we concatenate the multi-parametric MRI images and corresponding prostate masks predicted by coarse-segmentation block as the input to the classification network, aiming to use the results of prostate segmentation to facilitate classification of PCa. In section A of the results, we found that AUC of the classification network with coarse masks (0.991, 95% CI: 0.981–0.998) is substantially higher than that of the results without masks (0.946, 95% CI: 0.914–0.981).

Such performance gain is understandable, since the predicted masks enable the classification network to focus more on the prostate instead of background tissues and thus strengthen the ability of the classification network to diagnose PCa. To further validate this explanation, we visualized in the supplementary material (Figure S6) CRMs obtained by the classification network with and without using the coarse lesion masks. It shows that, when using the coarse lesion masks, the obtained CRMs are more focused on the prostate.

Secondly, the advantage of using classification to boost segmentation is usage of images with only image-level annotations to facilitate the training of a segmentation network, reducing the requirements for the pixel-level dense annotation. To demonstrate this advantage, we compared the proposed fine-segmentation model to a fully-supervised segmentation model (coarse-segmentation model), which has the same architecture as fine-segmentation and is trained only on images with pixel-level labels, but without using any lesion localization maps. The IOU of the fine segmentation increases compared to coarse segmentation (improved from 84.5% to 87.8%,  $p < 0.01$ ). In Figure 4, a comparison of the segmentation results of the coarse and fine segmentation network is shown. Moreover, the segmentation results based on Unet++ shown in the supplementary material (Figures S8 and S9) also demonstrate that the image-level information obtained by the classification network can help for the segmentation task.

## 4.2 | Advantages of multi-parametric MRI for PCa classification

As reported in Prostate Imaging Reporting and Data System Version 2.0 (PI-RADS 2.0) (a standardized PCa reporting system published in 2015 by the American College of Radiology), T2W and ADC from DWI are two most recommended sequences for PCa classification. Specifically, it pointed out that T2w which reflects the anatomy of a prostate is useful for delineating suspicious lesions because of their shorter T2 relaxation time compared with normal glandular tissue. And DWI which reflects the degree of water diffusion has low ADC values in cancerous regions due to the tightly packed cells. Moreover, several studies have shown significant increases in both sensitivity and specificity by combining both ADC and T2w.<sup>2,5,18</sup> Thus, in this work, we also utilize the ADC and T2w images. As can be noticed from Figure 7, the AUC of mp-MRI data (AUC: 0.991) was higher than AUC obtained only when using ADC (AUC: 0.956,  $p < 0.05$ ) or T2w (AUC: 0.929,  $p < 0.05$ ) images. In addition, we performed an ablation experiment with and without using T1 weighted images as an input for classification, and the results are shown in the supplementary material (Figure S11). The performance of classification was not improved when using

T1w images (98.6% with T1w, T2w and ADC, and 99.1% with T2w and ADC). One of potential reasons for not increasing AUC with additional T1w images might be similar anatomical information provided by T1w and T2w images.

Additionally, we evaluated the contribution of the co-training block in our model. As mentioned above, the function of the co-training block is to ensure consistency of CRMs derived using different modalities. We present the CRM obtained using only T2w or ADC in the supplementary material (Figure S12). It can be noted from this figure that the CRMs obtained either with T2w or ADC are different. We believe that the CNN of classification sees not only PCa-relevant patterns but also irrelevant visual patterns when distinguishing slices containing PCa from the normal ones. In addition, the irrelevant patterns are different for CNN ADC and CNN T2w, and in turn result in different CRMs. To address this problem, we enforce the CNN models of ADC and T2w to generate consistent prediction labels and similar CRMs. As the irrelevant visual patterns learned from CNN ADC are different from those learned from CNN T2w while PCa-relevant patterns learned from both models are similar, the enforcement can greatly reduce the amount of irrelevant visual patterns without losing PCa-relevant patterns.

To evaluate the loss function of the segmentation network, we trained our model using various loss functions, including generalized Dice loss (GDL), weighted cross-entropy (WCE) loss, the hybrid loss that combined GDL and WCE loss, and herein proposed loss (GDL + WCE loss + rank loss). The details of hyper-parameter evaluation for MC-DSCN are provided in the supplementary material.

There are some limitations to this study. Firstly, this new method is only validated on the ADC and T2w images. In future work, we plan to evaluate our method for a combination of more modalities such as Ktrans and DCE-MRI. Secondly, the network is based on 2D co-registered ADC-T2w images. Although, for all the datasets multi-slice 2D images cover a 3D volume for each patient, 2D slices from the same patient were treated independently in this work. Using a network trained on 3D data would allow to capture the spatial dependencies between slices of 3D data, leading to better performance in tasks that require 3D contextual information and can potentially extract more information compared to 2D networks. Our future plan is to extend our current model to multi-view CNNs for joint analysis of 3D MRI datasets. In addition, fusion methods for multimodal images also should be explored. Secondly, in this study, we only provide the malignant and nonmalignant classification for the PCa diagnosis. In future work, we would like to include the Gleason score for a more accurate diagnosis of PCa. Finally, this study's registration method for T2w and ADC images does not consider motion-induced misalignments.

## 5 | CONCLUSION

A new mutually communicated MC-DSCN is proposed for prostate segmentation and PCa classification based on mp-MRI. This network can effectively learn T2w and ADC image features by completing the prostate segmentation and PCa classification task and achieving mutual guidance and promotion. The proposed network has a certain value for application to prostate segmentation, PCa classification, and simple lesion detection in a clinical context.

## ACKNOWLEDGMENTS

We gratefully thank Chongxin Bai, Weida Xie, and Zhao Li for their advice on experimental design and the timely help in analyzing the large number of samples. We also gratefully acknowledge the financial support by the National Major Scientific Research Equipment Development Project of China (81627901), the National key of R&D Program of China (Grant 2018YFC0115000, 2016YFC1304702), National Natural Science Foundation of China (11575287, 11705274), and the Chinese Academy of Sciences (YZ201677).

## CONFLICT OF INTEREST STATEMENT

The authors have no relevant conflicts of interest to disclose.

## DATA AVAILABILITY STATEMENT

The data in this study can be requested from the corresponding author.

## REFERENCES

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin.* 2016;66(1):7-30. <https://doi.org/10.3322/caac.21332>
2. Peng Y, Jiang Y, Yang C, et al. Quantitative analysis of multiparametric prostate MR images: differentiation between prostate cancer and normal tissue and correlation with Gleason score - a computer-aided diagnosis development study. *Radiology.* 2013;267(3):787-796. <https://doi.org/10.1148/radiol.13121454>
3. De Rooij M, Hamoen EHJ, Fütterer JJ, Barentsz JO, Rovers MM. Accuracy of multiparametric MRI for prostate cancer detection: a meta-analysis. *Am J Roentgenol.* 2014;202(2):343-351. <https://doi.org/10.2214/AJR.13.11046>
4. Hoeks CMA, Barentsz JO, Hambrock T, et al. Prostate cancer: multiparametric MR imaging for detection, localization, and staging. *Radiology.* 2011;261(1):46-66. <https://doi.org/10.1148/radiol.11091822>
5. Fehr D, Veeraraghavan H, Wibmer A, et al. Automatic classification of prostate cancer Gleason scores from multiparametric magnetic resonance images. *Proc Natl Acad Sci U S A.* 2015;112(46):E6265-E6273. <https://doi.org/10.1073/pnas.1505935112>
6. Gibbons M, Starobinets O, Simko JP, Kurhanewicz J, Carroll PR, Noworolski SM. Identification of prostate cancer using multiparametric MR imaging characteristics of prostate tissues referenced to whole mount histopathology. *Magn Reson Imaging.* 2022;85(2):251-261. <https://doi.org/10.1016/j.mri.2021.10.008>
7. Turkbey B, Choyke P L. Multiparametric MRI and prostate cancer diagnosis and risk stratification. *Curr Opin Urol.* 2012; 261(3):36-46. <https://doi.org/10.1097/MOU.0b013e32835481c2>
8. Sankineni S, Osman M, Choyke PL. Functional MRI in prostate cancer detection. *Biomed Res Int.* 2014;2014. <https://doi.org/10.1155/2014/590638>
9. Valerio M, Donaldson I, Emberton M, et al. Detection of clinically significant prostate cancer using magnetic resonance imaging-ultrasound fusion targeted biopsy: a systematic review. *Eur Urol.* 2015;68:8-19. <https://doi.org/10.1016/j.eururo.2014.10.026>
10. Majdi MS, Keerthivasan MB, Rutt BK, Zahr NM, Rodriguez JJ, Saranathan M. Automated thalamic nuclei segmentation using multi-planar cascaded convolutional neural networks. *Magn Reson Imaging.* 2020;73(7):45-54. <https://doi.org/10.1016/j.mri.2020.08.005>
11. Kar BJ, Cohen M V, McQuiston SP, Malozzi CM. A deep-learning semantic segmentation approach to fully automated MRI-based left-ventricular deformation analysis in cardiotoxicity. *Magn Reson Imaging.* 2021;78(3):127-139. <https://doi.org/10.1016/j.mri.2021.01.005>
12. Wang Y, Yu B, Zhong F, et al. MRI-based texture analysis of the primary tumor for pre-treatment prediction of bone metastases in prostate cancer. *Magn Reson Imaging.* 2019;60(11):76-84. <https://doi.org/10.1016/j.mri.2019.03.007>
13. Wang S, Burt K, Turkbey B, Choyke P, Summers RM. Computer aided-diagnosis of prostate cancer on multiparametric MRI: a technical review of current research. *Biomed Res Int.* 2014;2014. <https://doi.org/10.1155/2014/789561>
14. Tian C, Fei L, Zheng W, Xu Y, Zuo W, Lin CW. Deep learning on image denoising: an overview. *Neural Networks.* 2020;131:251-275. <https://doi.org/10.1016/j.neunet.2020.07.025>
15. Zhang Y, Gorris JM, Dong Z. Deep learning in medical image analysis. *J Imaging.* 2021;7(4):31-38. <https://doi.org/10.3390/jimaging7040074>
16. Freidlin RZ, Agarwal HK, Sankineni S, et al. Application of an unsupervised multi-characteristic framework for intermediate-high risk prostate cancer localization using diffusion-weighted MRI. *Magn Reson Imaging.* 2016;34(9):1227-1234. <https://doi.org/10.1016/j.mri.2016.06.004>
17. Song Y, Zhang YD, Yan X, et al. Computer-aided diagnosis of prostate cancer using a deep convolutional neural network from multiparametric MRI. *J Magn Reson Imaging.* 2018;48(6):1570-1577. <https://doi.org/10.1002/jmri.26047>
18. Yang X, Liu C, Wang Z, et al. Co-trained convolutional neural networks for automated detection of prostate cancer in multiparametric MRI. *Med Image Anal.* 2017;42:212-227. <https://doi.org/10.1016/j.media.2017.08.006>
19. Sanford T, Harmon SA, Turkbey EB, et al. Deep-learning-based artificial intelligence for PI-RADS classification to assist multiparametric prostate MRI interpretation: a development study. *J Magn Reson Imaging.* 2020;52(5):1499-1507. <https://doi.org/10.1002/jmri.27204>
20. Zhu Y, Wei R, Gao G, et al. Fully automatic segmentation on prostate MR images based on cascaded fully convolution network. *J Magn Reson Imaging.* 2019;49(4):1149-1156. <https://doi.org/10.1002/jmri.26337>
21. Alkadi R, Taher F, El-baz A, Werghi N. A deep learning-based approach for the detection and localization of prostate cancer in T2 magnetic resonance images. *J Digit Imaging.* 2019;32(5):793-807. <https://doi.org/10.1007/s10278-018-0160-1>
22. Ghavami N, Hu Y, Gibson E, et al. Automatic segmentation of prostate MRI using convolutional neural networks: investigating the impact of network architecture on the accuracy of volume measurement and MRI-ultrasound registration. *Med Image Anal.* 2019;58:101558. <https://doi.org/10.1016/j.media.2019.101558>
23. Wang B, Lei Y, Tian S, et al. Deeply supervised 3D fully convolutional networks with group dilated convolution for automatic MRI prostate segmentation. *Med Phys.* 2019;46(4):1707-1718. <https://doi.org/10.1002/mp.13416>
24. Xie Y, Zhang J, Xia Y, Shen C. A mutual bootstrapping model for automated skin lesion segmentation and classification. *IEEE*

- Trans Med Imaging*. 2020;39(7):2482-2493. <https://doi.org/10.1109/TMI.2020.2972964>
25. Chen S, Wang Z, Shi J, Liu B, Yu N. A multi-task framework with feature passing module for skin lesion classification and segmentation. *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA*; 2018, 1126–1129. <https://doi.org/10.1109/ISBI.2018.8363769>
  26. González-Díaz I. DermaKNet: Incorporating the knowledge of dermatologists to convolutional neural networks for skin lesion diagnosis. *IEEE J Biomed Heal Informatics*. 2019;23(2):547-559. <https://doi.org/10.1109/JBHI.2018.2806962>
  27. Schaefer S, McPhail T, Warren J. Image deformation using moving least squares. *ACM SIGGRAPH 2006 Pap SIGGRAPH '06*. Published online; 2006: 533-540. <https://doi.org/10.1145/1179352.11441920>
  28. Levin D. The approximation power of moving least-squares. *Math Comput*, 1998, 67(224): 1517–1531
  29. Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*; 2015;9351:319-340. [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28)
  30. He K, Zhang X, Ren S, Sun J. Identity mappings in deep residual networks. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2016;9908:630-645. [https://doi.org/10.1007/978-3-319-46493-0\\_38](https://doi.org/10.1007/978-3-319-46493-0_38)
  31. Sudre CH, Li W, Vercauteren T, Ourselin S, Jorge Cardoso M. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2017;10553:240-248. [https://doi.org/10.1007/978-3-319-67558-9\\_28](https://doi.org/10.1007/978-3-319-67558-9_28)
  32. Kingma DP, Ba JL. Adam: a method for stochastic optimization. *3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc*. Published online; 2015:1-15
  33. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: a nested U-Net architecture for medical image segmentation. *Deep Learn Med Image Anal Multimodal Learn Clin Decis Support*. 2018;9:3-11. [https://doi.org/10.1007/978-3-030-00889-5\\_1](https://doi.org/10.1007/978-3-030-00889-5_1)
  34. Chen LC, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 2018;11211 LNCS:833-851. [https://doi.org/10.1007/978-3-030-01234-2\\_49](https://doi.org/10.1007/978-3-030-01234-2_49)
  35. Weinreb JC, Barentsz JO, Choyke PL, et al. PI-RADS prostate imaging - reporting and data system: 2015, Version 2. *Eur Urol*. 2016;69(1):16-40. <https://doi.org/10.1016/j.eururo.2015.08.052>
  36. Gibbs P, Tozer DJ, Liney GP, Turnbull LW. Comparison of quantitative T2 mapping and diffusion-weighted imaging in the normal and pathologic prostate. *Magn Reson Med*. 2001;46(6):1054-1058. <https://doi.org/10.1002/mrm.1298>

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Liu K, Li P, Otikovs M, et al. Mutually communicated model based on multi-parametric MRI for automated segmentation and classification of prostate cancer. *Med Phys*. 2023;50:3445–3458. <https://doi.org/10.1002/mp.16343>