

Received September 4, 2021, accepted September 18, 2021, date of publication September 22, 2021,
date of current version October 27, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3114733

ProCDet: A New Method for Prostate Cancer Detection Based on MR Images

YUEJING QIAN^{ID}¹, ZENGYOU ZHANG¹, AND BO WANG^{ID}²

¹Zhejiang Industry and Trade Vocational College, Wenzhou, Zhejiang 313103, China

²Zhejiang College of Security Technology, Wenzhou, Zhejiang 325000, China

Corresponding author: Bo Wang (zjaf_wb@126.com)

This work was supported by the Natural Science Foundation of Zhejiang Province under Grant LGF19F020004, and in part by Wenzhou Science and Technology Project under Grant G20190022.

ABSTRACT Prostate cancer is a malignant tumor that occurs in the male prostate. Prostate cancer lesions have the characteristics of small size and blurry outline, which is a challenge to design a robust prostate cancer detection method. At present, clinical diagnosis of prostate cancer is mainly based on magnetic resonance (MR) imaging. However, it is difficult to obtain prostate cancer data, and the data with true values is also very limited, which further increases the difficulty of prostate cancer detection methods based on MR images. To solve these problems, this paper designs a new method of prostate cancer detection based on MR images, which is recorded as ProCDet. The method consists of three modules: registration of prostate MR images, segmentation of prostate, and segmentation of prostate cancer lesions. First, the registration between different sequences of MR images is performed to find the spatial relationship between the different sequences. Then, the designed prostate segmentation network based on the attention mechanism is used to segment the prostate to remove the interference of background information. Finally, a 3D prostate cancer lesion segmentation network based on Focal Tversky Loss is applied to determine the specific location of prostate cancer. Moreover, in order to take full advantage of unlabeled prostate data, this paper designs a self-supervised learning method to improve the accuracy of prostate cancer detection. The proposed ProCDet has been experimentally verified on the ProstateX dataset. When the average number of false-positive lesions per patient is 0.6275, the true-positive rate is 91.82%. Experimental results show that the ProCDet can obtain competitive detection performance.

INDEX TERMS Prostate cancer detection, MR image, image registration, self-supervised learning, prostate segmentation.

I. INTRODUCTION

The prostate belongs to the male reproductive system and is a gland unique to men. The prostate is mainly composed of three regions: the transition zone, the central zone and the peripheral zone [1]. The peripheral zone is the largest part of the prostate, and the central zone is the second largest part of the prostate. The peripheral zone and the central zone together account for about 95% of the total volume of the prostate. Prostate cancer is a malignant tumor that occurs in the prostate. Approximately 70% of the probability of prostate cancer occurs in the peripheral zone, and only 2.5% of the probability of prostate cancer occurs in the central zone. Transition zone is the smallest part of the prostate, accounting

for about 5% of the total volume, but there is a 20% probability that prostate cancer will occur in this region [2].

According to related research by the American Cancer Society, in 2020, about 17,000 people in the United States have been diagnosed with prostate cancer, and about 30,000 people died of prostate cancer. Among men, the prevalence of prostate cancer is second only to skin cancer, and the mortality rate is second only to lung cancer [3]. Prostate cancer patients may have problems related to the urinary system and reproductive system, such as dysuria, sexual dysfunction, etc., and may even metastasize to other parts of the body with the blood and lymph [4].

Early detection and diagnosis of prostate cancer can effectively prevent prostate cancer from developing into advanced metastatic cancer, and can greatly improve the survival rate of patients. The current mainstream diagnosis process for

The associate editor coordinating the review of this manuscript and approving it for publication was Xiaojie Su^{ID}.

prostate cancer is to adopt prostate specific antigen (PSA) and magnetic resonance imaging (MRI) to pre-diagnose, and then perform biopsy pathological examination to confirm [5]. PSA blood test is a frequently used method for the diagnosis of prostate cancer. Through blood test, check the PSA content in the blood, if it is higher than a certain threshold, it means that the patient is very likely to develop prostate cancer. However, the sensitivity and specificity of the PSA test are not very high, and the specific location of the lesion cannot be known only through the PSA blood test [6]. Diagnosis through medical imaging can obtain a more comprehensive diagnosis result, thereby avoiding some unnecessary biopsy and surgery [7]. MRI is a commonly used imaging method in the diagnosis of prostate cancer. Through this method, doctors can clearly observe the prostate structure and pathological information, thereby greatly improving the sensitivity and specificity of diagnosis [8]. However, there are two problems with manual reading: 1) MR images are multi-modal images, which contain multiple sequences, such as T2 weighted sequence, DWI sequence, ADC sequence, KTrans sequence. And prostate cancer has different characteristics on different sequences. When diagnosing prostate cancer, doctors need to make judgments based on the characteristics of the lesions in different sequences, which requires doctors to have strong professional knowledge and rich clinical experience; 2) Manual reading takes a long time and is easily affected by subjective factors. Under high-intensity work pressure, the risk of missed or misdiagnosed will gradually increase [9]. Besides, for the same lesion, the diagnosis results given by different doctors may be different, that is, the results of manual reading are not repeatable.

To realize the automatic diagnosis of prostate cancer, this paper proposes a new method for detecting prostate cancer based on MR images, which is recorded as ProCDet. ProCDet can automatically segment the prostate from the MR image, and further obtain information such as the location, boundary and volume of the prostate cancer lesion. This not only improves the doctor's reading efficiency, but also greatly relieves the doctor's work pressure. In general, our technical contributions in this work mainly include the following three points.

(1) A new channel attention mechanism is designed and integrated into the prostate segmentation network based on UNet to improve the accuracy of prostate segmentation.

(2) A self-supervised learning network structure is proposed, and the designed local disturbance and combined disturbance transformer are employed to learn the shape, boundary, texture and global spatial distribution of the organization.

(3) The self-supervised learning network model is adopted as a pre-training model, and the 3D prostate cancer lesion segmentation network based on Focal Tversky Loss is used to determine the specific location of the prostate cancer.

II. RELATED WORK

With the development of medical imaging digital technology and the continuous improvement of computer performance, since the early 1990s, academia and business circles have set off a research climax on the computer-assisted detection and diagnosis of prostate cancer [10]. We can roughly divide these prostate cancer detection methods into two categories, one is non-CNN detection methods, and the other is CNN-based detection methods.

Although prostate cancer detection methods based on non-CNN have certain limitations, they have also made some progress [11]–[15]. Specifically, Barentsz *et al.*, proposed a fully automated computer-aided detection method for prostate cancer detection. This method uses the Hessian-based blob detection algorithm to perform voxel classification on multiple scales. Then, the candidate detection frame is characterized by performing histogram analysis on the multi-parameter MR image. Finally, a two-stage classification method is employed to classify the obtained feature set between benign and malignant [16]. In diffusion-weighted MR images, Firjani *et al.*, segmented the prostate according to the log-likelihood function. Then a non-rigid registration algorithm is applied to solve the local deformation caused by the patient's breathing and local motion. After that, based on the four appearance features extracted from the registered images, a classifier based on K nearest neighbors is utilized to classify the prostate as benign or malignant [17]. Reda *et al.*, first segmented the prostate based on a level set model; then used a generalized Gauss-Markov random field image model for normalization and optimization. Lastly, a stacked non-negative constraint algorithm is applied to train a deep learning autoencoder network to classify prostate tumors as benign or malignant [18]. Chung *et al.*, proposed a new method for automatically detecting prostate cancer using a radiation-driven conditional random field framework. In addition to using a complete set of voxel-level quantitative radiological features, they also utilized the relationship between the spatial and radiological features of voxel points to better detect prostate cancer [19]. Khalvati *et al.* [20] performed feature selection and analysis on each modality in the prostate MR image, and then constructed a texture feature model according to the best features of each modality to realize the detection of prostate cancer. Giannini *et al.*, developed a two-stage fully automated prostate cancer detection system. The system first creates a malignant probability map of all voxels in the prostate; then performs a candidate segmentation step to highlight the suspicious area, thereby achieving the location of prostate cancer [21]. Khalvati *et al.*, proposed a bag of bags nested multi-instance learning (BoB MIL) algorithm, in which the high-level bags (or parent bags) each contain multiple smaller instance bags. They used magnetic resonance imaging data to apply the BoB MIL algorithm to the problem of prostate cancer detection [22]. Taking into account the difference between the lesion area and the non-lesion area of prostate cancer. Du *et al.* [23]

first segmented the prostate and calculated some statistical data in the prostate area to find the difference between prostate cancer and non-cancer, and finally employ an atlas matching method to detect prostate cancer. Carina *et al.*, extracted image intensity, gradient, gradient direction and distance features from T2-weighted (T2W), diffusion-weighted imaging (DWI) and apparent diffusion coefficient (ADC) MRI series. These features are then used as input to the secondary discriminant analysis model to detect prostate cancer [24]. Wang *et al.* [25] proposed a stack-based ensemble learning method for the detection of prostate cancer, which can simultaneously construct a diagnostic model and extract interpretable diagnostic rules.

With the rapid development of deep learning in recent years, many CNN-based methods have also achieved better results in prostate cancer detection [26]–[29]. For example, Wang *et al.*, developed a computer-assisted prostate cancer detection and diagnosis system driven by a new type of deep convolutional neural network (CNN). The system first detects the prostate in a multi-parameter magnetic resonance imaging sequence. Then employ CNN to identify images containing prostate cancer to generate a heat map of the cancer. In the end, multi-modal CNN features and SVM classifiers are utilized to evaluate the aggressiveness level of each local lesion [30]. Le *et al.*, presented an automatic detection method for prostate cancer based on multi-modal CNN. Among them, multi-modal CNN effectively merges ADC image and T2W image, and designs a similarity loss function to force the same feature to be extracted from ADC and T2W [31]. The prostate cancer detection network proposed by Wang *et al.* [32] consists of two sub-networks, a new tissue deformation network for prostate detection and multi-modal registration, and a dual-path convolutional neural network for prostate cancer detection. Cao *et al.*, designed an improved CNN to detect prostate cancer lesions and accurately segment the lesion tissue. Specifically, they adopted focus loss to overcome the imbalance between cancerous and non-cancerous regions to improve lesion detection performance, and designed selective dense conditional random fields for post-processing to refine the segmented lesion mask [33]. The 2.5D method presented by Ruba *et al.*, can simultaneously perform prostate segmentation and prostate cancer diagnosis. They use deep learning methods to extract high-level features from the original T2W image, and then remap the features to the input to assign the predicted label to each pixel [34]. Xin *et al.*, designed a joint training CNN structure, which consists of two parallel convolutional networks for ADC and T2W images. And in a weakly supervised manner, a single pattern of images is utilized to train each network, so that distinguished visual lesion patterns can be effectively learned from the prostate and surrounding tissues. Finally, the CNN features of each mode are fed into the SVM classifier in series to obtain the confidence of prostate cancer [35]. Ruba *et al.*, trained a deep convolutional encoder-decoder architecture to segment prostate and malignant lesions simultaneously. Besides, in order to

incorporate the 3D context spatial information provided by the MRI series, they proposed a 3D sliding window method, which uses 3D information while retaining the complexity of the 2D domain [36]. Cao *et al.* [37] designed a novel multi-class CNN, FocalNet, to jointly detect prostate cancer lesions and apply the gleason score to predict its aggressiveness. Tsehay *et al.* [38] adopted the most advanced edge detector network architecture to detect prostate cancer, which takes the image as input and generates an image probability map. Independent of the hand-made features, Iqbal *et al.* [39] combined the long short-term memory network and residual network in deep learning for the detection of prostate cancer. Duran-Lopez *et al.*, presented a method for detecting prostate cancer based on deep learning. This method first applies different filters for block sampling and preprocessing, then adopts a block scoring algorithm to remove useless regions from the tissue, and finally uses this image block as the input of a custom CNN to obtain a heat map that can reflect prostate cancer [40]. Yu *et al.* [41] proposed a new method for prostate cancer lesion detection and segmentation, which has both semantic branch and instance branch, and utilizes attention modules to better combine local and global image features.

III. THE PROPOSED METHOD

Because the characteristics of prostate cancer lesions are not obvious, and the features of different sequences are also different. If the diagnosis is made only on a single sequence, it may be misdiagnosed due to insufficient information. However, if diagnosis is performed on many sequences, there will be a lot of redundant information with a high probability, which is not conducive to the learning of the network model. According to related investigations and experiments, the sequences used in this paper are the T2W sequence, ADC sequence and DWI sequence in the patient's MR image. The schematic flow chart of the proposed prostate cancer detection (ProCDet) method is shown in FIGURE 1.

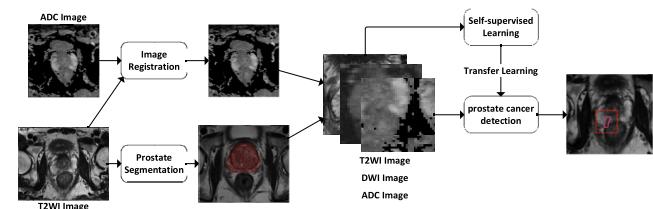


FIGURE 1. Schematic diagram of the processing flow of the ProCDet method.

The first step of the ProCDet method is the registration of MR images. Registration can correct the problem of misalignment between different images caused by tissue deformation during image acquisition. The second step of the ProCDet method is the segmentation of the prostate. Since the proportion of prostate cancer lesions in the entire MR image is small, the prostate segmentation network needs to be applied to segment the prostate, so that most of the irrelevant background information can be removed. To make

the most of the unlabeled data in the dataset, we adopted a self-supervised learning method to extract the features of the prostate MR image. And use self-supervised learning network parameters to initialize the prostate cancer lesion detection network to alleviate the overfitting problem caused by small sample datasets. The last step of the ProCDet method is to utilize the prostate cancer lesion segmentation network to segment the lesions, so as to realize the automatic detection of prostate cancer. The registration of images, segmentation of prostate, self-supervised learning network, and detection of prostate cancer will be introduced in detail below.

A. REGISTRATION OF MR IMAGES

During the image acquisition process, the patient may move to a certain extent, which makes the MR image different sequences may not be aligned. To improve the detection performance of prostate cancer, it is necessary to perform registration of prostate MR images.

According to whether the image pair to be registered belongs to the same patient and whether it belongs to the same image type, the registration of medical images can be mainly divided into three types. The three types are the registration of different modal images of the same patient, the registration of the same modal images of different patients, and the registration of the same modalities of the same patient collected at different times. Since the time for acquiring different sequences of MR images is different, the registration type in this article belongs to the third type, that is, the registration of the same modalities of the same patient acquired at different times.

In the diagnosis of prostate cancer, DWI images, ADC images, and T2W images play a vital role. In general, prostate cancer lesions show low signal in ADC images and T2W images, and high signal in DWI images. Since the ADC image is calculated from the DWI image, only the registration between the ADC image and the T2W image or the registration between the DWI image and the T2W image needs to be considered. This paper chooses the registration between ADC image and T2W image, which belongs to the registration between multi-modal images. Since the registration work proposed by Florkow *et al.* [42] is also the registration between multi-modal images, we draw on their research results. Florkow *et al.*, use mutual information as the objective function and perform B-spline transformation on the image. To prevent large deformation of the image during the transformation, they added a penalty for stiffness in the objective function. To make the registration easier, we employ two resolutions to enter the registration sequentially. First, the low-resolution ADC image is registered with the high-resolution T2W image to obtain the registered high-resolution ADC image. Then, the high-resolution T2W image is registered with the high-resolution ADC image to obtain the registered T2W image. An example of the result of the registration process of the prostate MR image is shown in FIGURE 2.

B. PROSTATE SEGMENTATION

The precise segmentation of the prostate is conducive to confining the prostate cancer lesions to the prostate region. The prostate MR image contains a variety of image sequences. Since the T2W image is a structural diagram and has a high resolution, the tissue structure can be clearly observed in the T2W. Therefore, the sequence used by the prostate segmentation module is a T2W image.

In order to better segment the prostate, this paper designs a U-Net [43] segmentation network based on the attention mechanism, and its network structure is shown in FIGURE 3.

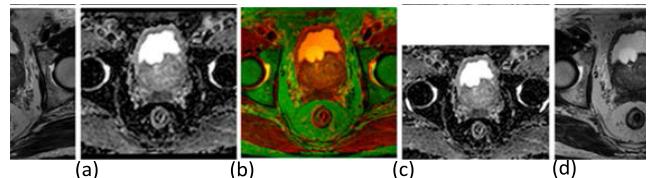


FIGURE 2. An example slices for image registration. Among them, (a) ADC slice; (b) T2W slice; (c) Result of registered ADC; (d) A registered ADC slice overlaid with a T2W slice.

The structure diagram of “Upsample Block” in FIGURE 3 is shown in FIGURE 4. Among them, the attention structure uses the channel attention mechanism CAB [44] proposed by Yu *et al.*, in 2018, and we have improved on it. Specifically, the CAB structure performs a global average pooling operation on two adjacent layers of feature maps to extract semantic information, ignoring the rich spatial features contained in the low-level features, and to a certain extent causes the lack of spatial information. As a result, the prostate region shrinks inward during segmentation, which reduces the integrity of the target. In response to this shortcoming, we designed the NewCAB module, increased the global max pooling operation, and used it to extract the spatial information of the lower-level features to retain more image spatial texture information. And use the global average pooling operation to extract the semantic information of higher-level features. Then, the high-level semantic features and the low-level spatial features are merged to construct a weight vector with both spatial and semantic information. Its structure is shown in FIGURE 5.

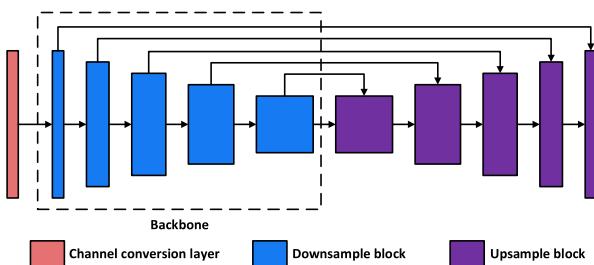


FIGURE 3. Schematic diagram of the structure of the prostate segmentation network.

The “Channel conversion layer” in FIGURE 3 is a 1×1 convolutional layer with 3 convolution kernels. Its purpose is

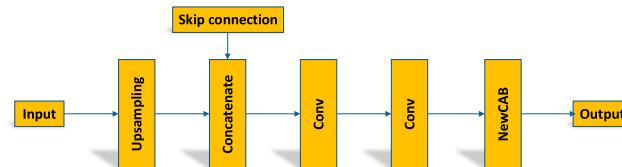


FIGURE 4. The structure diagram of Upsample bolock.

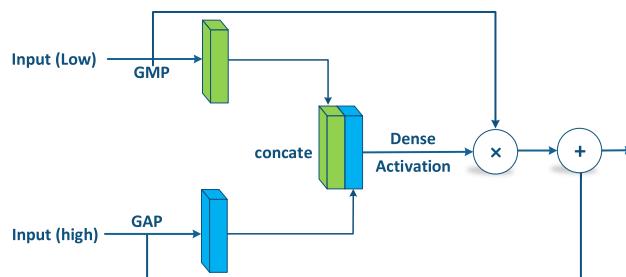


FIGURE 5. Schematic diagram of NewCAB structure.

to enable the designed prostate segmentation network to load the pre-trained model trained by ImageNet to realize transfer learning and accelerate the learning efficiency of the network. Furthermore, the input of the prostate segmentation network we designed is single-channel, and the number of input channels of the pre-training model on ImageNet is 3. Therefore, it is necessary to convert the number of channels input to the network from 1 to 3 before using the ImageNet pre-training model. The backbone network used in this paper is ResNet18, and the input size of the network is $512 \times 512 \times 1$. The parameters and FLOPs of the prostate segmentation network are 4.32×10^7 and 3.76×10^{10} , respectively.

When performing transfer learning, the network is prone to overfitting. To alleviate this problem, we use Stochastic Gradient Descent with Warm Restarts [45] (SGDR) as the optimizer. The SGDR optimizer combines the active annealing strategy with the periodic restart of the learning rate, which to a certain extent can help the network jump out of the local extreme point and reach the global optimum.

C. SELF-SUPERVISED LEARNING NETWORK

Unlike natural images, medical image data is mostly three-dimensional (3D) data, such as CT images, MR images, and so on. In addition, the amount of data of annotated medical images is very limited, and it is easy to cause network overfitting without transfer learning. However, the method of using a pre-trained network on ImageNet for transfer learning has the following two drawbacks: 1) Natural images and medical images are quite different, and this cross-domain migration is likely to cause the network to fail to extract effective features. 2) The prostate MR image is 3D data. If you use the pre-trained model on ImageNet, you must use a two-dimensional network to detect prostate cancer, but in this way the spatial information in the 3D data cannot be used.

To solve this problem, we adopted a self-supervised learning method, using a large amount of unlabeled data in the data

set to train a 3D pre-training model for transfer learning. The advantages of this 3D pre-training network are: 1) Since it is learned from the prostate MR image, it can better adapt to the detection of prostate cancer. 2) Since it is a 3D network, it can make better use of the spatial information in the 3D data. 3) This self-supervised learning method makes full use of unlabeled data and greatly reduces the cost of data labeling.

1) OVERALL FRAMEWORK

The framework of the self-supervised learning network designed in this paper is shown in FIGURE 6. The basic idea of the framework is that the input data can be restored to the original image after passing through the encoder-decoder network. Before entering the network, the input data needs to be disturbed by a converter. This disturbance increases the learning difficulty of the network, helps the network to extract better features, and alleviates the network overfitting. The function of the encoder-decoder is to restore the deformed data to the original image. After the self-supervised network learning is completed, a pre-trained model can be obtained. The pre-training model can be applied to the target task by means of transfer learning. Among them, the loss function L generally adopts the mean square error loss function.

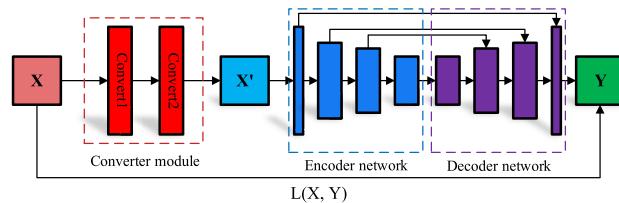


FIGURE 6. Schematic diagram of the framework of the designed self-supervised learning method.

2) CONVERTER

The converter used in this paper uses two different conversion methods: local disturbance and combined disturbance. We combine these two transformations and let the model learn from multiple angles to make the network model more robust. It will be described in detail below.

a: LOCAL DISTURBANCE

Randomly select some small windows in the image, scramble the pixels in this small window, and then let the model learn to restore them to the original image. When choosing the size of the window, it is necessary to ensure that the size of the window is smaller than the receptive field of the encoder-decoder network, so that the network can learn more about the feature representation of recovering local pixels. The size of the window can control the degree of deformation, and the effect diagram is shown in FIGURE 7. Unlike using noise to replace local pixels, disrupting local pixels can preserve the overall grayscale distribution of the image, which is more conducive to network learning. When selecting an appropriate window size, disrupting local pixels can increase the diversity of the data, and at the same time will not have a greater impact on

the overall structure of the data. This allows the model to learn more about the shape, boundary and texture of the organ.

b: COMBINED DISTURBANCE

Randomly select some windows with different sizes and different aspect ratios in the image, and then superimpose these windows together to form a complex-shaped window. Keep the data in this window and set the gray value of the pixels outside the window to 0. The operation process is shown in FIGURE 8. The network can learn global geometric information and the spatial distribution of organs by complementing the content outside the window. To limit the difficulty to a reasonable range and ensure that most of the data is preserved during transformation, we limit the area of these windows to more than three-quarters of the full image size.

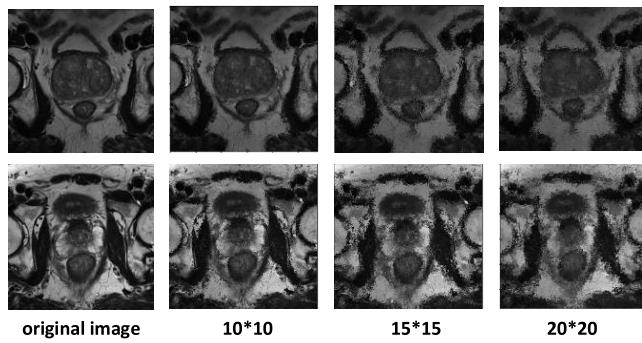


FIGURE 7. An example image of local disturbance.

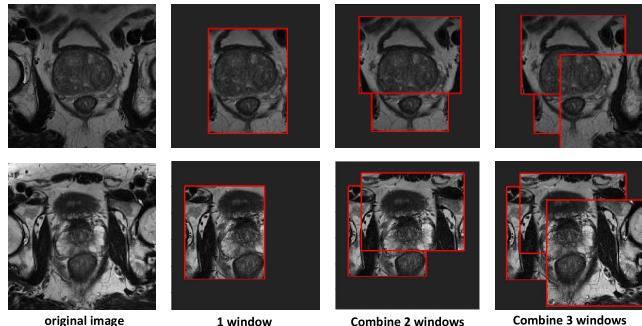


FIGURE 8. An example image of combined disturbances.

3) ENCODER-DECODER NETWORK

The encoder-decoder network in the self-supervised learning framework designed in this paper can change the structure of the network according to the actual needs of the problem. After pre-training, the encoder can be used for object classification tasks, and the encoder and decoder together can be used for semantic segmentation tasks. The network can be a 2D network or a 3D network. Since the task of this article is the detection of prostate cancer, it is necessary to convert the detection task into a segmentation task. We adopt a 3D segmentation network to segment prostate cancer to make better use of the spatial context information in the 3D data. The structure diagram of the 3D segmentation network we designed is shown in FIGURE 9.

During training, firstly, each sequence in the MR image is registered according to the T2W sequence, and then the prostate mask obtained by the prostate segmentation module is used to remove irrelevant information outside the prostate region. In the central region of the T2W sequence, ADC sequence and DWI sequence, image blocks of size $32 \times 128 \times 128$ are respectively cropped, and then these three image blocks are respectively used as the three channels of the input image. The shape of the input data of the network model is $(32, 128, 128, 3)$. Since the self-supervised network needs to restore the transformed input data, the shape of the network output data is also $(32, 128, 128, 3)$. The parameters and FLOPs of the prostate cancer segmentation network are 13.96×10^7 and 25.84×10^{10} , respectively.

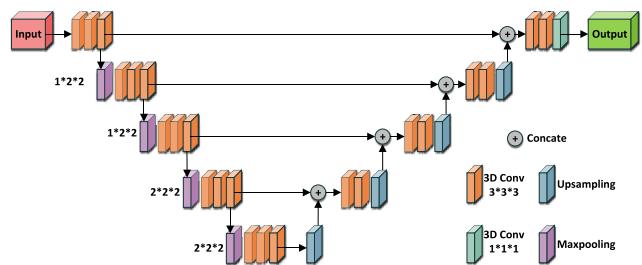


FIGURE 9. Schematic diagram of the structure of the prostate cancer segmentation network

D. PROSTATE CANCER DETECTION

The prostate cancer detection module is obtained by migration learning from the self-supervised learning network. Therefore, the structure of the prostate cancer detection network needs to have a similar structure to the self-supervised learning network, that is, the detection network needs to be converted into a segmentation network. Since the segmentation network only needs to output the mask of the prostate cancer lesion, it is only necessary to change the shape of the output layer of the network to $(32, 128, 128, 1)$ based on the self-supervised learning network.

The segmentation network based on medical images generally uses Dice loss [46] to calculate the loss during training, and its calculation formula is shown in Eq.(1).

$$D = 1 - \frac{2 \sum_i^N p_i g_i + \varepsilon}{\sum_i^N p_i^2 + \sum_i^N g_i^2 + \varepsilon} \quad (1)$$

Among them, N represents the number of pixels in the image, p_i represents the gray value of the i-th pixel in the prediction result, g_i represents the gray value of the i-th pixel in the ground truth, and ε represents the smoothing coefficient.

Dice loss does not fully consider the issue of recall and precision. Because prostate cancer lesions only occupy a small part of the entire image, it will lead to low recall. However, the clinical requirement is that the recall rate should be high, and the precision rate can be appropriately lower. Abraham et al., proposed Tversky loss [47], which is a generalized dice loss. This loss function fully considers the

relationship between false negatives (FN) and false positives (FP). The calculation formula of Tversky loss is as shown in Eq.(2).

$$TI_c = 1 - \frac{2 \sum_i^N p_{ic}g_{ic} + \varepsilon}{\sum_i^N p_{ic}g_{ic} + \alpha \sum_i^N p_{ic}\bar{g}_{ic} + \beta \sum_i^N p_{ic}\bar{g}_{ic} + \varepsilon} \quad (2)$$

p_{ic} represents the possibility that pixel i belongs to the lesion region in the prediction result; \bar{g}_{ic} represents the possibility that pixel i in the prediction result belongs to a non-lesion region; g_{ic} represents the possibility that the pixel i in the ground truth belongs to the lesion region; \bar{g}_{ic} represents the possibility that the pixel i in the ground truth belongs to the non-lesion region. The influence of FN and FP on the results can be adjusted by modifying the hyperparameters α and β .

Another problem with Dice loss is that it does not work well when segmenting small objects. This is because the inaccurate segmentation of small objects contributes less to loss. To this end, we adopted a focal Tversky loss [47]. There is a hyperparameter γ in this loss. By adjusting this hyperparameter, the contribution of simple samples and difficult samples to loss can be controlled. The calculation formula of focal Tversky loss is as shown in Eq.(3).

$$FTL_c = \sum_c (1 - TI_c)^{1/\gamma} \quad (3)$$

When $\gamma > 1$, the loss function pays more attention to the samples with poor segmentation effect. By setting a higher value of α , the network can reduce the prediction region of FN. When the parameters α and β are set to 0.5, γ is set to 1, the focal Tversky loss becomes dice loss.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

In this paper, the optimizer is Adam optimizer, the batch size is set to 2, the initial learning rate is 2e-5, and the training epoch is 60.

A. DATA

The experimental data used in this paper comes from a public dataset: ProstateX [48], [49]. ProstateX is a medical image recognition competition jointly initiated by the American Association of Physicists in Medicine (AAPM), the International Society for Optics and Photonics (SPIE) and the National Cancer Institute (NCI). The MR images in the ProstateX dataset contain a variety of sequences, such as T2 weighted (T2 W) images, proton density weighted (PD-W) images, dynamic enhancement (DCE) images, and diffusion weighted (DW) images. The label file provided by ProstateX includes the location of the lesion and the benign and malignant information. The training set of ProstateX contains 204 patients with a total of 330 lesions, and the testing set contains 142 patients with a total of 208 lesions. However, the PROSTATEx dataset only provides the location information of the lesion, and does not provide the mask information of the lesion. To meet the needs of the experiment in this paper, we invited professional radiologists to

label the prostate cancer lesion masks on the MR images of 346 patients in the ProstateX dataset. Meanwhile, in order to verify the performance of prostate segmentation model, we randomly selected 35 patients' MR images to label their prostate masks.

PROMISE12 [50] is a prostate segmentation competition organized by the International Medical Image Processing 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.

B. EVALUATION CRITERIA

The evaluation indicators used in the prostate segmentation task are Dice Similarity Coefficient (DSC), Positive Prediction Value (PPV) and Sensitivity (SEN). Their calculation formulas are as shown in Eq.(4), Eq.(5) and Eq.(6) respectively. Among them, Gt represents the ground truth, and Seg represents the predicted result.

$$DSC = \frac{2 \times V(Gt \cap Seg)}{V(Gt) + V(Seg)} \quad (4)$$

$$PPV = \frac{V(Gt \cap Seg)}{V(Seg)} \quad (5)$$

$$SEN = \frac{V(Gt \cap Seg)}{V(Gt)} \quad (6)$$

To quantitatively measure the performance of the proposed prostate cancer detection method, the true positive rate (TPR) and the number of the false positive per image (FPPI) were used as the evaluation indicators of the detection model [51].

C. PERFORMANCE OF PROSTATE SEGMENTATION

Since the ProstateX dataset does not label the prostate mask information, this paper can only train the prostate segmentation model on the PROMISE12 dataset. To verify the performance of the prostate segmentation model on the ProstateX dataset. We evaluated a testing set consisting of 35 samples with prostate masks in the ProstateX dataset. The results of the five-fold cross-validation are shown in TABLE 1. As can be seen from TABLE 1, the DSC of the trained prostate segmentation model on the ProstateX dataset reached 89.12%. This can basically meet the requirements of prostate cancer detection for prostate positioning.

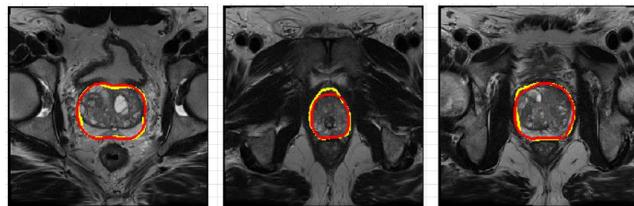
Besides, to show the effect of the prostate segmentation method more intuitively, we randomly selected three samples from 35 samples for visualization. The visualization results are shown in FIGURE 10. It can be seen from FIGURE 10 that the designed prostate segmentation network can segment the prostate better.

D. ABLATION STUDY

To prove the effectiveness of the various components in the prostate cancer detection method proposed in this paper, we adopted a five-fold cross-validation method for evaluation. The following will verify the effectiveness of image

TABLE 1. The performance of the prostate segmentation model on the ProstateX testing set.

| 5-fold cross validation | DSC | SEN | PPV |
|-------------------------|--------|--------|--------|
| 0 | 0.8951 | 0.9265 | 0.8802 |
| 1 | 0.8826 | 0.9446 | 0.8423 |
| 2 | 0.8927 | 0.9366 | 0.8679 |
| 3 | 0.8945 | 0.9281 | 0.8751 |
| 4 | 0.8909 | 0.9316 | 0.8589 |
| Average | 0.8912 | 0.9335 | 0.8649 |

**FIGURE 10.** An example image of the visualization results of prostate segmentation.

registration, 3D segmentation network and self-supervised learning.

1) THE EFFECTIVENESS OF IMAGE REGISTRATION

Because the patient may move to a certain extent during the image acquisition process, different sequences of MR images may be misaligned, which is a challenge for the automatic detection of prostate cancer. To verify the influence of image registration on prostate cancer detection methods, we use 2D-UNet as the basic network, ResNet18 as the 2D-UNet encoder, and ImageNet as the pre-training model. The following experiments are designed, and the experimental results are shown in TABLE 2. Among them, W-IR means that image registration is applied, and WO-IR means that image registration is not used.

Through the experimental results of WO-IR and W-IR in TABLE 2, it can be seen that although FPPI has increased after image registration, the detection rate of prostate cancer has been greatly improved. In particular, the TPR of the 2-fold experiment results increased by nearly 8%. On the whole, the registration of each sequence of the prostate MR image is beneficial to the detection of prostate cancer.

2) THE EFFECTIVENESS OF THE 3D SEGMENTATION NETWORK

We know that 3D MR images contain rich context information. Using a 3D segmentation network can make full use of the spatial information between the upper and lower slices, while using a 2D segmentation network will lose this spatial information. To prove the effectiveness of the 3D segmentation network, we compared the performance of the 2D segmentation network and the 3D segmentation network. In TABLE 3, comparing the experimental results of 2D-UNet and 3D-UNet in the last row, it can be found that the TPR of

TABLE 2. The effect of image registration on the detection performance of prostate cancer.

| 5-fold cross validation | WO-IR | | W-IR | |
|-------------------------|--------|--------|--------|--------|
| | TPR | FPPI | TPR | FPPI |
| 0 | 0.8458 | 0.9524 | 0.8906 | 1.2757 |
| 1 | 0.8608 | 1.5771 | 0.8757 | 1.0852 |
| 2 | 0.8160 | 0.6481 | 0.8911 | 1.4075 |
| 3 | 0.8309 | 0.8783 | 0.8903 | 1.3186 |
| 4 | 0.8487 | 1.1327 | 0.8881 | 1.1632 |
| Average | 0.8404 | 1.0377 | 0.8871 | 1.2501 |

3D-UNet is 1.5% higher than that of 2D-UNet. This proves from an experimental point of view that 3D-UNet is better than 2D-UNet for the detection of prostate cancer.

TABLE 3. Comparison of the performance of 2D/3D U-Net in prostate cancer detection. It should be noted that the 2D UNet is the same as the W-IR in TABLE 2.

| 5-fold cross validation | 2D UNet | | 3D UNet | |
|-------------------------|---------|--------|---------|--------|
| | TPR | FPPI | TPR | FPPI |
| 0 | 0.8906 | 1.2757 | 0.9011 | 1.0759 |
| 1 | 0.8757 | 1.0852 | 0.9005 | 1.0520 |
| 2 | 0.8911 | 1.4075 | 0.9002 | 1.0494 |
| 3 | 0.8903 | 1.3186 | 0.9155 | 1.0949 |
| 4 | 0.8881 | 1.1632 | 0.8932 | 0.9166 |
| Average | 0.8871 | 1.2501 | 0.9021 | 1.0378 |

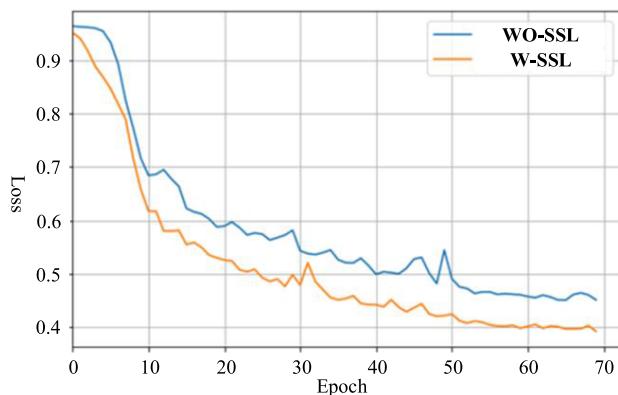
3) THE EFFECTIVENESS OF SELF-SUPERVISED LEARNING

In order to verify the effectiveness of the self-supervised learning network, experiments are carried out based on 3D UNet. The network that uses the self-supervised learning network for transfer learning is called W-SSL, and the network that does not use the self-supervised learning network for transfer learning is called WO-SSL. The loss change in the training process is an important indicator that reflects the network training situation. The training loss changes of the two networks are shown in FIGURE 11. It can be seen from the FIGURE 11 that W-SSL converges faster than WO-SSL, and the loss of W-SSL is smaller than WO-SSL at the end of the network training. This shows that the use of self-supervised learning network for transfer learning is beneficial to the training of the network.

TABLE 4 shows the performance of WO-SSL and W-SSL on the ProstateX testing set. By comparing the experimental results of WO-SSL and W-SSL in the last row of TABLE 4, it can be seen that W-SSL is better than WO-SSL in both TPR and FPPI indicators. This further illustrates the effectiveness of the designed self-supervised learning network.

E. EXPERIMENTAL COMPARISON

In order to verify the superiority of the prostate cancer detection method proposed in this paper, we compare it with the

**FIGURE 11.** The loss curve of the self-supervised learning network.**TABLE 4.** The effect of self-supervised learning network on prostate cancer detection performance. It should be noted that WO-SSL is the same as the 3D UNet in TABLE 3.

| 5-fold cross validation | WO-SSL | | W-SSL | |
|-------------------------|--------|--------|--------|--------|
| | TPR | FPPI | TPR | FPPI |
| 0 | 0.9011 | 1.0759 | 0.9155 | 0.6323 |
| 1 | 0.9005 | 1.0520 | 0.9267 | 0.6196 |
| 2 | 0.9002 | 1.0494 | 0.9233 | 0.6649 |
| 3 | 0.9155 | 1.0949 | 0.9148 | 0.5914 |
| 4 | 0.8932 | 0.9166 | 0.9106 | 0.6295 |
| Average | 0.9021 | 1.0378 | 0.9182 | 0.6275 |

TABLE 5. Comparison experiments with other prostate cancer detection methods. Among them, "Custom" means that the dataset is a non-public dataset constructed by them.

| Methods | dataset | data split | num | TPR | FPPI |
|---------------------|-----------|---------------|------|--------|------|
| Barents et al. [16] | Custom | three-fold | 200 | 0.6500 | 3.00 |
| Firjani et al. [17] | Custom | 13:0:15 | 28 | 0.7832 | 2.59 |
| Gianni et al. [21] | Custom | leave-one-out | 56 | 0.9700 | 3.00 |
| Ru et al. [33] | Custom | five-fold | 397 | 0.7510 | 1.00 |
| Le et al. [31] | Custom | 324:0:315 | 639 | 0.8985 | 1.00 |
| Wang et al. [32] | ProstateX | five-fold | 346 | 0.8978 | 1.00 |
| Cao et al. [37] | ProstateX | five-fold | 346 | 0.8790 | 1.00 |
| Yang et al. [35] | Custom | 88:0:72 | 160 | 0.9200 | 1.00 |
| Yu et al. [41] | Custom | 954:428:120 | 1502 | 0.8900 | 0.94 |
| Iqbal et al. [39] | ProstateX | five-fold | 346 | 0.9205 | 6.61 |
| Sedghi et al. [14] | ProstateX | five-fold | 346 | 0.7995 | 1.87 |
| Alkadi et al. [36] | ProstateX | five-fold | 346 | 0.8932 | 2.54 |
| Vente et al. [15] | ProstateX | five-fold | 346 | 0.9019 | 1.03 |
| Ours | ProstateX | five-fold | 346 | 0.9182 | 0.63 |

prostate cancer detection methods published in recent years. TABLE 5 shows the experimental comparison results of our method and other prostate cancer detection methods.

It is not difficult to see from TABLE 5, whether it is compared with the traditional machine learning-based method proposed in literature [16], [17], or compared with the

CNN-based method proposed in literature [31], [32], [37]. The prostate cancer detection methods proposed in this paper have shown excellent performance. It should be noted that although the TPR of the prostate cancer detection method proposed by Giannini *et al.*, and Yang *et al.*, is higher than ours, their FPPI is far worse than ours. Especially the method proposed by Giannini *et al.*, whose TPR is 5 percentage points higher than ours, but its FPPI is nearly 5 times higher than ours.

V. CONCLUSION

The method of directly detecting the prostate MR image layer by layer, although the model pre-trained on ImageNet can be used for transfer learning, but the model trained on the natural image is transferred to the medical image, which belongs to cross-domain learning. This is likely to cause the learned features to not be well adapted to the medical image model. Moreover, the prostate MR image is 3D data, which contains rich spatial context information, but the 2D network cannot use the spatial information. In addition, it is very difficult to obtain medical images, especially medical images with annotations are more difficult to obtain, which will result in a very limited sample size in the data set. If you directly train a 3D network using a dataset with a small sample size, it is easy to cause the network to overfit. To solve this problem, we designed a self-supervised learning network and used the unlabeled data in the ProstateX dataset for training to obtain a 3D pre-trained network model in the same domain as the medical image. This 3D pre-training model can make full use of these unlabeled data, allowing the network to learn the characteristic representations of the internal structure of the prostate, so that it can better adapt to the task of detecting prostate cancer. The main limitations of the proposed prostate cancer detection method are as follows: 1) The labeling information of the training data is required to be high, that is, the mask information of the prostate cancer lesion is required; 2) The self-supervised learning network used for prostate cancer segmentation is a three-dimensional network with a large amount of calculation and a long training time.

REFERENCES

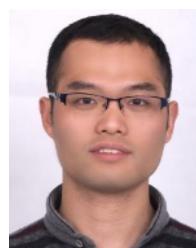
- [1] J. E. McNeal, "The zonal anatomy of the prostate," *Prostate*, vol. 2, no. 1, pp. 35–49, 1981, doi: [10.1002/pros.2990020105](https://doi.org/10.1002/pros.2990020105).
- [2] J. E. McNeal, E. A. Redwine, F. S. Freiha, and T. A. Stamey, "Zonal distribution of prostatic adenocarcinoma: Correlation with histologic pattern and direction of spread," *Amer. J. Surgical Pathol.*, vol. 12, no. 12, pp. 897–906, Dec. 1988, doi: [10.1097/00000478-198812000-00001](https://doi.org/10.1097/00000478-198812000-00001).
- [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](https://doi.org/10.3322/caac.21590).
- [4] I. W. van der Cruijsen-Koeter, A. N. Vis, M. J. Roobol, M. F. Wildhagen, H. J. de Koning, T. H. van der Kwast, and F. H. Schröder, "Comparison of screen detected and clinically diagnosed prostate cancer in the European randomized study of screening for prostate cancer, section Rotterdam," *J. Urol.*, vol. 174, no. 1, pp. 121–125, Jul. 2005, doi: [10.1097/01.ju.0000162061.40533.0f](https://doi.org/10.1097/01.ju.0000162061.40533.0f).
- [5] D. N. Costa, I. Pedrosa, F. Donato, C. G. Roehrborn, and N. M. Rofsky, "MR imaging–transrectal US fusion for targeted prostate biopsies: Implications for diagnosis and clinical management," *RadioGraphics*, vol. 35, no. 3, pp. 696–708, May 2015, doi: [10.1148/radio.2015140058](https://doi.org/10.1148/radio.2015140058).

- [6] 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](https://doi.org/10.1093/jnci/djr500).
- [7] J. C. Weinreb, J. O. Barentsz, P. L. Choyke, F. Cornud, M. A. Haider, K. J. Macura, D. Margolis, M. D. Schnall, F. Shtern, C. M. Tempany, H. C. Thoeny, and S. Verma, "PI-RADS prostate imaging—Reporting and data system: 2015, version 2," *Eur. Urol.*, vol. 69, no. 1, pp. 16–40, Jan. 2016, doi: [10.1016/j.euro.2015.08.052](https://doi.org/10.1016/j.euro.2015.08.052).
- [8] F. Cornud, N. B. Delongchamps, P. Mozer, F. Beuvon, A. Schull, N. Muradyan, and M. Peyromaire, "Value of multiparametric MRI in the work-up of prostate cancer," *Current Urol. Rep.*, vol. 13, no. 1, pp. 82–92, Feb. 2012, doi: [10.1007/s11934-011-0231-z](https://doi.org/10.1007/s11934-011-0231-z).
- [9] O. Ruprecht, P. Weisser, B. Bodelle, H. Ackermann, and T. J. Vogl, "MRI of the prostate: Interobserver agreement compared with histopathologic outcome after radical prostatectomy," *Eur. J. Radiol.*, vol. 81, no. 3, pp. 456–460, 2012, doi: [10.1016/j.ejrad.2010.12.076](https://doi.org/10.1016/j.ejrad.2010.12.076).
- [10] G. Lemaitre, R. Martí, J. Freixenet, J. C. Vilanova, P. M. Walker, and F. Meriaudeau, "Computer-aided detection and diagnosis for prostate cancer based on mono and multi-parametric MRI: A review," *Comput. Biol. Med.*, vol. 60, pp. 8–31, May 2015, doi: [10.1016/j.combiomed.2015.02.009](https://doi.org/10.1016/j.combiomed.2015.02.009).
- [11] W. Du, S. Wang, A. Oto, and Y. Peng, "Graph-based prostate extraction in T2-weighted images for prostate cancer detection," in *Proc. 12th Int. Conf. Fuzzy Syst. Knowl. Discovery (FSKD)*, Aug. 2015, pp. 1225–1229, doi: [10.1109/FSKD.2015.7382117](https://doi.org/10.1109/FSKD.2015.7382117).
- [12] N. Lay, Y. Tsehay, Y. Sumathipala, R. Cheng, S. Gaur, C. Smith, A. Barbu, L. Lu, B. Turkbey, P. L. Choyke, P. Pinto, and R. M. Summers, "A decomposable model for the detection of prostate cancer in multi-parametric MRI," in *Medical Image Computing and Computer Assisted Intervention—MICCAI*, 2018, pp. 930–939.
- [13] G. Lemaitre, R. Martí, M. Rastgoo, and F. Meriaudeau, "Computer-aided detection for prostate cancer detection based on multi-parametric magnetic resonance imaging," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 3138–3141, doi: [10.1109/EMBC.2017.8037522](https://doi.org/10.1109/EMBC.2017.8037522).
- [14] A. Sedghi, A. Mehrtash, A. Jamzad, A. Amalou, W. M. Wells, T. Kapur, J. T. Kwak, B. Turkbey, P. Choyke, P. Pinto, B. Wood, S. Xu, P. Abolmaesumi, and P. Mousavi, "Improving detection of prostate cancer foci via information fusion of MRI and temporal enhanced ultrasound," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 15, no. 7, pp. 1215–1223, Jul. 2020, doi: [10.1007/s11548-020-02172-5](https://doi.org/10.1007/s11548-020-02172-5).
- [15] C. D. Vente, P. Vos, M. Hosseiniزاده, J. Pluim, and M. Veta, "Deep learning regression for prostate cancer detection and grading in bi-parametric MRI," *IEEE Trans. Biomed. Eng.*, vol. 68, no. 2, pp. 374–383, Feb. 2021, doi: [10.1109/TBME.2020.2993528](https://doi.org/10.1109/TBME.2020.2993528).
- [16] P. Vos, J. Barentsz, N. Karssemeijer, and H. J. Huisman, "Automatic computer-aided detection of prostate cancer based on multiparametric magnetic resonance image analysis," *Phys. Med. Biol.*, vol. 57, no. 6, pp. 1527–1542, 2012, doi: [10.1088/0031-9155/57/6/1527](https://doi.org/10.1088/0031-9155/57/6/1527).
- [17] A. Firjani, F. Khalifa, A. Elnakib, G. Gimel'farb, M. A. El-Ghar, A. Elmaghriby, and A. El-Baz, "A novel image-based approach for early detection of prostate cancer," in *Proc. 19th IEEE Int. Conf. Image Process.*, Sep. 2012, pp. 2849–2852, doi: [10.1109/ICIP.2012.6467493](https://doi.org/10.1109/ICIP.2012.6467493).
- [18] I. Reda, A. Shalaby, F. Khalifa, M. Elmogy, A. Aboulfotouh, M. A. El-Ghar, E. Hosseini-Asl, N. Werghi, R. Keynton, and A. El-Baz, "Computer-aided diagnostic tool for early detection of prostate cancer," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2016, pp. 2668–2672, doi: [10.1109/ICIP.2016.7532843](https://doi.org/10.1109/ICIP.2016.7532843).
- [19] A. G. Chung, F. Khalvati, M. J. Shafiee, M. A. Haider, and A. Wong, "Prostate cancer detection via a quantitative radiomics-driven conditional random field framework," *IEEE Access*, vol. 3, pp. 2531–2541, 2015, doi: [10.1109/ACCESS.2015.2502220](https://doi.org/10.1109/ACCESS.2015.2502220).
- [20] F. Khalvati, A. Wong, and M. A. Haider, "Automated prostate cancer detection via comprehensive multi-parametric magnetic resonance imaging texture feature models," *BMC Med. Imag.*, vol. 15, no. 1, p. 27, Dec. 2015, doi: [10.1186/s12880-015-0069-9](https://doi.org/10.1186/s12880-015-0069-9).
- [21] V. Giannini, S. Mazzetti, A. Vignati, F. Russo, E. Bollito, F. Porpiglia, M. Stasi, and D. Regge, "A fully automatic computer aided diagnosis system for peripheral zone prostate cancer detection using multi-parametric magnetic resonance imaging," *Computerized Med. Imag. Graph.*, vol. 46, pp. 219–226, Dec. 2015, doi: [10.1016/j.compmedimag.2015.09.001](https://doi.org/10.1016/j.compmedimag.2015.09.001).
- [22] F. Khalvati, J. Zhang, A. Wong, and M. A. Haider, "Bag of bags: Nested multi instance classification for prostate cancer detection," in *Proc. 15th IEEE Int. Conf. Mach. Learn. Appl. (ICMLA)*, Dec. 2016, pp. 146–151, doi: [10.1109/ICMLA.2016.0032](https://doi.org/10.1109/ICMLA.2016.0032).
- [23] W. Du, Y.-P. Liu, S. Wang, Y. Peng, and A. Oto, "Features extraction of prostate with graph spectral method for prostate cancer detection," in *Proc. 17th IEEE/ACIS Int. Conf. Softw. Eng., Artif. Intell., Netw. Parallel/Distrib. Comput. (SNPD)*, May 2016, pp. 663–668, doi: [10.1109/SNPD.2016.7515975](https://doi.org/10.1109/SNPD.2016.7515975).
- [24] C. Jensen, A. S. Korsager, L. Boesen, L. R. Østergaard, and J. Carl, "Computer aided detection of prostate cancer on biparametric MRI using a quadratic discriminant model," in *Image Analysis*, 2017, pp. 161–171.
- [25] Y. Wang, D. Wang, N. Geng, Y. Wang, Y. Yin, and Y. Jin, "Stacking-based ensemble learning of decision trees for interpretable prostate cancer detection," *Appl. Soft Comput.*, vol. 77, pp. 188–204, Apr. 2019, doi: [10.1016/j.asoc.2019.01.015](https://doi.org/10.1016/j.asoc.2019.01.015).
- [26] J. C. Y. Seah, J. S. N. Tang, and A. Kitchen, "Detection of prostate cancer on multiparametric MRI," *Proc. SPIE*, vol. 10134, Mar. 2017, Art. no. 1013429, doi: [10.1117/12.2277122](https://doi.org/10.1117/12.2277122).
- [27] A. P. Kiraly, C. A. Nader, A. Tuysuzoglu, R. Grimm, B. Kiefer, N. El-Zehiry, and A. Kamen, "Deep convolutional encoder-decoders for prostate cancer detection and classification," in *Medical Image Computing and Computer Assisted Intervention—MICCAI*, 2017, pp. 489–497.
- [28] P. Lapa, I. Gonçalves, L. Rundo, and M. Castelli, "Semantic learning machine improves the CNN-based detection of prostate cancer in non-contrast-enhanced MRI," in *Proc. Genetic Evol. Comput. Conf. Companion*, Jul. 2019, pp. 1837–1845, doi: [10.1145/3319619.3326864](https://doi.org/10.1145/3319619.3326864).
- [29] N. Dhengre, S. Sinha, B. Chinni, V. Dogra, and N. Rao, "Computer aided detection of prostate cancer using multiwavelength photoacoustic data with convolutional neural network," *Biomed. Signal Process. Control*, vol. 60, Jul. 2020, Art. no. 101952, doi: [10.1016/j.bspc.2020.101952](https://doi.org/10.1016/j.bspc.2020.101952).
- [30] Z. Wang, C. Liu, X. Bai, and X. Yang, "DeepCADx: Automated prostate cancer detection and diagnosis in mp-MRI based on multimodal convolutional neural networks," in *Proc. 25th ACM Int. Conf. Multimedia*, Oct. 2017, pp. 1229–1230, doi: [10.1145/3123266.3127914](https://doi.org/10.1145/3123266.3127914).
- [31] M. H. Le, J. Chen, L. Wang, Z. Wang, W. Liu, K.-T. Cheng, and X. Yang, "Automated diagnosis of prostate cancer in multi-parametric MRI based on multimodal convolutional neural networks," *Phys. Med. Biol.*, vol. 62, no. 16, pp. 6497–6514, Jul. 2017, doi: [10.1088/1361-6560/aa7731](https://doi.org/10.1088/1361-6560/aa7731).
- [32] Z. Wang, C. Liu, D. Cheng, L. Wang, X. Yang, and K.-T. Cheng, "Automated detection of clinically significant prostate cancer in mp-MRI images based on an end-to-end deep neural network," *IEEE Trans. Med. Imag.*, vol. 37, no. 5, pp. 1127–1139, May 2018, doi: [10.1109/TMI.2017.2789181](https://doi.org/10.1109/TMI.2017.2789181).
- [33] R. Cao, X. Zhong, S. Shakeri, A. M. Bajgiran, S. A. Mirak, D. Enzmann, S. S. Raman, and K. Sung, "Prostate cancer detection and segmentation in multi-parametric MRI via CNN and conditional random field," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 1900–1904, doi: [10.1109/ISBI.2019.8759584](https://doi.org/10.1109/ISBI.2019.8759584).
- [34] R. Alkadi, A. El-Baz, F. Taher, and N. Werghi, "A 2.5D deep learning-based approach for prostate cancer detection on T2-weighted magnetic resonance imaging," in *Proc. Eur. Conf. Comput. Vis. (ECCV) Workshop*, Sep. 2019, pp. 734–739.
- [35] X. Yang, C. Liu, Z. Wang, J. Yang, H. L. Min, L. Wang, and K. T. T. Cheng, "Co-trained convolutional neural networks for automated detection of prostate cancer in multi-parametric MRI," *Med. Image Anal.*, vol. 42, pp. 212–227, Dec. 2017, doi: [10.1016/j.media.2017.08.006](https://doi.org/10.1016/j.media.2017.08.006).
- [36] R. Alkadi, F. Taher, A. El-baz, and N. Werghi, "A deep learning-based approach for the detection and localization of prostate cancer in T2 magnetic resonance images," *J. Digit. Imag.*, vol. 32, no. 5, pp. 793–807, Oct. 2019, doi: [10.1007/s10278-018-0160-1](https://doi.org/10.1007/s10278-018-0160-1).
- [37] R. Cao, A. M. Bajgiran, S. A. Mirak, S. Shakeri, X. Zhong, D. Enzmann, S. Raman, and K. Sung, "Joint prostate cancer detection and Gleason score prediction in mp-MRI via FocalNet," *IEEE Trans. Med. Imag.*, vol. 38, no. 11, pp. 2496–2506, Nov. 2019, doi: [10.1109/TMI.2019.2901928](https://doi.org/10.1109/TMI.2019.2901928).
- [38] Y. K. Tsehay, N. S. Lay, H. R. Roth, X. Wang, J. T. Kwak, B. I. Turkbey, P. A. Pinto, B. J. Wood, and R. M. Summers, "Convolutional neural network based deep-learning architecture for prostate cancer detection on multiparametric magnetic resonance images," in *Proc. SPIE*, vol. 10134, Mar. 2017, Art. no. 1013405, doi: [10.1117/12.2254423](https://doi.org/10.1117/12.2254423).
- [39] S. Iqbal, G. F. Siddiqui, A. Rehman, L. Hussain, T. Saba, U. Tariq, and A. A. Abbasi, "Prostate cancer detection using deep learning and traditional techniques," *IEEE Access*, vol. 9, pp. 27085–27100, 2021, doi: [10.1109/ACCESS.2021.3057654](https://doi.org/10.1109/ACCESS.2021.3057654).

- [40] L. Duran-Lopez, J. P. Dominguez-Morales, A. F. Conde-Martin, S. Vicente-Diaz, and A. Linares-Barranco, "PROMETEO: A CNN-based computer-aided diagnosis system for WSI prostate cancer detection," *IEEE Access*, vol. 8, pp. 128613–128628, 2020, doi: [10.1109/ACCESS.2020.3008868](https://doi.org/10.1109/ACCESS.2020.3008868).
- [41] X. Yu, B. Lou, D. Zhang, D. Winkel, N. Arrahmane, M. Diallo, T. Meng, H. von Busch, R. Grimm, B. Kiefer, and D. Comaniciu, "Deep attentive panoptic model for prostate cancer detection using biparametric MRI scans," in *Medical Image Computing and Computer Assisted Intervention—MICCAI*, 2020, pp. 594–604.
- [42] M. C. Florkow, F. Zijlstra, K. Willemsen, M. Maspero, C. A. T. Berg, L. G. W. Kerkmeijer, R. M. Castlein, H. Weinans, M. A. Viergever, M. Stralen, and P. R. Seevinck, "Deep learning-based MR-to-CT synthesis: The influence of varying gradient echo-based MR images as input channels," *Magn. Reson. Med.*, vol. 83, no. 4, pp. 1429–1441, Apr. 2020, doi: [10.1002/mrm.28008](https://doi.org/10.1002/mrm.28008).
- [43] 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, pp. 234–241.
- [44] 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](https://doi.org/10.1109/CVPR.2018.00199).
- [45] I. Loshchilov and F. Hutter, "SGDR: Stochastic gradient descent with warm restarts," in *Proc. 5th Int. Conf. Learn. Represent.*, (ICLR), Toulon, France, Apr. 2017, pp. 1–16.
- [46] 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](https://doi.org/10.1109/3DV.2016.79).
- [47] N. Abraham and N. M. Khan, "A novel focal Tversky loss function with improved attention U-Net for lesion segmentation," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 683–687, doi: [10.1109/ISBI.2019.8759329](https://doi.org/10.1109/ISBI.2019.8759329).
- [48] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle, and L. Tarbox, "The cancer imaging archive (TCIA): Maintaining and operating a public information repository," *J. Digit. Imag.*, vol. 26, no. 6, pp. 1045–1057, 2013, doi: [10.1007/s10278-013-9622-7](https://doi.org/10.1007/s10278-013-9622-7).
- [49] 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](https://doi.org/10.1109/TMI.2014.2303821).
- [50] 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](https://doi.org/10.1016/j.media.2013.12.002).
- [51] H. Cao, S. Pu, W. Tan, and J. Tong, "Breast mass detection in digital mammography based on anchor-free architecture," *Comput. Methods Programs Biomed.*, vol. 205, Jun. 2021, Art. no. 106033, doi: [10.1016/j.cmpb.2021.106033](https://doi.org/10.1016/j.cmpb.2021.106033).



YUEJING QIAN received the master's degree in computer technology from Hangzhou Dianzi University, Zhejiang. He is currently pursuing the Ph.D. degree in computer application with Huazhong University of Science and Technology. His current research interests include medical image processing and deep learning.



ZENGYOU ZHANG received the master's degree in software engineering from Huazhong University of Science and Technology, Wuhan. He is currently an Associate Professor with Zhejiang Industry and Trade Vocational College. His current research interests include data analysis and data processing.



BO WANG received the master's degree in computer technology from Hangzhou Dianzi University, Hangzhou. He is currently an Associate Professor with Zhejiang College of Security Technology. His current research interests include graph and image processing and medical big data.

• • •