

Semi-Automatic Classification of Prostate Cancer from Multi-Parametric MR Imaging Using Novel Image Sequences and 3D Convolutional Neural Networks

Li B¹, Oka R², Xuan P³, Yoshimura Y⁴ and Nakaguchi T⁵

¹Graduate School of Science and Technology, Chiba University, 1-33, Yayoicho, Inage Ward, Chiba-shi, Chiba, Chiba, 263-0024, current PhD student, Japan

²Toho University Sakura Medical Center, 564-1 Shimoshizu, Sakura, Chiba, Sakura, 285-0841, M.D, Japan

³School of Computer Science and Technology, Heilongjiang University, 74 Xuefu Road, Nangang District, Harbin, 150080, Professor, China

⁴Toyama University School of Medicine, 3190 Gofuku, Toyama, Gofuku, 930-8555, PhD, Japan

⁵Center for Frontier Medical Engineering, Chiba University, 1-33, Yayoicho, Inage Ward, Chiba-shi, Chiba, 263-0024, Professor, Japan

*Corresponding author:

Bochong Li,
Graduate School of Science and Technology, Chiba University, 1-33, Yayoicho, Inage Ward, Chiba-shi, Chiba, Chiba, 263-0024, current PhD student, Japan, Tel: 08081684298, E-mail: Bochonglok@gmail.com

Received: 28 Sep 2021

Accepted: 12 Oct 2021

Published: 18 Oct 2021

Copyright:

©2021 Li B, This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and build upon your work non-commercially.

Citation:

Li B, Semi-Automatic Classification of Prostate Cancer from Multi-Parametric MR Imaging Using Novel Image Sequences and 3D Convolutional Neural Networks. J Clin Med Img. 2021; V5(17): 1-9

Keywords:

Prostate cancer; Computer-aided detection; Magnetic resonance imaging; Machine learning

1. Abstract

The role of multi-parametric Magnetic Resonance Imaging (mp-MRI) is becoming increasingly important in the diagnosis of clinical severity of Prostate Cancer (PCa). However, mp-MRI images usually contain several unaligned 3D sequences, such as DWI image sequences and T2-weighted image sequences, and there are many images among the entirety of 3D sequence images that do not contain cancerous tissue, which affects the accuracy of large-scale prostate cancer detection. Therefore, there is a great need for a method that uses computer-aided accurate detection of mp-MRI images and minimizes the influence of useless features. Our proposed PCa detection method is divided into three stages: (i) multimodal image alignment, (ii) automatic cropping of the sequence images to the entire prostate region, and finally (iii) combining multiple modal images of each patient into novel 3D sequences, and using 3D-CNN networks to learn the newly composed 3D sequences by different modal alignments. We arrange the different modal methods to make the model fully learn the cancerous tissue features, and then predict the clinical severity of PCa and generate a 3D cancer response map for the 3D sequence images from the

last convolution layer of the network. The prediction results and 3D response map helps to understand the features that the model focuses on during the process of 3D-CNN feature learning. We applied our method to Toho hospital prostate cancer patient data, the AUC (=0.85) results are significantly higher than other methods.

2. Introduction

Prostate cancer [1] is currently one of the deadliest cancers in men, with a very high incidence and death rate each year. According to the World Health Organization, in 2020, about 1.41 million people will suffer from prostate cancer and 380,000 will die from it [2]. Early diagnosis and treatment of prostate cancer can be highly effective in preventing the development of cancerous tissue and metastasis to advanced prostate cancer, effectively improving the five-year survival rate of prostate cancer patients and reducing patients' suffering.

The diagnosis of PCa is currently made clinically by Prostate-Specific Antigen (PSA) [3] blood test and Digital Rectal Examination (DRE) [4], followed by Transrectal Ultrasound (TRUS) biopsy if the PSA test result is positive. However, due to the limited number

of biopsy samples and/or the low ultrasound resolution of TRUS [5], lesions may be missed or the Gleason Score (GS) determined from the biopsy sample may differ in repeat biopsies and sometimes from the score determined by radical prostatectomy. Moreover, prostate cancer is classified as clinically severe or clinically non-severe based on the GS, which is currently ≤ 7 for clinically non-severe prostate cancer and ≥ 8 for clinically severe prostate cancer. According to recent studies [3][4], the diagnosis of prostate cancer using PCa and biopsy has low sensitivity and specificity, which can lead to underdiagnoses and overtreatment, thus causing unnecessary suffering to patients.

Recent studies have demonstrated that multi-parametric Magnetic Resonance Imaging (mp-MRI) [6-8] can provide a simpler, non-invasive and more accurate method of detecting prostate cancer. Through combining MRI modality images, these previous studies showed that mp-MRI images have a higher detection rate and better sensitivity and specificity for prostate cancer; because of the non-invasive and highly detectable nature of MRI, more and more studies focus on the classification of prostate cancer clinical severity under multiple modalities [9]. However, it is very difficult to manually perform operations such as classification and judgment of mp-MRI because there is a large number of images for each patient, which requires a lot of time and expertise of the radiologist for judging and interpretation analysis; In addition, due to the subjectiveness of the radiologist, there will be low sensitivity and specificity in analyzing and judging the images [10], especially in the articulation of different regions of the prostate. Therefore, there is a need for a computer-assisted prostate cancer classification method that can reduce the time required to classify prostate cancer and improve the specificity and sensitivity of prostate cancer diagnosis.

In recent studies [11-17, 12, 18], methods were developed for automatic prostate cancer detection, diagnosis and classification. Currently, the prostate cancer diagnosis method consists of three main parts: first, data pre-processing, (cropping the overall prostate image to the prostate region or specific cancer site region); second, inputting the pre-processed image into a deep learning network for feature learning to obtain a feature map of the prostate; and finally, output the results of cancer grade according to the voxels in the learned feature map. The first computer-aided diagnosis system for prostate cancer, designed by Chen et al. [19], extracts pixel features from T2 images by matrix and discrete cosine transform, and then uses SVM classifier to classify the peripheral regions of the prostate. In addition, Langer et al.[20] classified the PZ part of the prostate using DCE map, and Tiwari et al. [21] designed a classification system using semi-supervised multi-modal data. However, these studies separated different regions of the prostate, resulting in cancer at the junction of different regions to be easily missed and global features of the prostate to be ignored. Many recent stud-

ies have focused on improving neural network models, but it is known that deep learning is still a near black box [22] system, and the intermediate learning process is difficult to understand. Therefore, there is the field of explainable deep learning, including CAM (Class Activation Mapping) [23] technology, which uses feature visualization to explore the working mechanism of deep convolutional neural networks and the basis of judgment. However, when implementing CAM, it is necessary to change the structure of the network itself; thus, Grad-cam has been investigated on the basis of CAM [24]. Grad-cam can be implemented without changing the structure of the network itself, and can extract the heat map of features of any layer, and a recent study investigated Grad-cam++ [25] to optimize the results of Grad-cam and make the positioning more accurate.

In this paper, we design a novel method for prostate cancer classification based on fusing image features under multiple modalities to enable the classification of clinical severity of prostate cancer with a single input rather than using a costly multiple input method with complex training. Specifically, we align the T2 and DWI images of the same patient to align the prostate region in space, crop the whole MRI image to the prostate region, fuse the aligned images with the T2 and DWI images to form a new 3D image sequence, and then input the new 3D sequence into the 3D-CNN network for feature learning. Finally, we output the features for prostate cancer severity classification and visualize the learning interest points of the network using the improved 3D-Grad-cam.

In this study, there are three main contributions:

- We developed a novel 3D-CNN input method that maintains the advantage of low training cost for single input and the advantage of multi-modal feature fusion of previous multi-input models, such that the model can fully fuse multi-modal features and facilitate network prediction with a single input.
- We improve the category activation map based on CAM by using the category activation map in a 3D image sequence to obtain a 3D-Grad-cam to facilitate our visualization of the network learning process.
- We performed an extensive experimental evaluation and comparison and used different 3D-CNN models and different sampling methods for 3DCNN models, and the AUC, sensitivity and specificity of this method on a test dataset were 0.85, 0.88; and 0.88, respectively.

The rest of the paper is structured as follows. The following section focuses on the proposed method and the dataset used for the experiments, the third section presents the experimental results and compares the baseline with the latest methods, the discussion is presented in the fourth section, and finally, the conclusions are presented.

3. Method

We predominantly used DWI and T2 image sequences from mp-MRI images in this study. Our main goal was to classify patients with prostate cancer as clinically severe and clinically less severe. Figure 1 illustrates the main framework of our proposed method, which has 3 main parts. First, we rigidly aligned [26, 27] the T2 images with the DWI images in the planar spatial domain to correct the misalignment of the prostate region due to different modal

image sequences and biases in the acquisition process, we then cropped out each T2 image with the DWI image containing the entire prostate region using an automatic method used for prostate region boundary detection, and then the cropped images were pixel normalized. Third, we used the aligned and cropped T2 and DWI images to create a new 3D image sequence of the prostate, and we fed the new 3D image sequence into the 3D-CNN network and obtained two outputs. The details of each step are presented in the following sections.

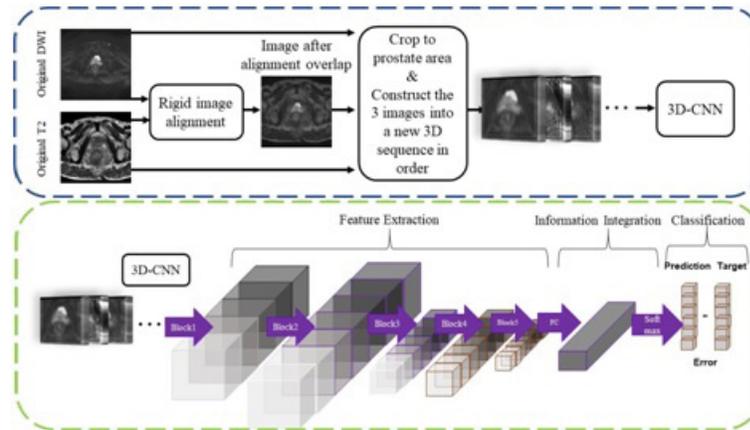


Figure 1: The framework of the proposed method consists of four key steps: (1) rigid multi-parameter (DWI, T2) image alignment, (2) prostate region cropping, and (3) building anew 3D image sequence for input into a 3D-CNN network.

3.1. Rigid Alignment of DWI and T2 Images

In previous studies [12,6,8], it has been demonstrated that in prostate mp-MRI, different MRI sequences are deterministic for prostate cancer detection and classification results, but the sensitivity of detection results under single modality images is limited, so there is a need to use multiple MRI sequences to make judgments and fully utilize the characteristics of cancerous tissues under different modalities. Among all mp-MRI sequences, T2 images are more favorable for prostate cancer detection and diagnosis based on previous studies, but the sensitivity of T2 images is low [4, 28]. DWI images show the extent of water diffusion in the prostate due to the tight accumulation of cancer cells, and any changes in the prostate cancer can be detected more easily in DWI images; thus, DWI is another recommended image use for diagnosis. However, DWI does not completely represent the prostate lesions [28, 29], so there are many studies combining DWI with T2 images to achieve better sensitivity and specificity [6,8]. In the present study, we use DWI and T2 sequences in mp-MRI for prostate cancer classification.

One of the keys to accurately combine the DWI and T2 image features is to align the DWI and T2 sequences, which can effectively eliminate the small variations between different sequences caused by external factors during mp-MRI acquisition [11]. In this study, we use the rigid 2D medical image alignment algorithm based on mutual information to maximize the mutual information between

the reference image and the target image without changing the shape information of the cancerous region, and we use DWI as the target image and T2 as the reference image. In this study, we use DWI as the target image and T2 as the reference image. We use the best available medical image alignment algorithms, ANTs SyN [30], to align the images. The image alignment strategy generally starts with an initial globally aligned linear transformation, and the linear changes available in ANTs are optimized for mean squared deviation, correlated similarity measures, each of which are optimized for translation and rotation.

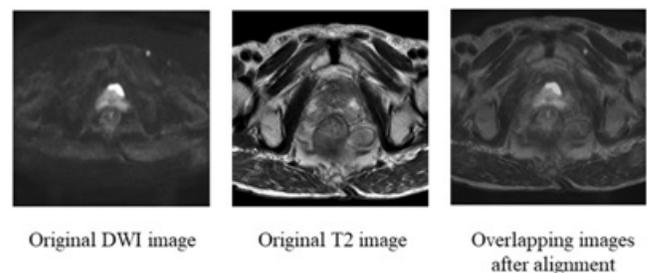


Figure 2: Examples of alignment of DWI and T2 images are shown (1) original T2 image, (2) original DWI image, and (3) aligned T2 image after overlay with DWI.

3.2. Prostate Area Cropping

After alignment, we use a basic regression CNN network to crop each image into a square region containing the prostate region,

Figure 2 shows the architecture of our CNN model for automatically cropping the prostate region. We take the original image for training, the bounding box of the prostate region is marked manually, and the model outputs three parameters, the center coordinates of the square region (x, y) , and the length l . The activation functions of all layers are tanh functions, and the corresponding loss function of our model is:

$$\text{loss} = \frac{1}{3}([\tanh o_1 - x_t] + [\tanh o_2 - y_t] + [\tanh o_3 - l_t]) \quad (1)$$

o_1, o_2 are the x and y coordinates, respectively, and o_3 is the length. In the present study, there have been more complex target detection networks, such as R-CNN [31] or automatic segmentation networks [32], but in our experiments, a simple regression CNN network has been able to achieve the detection of prostate square area more accurately, and the surrounding tissues outside the prostate square area did not have any effect on the detection of prostate cancer.

3.3. New Sequence-Based 3D-CNN Network

In the previous steps, we obtained the new aligned DWI and T2 images. We arrange the aligned and cropped prostate T2, DWI images and overlay images in order to form a new 3D image sequence. In the next experiments, we resample the new 3D image sequence of each patient 6 times and input it into the 3D-CNN model to meet the training needs of the 3D-CNN model and obtain two outputs: (1) 3D class activation map, where the values of the pixels in the map represent the importance of the model to focus on this region; and (2) high-dimensional semantic feature vectors, through which the 3D image sequence is classified. We use this novel 3D-CNN input and training method has three advantages: feature fusion, reinforcement features, and influence weight. (1) Feature fusion: with the 3D convolution kernel process and operation of spatial convolution of the image sequence, the convolution kernel will convolve the single image adjacent to the z -axis in the

sequence image, and the features of the single image adjacent to the z -axis will be calculated by the convolution kernel and extracted as high-dimensional vectors. This operation is good for fusing all the adjacent image features, and can replace the traditional method by a multiple input multi-modal image

method. We form the images of different modalities into a new 3d image sequence, so that neighboring images of each image in the sequence are images of different modalities. This method is a very cost-effective way to fuse the image features of different modalities. (2) Reinforcement features: in building a new 3d image sequence, we build the images of different modalities in different order to create a new image sequence; thus, the adjacent image modalities are often different, and the cancerous tissue features will be different under images of different modalities. The operation of 3D convolution kernel will make the model remember the features of cancerous tissue under the images of different modalities, which can enhance the learning of cancerous features. (3) Influence weight: in the learning of 3d convolution, the features of the image sequence are gradually high-dimensionalized, and the high-dimensional vector contains the features of cancerous tissue in the full. The linkage layer is expanded and the proportion of high-dimensional vectors containing cancerous tissue features to the total vectors increases, which can increase the accuracy of the prediction output. In the following, we provide detailed information of each step. In previous studies, the input to the 3D-CNN network was usually a sequence of images of a patient with a particular MRI modality, and we fused the images of different modalities into one sequence into the network, The features of the image columns are extracted by 3D-CNN, and the features of the z -axis will be observable in the z -axis direction because each image of the new sequence has the most evident cancerous tissue. In section3, we use the best available 3D-CNN models for comparison experiments.

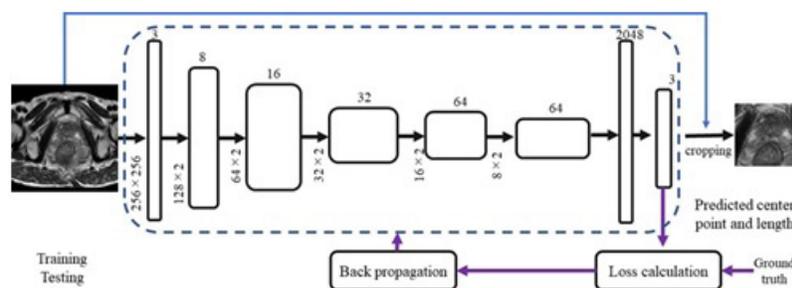


Figure 3: The prostate region detection and cropping used in this paper, in the convolutional neural network based prostate region detection, each rectangular box in the figure represents a feature map vector and shows the dimensional information of the feature map, the lower left corner of each feature map shows the length and width of the feature map, and the top shows the number of channels of the feature map, after this network three output parameters can be obtained, the center coordinates (x,y) of the detected square region containing the prostate, and the length of the square region L .

3.4. Implementations

All experiments in this study were conducted on a Windows computer using Python 3.6, with an Nvidia TITAN RTX graphics card and 24GB of RAM, on an Intel(R) Core(TM) i7-9700K 3.60GHz CPU. Pytorch [33] was used as the model backend to build the network architecture in all experiments. We use cross-entropy [34] loss function as the loss function, train 2000 epochs with batch size 2, and the model converges at 500 epochs. We use adam [35] as the model optimizer and set the learning rate to be automatically adjusted; the initial learning rate is 1e-5, the learning rate is multiplied by 0.1 every 50 epochs, and the input images are flipped at random level with a probability of 0.5 during training. The data are regularized using regularization, and all data are randomly divided into training validation and test sets with a ratio of 50:30:20, the network The network model is set to the model that preserves the best results.

4. Experiments

4.1. Setup

We collected T2 and DWI images of the prostate, which were used to train the model and evaluate the performance of the model. The prostate MRI data used in this paper consisted of 129 samples from Toho University Medical Center, Japan (dataset A), and 121 samples from the 2017 SPIE-AAPM-NCI PROSTATEx challenge dataset (dataset B), PROSTATEx Challenge [36] ("SPIE-AAPM-NCI Prostate MR Classification Challenge") was held in conjunction with the 2017 SPIE Medical Imaging Symposium and focuses on quantitative image analysis methods for diagnostic

purposes and clinically meaningful prostate cancer classification. The two datasets used, with data from different sites, were collected using different devices. We performed regularization preprocessing on these two datasets as in the previous step. The method proposed in this paper is mainly used to predict high and low risk of early prostate cancer (according to the Gleason score, a score greater than or equal to 8 is considered as greater than or equal to 8 is considered as clinically severe and a score less than or equal to 7 is considered as clinically insignificant). We use three main evaluation criteria to assess the performance of the model: AUC (area under curve) value, sensitivity and specificity, with AUC being defined as the area under the ROC curve. Sensitivity (Se), called True-Positive Fraction (TPF; or True-Positive Rate (TPR)), is the probability that a diagnostic test is correctly diagnosed as positive in a case group. Specificity (Sp), called the True-Negative Rate (TNF; or True-Negative Rate (TNR)): is the probability that the diagnostic test is correctly diagnosed as negative in the control group, False-negative rate (False-Negative Fraction, FNF; or False-Negative Rate, FNR) is the probability that the diagnostic test is negative in the case group, which will lead to delayed disease and treatment. False-Positive Fraction (FPF; or False Positive

Rate (FPR)) is the probability that a diagnostic test is incorrectly diagnosed as positive in the control group. A false positive will result in incorrect treatment and patients sometimes suffer from risky confirmatory tests.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{Sensitivity} = \frac{TN}{FP+TN} \quad (3)$$

In the above equation, TP, TN, FP, and FN; represent true positive, true negative, false positive and false negative respectively.

In the following sections, experiments are conducted to evaluate the performance of the proposed method in this paper. Table 1 shows the comparison experiments for different 3D-CNN models, Table 2 shows the comparison experiments for different newly order input sequences, and Table 3 shows the comparison experiments for different modalities of the integral image sequence, and the comparison with the recent method is shown in Table 4.

Table 1: Comparison with 3D CNN methodology

Methods	Sensitivity	Specificity	AUC	CI 95%	Paramters
C3D	0.83	0.79	0.81	0.80-0.83	78M
3DSqueezeNet	0.73	0.68	0.7	0.72-0.78	2.15M
3DMobileNet	0.74	0.67	0.69	0.73-0.75	8.22M
3DShuffleNet	0.74	0.65	0.68	0.74-0.76	6.64M
3DResnet50	0.88	0.88	0.85	0.85-0.87	44.24M
3DResnet101	0.88	0.84	0.83	0.84-0.87	83.29M
ResNext101	0.83	0.75	0.81	0.76-0.82	48.34

4.2. Comparison with the Classic 3D CNN Networks

In the first step of the experiment, we input the new 3D image sequences into different 3D-CNN models and uniformly use the pre-training weights of ucf-101[37]. Table 1 shows the results of all the 3D-CNN models when processing the new standard sequence images. All the models are from [38]. From the comparisons, we find that although the model parameters of ShuffleNet are very limited, ResNet50 achieves the best AUC value in the test set. Sensitivity, specificity and AUC values reached 0.88, 0.88 and 0.85, respectively.

4.3. Comparison of Different Input Order

In the second step of the experiment, we input the images obtained in the previous steps into the model in different orders. In the previous step of the experiment (section 3.2), the order of the single modal images in our new input image sequence is DWI, T2, and then overlap. In this section, we divide the order in the new image sequence into four different orders. Order one: DWI, T2, and then overlap as a set of resampling three times; order two: T2 resampling three times, DWI resampling three times, and then overlap resampling three times; order three: DWI resampling three times, T2 resampling three times, and then overlap resampling three

times; and order four: overlap resampling three times, T2 resampling three times, and then DWI resampling three times. In the Experiment in section 3.2 3DResNet50 achieved the best performance, we input different input sequences into the 3DResNet50 network, and in table 2, we can see in the best results are produced by order 1.

Table 2: Comparison experiments with different sequence order

Methods	Sensitivity	Specificity	AUC	CI 95%	Parameters
Order1	0.88	0.88	0.85	0.85-0.87	44.24M
Order2	0.84	0.84	0.82	0.86-0.88	–
Order3	0.84	0.84	0.81	0.82-0.84	–
Order4	0.88	0.84	0.84	0.79-0.84	–

Table 3: Comparison experiments of the integral 3D sequence.

Methods	Sensitivity	Specificity	AUC	CI 95%	Parameters
T2(384)	0.68	0.63	0.68	0.66-0.688	44.24M
T2(128)	0.74	0.71	0.72	0.71-0.74	–
DWI(384)	0.72	0.71	0.71	0.69-0.72	–
DWI(128)	0.76	0.73	0.74	0.73-0.76	–

4.4. Comparison Experiments of the Integral 3D Sequence

In our paper, we propose a new 3D-CNN sequence. In the experiment of this section, we compare this new sequence with the original integrity image sequence (Table 3). we selected the integrity unprocessed image sequence of each patient (T2 followed by DWI), and then cropped the original 512×512 image. The processed images were fed into the 3DResnet50 CNN network, and the Table 3 shows that the input of the original complete image sequence is not as good as the results of our proposed method.

Table 4: Comparison with recent methodology.

Methods	Sensitivity	Specificity	AUC	CI 95%	Parameters
Zhong et al.2018[18]	0.636	0.8	0.723	0.58-0.88	–
Aldoj et al.2019[19]	0.74	0.7	0.78	–	–
Chen et al.2017[17]	0.78	0.83	0.83	–	–
Our method	0.88	0.88	0.85	0.85-0.87	44.24M

4.5. Comparison with State-of-the-Art Methodology

We also compared our proposed method with state-of-the-art methods, including the one proposed by Aldoj et al. [39] in 2020 for prostate cancer classification using multichannel convolutional neural networks on multimodal

MRI images. Their method method takes images of three modal-

ities, ADC, DWI and T2, as input, and inputs each modality to a different channel. There are 11 layers of 3D convolution with a convolution kernel of 333,

an ensemble step of $2 \times 2 \times 2$, and two fully connected layers. Because the method chooses data of three modes in the experiment, we only choose the results of two modes from Aldoj et al. [39] as input in order to balance the comparison of experimental results. We can see that the sensitivity, specificity, and AUC values of our method are 0.14, 0.18, and 0.07 higher than the same 2-modality image inputs, respectively. A recent study by Zhong et al. [40] used deep migration learning for prostate cancer classification based on multimodal MRI images. It proposed to feed both T2 and ADC modality images into a deep migratory learning network for feature extraction and obtain prediction results after a fully connected layer. In the comparison experiments of Zhong et al. [40], they compared results using unimodal and bimodal image inputs. Here, for objectivity in the comparison experiments, we only selected the results of their comparison experiments with bimodal inputs, and find that the sensitivity, specificity, and AUC values of our model

improved by 0.144, 0.08, and 0.127, respectively. Chen et al. [41] proposed an approach to classify the clinical severity of prostate cancer using migration learning on the basis of multimodal MRI; the authors mainly used migration

learning and pre-trained weights obtained after training on Imagenet, and conducted experiments using InceptionV3. The sensitivity, specificity and AUC values of our method were 0.1, 0.05 and 0.002 higher than those of Chen et al. [41].

4.6. 3D-CNN Learning Process Visualization

There have been many previous studies [25][24] on deep learning model explanation and also on deep learning visualization, among which, the most well-known is CAM. CAM shows the basis of its decision in the form of a

heat map when a model is needed to explain the reason for its classification, as in informing where there are focal points in the map. For a deep convolutional neural network, after multiple convolution and pooling, the last convolutional layer contains the richest spatial and semantic information, and the next layers are the fully connected layer and softmax layer, which contain information that is difficult for humans to understand and display in a visual way. Therefore, in order to provide a reasonable explanation of the classification results of the convolutional neural network, it is necessary to make full use of the last convolutional layer, and CAM draws on the idea of the well-known paper on Network in Network [24], which uses GAP (Global Average Pooling) to replace the fully connected layer. The GAP can be considered as a special average pooling layer, except that its pool size is as large as the whole feature map, which is actually the average value of all pixels in each feature map. This greatly limits its use. If the model is al-

ready online or the training cost is very high, it is almost impossible to retrain it. The basic idea of Grad-CAM is the same as that of CAM, which is to obtain the weights of each pair of feature maps and then find a weighted sum. CAM replaces the fully connected layers with GAP layers and retrains the weights, whereas Grad-CAM takes a different approach and uses the global average of gradients to calculate the weights. Although Grad-CAM and other similar methods are effective, they have limitations, such as the localization of multiple similar targets at the same time, even for a single object, Grad-CAM cannot localize it completely. Based on Grad-CAM, [25] proposed Grad-CAM++, which improved the previous method, with the main contribution that ++ introduced a

pixel-level weighting of the output gradient for a specific location. This method provides a measure of the importance of each pixel in the feature map and more importantly, they derive closed-form solutions while obtaining higher-order exact representations, including softmax and exponential activation outputs. Our method requires one backpropagation, so the computational effort is consistent with the previous gradient-based method, but the results are more effective. And it can be extended in the field of 3D deep learning visualization. In this paper, we use Grad-Cam++ in 3d image sequence, in the Figure 4, we can find the model accurately focus on the cancerous tissue by focus-map and heat-map, and learn the feature details of the cancerous tissue.

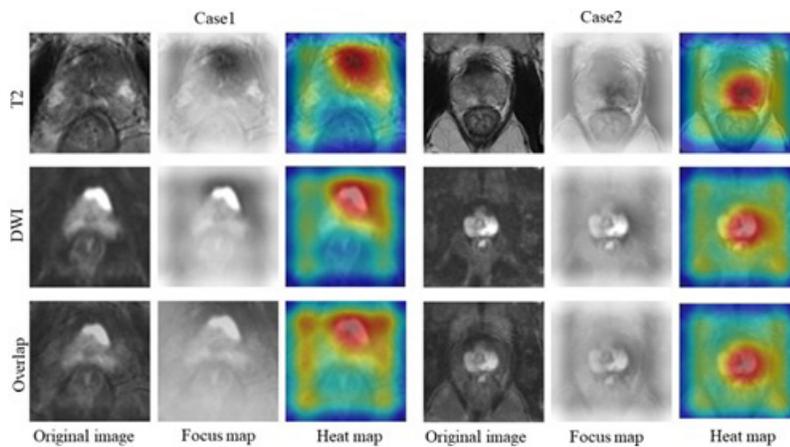


Figure 4: 3D-CNN model is visualized for learning, the figure shows 2 cases, the first column of each case is the original image, the second column is the focus map obtained by calculation, and the third column is the heat map obtained by 3D Grad-cam++ calculation

5. Discussion

There are few studies that use 3D-CNN networks to classify the clinical severity of the prostate. The main reason for that is that the cancerous tissue portion of the patient's whole prostate sequence often accounts for only a small part of the entire prostate image sequence. Due to the very small size of the cancerous tissue, although 3D-CNN can learn the features of the sequence images better than 2D-CNN, it is also difficult to learn the features of the cancerous tissue adequately with very small targets, and it is easy for the large number of useless features in the prostate cancer image sequence to affect the model learning results. So we propose the method in this paper, which solves this problem perfectly, but it is difficult to determine an optimal sequence length in constructing a new image sequence, the original image sequence length is determined by the original sequence, but the newly constructed image sequence does not have a perfect graph column length. In this paper, we have explored different alignment methods in constructing the sequence as much as possible, and in future experiments, the sequence length will be investigated in order to find an optimal sequence length.

6. Conclusion

In this paper, we propose a novel method for constructing 3D-CNN

sequences and use the newly constructed 3D image sequences to input different 3D-CNN models for comparison experiments, compare the results after different

fine-tuning based on the basic constructed method, and finally compare the results of other 3D-CNN methods. The results conclude that our proposed method has the best AUC value of 0.85 and using the improved 3d model visualization method shows the focus of model learning.

References

1. Mohler J, Bahnson RR, Boston B, Busby JE, D'Amico A, Eastham JA et al. Prostate cancer. *Journal of the National Comprehensive Cancer Network*. 2010; 8: 162-200.
2. Gillessen S, Attard G, Beer TM, Beltran H, Bjartell A, Bossi A et al. Management of patients with advanced prostate cancer: report of the advanced prostate cancer consensus conference 2019. *European urology*. 2020; 77: 508-547.
3. Weinreb JC, Barentsz JO, Choyke PL, Cornud F, Haider MA, Macura KJ et al. Pi-rads prostate imaging reporting and data system: 2015, version 2. *European urol- og*. 2016; 69: 16-40.
4. Rooij MD, Hamoen EHJ, Futter JJ, Barentsz JO, Maroeska MR. Accuracy of multiparametric mri for prostate cancer detection: a meta-analysis. *American Journal of Roentgenology*. 2014; 202: 343-351.

5. Schroder FH, Hugosson J, Roobol MJ, Tammela TLJ, Vera Nelen SC, Kwiatkowski M et al. Screening and prostate-cancer mortality in a randomized european study. *New England journal of medicine*. 2009; 360: 1320-28.
6. Fehr D, Veeraraghavan H, Wibmer A, Gondo T, Matsumoto K, Vargas HA et al. Automatic classification of prostate cancer gleason scores from multiparametric magnetic resonance images. *Proceedings of the National Academy of Sciences*. 2015; 360: 1320-28.
7. Turkbey B, Choyke PL. Multiparametric mri and prostate cancer diagnosis and risk stratification. *Current opinion in urology*. 2012; 22: 310.
8. Peng Y, Jiang Y, Yang C, Brown JB, Antic T, Sethi I 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: 787-96.
9. Turkbey B, Xu S, Kruecker J, Locklin J, Pang Y, Bernardo M et al. Documenting the location of prostate biopsies with image fusion. *BJU international*. 2011; 107: 53.
10. Valerio M, Donaldson I, Emberton M, Ehdai B, Hadaschik BA, Marks LS et al. Detection of clinically significant prostate cancer using magnetic resonance imaging {ultrasound fusion targeted biopsy: a systematic review. *European urology*. 2015; 68: 8-19.
11. Liu P, Wang S, Turkbey B, Grant K, Pinto P, Choyke P et al. A prostate cancer computer-aided diagnosis system using multimodal magnetic resonance imaging and targeted biopsy labels. In *Medical Imaging 2013: Computer-Aided Diagnosis*, volume 8670, page 86701G. International Society for Optics and Photonics, 2013.
12. Guillaume Lemaitre. Computer-aided diagnosis for prostate cancer using multi-parametric magnetic resonance imaging. PhD thesis, Universitat de Girona. Escola Polit`ecnica Superior. 2016.
13. Geert JS Litjens, Pieter C Vos, Jelle O Barentsz, Nico Karssemeijer, Henkjan J Huisman Automatic computer aided detection of abnormalities in multi-parametric prostate mri. In *Medical Imaging 2011: Computer-Aided Diagnosis*, volume 7963, page 79630T. International Society for Optics and Photonics, 2011.
14. Geert JS Litjens, Jelle O Barentsz, Nico Karssemeijer, and Henkjan J Huisman. Automated computer-aided detection of prostate cancer in mr images: from a whole-organ to a zone-based approach. In *Medical Imaging 2012: Computer-Aided Diagnosis*, volume 8315, page 83150G. International Society for Optics and Photonics. 2012.
15. Artan Y, Haider MA, Langer DL, Van der Kwast TH, Evans AJ, Yang Y et al. Prostate cancer localization with multispectral mri using cost-sensitive support vector machines and conditional random fields. *IEEE Transactions on Image Processing*. 2010; 19: 2444-2455.
16. Niaf E, Rouviere O, Mege-Lechevallier F, Bratan F, Lartizien C. Computer-aided diagnosis of prostate cancer in the peripheral zone using multiparametric mri. *Physics in Medicine & Biology*. 2012; 57: 3833.
17. Tiwari P, Kurhanewicz J, Madabhushi A. Multikernel graph embedding for detection, gleason grading of prostate cancer via mri/mrs. *Medical image analysis*. 2013; 17: 219-235.
18. 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 research international*. 2014; 2014.
19. Chan J, Wells VIII, Mulkern RV, Haker S, Zhang J, Zou KH et al. Detection of prostate cancer by integration of line-scan diffusion, t2-mapping and t2-weighted magnetic resonance imaging; a multi-channel statistical classifier. *Medical physics*. 2003; 30: 2390-8.
20. Langer DL, Kwast TH, Evans AJ, Trachtenberg J, Wilson BC, Haider MA et al. Prostate cancer detection with multi-parametric mri: Logistic regression analysis of quantitative t2, diffusion-weighted imaging, and dynamic contrastenhancedmri. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*. 2009; 30: 327-334.
21. Tiwari P, Viswanath S, Kurhanewicz J, Sridhar A, Madabhushi A. Multimodal wavelet embedding representation for data combination (maweric): integrating magnetic resonance imaging and spectroscopy for prostate cancer detection. *NMR in Biomedicine*. 2012; 25: 607-19.
22. Castelvechi D. Can we open the black box of ai? *Nature News*. 2016; 538: 20-3.
23. Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. Learning deep features for discriminative localization. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016; 2921.
24. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*. 2017; 618.
25. Chattopadhyay A, Sarkar A, Howlader P, Balasubramanian VN. Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. In *2018 IEEE winter conference on applications of computer vision (WACV)*. 2018; 839.
26. Zitova B, Flusser J. Image registration methods: a survey. *Image and vision computing*. 2003; 21: 977-1000.
27. Hill DL, Batchelor PG, Holden M, Hawkes DJ. Medical image registration. *Physics in medicine & biology*. 2001; 46: 1-45.
28. Sankineni S, Osman M, Choyke PL. Functional mri in prostate cancer detection. *BioMed research international*. 2014; 2014: 590638.
29. Gibbs P, Tozer DJ, Liney GP, Turnbull LW. Comparison of quantitative t2 mapping and diusion-weighted imaging in the normal and pathologic prostate. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*. 2001; 46: 1054-8.
30. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC, et al. A reproducible evaluation of ants similarity metric performance in brain image registration. *Neuroimage*. 2011; 54: 2033-44.
31. Girshick R, r-cnn F. In *Proceedings of the IEEE international conference on computer vision*. 2015; 1440.
32. Yu L, Yang X, Chen H, Qin J, Heng PA. Volumetric convnets with mixed residual connections for automated prostate segmentation from 3d mr images. In *Thirty-first AAAI conference on artificial intelligence*. 2017.

33. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*. 2019; 32: 8026.
34. De Boer PT, Kroese DP, Mannor S, Rubinstein RY. A tutorial on the cross-entropy method. *Annals of operations research*. 2005; 134: 19.
35. Kingma DP, Ba J. Adam: A method for stochastic optimization. *arXiv preprint arXiv*. 2014; 1412.
36. Armato SG, Huisman H, Drukker K, Hadjiiski L, Kirby JS, Petrick N, et al. Prostatex challenges for computerized classification of prostate lesions from multiparametric magnetic resonance images. *Journal of Medical Imaging*. 2018; 5: 044501.
37. Karpathy A, Toderici G, Shetty S, Leung T, Sukthankar R, Fei-Fei L, et al. Large-scale video classification with convolutional neural networks. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*. 2014; 1725.
38. Kopuklu O, Kose N, Gunduz A, Rigoll G. Resource efficient 3d convolutional neural networks. In *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*. 2019.
39. Aldoj N, Lukas S, Dewey M, Penzkofer T. Semiautomatic classification of prostate cancer on multi-parametric mr imaging using a multi-channel 3d convolutional neural network. *European radiology*. 2020; 30: 1243.
40. Zhong X, Cao R, Shakeri S, Scalzo F, Lee Y, Enzmann DR, et al. Deep transfer learning-based prostate cancer classification using 3 tesla multi-parametric mri. *Abdominal Radiology*. 2019; 44: 2030-2039.
41. Chen Q, Xu X, Hu S, Li X, Zou Q, Li Y, et al. A transfer learning approach for classification of clinical significant prostate cancers from mpMRI scans. In *Medical Imaging 2017: Computer-Aided Diagnosis*. 2017; 10134: 101344.