An artificial intelligence model based on transrectal ultrasound images of biopsy needle tract tissues to differentiate prostate cancer

Shiyu Li1,1, Xiuqin Ye2,2, Hongtian Tian3, Zhimin Ding3, Chen Cui2, Siyuan Shi2, Yang Yang2, Guoqiu Li2, Jing Chen2, Ziwei Lin2, Zhipeng Ni2, Jinfeng Xu3,∗, Fajin Dong2,†

1Department of Ultrasound, The Second Clinical Medical College of Jinan University, China
2Department of Ultrasound, The Second Clinical Medical College, Jinan University, Shenzhen People’s Hospital, Shenzhen, Guangdong 518020, China
3Corresponding author. Department of Ultrasound, The Second Clinical Medical College, Jinan University, Shenzhen People’s Hospital, No. 1017 Dongmen north Road, Shenzhen, Guangdong 518020, China. J.X.: e-mail: xujinfeng@yahoo.com; F.D.: e-mail: dongfajin@szhospital.com

1Shiyu Li and Xiuqin Ye shared the first authorship.

Abstract

Purpose: We aimed to develop an artificial intelligence (AI) model based on transrectal ultrasonography (TRUS) images of biopsy needle tract (BNT) tissues for predicting prostate cancer (PCa) and to compare the PCa diagnostic performance of the radiologist model and clinical model.

Methods: A total of 1696 2D prostate TRUS images were involved from 142 patients between July 2021 and May 2022. The ResNet50 network model was utilized to train classification models with different input methods: original image (Whole model), BNT (Needle model), and combined image [Feature Pyramid Networks (FPN) model]. The training set, validation set, and test set were randomly assigned, then randomized 5-fold cross-validation between the training set and validation set was performed. The diagnostic effectiveness of AI models and image combination was accessed by an independent testing set. Then, the optimal AI model and image combination were selected to compare the diagnostic efficacy with that of senior radiologists and the clinical model.

Results: In the test set, the area under the curve, specificity, and sensitivity of the FPN model were 0.934, 0.966, and 0.829, respectively; the diagnostic efficacy was improved compared with the Whole and Needle models, with statistically significant differences (P < 0.05), and was better than that of senior radiologists (area under the curve: 0.667). The FPN model detected more PCa compared with senior physicians (82.9% vs. 55.8%), with a 61.3% decrease in the false-positive rate and a 23.2% increase in overall accuracy (0.887 vs. 0.655).

Conclusion: The proposed FPN model can offer a new method for prostate tissue classification, improve the diagnostic performance, and may be a helpful tool to guide prostate biopsy.

Key messages

What is already known on this topic
The application of artificial intelligence in transrectal ultrasound can assist in the diagnosis of prostate cancer (PCa).

What this study adds
Differing from previous studies, our study combined magnetic resonance imaging fusion targeted biopsy and systematic biopsy, in analyzing segmented needle tract ultrasound images of prostate biopsies. Excluding the interference of inaccurate region of interest outlining, a deep learning PCa diagnostic model with better diagnostic efficacy was constructed.

How this study might affect research, practice, or policy
Through our model screening, a greater number of suspected PCa patients who need further biopsy can receive the proper diagnosis and treatment, avoiding unnecessary biopsy in patients without PCa.

Keywords: artificial intelligence; deep learning; prostate cancer; transrectal ultrasound

Introduction

Prostate cancer (PCa) is the second most common cancer among men, ranking fourth among all cancers in humans [1]. Conventional imaging examinations for PCa detection include computed tomography (CT), magnetic resonance imaging (MRI), as well as ultrasound [2].

MRI prostate scanning is the most sensitive imaging technique for the diagnosis of PCa; however, the specificity is limited. Some benign conditions may mimic the presence of prostate lesions, which can cause high Prostate Imaging Reporting and Data System (PI-RADS) scores for nodules that result in unnecessary prostate biopsies [3–5]. Transrectal ultrasonography (TRUS)-guided random systematic biopsies is a routine method for obtaining prostate pathology. With the development of fusion imaging, more significant PCas can be detected by fusion core needle biopsy, but fusion biopsy is no substitute for systematic
biopsy [6]. As for TRUS-MRI fusion targeting biopsy, it is the standard clinical procedure for diagnosing PCAs in patients with elevated prostate-specific antigen (PSA) levels or suspicious digital rectal examinations (DREs) [7]. If fusion targeting biopsy and systematic biopsy are performed simultaneously, a dilemma can be posed for the clinical diagnosis and treatment process of PCAs. It is a time-consuming, laborious, and costly procedure that requires a high level of expertise on the part of the radiologists. Meanwhile, the pain and discomfort caused by a 12-needle systematic puncture may have a negative physical and psychological impact on the patient [8].

The application of artificial intelligence (AI) to TRUS prostate images provides a possible way to improve the diagnostic performance of TRUS for PCAs [9–13]. In addition, applying ultrasound only might offer a choice for early diagnosis, which can possibly spare patients from paying high bills and also long-term appointments. A hypothesis is proposed that deep learning (DL) based on ultrasound images can achieve our intended purpose.

DL, a rapidly developing branch of AI, has shown great advantages in various medical image analysis tasks [14]. PCa nodules appear as hypoechoic masses with poorly defined borders; thus, there are difficulties in defining the verge of the region of interest (ROI), which may result in misunderstanding of malignant features. Misunderstanding fosters an obstacle to acceptably accurate detection and classification of PCa. A possible solution to this challenge is the application of segmented needle tract ultrasound images of prostate biopsies, which have their corresponding independent pathology results. Our study established a classification model through DL based on needle tract ultrasound images as well as TRUS images before the biopsies. Since we accomplished image alignment and then automatically intercepted needle tract images, this effectively eliminated the subjectivity of target area outlining. We obtained a model for PCa classification with higher accuracy and robustness.

Materials and methods

Data acquisition and preparation

This retrospective study was approved by the institutional review board with waiver of informed consent. In total, 1696 2D prostate TRUS images acquired from 142 patients from July 2021 to May 2022 were involved in this study (Table 1).

The images were classified based on the biopsy results. In total, 835 benign prostate TRUS images and 898 malignant images were integrated into this study.

Enrollment criteria for patients in the study were as follows: (i) image that is clear (prostate tissue is clearly visible and there are no artifacts in the entire image); (ii) image of the matching needle track [the images were basically consistent except for the biopsy needle tract (BNT)]; and (iii) BNT tissue with pathological confirmation (pathology results are provided for each tissue strip). The following patients were excluded: (i) image that is blurry; (ii) no corresponding needle track image; and (iii) no corresponding pathological result.

Instrumentation and image acquisition

Instrumentation

A LOGIQ E9 color Doppler US instrument (GE Healthcare, Little Chalfont, UK) and an ICS-9-D cavitary transducer were used. Patients were instructed to remain in the left decubitus position with both hands on their knees, so that the knees were bent to the chest wall as far as possible. A TSK 18-gauge automatic biopsy gun (TSK Laboratory, Oisterwijk, the Netherlands) was used for TRUS-guided prostate biopsy. TRUS-MRI fusion imaging-guided prostate targeted biopsy was performed. Two biopsy tissue specimens in the suspected lesion area were sampled. Additionally, the traditional 12-core systematic biopsy was also performed [15].

Image acquisition

After the biopsy area was determined, the corresponding tissue was obtained by the biopsy needle. With the machine’s built-in playback function, the post-biopsy needle track image with the biopsy needle track and the corresponding image from before the biopsy needle entered the prostate were obtained. Since the acquired pre- and post-biopsy ultrasound images are kept in the same section, the post-biopsy ultrasound image compares to the pre-biopsy ultrasound image with only one additional strong echo image of the biopsy needle. Each biopsy specimen was stored in a 10% formaldehyde solution and ultimately sent to the pathological department for pathological diagnosis so that each tissue could receive a separate pathological result.

Image analysis

The test set images, containing both the original images acquired before the biopsies and the images of the needle tracks during the procedure, were evaluated by three radiologists (15 years, 5 years, and <1 year of experience). Evaluation indicators included: (i) whether there are suspicious lesions (poorly defined borders, irregular morphology, abundant blood flow, etc.) or not; (ii) whether the capsule of the prostate is smooth or interrupted; and (iii) ultrasound manifestations of PCa such as the relationship between the prostate and adjacent tissues. The radiologists made diagnoses based on the aforementioned indicators. Then the diagnostic performance of the radiologists was calculated.

Statistical analysis

The statistical software adopted was R version 3.6.3 (Copyright © 2020 The R Foundation for Statistical Computing). Referring to the pathology results, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy were the parameters chosen to determine the diagnostic performance of all radiologists and AI diagnostic models. We evaluated the parameters using receiver operating characteristics (ROC) and measured area under the curve (AUC). Counting data are expressed as ratios, and chi-square tests were used to compare diagnostic results from different diagnostic methods for the same lesion. A P value of <0.05 is considered statistically significant.

Deep learning architecture

We began with a pair of pre-biopsy and post-biopsy images with the corresponding moving labels. First, all images were interpolated with bilinear interpolation to 256 × 256 pixels. Then, the Symmetric Normalization (SyN) [16] registration algorithm, the most widely used algorithm in medical image registration, was applied to register the image, following which the obtained deformation field was interpolated over the post-biopsy image in order to obtain the registered image. Similarly, the registered label was obtained by performing deformation interpolation on the moving label obtained with the deformation field (Fig. 1).

To make DL networks focus more on the needle track region, we took the needle track region as prior knowledge and used an additional segmentation branch to guide it. The entire model structure was based on the Feature Pyramid Networks (FPN) [17] network. FPN is a network architecture that employs a feature pyramid to address the challenge of multiscale variations in
Table 1. Patient characteristics.

<table>
<thead>
<tr>
<th></th>
<th>Malignant (n = 72)</th>
<th>Benign (n = 70)</th>
<th>Total (n = 142)</th>
<th>P overall</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age</strong></td>
<td>Mean (SD)</td>
<td>70.0 (7.69)</td>
<td>65.7 (6.96)</td>
<td>67.8 (7.63)</td>
</tr>
<tr>
<td></td>
<td>Median [min, max]</td>
<td>70.5 [50.0, 85.0]</td>
<td>64.0 [52.0, 83.0]</td>
<td>67.0 [50.0, 85.0]</td>
</tr>
<tr>
<td><strong>Medical history</strong></td>
<td>Elevated PSA levels</td>
<td>57 (79.2%)</td>
<td>54 (77.1%)</td>
<td>111 (78.2%)</td>
</tr>
<tr>
<td></td>
<td>Hyperplasia of prostate gland</td>
<td>9 (12.5%)</td>
<td>13 (18.6%)</td>
<td>22 (15.5%)</td>
</tr>
<tr>
<td></td>
<td>Prostate lesions</td>
<td>6 (8.3%)</td>
<td>3 (4.3%)</td>
<td>9 (6.3%)</td>
</tr>
<tr>
<td><strong>PSA</strong></td>
<td>Mean (SD)</td>
<td>185 (913)</td>
<td>16.9 (25.3)</td>
<td>102 (653)</td>
</tr>
<tr>
<td></td>
<td>Median [min, max]</td>
<td>17.1 [2.38, 7680]</td>
<td>9.30 [1.06, 145]</td>
<td>12.1 [1.06, 7680]</td>
</tr>
<tr>
<td><strong>PSA/PSA</strong></td>
<td>Mean (SD)</td>
<td>8.24 (14.3)</td>
<td>3.03 (6.95)</td>
<td>5.67 (11.6)</td>
</tr>
<tr>
<td></td>
<td>Median [min, max]</td>
<td>1.90 [0.0100, 50.0]</td>
<td>1.27 [0.190, 50.0]</td>
<td>1.67 [0.0100, 50.0]</td>
</tr>
<tr>
<td><strong>fPSA/PSA</strong></td>
<td>Mean (SD)</td>
<td>0.128 (0.0872)</td>
<td>0.153 (0.0668)</td>
<td>0.140 (0.0786)</td>
</tr>
<tr>
<td></td>
<td>Median [min, max]</td>
<td>0.106 [0.00500, 0.680]</td>
<td>0.149 [0.0470, 0.476]</td>
<td>0.126 [0.00500, 0.680]</td>
</tr>
<tr>
<td><strong>PSAD</strong></td>
<td>Mean (SD)</td>
<td>2.72 (8.98)</td>
<td>0.297 (0.446)</td>
<td>1.53 (5.50)</td>
</tr>
<tr>
<td></td>
<td>Median [min, max]</td>
<td>0.543 [0.0472, 57.6]</td>
<td>0.180 [0.0123, 3.28]</td>
<td>0.257 [0.0123, 57.6]</td>
</tr>
<tr>
<td><strong>Volume</strong></td>
<td>Mean (SD)</td>
<td>44.5 (20.8)</td>
<td>64.1 (40.3)</td>
<td>54.2 (33.3)</td>
</tr>
<tr>
<td></td>
<td>Median [min, max]</td>
<td>43.4 [15.7, 151]</td>
<td>57.0 [10.2, 219]</td>
<td>48.9 [10.2, 219]</td>
</tr>
<tr>
<td><strong>PI-RADS</strong></td>
<td>2</td>
<td>2 (2.8%)</td>
<td>13 (18.6%)</td>
<td>15 (10.6%)</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>8 (11.1%)</td>
<td>39 (55.7%)</td>
<td>47 (33.1%)</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>21 (29.2%)</td>
<td>12 (17.1%)</td>
<td>33 (23.2%)</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>40 (55.6%)</td>
<td>4 (5.7%)</td>
<td>44 (31.0%)</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
<td>1 (1.4%)</td>
<td>2 (2.9%)</td>
<td>3 (2.1%)</td>
</tr>
</tbody>
</table>

*N/A, not available.*

object detection while adding minimal computational overhead. The input image was down-sampled five times by using the encoder block, and the obtained feature layers were marked as C1, C2, C3, C4, and C5; then, a classifier head was added for benign and malignant classification. The classifier head consisted of global average pooling, dropout, and fully connected layers. In the additional segmentation branch, for the obtained C5 feature layer, nearest neighbor interpolation was used for continuous up-sampling, while in order to fuse the C2, C3, and C4 feature layers with richer spatial information, they were added to the feature map of the corresponding size to obtain the fusion feature layer P2, P3, P4, P5, after changing the channel with a 1×1 convolution. For the obtained fusion feature layer, first we used up-sampling to interpolate P3, P4, and P5 by two times, four times, and eight times, respectively, to transform their sizes to be the same. Then we used a segment block, which was composed of convolution, Group Norm (GN), and Rectified linear unit (ReLU), for further fusion. After the features were fused, element-wise addition was performed to make a merge block. Finally, the feature map was interpolated back to the original image size by using 4× up-sampling, and a 1×1 single-channel convolution was used as the segment head for segmentation (Fig. 2).

**Experiment**

Our experiments used ResNet50 as the backbone structure. For the classification branch, binary cross entropy (BCE) was used as the loss function, and for the additional segmentation branch, Dice was used as the loss function. The two were weighted with a weighting factor of 0.7 and 0.3, respectively. Fivefold cross-validation was performed on the entire dataset by using the ADAM optimizer [18], 120 epoch iterations, and the learning rate was set to 1e−4. To reduce overfitting, various data augmentations were performed on the input image, including random horizontal flipping, scaling, rotation, translation, and shearing. The epoch with the highest accuracy on the validation set during training was saved and tested on the test set.

In order to verify the classification validity of the needle track region, the needle track region was separately input into the network for classification. To alleviate the inaccurate selection of the needle track region caused by the registration, we carried out special enhancements which were different from those to the needle track region aforementioned. The obtained registered needle track label was randomly rotated, translated, and widened (Fig. 3). Then it was used to take the BNT out from the original image as the input of the DL network (Needle model). At the same time, we used the original image without the needle tract (Whole model) and the original image combined with BNT (FPN model) as the input of the network, respectively, for training.

**Clinical model for comparison**

To verify the effectiveness of DL-assisted diagnosis of PCa, we constructed clinical models for comparison. Some clinical characteristics including the age, PSA, free PSA (fPSA), the ratio of PSA to fPSA, PSA density (PSAD), volume, and the score of PI-RADS were applied to construct clinical models. Lasso regression was performed to select the independent factors.

**Results**

**Baseline characteristics and assessment criteria**

A total of 1696 pairs of prostate TRUS images before and after biopsy (including malignant patients: n = 72; images: n = 898) from...
142 patients were included in this study. The training set, validation set, and test set were randomly assigned, with randomized 5-fold cross-validation being performed between the training set and validation set (including 666 malignant images, 49.3%), and the test set included 345 images (including 199 malignant images, 57.68%) (Fig. 4). Table 1 shows detailed patient characteristics.

**Performance of the artificial intelligence system**

In the test set, the optimal diagnostic efficacy was achieved when the whole model thresholds were 0.351, and the AUC (95% CI), specificity, sensitivity, accuracy, NPV, and PPV for the diagnosis of PCa were 0.908 (0.878–0.939), 0.938, 0.832, 0.737, and 0.943, respectively. The optimal diagnostic efficacy was achieved when the Needle model thresholds were 0.361, and the AUC (95% CI), specificity, sensitivity, accuracy, NPV, and PPV for the diagnosis of PCa were 0.905 (0.875–0.936), 0.918, 0.764, 0.829, 0.740, and 0.927, respectively. When the original images combined with BNT were input as the FPN model, optimal diagnostic efficacy was achieved when the thresholds were 0.464, and the AUC (95% CI), specificity, sensitivity, accuracy, NPV, and PPV for diagnosing PCa were 0.934 (0.908–0.960), 0.966, 0.829, 0.887, 0.806, and 0.971, respectively; the diagnostic efficacy was improved compared with the Whole and Needle models, with statistically significant differences (P < 0.05) (Table 2, Fig. 5).

**Performance of the clinical model**

Lasso regression was performed to select the independent factors. Age, volume, PI-RADS score, and PSA were selected as the independent indicators for the clinical model (Fig. 6).

The AUC (95% CI), specificity, sensitivity, accuracy, NPV, and PPV of the diagnosis of PCa achieved 0.911 (0.8628–0.9644), 0.882, 0.901, 0.892, 0.896, and 0.889, respectively.

**Performance of the radiologists**

In the test set, the AUC (95% CI), specificity, sensitivity, accuracy, NPV, and PPV for the diagnosis of PCa by the senior radiologist were 0.673 (0.625–0.721), 0.788, 0.558, 0.655, 0.567, and 0.781, respectively. The AUC (95% CI), specificity, sensitivity, accuracy, NPV, and PPV for the diagnosis of PCa of the junior radiologist were 0.539 (0.496–0.583), 0.822, 0.256, 0.496, 0.448 and 0.662, respectively. The diagnostic efficacy of the senior radiologist was significantly higher than that of the intermediate and junior radiologists, with a statistically significant difference (P < 0.05), while the diagnostic efficacy of the intermediate experienced radiologist was slightly higher than that of the junior radiologist in the test set. However, there was no statistically significant difference in this latter case (P > 0.05) (Table 2, Fig. 5).

**Comparison of diagnostic efficacy of the radiologists, clinical model, and artificial intelligence system**

In the test set, AI diagnostic efficacy was significantly higher than that of all radiologists of different seniority (P < 0.05), and the FPN model detected more PCa compared with senior radiologists [82.9% (165/199) vs. 55.8% (111/199)], with a 61.3% decrease in the false-positive rate and a 23.2% increase in overall accuracy (0.887 vs. 0.655).
Figure 2  Schematic illustration of the proposed classification network with auxiliary mask guide in this work.

Figure 3  Data enhancement methods for an individual needle track region.
### Table 2. Results of the AI system and radiologists in the testing set.

<table>
<thead>
<tr>
<th>Model</th>
<th>AUC</th>
<th>Specificity</th>
<th>Sensitivity</th>
<th>Accuracy</th>
<th>NPV</th>
<th>PPV</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Whole</td>
<td>0.905</td>
<td>0.938</td>
<td>0.754</td>
<td>0.832</td>
<td>0.737</td>
<td>0.943</td>
<td>0.011</td>
</tr>
<tr>
<td>Needle</td>
<td>0.905</td>
<td>0.918</td>
<td>0.764</td>
<td>0.829</td>
<td>0.740</td>
<td>0.927</td>
<td>0.029</td>
</tr>
<tr>
<td>FPN</td>
<td>0.934</td>
<td>0.966</td>
<td>0.829</td>
<td>0.887</td>
<td>0.806</td>
<td>0.971</td>
<td>*</td>
</tr>
<tr>
<td>SR</td>
<td>0.673</td>
<td>0.788</td>
<td>0.558</td>
<td>0.655</td>
<td>0.567</td>
<td>0.781</td>
<td>*</td>
</tr>
<tr>
<td>IR</td>
<td>0.539</td>
<td>0.822</td>
<td>0.256</td>
<td>0.496</td>
<td>0.448</td>
<td>0.662</td>
<td>*</td>
</tr>
<tr>
<td>JR</td>
<td>0.505</td>
<td>0.452</td>
<td>0.558</td>
<td>0.513</td>
<td>0.429</td>
<td>0.581</td>
<td>*</td>
</tr>
<tr>
<td>Clinical</td>
<td>0.911</td>
<td>0.882</td>
<td>0.901</td>
<td>0.892</td>
<td>0.896</td>
<td>0.882</td>
<td>*</td>
</tr>
</tbody>
</table>

*Whole model: a model constructed with the original image as input. Needle model: a model constructed with the BNT image as input. FPN model: a model constructed with the original image combined with the BNT image as input. SR, senior radiologist. IR, intermediate radiologists. JR, junior radiologists.

#### Discussion

TRUS has become a significant imaging tool for early screening because of its economic and convenience advantages. However, the accuracy and specificity are relatively low [19, 20], therefore a great number of scholars have tried to enhance the diagnostic efficiency of PCa of TRUS by applying computer-aided diagnosis (CAD). The lack of typical features in PCa ultrasound images fosters an obstacle for lesion labeling, leading to the stagnation of research on prostate AI. A possible solution to this challenge is the application of the BNT images. Our study established an effective classification FPN model based on segmented needle tract images as well as the images before biopsy for PCa diagnosis. Its diagnostic capability is superior to the BNT model, Whole model, and radiologists.

In our experiment, BNT images were obtained by registering prostate images before and after biopsy. They were arranged as input to participate in model training for the first time. Based on the original image and individual BNTs, the network model was able to distinguish the benignity and malignancy of prostate ultrasound images to the following degree: AUC, specificity, sensitivity, accuracy: 0.908, 0.938, 0.754, 0.832 and 0.905, 0.918, 0.764, 0.829, respectively. Results indicate that although the location of the lesion in the image was not marked, the model was able to effectively differentiate benign and malignant prostate TRUS images...
AI model to differentiate prostate cancer

Figure 5 Comparison of diagnostic efficacy between the AI system and radiologists in the testing set.

Figure 6 Lasso regression of the clinical model.

through autonomous analysis, learning, and generalization based on the given labels. Furthermore, we also trained the FPN model by combining the BNT and the original image as input. Compared with the model using only a single image, the AUC, specificity, sensitivity, accuracy, and PPV have been greatly improved, reaching 0.934 (0.908–0.960), 0.966, 0.829, 0.887, 0.806, and 0.971, respectively.

As for the clinical model, before performing a prostate biopsy, doctors evaluate patients based on their clinical presentation, including the level of PSA, the results of a positive rectal
examination, and the PI-RADS score. These indicators are usually important factors in determining the risk of PCa. Therefore, these patients are more likely to be diagnosed with malignant tumors when undergoing a prostate biopsy. Since the clinical model can predict whether a patient has cancer based on these indicators, the clinical model we constructed showed a large selection bias which leads to a higher AUC.

Compared with previous studies, Gómez-Ferrer and Arlandis [21] extracted 3 still images from the biopsy area (288 patients, 4725 images). The texture features of ultrasonic images were obtained by "simple mapping" on gray and spatial gray correlation matrices. Finally, two methods based on nearest neighbor and a Markov hidden model were used for classification. The nearest neighbor of the ROC curve was 59.7%, and the classification of the Markov hidden model was 61.6%, the ROC curve area of cooccurrence matrix was 60.1% for nearest neighbor, and for the Markov hidden model it was 60.0% [21]. Huang et al. [22], by manually labeling the prostate TRUS images of the needle track before biopsy (48 patients, 342 images), extracted the texture features from the images after pretreatment and performed linear combination. Finally, the fused texture features were provided to a support vector machine (SVM) classifier for classification. The accuracy, sensitivity, and specificity for the diagnosis of PCa by this method were 70.93%, 70.00%, and 71.74%, respectively [22].

In this study, we combined MRI fusion targeted biopsy and systematic biopsy, extracting and segmenting the prostate tissue ultrasound image before biopsy and combining it with the original pre-biopsy ultrasound image. The FPMM model learning process becomes more precise due to the correspondence of each needle pathway image with its corresponding known pathological result. Additionally, we have included the entire pre-biopsy image into the model, which allows the learning of image details related to the surrounding tissue of the needle pathway. These foster significant advances in the accuracy of diagnosing PCa. It is similar to the clinical diagnosis of PCa by MRI. [Literature reports range from 73% to 96% for sensitivities and 0.50% to 93% for specificities [23–32].] As a result of the high PPV rates, the model is capable of detecting PCa patients among men undergoing PCa screening.

However, there are some limitations in our study. The sensitivity of the model in identifying malignant prostate tissues in the images was not adequate. Some patients with malignant tumors could still be missed in terms of 2D grayscale images only. Therefore, the model still needs to be improved in terms of sensitivity, which may be related to the lack of specificity and sensitivity of 2D grayscale ultrasound imaging for PCa lesions.

**Conclusion**

Through our model screening, a greater number of suspected PCa patients who need further biopsy can receive the proper diagnosis and treatment, avoiding unnecessary biopsy in patients without PCa.

**Acknowledgements**

Thanks for the support and help of our D team of the Ultrasound Department, Shenzhen People’s Hospital.

**Funding**

This project was supported by Commission of Science and Technology of Shenzhen, grant nos. GJHZ20200731095401004 and GJHZ20210705142204014.

**Conflict of interest statement:** We declare that we have no financial and personal relationships with other people or organizations that can inappropriate influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position presented in, or the review of, the manuscript entitled.

**Data availability**

The data are de-identified ultrasound image data and clinical profile data of the participants, and the data can be accessed for details via: xujinfeng@yahoo.com. Please ensure that appropriate permissions are obtained prior to accessing the data. The period of availability of the data is not specified. Reuse of the data is not permitted and no additional information is available.

**Author contributions**


**References**

29. Han C, Liu S, Qin XB. et al. MRI combined with PSA density in detecting clinically significant prostate cancer in patients with PSA serum levels of 4–10ng/mL: biparametric versus multiparametric MRI. *Diagn Interv Imaging* 2020; **101**: 235–44.