A Deep Learning Approach for Targeted Contrast-Enhanced Ultrasound Based Prostate Cancer Detection

Yujie Feng, Fan Yang, Xichuan Zhou *Member, IEEE,* Yanli Guo, Fang Tang, Fengbo Ren, Jishun Guo, Shuiwang Ji *Member, IEEE*

Abstract—The important role of angiogenesis in cancer development has driven many researchers to investigate the prospects of noninvasive cancer diagnosis based on the technology of contrast-enhanced ultrasound (CEUS) imaging. This paper presents a deep learning framework to detect prostate cancer in the sequential CEUS images. The proposed method uniformly extracts features from both the spatial and the temporal dimensions by performing three-dimensional convolution operations, which captures the dynamic information of the perfusion process encoded in multiple adjacent frames for prostate cancer detection. The deep learning models were trained and validated against expert delineations over the CEUS images recorded using two types of contrast agents, i.e. the anti-PSMA based agent targeted to prostate cancer cells and the non-targeted blank agent. Experiments showed that the deep learning method achieved over 91% specificity and 90% average accuracy over the targeted CEUS images for prostate cancer detection, which was superior ($p < 0.05$) than previously reported approaches and implementations.

Index Terms—Contrast-enhanced ultrasound, prostate cancer detection, convolutional neural network, targeted agent

I. Introduction

PROSTATE cancer is one of the most common forms of cancer which accounts for 26% of all cancer diagnoses for American men [1]. Till now, prostate cancer is still diagnosed using systematic biopsies which consist of taking a dozen of specimens from the prostate using a core needle. Systematic biopsy is invasive and has low sensitivity; moreover, it carries a risk of infection, sepsis and bleeding [2], [3]. A noninvasive imaging technology to detect prostate caner at an early stage could improve both prostate cancer diagnosis and treatment. As a novel ultrasound technology, the CEUS can provide useful modality for visualizing the dynamic patterns of the blood flow, allowing medical experts to detect angiogenesis for cancer diagnosis [4].

So far the researches of ultrasound based cancer detection were mostly based on the traditional B-mode ultrasound and the color Doppler which could provide information on blood flow in the cancer region [5], [6]. However, while there is no doubt of its usefulness, the color Doppler is less suitable for cancer detection than the emerging CEUS technology, because the border of the color-coded velocity image strongly depends on user-controlled settings such as the Doppler gain and velocity range [7], which requires personal professional expertise.

Due to its real-time availability at the bedside, ultrasound offers a practical and cost-effective method for prostate imaging. The CEUS is especially suitful for visualizing tumor angiogenesis because of its ability of highlighting the blood flow in microvasculature. Technically, the sequence of the CEUS images records both the spatial and the temporal dynamic patterns of the micro-bubble agent carried by blood flow, which has been proven crucial for diagnosing cancer [8]. Fig. 1 shows the sequence of the CEUS images captured at the first a few seconds after injecting the contrast agent in the blood vessel. The region surrounded by red dots indicates the cluster of cancer tissues. The label information of cancer region is provided by clinical experts according to domain knowledge and pathological test results. Coherent with clinical findings [4], a distinct temporal pattern of climbing intensity can be seen in the cancer region in Fig1. The different blood perfusion patterns of cancer and normal images are the intuitive and clinical ground for machine learning based cancer detection.

So far, many researches of the CEUS based prostate cancer detection were achieved by quantifying different parameters of the measured time intensity curve (TIC). Several methods were proposed for prostate cancer detection by the assessment of tissue perfusion [9]–[12]; More recently, contrast ultrasound dispersion imaging was proposed as an alternative method for prostate cancer detection [13]–[15]. Kuenen estimated the dispersion by fitting each TIC with a convective-dispersion model in the time domain [13]. Later, indirect dispersion estimation approaches were proposed, and the similarity between neighboring TICs were computed using different linear measures such as spectral coherence and temporal correlation [14], [15]. Gasnier and Schalk proposed respective approaches for quantifying tumor vascularity in 3D-CEUS data [16], [17]. Other parameters such as the dispersion coefficient, velocity and mutual information were also estimated based on the CEUS data to detect prostate cancer [18]–[20]. Clinical research of the parametric approaches for prostate cancer diagnosis has also been performed [21]. However, until now,
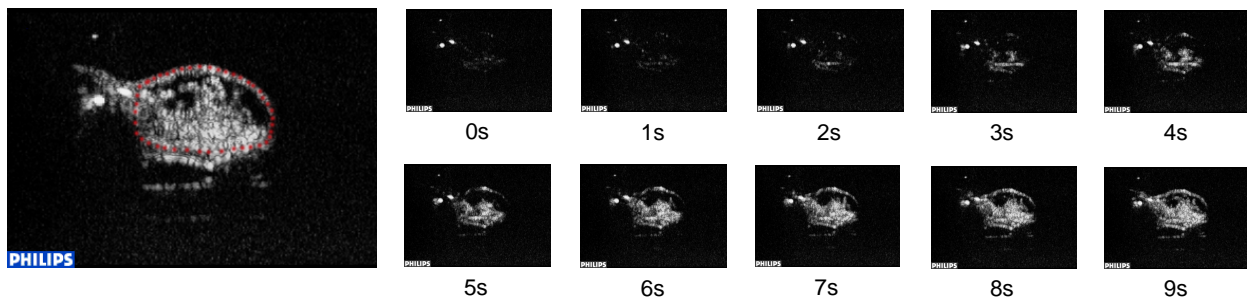
Fig. 1. Example of the CEUS video frames examined for cancer detection. The region of interest surrounded using red dots indicates the prostate cancer.

the parameter based CEUS quantification has not shown reliable cancer detection results in the prostate [22]. It is still a challenge to automatically detect and segment cancer via CEUS imaging [8], partly because the dynamic of the time-intensity-curve varies with respect to different subjects, and different locations, types and shapes of the cancer tissues.

Different from the parameter based approaches, the recent development of the machine learning approaches has provided another tool set for analyzing ultrasound images for prostate cancer detection. Several learning based approaches including the nearest neighbor classifier and the neural networks were applied to detect tumor in static CEUS images [23], [24]. The latest research attempted to apply the recurrent neural network (RNN) for ultrasound based tumor segmentation [27]; however, due to the computational complexity of its fully-connected recurrent layer, the classic recurrent neural network could only process static ultrasound images. On the other hand, given the dynamic characteristic of the CEUS data, it is preferable to extract the spatial-and-temporal features uniformly from the CEUS video for cancer detection.

To address this challenge, this paper presents a deep learning framework, which is based on the three-dimensional convolutional neural network (3D-CNN), to extract the spatial-temporal features uniformly from the sequential CEUS images for cancer detection. The convolutional neural networks are a type of deep learning models in which trainable filters and local pooling operations are applied alternatingly on the input images [28], resulting in a hierarchy of two-dimensional feature matrices [31]. To capture the temporal patterns in the CEUS image sequences, a third dimension is incorporated in the CNN model [30]. Since the 3D-convolutional feature extractors operate in both the spatial and the temporal dimensions, the 3D-CNN is able to capture the dynamic CEUS patterns reflecting angiogenesis. To detect and segment the cancer tissues in the ultrasound video images, we split the video frames into small image tensors (Fig. 2), and each small image tensor is treated as a sample. After training the proposed models using samples extracted from nearly forty gigabytes of CEUS video clips, the 3D-CNN could accurately detect the prostate cancer tissues in the video images of unknown subjects. Due to its ability to extract features from three-dimensional image voxels, the 3D-CNN has being recently

applied for analyzing the MR and CT images [25], [26], but few attempts have been tried to use the 3D-CNN for CEUS based cancer detection.

Recent research shows that, by taking advantage of agent material targeted to prostate cancer cells, the CEUS may potentially provide higher sensitivity and accuracy for diagnosing prostate cancer [32]. As a preclinical study, our research evaluated the CEUS data acquired with the anti-PSMA (prostate specific membrane antigen) based contrast agent which was targeted to the prostate cancer cells. Experimental results showed that the 3D-CNN method outperformed other learning based and parameter based approaches, and achieved over 90% average accuracy for prostate cancer detection. Moreover, since the targeted agent led to 26% higher increase of intensity in the region of cancer tissues than the blank non-targeted agent, the examined learning based approaches generally showed $6.70\% - 18.77\%$ higher accuracy using the targeted contrast agent in our experiments.

The rest of this paper presents the deep learning framework we built for prostate cancer detection, which is organized as follows. We describe the material and data collected in Section 2. The spatial-temporal convolution operation and the 3D-CNN architecture are presented in Section 3. The experimental results are reported in Section 4. We conclude in Section 5 with discussions.

II. MATERIAL AND DATA

Different from standard B-mode ultrasound or color Doppler imaging, the CEUS uses gas-filled bubbles as contrast agents to enhance the reflection of the ultrasound waves. The bubbles are injected intravenously into the mouses systemic circulation at second 0, and the ultrasound waves are directed on the region of interest (ROI). When the contrast agent in the blood flow passes the imaging window, as shown in Fig. 1, the bubbles reflect a unique echo that stands in sharp contrast to the surrounding tissue due to the orders of magnitude mismatch between bubble and tissue echogenicity. The ultrasound system converts the strong echogenicity into contrast-enhanced sequential images to show the blood flow patterns in cancer and normal tissues.

In our research, two types of contrast agents were constructed to facilitate prostate cancer imaging, i.e. the targeted agent
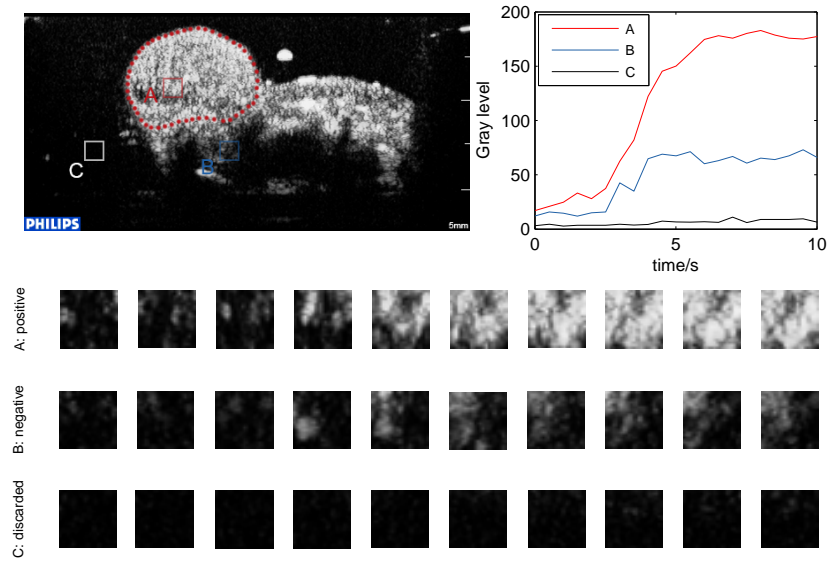
Fig. 2. Three examples of the examined CEUS tensor image samples. The right diagram shows the average time intensity curve associated with each sample.

TABLE I
THE CEUS SAMPLE SETS CONSTRUCTED FOR TRAINING,
VALIDATING AND TESTING THE DEEP LEARNING MODELS

Agent type	Train		Validation	Test	Total
	Negative	Positive			
Targeted	2479	2479	9372	7511	21841
Non-targeted	2769	2769	9715	10485	25738

made of anti-PSMA (prostate specific membrane antigen) and the blank contrast agent. Detailed information about the bubble agents constructed can be found in the experimental research [32]. Both types of bubble agents were of the size 487.60 ± 33.55 nm. To evaluate the proposed method, we performed CEUS on twenty different *anaesthetized* animal xenografts (lab mice implanted with human prostate cancer). For comparison, two respective CEUS tests with targeted and non-targeted agents were performed on each subject with thirty minutes' interval.

The CEUS images examined were captured by the iU22 ultrasound system (Philips, Amsterdam, The Netherlands), and the raw data were 505x246 AVI videos. Fig. 2 shows an example of the CEUS image for prostate cancer detection. The region surrounded by the red dots indicates the prostate cancer implanted under the xenograft's skin. The blood perfusion process captured in each test lasts about 10 seconds starting from the moment of injecting the contrast agent. Since the blood perfusion process was relatively slow, we extracted one frame per second to avoid information redundancy in adjacent frames. For cancer detection, each frame was split into a group of 23-pixel-by-23-pixel windows with a step size of five pixels. For each subject, ten sequential frames were extracted and reconstructed as a set of 23*23*10 image tensors. The three-dimensional tensor samples were later used for training and testing the proposed model. Each tensor sample was labeled

as a negative or positive cancer sample based on professional expertise and histopathological analysis results.

One consideration of CEUS based cancer detection is unbalanced data. Specifically, in the raw data set, the amount of negative samples (78277 samples) were significantly larger than the amount of positive samples (9073 samples). The unbalanced data might degrade the sensitivity and accuracy for cancer detection; therefore, the raw negative samples in the area with no perfusion were removed. Technically, 39721 negative samples with less than 3.5% intensity increase were discarded, leaving 21841 targeted samples and 25738 non-targeted samples. Respective models were built using the targeted and non-targeted CEUS samples for comparison. The training set contained equal number of positive and negative samples for balanced training. Detailed information of the sample sets constructed and examined are shown in Table I.

Fig. 2 shows three examples of the tensor samples constructed, including one positive cancer sample A, one negative sample B and one discarded sample C. As shown in Fig. 2, different temporal patterns of increased intensity can be found in the compared CEUS tensor samples. Coherent with the clinical observation that the prostate cancer tissues usually have richer blood flow [4], the positive sample A has higher intensity increase compared to the negative sample B. The different temporal patterns of blood flow between the normal and cancer tissues are the intuitive and medical basis of the learning based methods for automated cancer detection.

III. METHOD

The proposed method attempts to detect prostate cancer in the ultrasound sequential images automatically. As described in the previous section, the frames of the CEUS videos were split into a set of three-dimensional tensor samples. Each sample was labeled as a positive or negative cancer sample. A 3D-
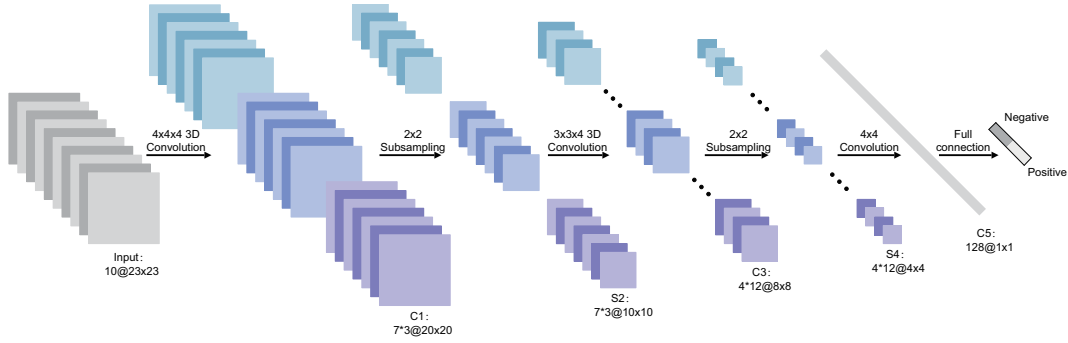
Fig. 3. The three dimensional convolutional neural network built for extracting and classifying the CEUS tensor samples. The framework consists of three convolutional layers, two sub-sampling pooling layers and one fully-connected classification layer.

convolutional neural network was built to uniformly extract the spatial and temporal features, and classify the tensor samples as cancer or normal. Fig. 3 shows the convolutional neural network we built, which consists of three types of layers, i.e. three convolutional layers, two sub-sampling pooling layers and one fully-connected classification layer.

In a standard two-dimensional convolutional neural network, the convolutional layers are two-dimensional image filters which may magnify and extract the discriminant features of the input images. Formally, the value of a neuron v_{ij}^{xy} at position (x, y) of the jth feature map in the ith layer is denoted as follows:

$$v_{ij}^{xy} = g(\sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} w_{ijm}^{pq} v_{(i-1)m}^{(x+p)(y+q)} + b_{ij}) \qquad (1)$$

where m indexes the feature maps in the $(i-1)$th layer connected to the current jth feature map, w_{ijm}^{pq} is the weight of position (p, q) connected to the mth feature map, P_i and Q_i are the height and the width of the two-dimensional convolution kernels, and b_{ij} is the bias of the jth feature map in the ith layer. The function $g(x) = \frac{1}{1+\exp(-x)}$ is the standard sigmoid activation function adopted in neural network models.

Since the CEUS samples are sequential images, it is natural to use a three-dimensional convolutional layer to extract the spatial-temporal features from the 3D-tensor samples. One advantage of the convolutional neural network is its ability to extract multiple feature maps using different convolutional kernels. Suppose the value of the neuron v_{ij}^{xyz} at position (x, y, z) of the jth feature map in the ith layer is given by

$$v_{ij}^{xyz} = g(\sum_m \sum_{p=0}^{P_i-1} \sum_{q=0}^{Q_i-1} \sum_{r=0}^{R_i-1} w_{ijm}^{pqr} v_{(i-1)m}^{(x+p)(y+q)(z+r)} + b_{ij}) \qquad (2)$$

where m, b_{ij}, P_i and Q_i have the same definitions as the 2D model, R_i is the size of the kernel along the temporal dimension, w_{ijm}^{pqr} is the value of position (p, q, r) connected to the mth feature map. Through 3D convolution, the CNN can extract the spatial and temporal information of CEUS data simultaneously.

The pooling layer in the convolutional neural network is a sub-sampling layer commonly constructed after the convolutional layer. The most common pooling operation is max-pooling earlier adopted in [30], which is used throughout this paper. The max-pooling operation can be written as follows:

$$\widehat{\alpha} = \max_{i,j,m} \alpha_{i,j,m} * u(n, n, n) \qquad (3)$$

where $u(n, n, n)$ is a three-dimensional window function used to define the compressing patch of the convolutional layer, and $\widehat{\alpha}$ is the maximum in the neighborhood.

After several convolutional and pooling layers, the classification task in the neural network is done via a fully-connected layer. As in regular neural networks, neurons in a fully-connected layer have full connections to all activations in the previous layer. Their activations can hence be computed with a matrix multiplication followed by a bias offset. For cancer detection, the output of the fully-connected layer consists of two neurons which represent positive and negative cancer samples respectively (Fig. 3).

An automatic learning process can be applied to estimate the parameters of the three-dimensional CNN. Our implementation details are based on those of the standard CNN as described in [33], [34]. The training algorithm is based on the stochastic diagonal Levenberg-Marquardt method [33], [34]. All the model parameters are randomly initialized as in [33], [34]. The learning rate adopted at the start of training process was 0.05. If no improvement was detected after 50 iterations, the learning rate was adaptive reduced, and the finial learning rate to stop the training process was 10^{-5}. The mini-batch size was 50. We used the l_2 regularization in training process to prevent over-fitting, and the regularization factor was 0.0005.

IV. EXPERIMENTS

Our experiments evaluated the proposed framework for prostate cancer detection based on both targeted and non-targeted CEUS data over 20 xenograft subjects. The experiments evaluated the targeted and non-targeted samples independently for comparison. As shown in Table I, the examined data sets were extracted from 38.2 gigabytes of raw video images. The video frames were split into three-dimensional tensor samples. After discarding the samples in which no perfusion was detected, about 47.57 thousand image sequences were arranged for training, validating and testing the proposed framework.

The CEUS data set consisted of two groups of subsets, i.e. the targeted group of 21844 samples and the non-targeted

TABLE II
ARCHITECTURE OF THE 3D-CNN BASED FRAMEWORK BUILT FOR AUTOMATED CANCER DETECTION.

No.	Layer type	Input size	Kernel size	Feature maps
1	convolutional	23*23*10	4*4*4	3
2	pooling	20*20*7	2*2*1	3
3	convolutional	10*10*7	3*3*4	12
4	pooling	8*8*4	2*2*1	12
5	convolutional	4*4*4	4*4*1	48
6	fully-connected	128*1	-	1

TABLE III
ACCURACIES OF DIFFERENT NETWORKS OVER THE VALIDATION SET.

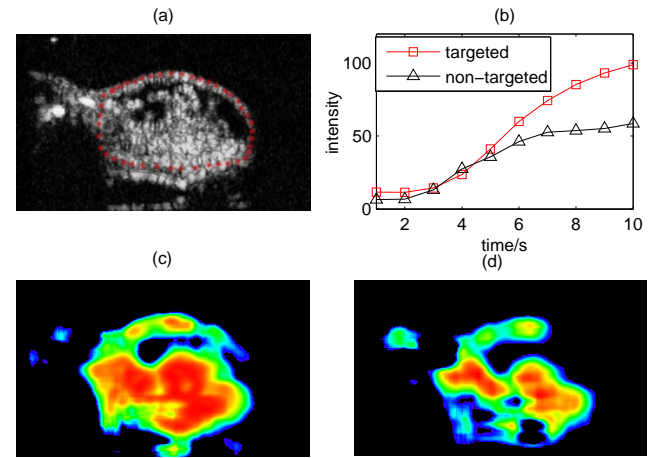| Number of kernels		Data sets	
Convolution 1	Convolution 2	Non-targeted	Targeted
2	4	81.41 ± 4.63	84.86 ± 4.90
	8	79.35 ± 5.10	85.64 ± 2.53
	12	80.88 ± 4.67	85.20 ± 3.01
3	4	80.32 ± 4.67	83.56 ± 4.30
	8	80.90 ± 5.20	84.99 ± 4.86
	12	$\mathbf{84.22 \pm 1.27}$	$\mathbf{89.90 \pm 1.14}$
4	4	79.25 ± 7.50	85.25 ± 6.03
	8	79.16 ± 5.51	85.23 ± 5.17
	12	80.28 ± 5.96	85.59 ± 4.38

Fig. 4. Comparison between the targeted and non-targeted CEUS images for cancer detection. (a) dot circle indicates the region of prostate cancer, (b) the average intensity of pixels in the cancer region. Lower figures are the color maps of estimated probability of being cancer over the (c) targeted and (d) non-targeted CEUS frames, with red indicating high probability and blue indicating low probability.

group of 25738 samples. Each group contained three subsets, i.e. the training set, validation set and test set. The training samples were used to fit the learning based models, and the validation samples were used to evaluate the models with different hyper-parameters. The measurements of performance, including sensitivity, specificity and the average accuracy were calculated using the test set.

A. Network Architecture and Hyperparameters

The network architecture of the proposed framework is listed in order in Table II. Specifically, it consists of six layers, i.e. three convolutional layers, two pooling layers and one fully-connected classification layer. The six layers are stacked as a pipeline, where each layer uses the output of the previous layer as input. Since the input data are three-dimensional image samples, the kernels used in the convolutional and pooling layers are also three dimensional.

One advantage of the convolutional neural network is its ability to extract more than one feature maps for each convolutional layer using different kernels. The number of kernels determines the number of features extracted for classification, which is an important trade-off between performance and computational complexity. Our experiment evaluated different networks with 2 to 12 kernels in the first and second convolutional layer. Table III lists the classification accuracies of the examined networks. Generally speaking, for the first

TABLE IV
ACCURACIES OF THE PROPOSED FRAMEWORK WITH DIFFERENT SIZES OF THE CONSTRUCTED CEUS SAMPLES.

Data sets	Size of the input image samples			
	21*21*10	23*23*10	25*25*10	27*27*10
Non-targeted	82.55 ± 0.76	$\mathbf{84.22}$ ± 1.27	84.07 ± 0.09	81.65 ± 1.05
Targeted	87.39 ± 0.23	$\mathbf{89.90}$ ± 1.14	87.71 ± 0.52	87.93 ± 0.09

convolutional layer, the networks with three kernels achieved the highest accuracy. And for the second convolutional layer, higher classification accuracies were achieved when twelve kernels were incorporated.

As described in Section 3, our experiment split the CEUS video frames into small image windows for cancer detection. Our experiment examined experiment settings of different window sizes. As shown in Table IV, the classification accuracy is relatively robust with different window sizes. Since smaller window sizes led to higher spatial resolution for cancer detection, we chose the window size of 23*23 in our research. And it was interesting to see that the targeted CEUS sets showed approximately two to four percent higher classification accuracy than the non-targeted sample sets.

B. Analysis of Targeted and Non-targeted Results

As the frontier of CEUS imaging research, the contrast agents made of material targeted to cancer cells has been recently applied for diagnosing the prostate cancer [35]. As a primary attempt for targeted CEUS based cancer detection, our experiments examined the anti-PSMA based targeted contrast agent, and compared it with the blank non-targeted agent for the application of prostate cancer detection.

Fig. 4 shows the change of intensity for the same subject injected with targeted and non-targeted contrast agents. Subgraph a illustrates the cancer region, and subgraph b compares the average increase of intensity in the cancer region. Though both types of the agents increase the average intensity in the cancer region, the targeted agent results in significantly higher average intensity than the non-targeted agent at the end of the perfusion process.

Experiment results showed that, in average, the targeted CEUS agent achieved 25.70% higher increase of intensity than the non-targeted agent after calibration, which was coherent with the earlier experimental research [32]. For the application

TABLE V
COMPARING WITH RECENT RESEARCHES OF THE CEUS BASED PROSTATE CANCER DETECTION

| Work	Parameters	Subjects	Sensitivity	Specificity	Accuracy
This	-	20	82.83	91.45	90.18
[15]	r	8	77.30	86.00	−
	ρ	8	78.10	81.60	−
[36]	r	21	77.90	82.40	−
	ρ	21	77.30	81.30	−
	κ	8	77.30	81.30	−
[18]	Pe	10	73.70	79.80	−
[19]	v	25	72.50	82.10	−
[?]	r	19	71.00	75.00	73.00
[38]	WIT	19	75.00	68.20	72.00
	κ,MTT,r,PT	19	79.00	80.00	81.00
[20]	I	23	71.00	71.30	−
	r	23	69.00	68.90	−
	ρ	23	67.90	67.30	−
	FWHM	23	62.20	66.80	−

Fig. 5. Illustration of the (a) input CEUS samples and the (b) output activations of the last convolutional layer of the 3D-CNN model. Both the input and output data are linearly embedded in the two-dimensional subspace for visualization using the Principal Component Analysis.

of automated cancer detection, higher increase of intensity in the cancer region could partly explain the higher accuracy achieved by the examined approaches over the targeted data sets.

The lower subgraphs of Fig. 4 illustrate the 3D-CNN estimated risk of prostate cancer. It seems that, the targeted CEUS agent could increase the sensitivity and specificity for prostate cancer detection. Specifically, the sensitivity, specificity and accuracy were 10.43%, 5.97% and 7.20% higher over the targeted test sample set than over the non-targeted sample set in our experiments.

C. Comparing with Other Approaches

We calculated the average sensitivity, specificity and accuracy of the proposed method over the test sample set, and compared the results with the state-of-the-art CEUS based approaches for prostate cancer detection. So far, most existing approaches were parameter based methods, which calculated and used different perfusion or dispersion related parameters for prostate cancer detection. The parameters included the temporal correlation (r), the spectral coherence (ρ), the kappa coefficient (κ), the mutual information (I), and the peak time (PT), wash-in time (WIT) and mean transit time (MTT) of the time intensity curve [15], [20], [36]. The recently examined parameters of the dispersion velocity (v) , the Péclet number (Pe) and the full-width at half maximum (FWHM) of the TIC were also compared [18], [19]. Besides single parameter based approaches, the multi-parametric models were also compared with the proposed method [38]. Table V lists the compared

TABLE VI
COMPARING WITH OTHER MACHINE LEARNING BASED APPROACHES

Methods	Sensitivity	Specificity	Accuracy
Proposed method	**82.98 ± 6.23**	**91.45 ± 6.75**	**90.18 ± 6.62**
Logistic	51.15 ± 4.05	75.21 ± 4.51	71.55 ± 5.87
kNN	62.98 ± 25.12	93.50 ± 4.60	88.43 ± 8.22
Random Forest	76.95 ± 23.05	92.05 ± 4.15	89.79 ± 7.23
FLDA	50.40 ± 6.20	80.68 ± 3.12	75.89 ± 5.57
J48	70.85 ± 22.65	83.50 ± 6.30	81.24 ± 4.10
Decision Table	67.53 ± 27.47	86.83 ± 5.27	83.53 ± 5.61
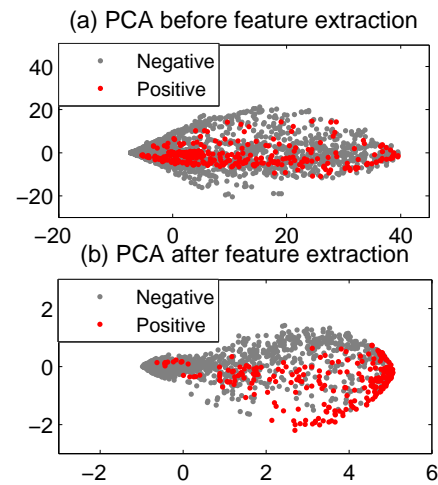
results. In summary, the proposed method showed 5.45% to 24.65% higher specificity and 3.83% to 20.63% higher sensitivity than the compared parameter based approaches. The reasons why the deep learning approach achieved higher performance may partly be because that, most parameter based methods were based on different assumptions about the perfusion or dispersion process. Since the deep learning approach doesn't require any assumptions about the perfusion process, it is more robust against different types of noises.

Besides the parameter based approaches, we also performed a group of experiments to compare the proposed deep learning framework with conventional machine learning approaches. Table VI lists the detection results over the targeted CEUS test sample set achieved by different learning models, i.e. the Logistic Regression, the k-Nearest Neighbor classifier, the Random Forest, the Fisher Linear Discriminative Analysis (FLDA), the J48 Decision Tree and the Decision Table. All the compared approaches used the same experimental setting. All models were trained using equal number of positive and negative samples. In summary, the 3D-CNN showed the highest sensitivity and average accuracy than the conventional machine learning approaches. The experiment was also performed over the non-targeted CEUS sample set. It was interesting to find that all the learning based approaches showed higher sensitivity ($3.17\% − 16.48\%$) and higher specificity ($6.15\% − 13.75\%$) over the targeted data than over the non-targeted data.

We also compared proposed method with the SVM, the naive Bayes and the neural network classifier. All the compared classifiers were trained using the same balanced data set. Experiment results show that, the compared approaches have comparable specificity as the 3D-CNN model; however, the sensitivity of the SVM (64.70%), the naive Bayes (54.98%) and the neural network (61.60%) were 18.13-27.85% lower than the 3D-CNN model. This may be caused by their high error rate over the positive cancer samples. It seems that the 3D-CNN model is more robust for CEUS based cancer detection than the conventional machine learning methods.

The higher performance achieved by the deep learning framework can be better understood by visualizing the nonlinearity of the data. Possibly due to the infiltration characteristic of cancer, the positive and negative cancer samples are highly non-separable in linear subspaces. Fig. 5 illustrates the original CEUS samples (gray dots) and the output of the third convolutional layer of the 3D-CNN (red dots). Both the input and the output data are vectorized and reduced to two dimensions using the Principle Component Analysis for visualization. Before feature extraction, the positive and negative CEUS samples tend to blend (Fig. 5.a); However, after the nonlinear 3D-CNN based feature extraction, the positive and negative samples (Fig. 5.b) become more separable, which may explain the relatively high classification accuracy achieved by the proposed deep learning framework.

It is worth noting that, the testing data set of our experiments were not balanced, which was coherent with clinical practice. To evaluate the proposed method, we calculated the receiver operating curve (ROC), and compared the area under the curve (AUC) of the examined learning based approaches. In general, the 3D-CNN (0.980) showed higher AUC measurement than all the compared machine learning approaches, including the SVM (0.875), the Logistic Regression (0.782), the Neural Network (0.950), the kNN (0.804), the Random Forest (0.891) and the Naive Bayes (0.790).

D. Computational Complexity

One important consideration for deep learning based approach is computational complexity. Generally speaking, the computational complexity of training a convolutional neural network is higher than training a conventional neural network. Fortunately, the weight-sharing strategy significantly reduces the number of parameters in the convolutional layers and saves computation time. Technically, our experiments were performed using a regular computer with single Intel Core I3-3240 3.4-GHz processor and 10-GB memory. The training process of a typical 3D-CNN (Table II) consumed about half an hour over the examined data sets. Recent research indicated that the training process of the CNN could be accelerated using off-the-shelf graphics processing units (GPUs). Since the proposed 3D-CNN is a spatial-temporal variant of the CNN, any GPU-based implementations of the CNN can be easily modified to support the proposed framework for large-scale implementations.

V. DISCUSSION AND CONCLUSION

Though noninvasive early detection of prostate cancer is of great importance for cancer diagnosis and treatment, at present, accurate cancer detection based on contrast enhanced ultrasound images is still challenging due to the complexity of the angiogenesis process. This paper presents a deep learning framework to detect prostate cancer in the CEUS videos. By uniformly extracting the features from both the spatial and the temporal dimensions, the proposed method achieves over 90% specificity and average accuracy, which is competitive with the state-of-the-art prostate cancer detection approaches.

As far as we know, this could be the first attempt to apply the CEUS with the targeted contrast agent for automated prostate cancer detection. We examined the recently invented anti-PSMA based agent targeted to prostate cancer cells, and compared it with the non-targeted blank agent. The experiment indicated that, the examined targeted agent increased the volume of agent aggregated in the cancer region. As a result, the targeted models achieved over ten percent higher sensitivity for prostate cancer detection. The quantitative cancer detection results may help to evaluate different contrast agents for cancer diagnosis.

In summary, this research is a preliminary attempt to apply deep learning for CEUS based cancer diagnosis. Experiment showed that a middle-size convolutional neural network could achieve relatively high sensitivity and accuracy for prostate cancer detection. So far, our research has been based on the CEUS videos recorded with experimental targeted contrast agent, which limited the size of available data. In the future, we plan to extend our research and take advantage of large-scale clinical CEUS data for prostate cancer diagnosis. With more qualified data, we can build more advanced deep learning models such as the ResNet and GoogLeNet for end-to-end prostate cancer detection.

REFERENCES

[1] R. Siegel et al., Cancer statistics, 2015, CA-Cancer J. Clin., vol. 65, no. 1, pp. 529, 2015.

[2] D.Rosario et al., Short term outcomes of prostate biopsy inmen tested for cancer by prostate specific antigen: Prospective evaluationwithin ProtecT study, Brit. Med. J., vol. 344, pp. 112, 2012, Art. no. d7894.

[3] B. Weber et al., A near-fatal case of sepsis with an antibiotic-resistant organism complicating a routine transrectal prostate biopsy in a health care worker, Can. Urol. Assoc. J., vol. 2, no. 5, pp. 543545, 2008.

[4] E. J. Halpern, Contrast-enhanced ultrasound imaging of prostate cancer, Reviews in urology, vol. 8 no. 29, 2006

[5] J T. Salonen, R. Salonen, Ultrasound B-mode imaging in observational studies of atherosclerotic progression, Circulation, vol. 87, no. 3, pp. 56-65, 1993.

[6] F. Helmcke, N C. Nanda, M C. Hsiung, Color Doppler assessment of mitral regurgitation with orthogonal planes Circulation, vol. 75, no. 1, pp. 175-183, 1987.

[7] M. Claudon, D. Cosgrove, T. Albrecht, Guidelines and good clinical practice recommendations for contrast enhanced ultrasound (CEUS)-update 2008, Ultraschall in der Medizin-European Journal of Ultrasound, vol. 29, no. 01, pp. 28-44, 2008.

[8] M X. Tang, H. Mulvana, T. Gauthier, Quantitative contrast-enhanced ultrasound imaging: A review of sources of variability, Interface Focus A Theme Supplement of Journal of the Royal Society Interface, vol. 1, no. 4, pp. 520-539, 2011.

[9] R. J. Eckersley, J. P. Sedelaar, M. J. K. Blomley, H. Wijkstra, N. M. deSouza, D. O. Cosgrove, and J. J. M. C. H. de la Rosette, Quan- titative microbubble enhanced transrectal ultrasound as a tool for monitoring hormonal treatment of prostate carcinoma, Prostate, vol. 51, pp. 256267, 2002.

[10] N. Elie, A. Kaliski, P. Pronneau, P. Opolon, A. Roche, and N. Lassau, Methodology for quantifying interactions between perfu- sion evaluated by DCE-US and hypoxia throughout tumor growth, Ultrasound Med. Biol., vol. 33, no. 4, pp. 549560, 2007.

[11] D Cosgrove and N Lassau, Imaging of perfusion using ultrasound, European journal of nuclear medicine and molecular imaging, vol. 37, no. 1, pp. 6585, 2010.

[12] A. Postema, M. Mischi, l. R. J. De, et al. Multiparametric ultrasound in the detection of prostate cancer: a systematic review, World Journal of Urology, vol. 33, no.11, pp. 11-19, 2015.

[13] M P J. Kuenen, M. Mischi, H. Wijkstra, Contrast-Ultrasound Diffusion Imaging for Localization of Prostate Cancer, IEEE Transactions on Medical Imaging, vol. 30, no. 8, pp. 1493, 2011.

[14] M. Mischi, M. P. J. Kuenen, and H. Wijkstra, Angiogenesis imaging by spatiotemporal analysis of ultrasound-contrast-agent disper- sion kinetics, IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 59, no. 4, pp. 621629, Apr. 2012.

[15] M. P. Kuenen, T. A. Saidov, H. Wijkstra, et al. Spatiotemporal correlation of ultrasound contrast agent dilution curves for angiogenesis localization by dispersion imaging, IEEE Transactions on Ultrasonics Ferroelectrics & Frequency Control, vol. 60, no. 12, pp. 2665-2669, 2013.

[16] A. Gasnier, R. Ardon, C. Ciofolo-Veit, and E. Leen, Assessing tumour vascularity with 3D contrast-enhanced ultrasound: a new semi-automated segmentation framework, Biomedical Imaging: From Nano to Macro, IEEE International Symposium on, 2010.

[17] S. Schalk, L. Demi, M. Smeenge, et al. Three-dimensional contrast-ultrasound dispersion imaging for prostate cancer localization, a feasibility study, Ultrasonics Symposium. IEEE, pp.616-619, 2014.

[18] R. J. Van Sloun, L. Demi, A. W. Postema, et al. Imaging of the dispersion coefficient of Ultrasound contrast agents by Wiener system identification for prostate cancer localization, Ultrasonics Symposium. IEEE, vol. 21, pp. 1-4, 2015.

[19] R J. Van Sloun , L. Demi, A W. Postema, Ultrasound-contrast-agent dispersion and velocity imaging for prostate cancer localization, Medical Image Analysis, vol. 35, pp. 610-619, 2017.

[20] S. Schalk, L. Demi, N. Bouhouch, et al. Contrast-enhanced Ultrasound Angiogenesis Imaging by Mutual Information Analysis for Prostate Cancer Localization,IEEE transactions on bio-medical engineering, vol. 64, no. 8, pp. 661-670, 2017.

[21] A. Postema, P. J. A. Frinking, M. Smeenge, et al. Dynamic contrasten-hanced ultrasound parametric imaging for the detection of prostate cancer, Bju International, vol. 117, no. 4, pp. 598-603, 2016.

[22] G. Russo, M. Mischi, W. Scheepens, J. de la Rosette, and H. Wijks- tra, Angiogenesis in prostate cancer: Onset, progression and imag- ing, BJU Int., vol. 110, pp. E794E808, 2012

[23] Q. Zhang, L. Yang, C. Li,and W. Wang, Contrast-enhanced ultrasound image segmentation of atherosclerotic plaques using spatial-temporal analysis and snakes, Systems and Informatics (ICSAI), International Conference on. IEEE, 2012.

[24] Ma, J., et al. "A pre-trained convolutional neural network based method for thyroid nodule diagnosis." Ultrasonics 73(2017):221.

[25] Dou, Q., et al. "Automatic Detection of Cerebral Microbleeds from MR Images via 3D Convolutional Neural Networks." IEEE Transactions on Medical Imaging 35.5(2016):1182-1195.

[26] Zou, Liang, et al. "3D CNN based Automatic Diagnosis of Attention Deficit Hyperactivity Disorder Using Functional and Structural MRI." IEEE Access PP.99(2017):1-1.

[27] Yang, X., Yu, L., Wu, L., Wang, Y., Ni, D., Qin, J., et al. . Fine-grained recurrent neural networks for automatic prostate segmentation in ultrasound images. AAAI2017, Thirty-First AAAI Conference on Artificial Intelligence.

[28] A. Krizhevsky, I. Sutskever, G E. Hinton, Imagenet classification with deep convolutional neural networks, Advances in neural information processing systems, pp. 1097-1105 2012.

[29] L A. Meinel, A H. Stolpen, K S. Berbaum, Breast MRI lesion classification: Improved performance of human readers with a backpropagation neural network computeraided diagnosis (CAD) system, Journal of magnetic resonance imaging, vol. 25, no. 1, pp. 89-95, 2007.

[30] S. Ji, W. Xu, M. Yang, 3D convolutional neural networks for human action recognition, IEEE transactions on pattern analysis and machine intelligence, vol. 35, no. 1, pp. 221-231, 2013.

[31] Zhou, Xichuan, et al. "Deep Learning With Grouped Features for Spatial Spectral Classification of Hyperspectral Images." IEEE Geoscience and Remote Sensing Letters 14.1(2016):97-101.

[32] X. Fan, L. Wang, Y.Guo, and Z. Tu, Ultrasonic nanobubbles carrying anti-PSMA nanobody: construction and application in prostate cancer-targeted imaging, PloS one vol. 10, no. 6, 2015.

[33] Y. LeCun, L. Bottou, G. Orr, and K. Muller, Efficient Backprop, Neural Networks: Tricks of the Trade, G. Orr and M. Klaus-Robert, eds., Springer, 1998.

[34] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, Gradient-Based Learning Applied to Document Recognition, Proc. IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998.

[35] X. Fan, L. Wang, Y. Guo, Experimental investigation of the penetration of ultrasound nanobubbles in a gastric cancer xenograft Nanotechnology, vol. 24, no. 32, July 2013.

[36] M. Mischi, H. Wijkstra, Contrast dispersion imaging for cancer localization, Conf Proc IEEE Eng Med Biol Soc, no. 2014, pp. 4268-4271, 2014

[37] R. J. Van Sloun, L. Demi, A. W. Postema, et al. Statistical characterization of Ultrasound-Contrast-agent velocity fields for prostate cancer localization, Ultrasonics Symposium. IEEE, pp. 1-4, 2016.

[38] R R. Wildeboer, A W. Postema, L. Demi, Multiparametric dynamic contrast-enhanced ultrasound imaging of prostate cancer, European Radiology, vol. 27, no. 8, pp. 1-9, 2017.