HHS Public Access
Author manuscript
Phys Med Biol. Author manuscript; available in PMC 2025 January 24.

Published in final edited form as:
Phys Med Biol. ; 65(20): 205003. doi:10.1088/1361-6560/aba410.

Automatic multi-needle localization in ultrasound images using large margin mask RCNN for ultrasound-guided prostate brachytherapy

Yupei Zhang,
Zhen Tian,
Yang Lei,
Tonghe Wang,
Pretesh Patel,
Ashesh B Jani,
Walter J Curran,
Tian Liu,
Xiaofeng Yang
Department of Radiation Oncology and Winship Cancer Institute, Emory University, Atlanta, GA, United States of America

Abstract

Multi-needle localization in ultrasound (US) images is a crucial step of treatment planning for US-guided prostate brachytherapy. However, current computer-aided technologies are mostly focused on single-needle digitization, while manual digitization is labor intensive and time consuming. In this paper, we proposed a deep learning-based workflow for fast automatic multi-needle digitization, including needle shaft detection and needle tip detection. The major workflow is composed of two components: a large margin mask R-CNN model (LMMask R-CNN), which adopts the lager margin loss to reformulate Mask R-CNN for needle shaft localization, and a needle based density-based spatial clustering of application with noise algorithm which integrates priors to model a needle in an iteration for a needle shaft refinement and tip detections. Besides, we use the skipping connection in neural network architecture to improve the supervision in hidden layers. Our workflow was evaluated on 23 patients who underwent US-guided high-dose-rate (HDR) prostrate brachytherapy with 339 needles being tested in total. Our method detected 98% of the needles with 0.091 ± 0.043 mm shaft error and 0.330 ± 0.363 mm tip error. Compared with only using Mask R-CNN and only using LMMask R-CNN, the proposed method gains a significant improvement on both shaft error and tip error. The proposed method automatically digitizes needles per patient with in a second. It streamlines the workflow of transrectal ultrasound-guided HDR prostate brachytherapy and paves the way for the development of real-time treatment planning system that is expected to further elevate the quality and outcome of HDR prostate brachytherapy.

xiaofeng.yang@emory.edu .

Keywords

multi-needle localization; mask RCNN; DBSCAN algorithm; ultrasound images; US-guided HDR prostate brachytherapy; deep learning

1. Introduction

Prostate cancer is one of the most common cancers affecting the male population, which accounts for nearly 20% of new cancer diagnoses in 2019 (Siegel et al 2019). High-dose-rate (HDR) prostate brachytherapy, emerged in the early 1990s, is currently offered to prostate cancer patients with intermediate- to high-risk disease as a boost treatment together with pelvic external beam radiation therapy, and to patients with low risk disease as monotherapy option (Yoshioka et al 2014). As an outpatient procedure, it gains a lot of attractions to radiation oncologists and patients because of its excellent dose conformity to target and sparing of adjacent organs, and its high-dose rate hypofractionation scheme which is preferable to prostate cancer having a postulated lower alpha-beta ratio than normal tissue (Challapalli et al 2012). Hence, HDR prostate brachytherapy is the most commonly used technique in clinical prostate cancer therapy.

An HDR prostrate brachytherapy usually involves: (1) inserting 12–20 needles into the prostate under the guidance of transrectal ultrasound (TRUS) imaging; (2) scanning patients by CT (and/or MRI) for anatomy delineations and needle digitization; (3) treatment planning to optimize dwell positions and dwell times of the radioactive source in implanted needles; and (4) delivering the planned treatment by dwelling the radioactive source through the needles via a remote afterloader. Although the current process has been widely adopted in clinics, there are still a few drawbacks in the first two steps. First, the CT scan (and/or MRI scan) introduces extra steps into the clinical workflow. It prolongs patients' length of stay and raises the medical cost. Moving patients on and off the scanner couch also increases the risk of needle shifts, potentially impairing the clinical outcome (Holly et al 2011). An ideal solution is using a single imaging modality for both needle insertion guidance and treatment planning (Yang et al 2017). Second, the clinicians, determine needle localizations and place needles in operating room (OR) solely based on their experiences, which could lead to a suboptimal needle pattern, leaving little room for treatment plan quality improvement later in the planning stage (Siauw et al 2012, Sadowski et al 2017, Meer et al 2018, Wang et al 2020). Real-time dosimetric information of treatment plan, if could be provided in OR, would help clinicians make objective judgment calls for needle placements. Finally, obtaining manually digitizing needles and needle tips is tedious, time-consuming and prone to human errors (Jung et al 2019). It poses a big efficiency challenge for real-time treatment planning in OR. An accurate and quick automatic needle localization method is of imminent need.

Efforts have been made to address these issues. TRUS-only (Schmid et al 2013) and MRI-only (Lindegaard et al 2016) based HDR prostate brachytherapy techniques have been developed to streamline the clinical practice. Compared with MRI-based counterparts, TRUS-based approach bears drastically less cost and is widely adopted. However, TRUS imaging often suffers from serious image noise, speckles and artifacts, which have similar appearances as the needle shown on the images, challenging the application of traditional object localization methods for needle digitization (Uherčík et al 2010, Younes et al 2018). Moreover, when multiple needles are inserted, the interplays between them may lead to partially overlapped trajectories on TRUS images (Yang et al 2014). Needles that are relatively farther away from the transrectal probe usually yield weaker signals, making them more susceptible to noise (Zhang et al 2020a, 2020b, 2020c). Besides, the rapid dose falloff from the radioactive source puts high penalties on inaccurate needle localizations. Tiong et al (2010) have reported that the accuracy of needle digitization should be within 3 mm for acceptable dosimetric uncertainty using CT technique. All these issues and facts challenge the realization of automatic multi-needle localization.

Many methods have been proposed to localize a single needle or a needle-like instrument employing the imaging physical properties (Daoud et al 2015), mathematical algorithms (Uherčík et al (2010) and machine learning models (Pourtaherian et al 2018). The distinct signal pattern in medical imaging caused by needle insertion were used to identify the needle from surrounding tissues. Daoud et al (2015) used the needle reflection pattern of circular ultrasound (US) waves, large amplitudes and repetitive reflection patterns, to identify the needle in the received signals. Since the needle has high intensity, Uherčík et al (2010) sought an intensity threshold by minimizing voxel classification error to separate the needle voxel from the image background, followed by the random sample consensus (RANSAC) algorithm for needle modeling. Besides the RANSAC, a few mathematical models have been proposed and investigated on needle modeling. The Kalman filter was used in the work of Zhao et al (2013), while the spatial parametric polynomial curve model was employed in the work of Younes et al (2018), to improve the performance of RANSAC. Zhou et al (2007) combined a novel line representation model with 3D Hough transform to detect a straight needle in US images. While mathematical models and algorithms are responsible for needle modeling in a region of interest (ROI), machine learning methods are often used to identify the needle ROI composed of needle pixels or voxels. Pourtaherian et al (2017) used the linear discriminant analysis and linear support vector machine (SVM) to classify the proposed voxel representations and then detected the medical instrument axis using the RANSAC. Naïve Bayes classifier (Younes et al 2018), multilayer perceptron network Geraldes and Rocha (2014), and probabilistic SVM (Beigi et al 2017) have also been employed to partition the images pixels or voxels into needle or non-needle type. In comparison with traditional machine learning models, deep learning models have become the benchmarks in image classification (Lecun et al 2015) and segmentation (Falk et al 2019), as well as objective detection (He et al 2017). Mwikirize et al (2018) used a fully convolutional network (FCN) to proposed candidate regions that are then fed into a fast region-based convolution neural network (R-CNN) for needle detection in 2D US. Pourtaherian et al (2018) proposed two convolutional neural networks (CNNs) to classify needle voxels or segment the needle from orthogonal 2D cross-sections from the 3D US volume. To localize a needle in MRI-targeted prostate biopsy, Mehrtash et al (2018) proposed a deep anisotropic 3D FCN with skip-connections in prostate MRI segmentation. However, without modifications or manual intervention, none of these methods could be used to detect multiple inserted needles in TRUS-only HDR prostate brachytherapy. This is due to the two challenges in multi-needle detection in US images: the low signal-to-noise ratio, and the overlapping ROIs of adjacent needles.

Recently, we proposed a multi-needle detection workflow for the TRUS-guided prostate brachytherapy consisting of an order-graph regularized sparse dictionary learning (ORDL), a ROIs seeking algorithm and an iterative ROI-RANSAC (Zhang et al 2020c). The ORDL-based method learned the dictionaries from the US images without needles to aid the needle detection. U-Net model was modified to segment needles in US images (Zhang et al 2020d). However, the ORDL-based method involves intensive computations, while the U-Net-based method needs postprocessing and human intervention to determine the needle tips.

In this study, we proposed a deep learning-based multi-needle localization workflow to accurately detect the needles in the US images in real-time. The proposed method trains a large margin mask R-CNN model (LMMask R-CNN) to localize needle shafts on patient's US images, and uses a needle based density-based spatial clustering of applications with noise (DBSCAN) method (Schubert et al 2017) to refine the shaft localizations and detect the needle tips. This approach is the first attempt at adapting mask R-CNN and the DBSCAN algorithm for multi-needle localization in the TRUS guided HDR prostate brachytherapy (Zhang et al 2020c). The designed network can achieve automatic needle digitization on US images within a second, which is promising for real-time needle tracking. Besides, using it as one of the building blocks, a real-time HDR planning system could be built to generate an intermediate treatment plan based on the current needle pattern to provide clinicians real-time dosimetric feedback to guide the needle placement. It also enables the completion of the entire HDR prostate procedure in OR, reducing unnecessary patient transportation and hence the risk of needle shift.

2. Materials and methods

2.1. Data acquisition and annotation

The data used in this study consists of $1024 \times 768 \times S \times 23$ 3D TRUS images collected from 23 patients who have received HDR prostate brachytherapy in our institution. Here, S is the slice number ranging from 26 to 40. All US images were acquired using a Hitachi Hi Vision Avius (Model: EZU-MT30-S1, Hitachi Medical Corporation, Tokyo, Japan) US system equipped with a transrectal probe (Model: HITACHI EUP-U533). These images were acquired in B-mode, with the settings: 7.5 MHz center frequency, 17 frames per second, thermal index <0.4, mechanical index 0.4, and 65 dB dynamic range. 12–19 (depending on patient prostate size) Nucletron ProGuide Sharp 5 F needles were inserted under TRUS guidance. The Needle length is 240 mm and has an outer diameter of 1.67 mm. Institutional review board approval was obtained; informed consent was not required for this Health Insurance Portability and Accountability Act (HIPAA)-complaint retrospective analysis.

VelocityAI 3.2.1 (Varian Medical Systems, Palo Alto, CA) was used to manually annotate all trace points of the inserted needles on all image slices by an experienced physicist. In our study, these manual annotations are adopted as ground truths for model learning and evaluation. Here, needle shafts are defined as these curves along needle trajectories, where a needle trajectory is a set of needle centers on all related slices; a needle tip is defined as the center point of the needle shaft at the most distal slice. A needle is identified by the white susceptible artifacts around its truth shaft in TRUS images. Besides, all susceptible artifacts from a needle shaft are continuous in multiple adjacent slices. Hence, the annotation tool allows the physicist to label several pivotal slices along needle shafts, and then automatically produces needle trajectories via linear interpolation. A needle was reconstructed by placing a circle with 1.67 mm diameter on each related slice along the needle trajectory. Mask images were finally adopted as learning targets to train the proposed model.

2.2. Large margin mask R-CNN

The proposed deep learning-based needle localization model produces numerous needle ROI proposals on the US images and then implements the three following parts on each proposal simultaneously: bounding-box center regression, object classification and pixel classification. Figure 1 shows the schematic workflow of training the proposed deep model. In ROI proposal stage, our model uses bounding boxes to generate many proposals that may contain needles. The bounding-box center regression step aims to minimize the deviation of centers of the bounding boxes from the ground truth. In object classification stage, we classify each ROI proposal into a needle or non-needle category. Meanwhile, US pixel classification, *i.e.* needle segmentation, is also performed to identify needle pixels. Data augmentation, such as rotation, flipping, and scaling, was used to enlarge the variety of the training data.

2.2.1. Mask R-CNN.—The proposed model is based on the classical Mask R-CNN which is a popular deep learning architecture for object detection (He et al 2017). Mask R-CNN aims to minimize the multi-task loss L on each sampled ROI, defined as:

$$L = \omega_1 L_{cls} + \omega_2 L_{box} + \omega_3 L_{mask} \tag{1}$$

where the classification loss L_{cls} and the bounding box loss L_{box} were defined as in the work of Girshick (2015); the mask loss L_{mask} were defined as in the work of He et al (2017). Mask R-CNN architecture includes a deep feature learning network (*i.e.* ResNet), the RPN and a ROIAlign layer for ROI proposals, an FCN for mask prediction, and another FCN for softmax classification and bounding box regression. In addition, the original Mask R-CNN set $\omega = [\omega_1, \omega_2, \omega_3] = [1, 1, 1]$ to have an equal balanced trade-off among these three losses in network training. To use Mask R-CNN for needle detection in US images, we reformulate the objective functions with needle priors, shown in equations (3)–(5).

2.2.2. Bounding-box center regression.—Since the bounding box used in our problem has a fixed size of 1.67 mm according to the known needle diameter, we pay more attention to the bounding-box center localization. To accurately localize the needle centers, the proposed model adopts the following two coordinates in box regression (Girshick et al 2014):

$$t_x = (x - x_c), t_y = (y - y_c), t_x^* = (x^* - x_c), t_y^* = (y^* - y_c) \tag{2}$$

where x and y are the predicted coordinates, while x^* and y^* are the ground truth coordinates. To build the mapping from predicted coordinates to ground truth coordinates, we employ the robust loss Girshick (2015) to redefine the bounding-box regression loss that is:

$$L_{box} = \sum_{i=1}^{N} \mathcal{L}\left(t_x^i - t_x^*\right) + \sum_{i=1}^{N} \mathcal{L}\left(t_y^i - t_y^*\right) \tag{3}$$

where N is the number of samples in mini-batches and

$$\mathcal{L}(u) = \begin{cases} 0.5u^2, & if |u| < 1 \\ |u| - 0.5, & otherwise \end{cases}$$

In comparison with the original box loss, equation (3) removes the effects of the fixed box size.

2.2.3. Object classification.—The softmax objective function used in Mask R-CNN is popular in multi-object recognition due to its efficient computation. However, our problem is a binary classification problem where a large margin makes the objects more discriminative without increasing computation time (Liu et al 2016, Elsayed et al 2018). The proposed method reformulates the classification loss L_{cls} using the following objective function:

$$L_{cls} = \sum_{i=1}^{N} \sum_{l} \max\left\{0, \gamma + \frac{f_o(x_i) - f_t(x_i)}{\varepsilon + \| \nabla_l f_o(x_i) - \nabla_l f_t(x_i) \|_2}\right\} \tag{4}$$

where l indicates the lth layer of the classification net, γ is a distance for the decision boundary, $\varepsilon = 1 \times 10^{-6}$ to remove the numerical issue, $f_o(x_i)$ and $f_t(x_i)$ represent the scores of classifying the i-th sample into the class o and its ground truth label t, respectively. Note that the loss in equation (4) collects the margin losses at all layers of the classification network to implement the deep supervision.

2.2.4. Pixel classification.—The mask segmentation aims to identify the needle pixels from the US images. This problem is a binary classification problem on all pixels of bounding box, so we adopt the large margin softmax loss to achieve accurate classification (Liu et al 2016). We reformulate the mask loss L_{mask} as:

$$L_{mask} = \sum_{i=1}^{N} -\log\left(\frac{e^{\|\boldsymbol{W}_{y_i}\|_2\|\mathbf{x}_i\|_2\varphi(\theta_{y_i})}}{e^{\|\boldsymbol{W}_{y_i}\|_2\|\mathbf{x}_i\|_2\varphi(\theta_{y_i})} + \sum_{j \neq y_i} e^{\|\boldsymbol{W}_j\|_2\|x_i\|_2\cos(\theta_j)}}\right) + \lambda\|\boldsymbol{W}\|_{\mathcal{F}}^2 \tag{5}$$

where y_i is the ground truth label of the sample x_i, $j \in \{0, 1\}$ where 0 and 1 denote a non-needle pixel and a needle pixel, respectively. For $\varphi(\theta)$, we adopt the specific one used in the work of (Liu et al 2016). Note that, for the smoothness, we introduce l_2-norm on all weights, *i.e.* both hidden layers and the output layer of the pixel classification networks (Martins and Astudillo 2016).

2.3. Needle based-DBSCAN

After LMMask R-CNN, we propose a needle based-DBSCAN algorithm to model each needle, refine shaft localizations and detect needle tips. DBSCAN is a popular density-based clustering algorithm that identifies a series of core points by density estimation and iteratively seeks clusters by direct density connection (Schubert et al 2017). Density estimation counts the number of neighbors within a circle of radius μ and direct density connection means that a core point can be reached from core points in the current cluster. Unlike k-means (Zhang et al 2020c) requiring a predefined cluster number, DBSCAN separates needle points by areas of low density. However, DBSCAN is susceptible to needle overlapping or noisy points generated from LMMask R-CNN. Our method overcomes this issue by modeling needle in each iteration, whose major steps are listed in table 1.

2.4. Model training and test

We used the 5-fold cross validation method on our US image dataset of the 23 patient cases to evaluate the proposed method. Specifically, we divided the dataset into 5 subsets, *i.e.* 3 subsets of 5 patients each and 2 subsets of 4 patients, and then ran five individual tests. In each test, we trained the proposed deep learning-based model using 4 of the 5 subsets and evaluated the trained model on the rest subset. That is, the model was trained on 18, 18, 18, 19, 19 patients and evaluated on 5, 5, 5, 4, 4 patients in the five individual tests, respectively. The quantitative evaluation results were calculated on all these tests. Note that the 5-fold cross validation was used for method evaluation, while the model could be retrained on all data for clinical practical use. For the hyperparameters in the proposed deep model, we set $\omega = [0.2, 0.6, 0.2]$ in equation (1) and $\lambda = 0.01$ in equation (5) empirically, while other parameters were set as that used in the corresponding references. In network learning, we set the learning rate to 1e−6 and terminated the training at 500 epochs where the decrease of the training error was observed to be negligible.

2.5. Evaluation metrics

For evaluations, we computed the metrics including needle shaft localization error, needle tip localization error and needle detection accuracy (Zhang et al 2020d). Needle shaft localization is defined as:

$$E_{shaft} = \frac{1}{N}\sum_{i=1}^{N}\left\|o_i - t_{i2}\right\|, \tag{6}$$

where N is the number of all center position in computations, o_i represents the predicted position for the i th center, and t_i is the ground truth position for the i th center. To ease our evaluation on the shaft error, we use 1.67 mm as a threshold and exclude the points that do not have a ground-truth point within 1.67 mm from the calculation. One reason is that it is difficult to determine the correct corresponding ground-truth points for these points, in the presence of multiple needles. These predicted positions whose deviation from the ground truth exceed 1.67 mm are considered as false localization points and hence involved in the calculation in equation (9). In practical clinical application, we will allow users to review our result and modify it if needed. Note that Mask R-CNN and LMMask R-CNN may produce multiple bounding boxes for a needle point. In this situation, the box with the highest possibility (He et al 2017) is adopted for evaluations. Needle tip localization error is defined as:

$$E_{tip} = \frac{1}{M}\sum_{i=1}^{M}\left|l_i - s_i\right| \tag{7}$$

where M is the number of needles in computations, l_i indicates the predicted length for the i th needle, and s_i is the ground truth length for the i th needle. Similarly, those predicted tips whose deviation from the ground truth exceed 6 mm are considered to be a false tip detection and involved by the calculation in equation (8). This 6 mm threshold was chosen based on our limited slice thickness (~2 mm for most of our cases) and a phantom study, which reported a range of [−5.8 mm, 3.4 mm] for manual tip identification errors on US images (Schmid et al 2013). Needle detection accuracy is defined as:

$$Acc = \frac{1}{M}\#\left\{\left|l_i - s_i\right| \leqslant \tau_i : i = 1, 2, \ldots, M\right\} \tag{8}$$

where τ_i was here set as 6 mm in this study and $\#\{A\}$ counts the elements in a set A. Besides, we calculated the common indexes on all shaft points from the perspective of retrieval (Zhang et al 2017), including

$$Precision = \frac{\#\{\{Detection\ \ Set\} \cap \{Ground\ \ Trut\ \ Set\}\}}{\#\{Detection\ \ Set\}},$$
$$Recall = \frac{\#\{\{Detection\ \ Set\} \cap \{Ground\ \ Truth\ \ Set\}\}}{\#\{Ground\ \ Truth\ \ Set\}},$$
$$\text{and}\, F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall},$$

(9)

to quantitatively evaluate the proposed method. We have also compared the proposed method with Mask R-CNN, U-Net and ORDL. Two-sample t-test was adopted to evaluate the significant difference between the detection result and the manual delineation, using the function 'ttest2' in MATLAB 2019b.[1]

3. Results

3.1. Visualizations

Figure 2 shows the 2D visualization of the needle localization results on 5 slices from a patient's US images. As shown on these 5 slices, Mask R-CNN results in 13 missed needle points and 8 false localization points; LMMask R-CNN results in 3 missed needle points and 5 false localization points; while our method leads to no missing and false points.

Figure 3 shows the 3D visualizations of the identified needles obtained from different methods. Note that both Mask R-CNN and LMMask R-CNN are only for needle segmentation and do not include any post-processing for tip identification. Hence the subfigures (b) and (c) in figure 4 do not show needle tip. Mask R-CNN loses 59 needle points in total and produced 16 false points in shaft localization; LMMask R-CNN leads to 41 missed needle points and 5 false points; while our method yields 2 missed needle points and no error points. Besides, we considered the most distal detected point in a need shaft as needle tip for both tip evaluations of Mask R-CNN and LMMask R-CNN.

3.2. Quantification evaluations

Table 2 summarizes the results in terms of our used metrics on the 339 needles from 23 patients, including the average needle shift localization error (E_{shaft}), the average tip localization error (E_{tip}) and the detection accuracy (Acc). Note that we manually identify the needle tips on the results obtained from Mask R-CNN and LMMask R-CNN in order to calculate tip localizations for evaluation purpose. In our study, LMMask R-CNN that uses large margin losses reduces the shaft error by about 0.07 mm and the tip error by about 0.29 mm in comparison with Mask R-CNN that uses the softmax loss. Besides, the large margin loss also helps to increase the detection accuracy by 6%. The refinement that uses the proposed needle based-DBSCAN furthermore reduces the shaft error and the tip error by about 0.02 and 0.16 mm, respectively. Finally, our method detected 98% needles with 0.091 mm shaft error and 0.330 mm tip error, which is statistically significantly better than Mask R-CNN and LMMask R-CNN (P-values <0.001 in T-tests). Table 2 also shows the results with setting τ_i in equation (8) as 3 mm.

Table 3 summarizes the statistical indexes calculated on all needle points from the perspective of needle point retrieval. Compared to Mask R-CNN, LMMask R-CNN results in higher precision, recall and F1. These results demonstrate the gain of using large margin losses and our algorithm in table 1.

[1] www.mathworks.com/

Figure 4 shows average needle shift localization errors for each patient. As can be seen, the average needle shaft error obtained by Mask R-CNN ranges from 0.043 to 0.46 mm, with the errors of 7 cases exceeding 0.3 mm. The shaft error of LMMask R-CNN ranges from 0.031 to 0.34 mm, with only one case exceeding 0.3 mm. In contrast, our method yielded a shaft error ranging from 0.022 to 0.165 mm. These results have demonstrated that our method consistently achieved smaller shaft error than Mask R-CNN and LMMask R-CNN for all the testing patients.

Figure 5(a) shows the distribution of all shaft localization centers, where the probability is computed by $\#\{\textit{localization error in the interval}\}/\#\{\textit{all localizations}\}$ and the interval is from 0 to 1 mm with step 0.1 mm. Our method has about 89% localizations within 0.4 mm errors, while Mask R-CNN and LMMask R-CNN have about 63% and 79% localizations respectively. The most probable shaft error of our method is 0.1 mm, while both those of Mask R-CNN and LMMask R-CNN are 0.2 mm.

Figure 5(b) lists the counts of detected needles of different tip detection errors. From the results, the three quartiles of the three methods are all localized at 0 mm. Mask R-CNN, LMMask R-CNN and our method has 74%, 82%, and 87% needle tips with 0 mm errors respectively. The tip detection errors of our method are all within 4 mm, while Mask R-CNN and LMMask R-CNN have 13 and 5 detected needle that have a tip error of 6 mm.

3.3. Inter-observer effect

To investigate the sensitivity of our method to the inter-observer variability of manual needle digitization, we asked the other two physicists (referred to as physicist 2 and 3) to manually delineate all trace points on the US images as well. We conducted our workflow repeatedly, using these two manual annotations and the mean positions averaged over all the three annotations as ground truths, respectively. Table 4 shows our evaluation results in terms of the average shaft error, the average tip error and the detection accuracy. As is shown, there are no significant differences between the results obtained by using these different manual annotations. Besides, table 4 also shows the results with setting τ_i in equation (8) as 3 mm. All P-values between the evaluations in tables 2 and 4 are greater than 0.05.

4. Discussions and conclusions

Accurately obtaining needle positions is very essential to accomplish reliable dosimetry of HDR prostate brachytherapy. Accurate multi-needle localization in TRUS images is a challenge due to image noise, speckles and artifacts. However, the manual needle-digitization is a labor-intensive task and prone to errors, which prevents real-time treatment planning and dosimetric evaluation during needle insertion.

This study presents an automatic method for accurate needle localization in US images by adopting a deep earning-based image detection model. Specifically, the proposed needle localization workflow is composed of a LMMask R-CNN and a needle based DBSCAN. The LMMask R-CNN aims to reduce the multi-objective loss: needle center regression, object classification, and pixel classification, where the used large margin loss is more discriminative in a binary classification problem. The proposed needle based DBSCAN integrates the needle prior knowledge to model a needle in an iteration, meanwhile refining needle shafts and detecting needle tips. Our workflow digitizes the inserted needles into shaft positions and tip positions automatically.

Our evaluations on 23 patients with 339 needles show that the proposed approach is significantly effective for multi-needle detection in US images. Comparing with Mask R-CNN, LMMask R-CNN detects 98% needles with 0.11 mm needle shaft error and 0.49 mm tip error, manifesting the effectiveness of the larger margin loss. After the refinement stage, our method delivers a better performance. i.e. 98% needles detected with 0.09 mm shaft error and 0.33 mm tip error, manifesting the effectiveness of needle based DBSCAN. In addition, the proposed method could tackle needle localization within about 0.6 s for a patient, while an experienced physicist usually takes about 15–20 min to manually digitize all catheters. It shows a potential of developing a real-time automatic needle tracking system.

In comparison with the state-the-art methods using ORDL (Zhang et al 2020c) and U-Net (Zhang et al 2020d), our proposed method achieves smaller shaft and tip errors and higher accuracy (*i.e.* $E_{shaft} = 0.09 \pm 0.04$ mm, $E_{tip} = 0.33 \pm 0.36$ mm and $Acc = 0.98$). On the dataset used in this study, the ORDL based method results in: $E_{shaft} = 0.19 \pm 0.13$ mm, $E_{tip} = 1.01 \pm 1.74$ mm and $Acc = 0.95$, while the U-Net based method delivers: $E_{shaft} = 0.29 \pm 0.23$ mm, $E_{tip} = 0.44 \pm 0.93$ mm and $Acc = 0.96$. Although the differences in the shaft and tip errors among these methods are not clinically significant, our improvement on accuracy percentage reduces the amount of the missing points and incorrectly identified points, and hence reduces the amount of post-processing and human intervention needed for correction. In terms of the computational efficiency, the proposed workflow takes about 0.6 s per patient on a NVIDIA GeForce RTX 2080Ti GPU, including the needle reconstructions and the post-processing. U-Net takes about 0.4 s per patient for needle reconstruction only but needs a post-processing step to obtain the needle tip. Although ORDL includes a post-processing step, it takes more than 36 s per patient.

Our proposed approach could be employed to streamline and speed-up the clinical workflow of US-guided HDR prostate brachytherapy and achieve real-time treatment planning in OR. Our method takes about 0.6 s per patient to automatically delineate the needles and identify needle tips. A prostate auto-segmentation method on US has also been developed recently, which could be done within less than 1 s (Lei et al 2019). Besides, plan optimization for prostate HDR brachytherapy can be automatically completed within 5 s, using a multi-criteria optimization algorithm (Breedveld et al 2019). With these three methods, we expect to achieve real-time automatic treatment planning in OR for prostate HDR brachytherapy in near future, obtaining a treatment plan from scratch within seconds. In this way, physicians could adjust the needle pattern in OR based on the feedback of achievable intermediate dose distribution. Potentially, it would enable treatment right after needle placement in OR room, which not only decreases the patient's overall stay and treatment cost, but also eliminate unnecessary patient transportation. All these benefits would ultimately lead to better clinical outcomes.

There are two limitations in our method. (1) Although the proposed method detects most needles with small shaft errors, the tip error still has a big improvement space in future work. (2) The aim of our proposed method is to achieve automatic needle digitization with an accuracy level comparable to manual delineation. Hence, we used the manual needles delineated by our experienced physicists as our ground truth for model training and evaluation. However, the manual needle digitization on US images tend to have relatively large errors. It was reported by Schmid et al (2013) that the error of the needle manual localization on US images of a prostate phantom is within 1 mm, and the tip identification error is in a range of (−5.8 mm, 3.4 mm) with a median error of 0.5 mm, compared to CT-based post-insertion manual delineation. This is limited by the image quality of US images and interferences between needles, and beyond the scope of this study.

Overall, the proposed method can achieve a clinically acceptable performance on multi-needle detection in TRUS images for HDR prostate brachytherapy. Since there is no assumption regarding the prostate images, this method can also be adapted for other object detection assignments in US images, such as implanted seeds in low-dose-rate brachytherapy or fiducial markers. The proposed method may also be adapted for needle localization on CT and MR images. CT images have good contrast and no needle shadow as appeared on US images when obturators are removed or replaced by dummy wires during scan, thus it can be less challenging. For MR images, depending on the MR sequences and whether the dummy markers are used, the contrast of needle may vary, which needs further evaluation with possible issues to be addressed (Dai et al 2020). We will test our method on these image modalities in our future work. On the other hand, more training samples could enhance the deep learning model to more effectively learn the needle patterns in the images. In the future, we will also evaluate the proposed deep-learning method on more patient's data for model training and conduct a more comprehensive evaluation on different scenarios.

Acknowledgments

This research is supported in part by the National Cancer Institute of the National Institutes of Health under Award Number R01CA215718 (XY), the Department of Defense (DoD) Prostate Cancer Research Program (PCRP) Award W81XWH-17-1-0438 (TL) and W81XWH-17-1-0439 (AJ) and Dunwoody Golf Club Prostate Cancer Research Award (XY), a philanthropic award provided by the Winship Cancer Institute of Emory University.

References

Beigi P, Rohling R, Salcudean T, Lessoway VA and Ng GC 2017 Detection of an invisible needle in ultrasound using a probabilistic SVM and time-domain features Ultrasonics 78 18–22 [PubMed: 28279882]

Breedveld S, Bennan AB, Aluwini S, Schaart DR, Kolkman-Deurloo I-K-K, Heijmen B J and Biology 2019 Fast automated multi-criteria planning for HDR brachytherapy explored for prostate cancer Phys. Med 64 205002

Challapalli A, Jones E, Harvey C, Hellawell G and Mangar S 2012 High dose rate prostate brachytherapy: an overview of the rationale, experience and emerging applications in the treatment of prostate cancer Br. J. Radiol 85 S18–S27 [PubMed: 23118099]

Dai X, Lei Y, Zhang Y, Qiu RL, Wang T, Dresser SA, Curran WJ, Patel P, Liu T and Yang X 2020 Automatic multi-catheter detection using deeply supervised convolutional neural network in MRI-guided HDR prostate brachytherapy Med. Phys (10.1002/mp.14307)

Daoud MI, Rohling RN, Salcudean SE and Abolmaesumi P 2015 Needle detection in curvilinear ultrasound images based on the reflection pattern of circular ultrasound waves Med. Phys 42 6221–33 [PubMed: 26520715]

Elsayed G, Krishnan D, Mobahi H, Regan K and Bengio S 2018 Large margin deep networks for classification Adv. Neural Inf. Process. Syst 2018 842–52

Falk T, Mai D, Bensch R, Çiçek Ö, Abdulkadir A, Marrakchi Y, Böhm A, Deubner J, Jäckel Z and Seiwald K 2019 U-Net: deep learning for cell counting, detection, and morphometry Nat. Methods 16 67–70 [PubMed: 30559429]

Geraldes AA and Rocha TS 2014 A neural network approach for flexible needle tracking in ultrasound images using Kalman filter IEEE RAS/EMBS Int. Conf. on Biomedical Robotics and Biomechatronics pp 70–5

Girshick R 2015 Fast r-cnn Proc. of the IEEE Int. Conf. on Computer Vision pp 1440–8

Girshick R, Donahue J, Darrell T and Malik J 2014 Rich feature hierarchies for accurate object detection and semantic segmentation Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition pp 580–7

He K, Gkioxari G, Dollár P and Girshick R 2017 Mask R-CNN Proc. of the IEEE Int. Conf. on Computer Vision pp 2961–9

Holly R, Morton GC, Sankreacha R, Law N, Cisecki T, Loblaw DA and Chung HT 2011 Use of cone-beam imaging to correct for catheter displacement in high dose-rate prostate brachytherapy Brachytherapy 10 299–305 [PubMed: 21190903]

Jung H, Shen C, Gonzalez Y, Albuquerque K, Jia X M and Biology 2019 Deep-learning assisted automatic digitization of interstitial needles in 3D CT image based high dose-rate brachytherapy of gynecological cancer Phys. Med. Biol 64 215003 [PubMed: 31470425]

Lecun Y, Bengio Y and Hinton G 2015 Deep learning Nature 521 436–44 [PubMed: 26017442]

Lei Y, Tian S, He X, Wang T, Wang B, Patel P, Jani AB, Mao H, Curran WJ and Liu T 2019 Ultrasound prostate segmentation based on multidirectional deeply supervised V-Net Med. Phys 46 3194–206 [PubMed: 31074513]

Lindegaard JC, Madsen ML, Traberg A, Meisner B, Nielsen SK, Tanderup K, Spejlborg H, Fokdal LU and Nørrevang O 2016 Individualised 3D printed vaginal template for MRI guided brachytherapy in locally advanced cervical cancer Radiother. Oncol 118 173–5 [PubMed: 26743833]

Liu W, Wen Y, Yu Z and Yang M 2016 Large-margin softmax loss for convolutional neural networks Int. Conf. Mach. Learn 2 7

Martins A and Astudillo R 2016 From softmax to sparsemax: a sparse model of attention and multi-label classification Int. Conf. Machine Learning (ICML) pp 1614–23

Meer M, Pieters BR, Niatsetski Y, Alderliesten T, Bel A and Bosman PAN 2018 Better and faster catheter position optimization in HDR brachytherapy for prostate cancer using multi-objective real-valued GOMEA Proc. of the Genetic and Evolutionary Computation Conf (Kyoto, Japan: Association for Computing Machinery) pp 1387–94

Mehrtash A, Ghafoorian M, Pernelle G, Ziaei A, Heslinga FG, Tuncali K, Fedorov A, Kikinis R, Tempany CM and Wells WM 2018 Automatic needle segmentation and localization in MRI with 3-D convolutional neural networks: application to MRI-targeted prostate biopsy IEEE Trans. Med. Imaging. 38 1026–36 [PubMed: 30334789]

Mwikirize C, Nosher JL and Hacihaliloglu I 2018 Convolution neural networks for real-time needle detection and localization in 2D ultrasound Int. J. Comput. Assist. Radiol. Surg 13 647–57 [PubMed: 29512006]

Pourtaherian A, Scholten HJ, Kusters L, Zinger S, Mihajlovic N, Kolen AF, Zuo F, Ng GC, Korsten HHM and de With PHN 2017 Medical instrument detection in 3-dimensional ultrasound data volumes IEEE Trans. Med. Imaging. 36 1664–75 [PubMed: 28410101]

Pourtaherian A, Zanjani FG, Zinger S, Mihajlovic N, Ng GC and Korsten HHM 2018 Robust and semantic needle detection in 3D ultrasound using orthogonal-plane convolutional neural networks Int. J. Comput. Assist. Radiol. Surg 13 1321–33 [PubMed: 29855770]

Sadowski KL, Van der Meer MC, Luong NH, Alderliesten T, Thierens D, van der Laarse R, Niatsetski Y, Bel A and Bosman PA 2017 Proc. of the Genetic and Evolutionary Computation Conf pp 1224–31

Schmid M, Crook JM, Batchelar D, Araujo C, Petrik D, Kim D and Halperin R 2013 A phantom study to assess accuracy of needle identification in real-time planning of ultrasound-guided high-dose-rate prostate implants Brachytherapy 12: 56–64 [PubMed: 22513104]

Schubert E, Sander J, Ester M, Kriegel HP and Xu X 2017 DBSCAN revisited, revisited: why and how you should (still) use DBSCAN ACM Trans. Database Syst 42 1–21

Siauw T, Cunha A, Berenson D, Atamturk A, Hsu IC, Goldberg K and Pouliot J 2012 NPIP: a skew line needle configuration optimization system for HDR brachytherapy Med. Phys 39 4339–46 [PubMed: 22830767]

Siegel RL, Miller KD and Jemal A 2019 Cancer statistics, 2019 CA Cancer J. Clin 69 7–34 [PubMed: 30620402]

Tiong A, Bydder S, Ebert M, Caswell N, Waterhouse D, Spry N, Camille P and Joseph D 2010 A small tolerance for catheter displacement in high–dose rate prostate brachytherapy is necessary and feasible Int. J. Radiat. Oncol. Biol. Phys 76 1066–72 [PubMed: 19616901]

Uherčík M, Kybic J, Liebgott H and Cachard C 2010 Model fitting using RANSAC for surgical tool localization in 3-D ultrasound images IEEE Trans. Biomed. Eng 57 1907–16 [PubMed: 20483680]

Wang T, Giles M, Press RH, Dai X, Jani AB, Rossi P, Lei Y, Curran WJ, Patel P and Liu T 2020 Multiparametric MRI-guided high-dose-rate prostate brachytherapy with focal dose boost to dominant intraprostatic lesions Proc. SPIE 11317 1131724

Yang X, Rossi P, Ogunleye T, Marcus DM, Jani AB, Mao H, Curran WJ and Liu T 2014 Prostate CT segmentation method based on nonrigid registration in ultrasound-guided CT-based HDR prostate brachytherapy Med. Phys 41 111915 [PubMed: 25370648]

Yang X, Rossi PJ, Jani AB, Mao H, Zhou Z, Curran WJ and Liu T 2017 Improved prostate delineation in prostate HDR brachytherapy with TRUS-CT deformable registration technology: a pilot study with MRI validation J. Appl. Clin. Med. Phys 18 202–10 [PubMed: 28291925]

Yoshioka Y, Suzuki O, Otani Y, Yoshida K, Nose T and Ogawa K 2014 High-dose-rate brachytherapy as monotherapy for prostate cancer: technique, rationale and perspective J. Contemp. Brachytherapy 6 91 [PubMed: 24790627]

Younes H, Voros S and Troccaz J 2018 Automatic needle localization in 3D ultrasound images for brachytherapy 2018 IEEE 15th Int. Symp. on Biomedical Imaging (ISBI. 2018) pp 1203–7

Zhang Y, Harms J, Lei Y, Wang T, Liu T, Jani AB, Curran WJ, Patel P and Yang X 2020a Weakly supervised multi-needle detection in 3D ultrasound images with bidirectional convolutional sparse coding Proc. SPIE 11319 1131914

Zhang Y, He X, Tian Z, Jeong J, Lei Y, Wang T, Zeng Q, Jani AB, Curran W and Patel P 2020b Multi-needle detection in 3D ultrasound images with sparse dictionary learning Proc. SPIE 11319 113190I

Zhang Y, He X, Tian Z, Jeong JJ, Lei Y, Wang T, Zeng Q, Jani AB, Curran WJ and Patel P 2020c Multi-needle detection in 3D ultrasound images using unsupervised order-graph regularized sparse dictionary learning IEEE Trans. Med. Imaging 39 2302–15 [PubMed: 31985414]

Zhang Y, Lei Y, Qiu RL, Wang T, Wang H, Jani AB, Curran WJ, Patel P, Liu T and Yang X 2020d Multi-needle localization with attention U-net in US-guided HDR prostate brachytherapy Med. Phys 47 2735–45 [PubMed: 32155666]

Zhang Y, Xiang M and Yang B 2017 Graph regularized nonnegative sparse coding using incoherent dictionary for approximate nearest neighbor search Pattern Recognit 70 75–88

Zhao Y, Cachard C and Liebgott H 2013 Automatic needle detection and tracking in 3D ultrasound using an ROI-based RANSAC and Kalman method Ultrason. Imaging 35 283–306 [PubMed: 24081726]

Zhou H, Qiu W, Ding M and Zhang S 2007 Automatic needle segmentation in 3D ultrasound images using 3D Hough transform Proc. SPIE 6789 67890R

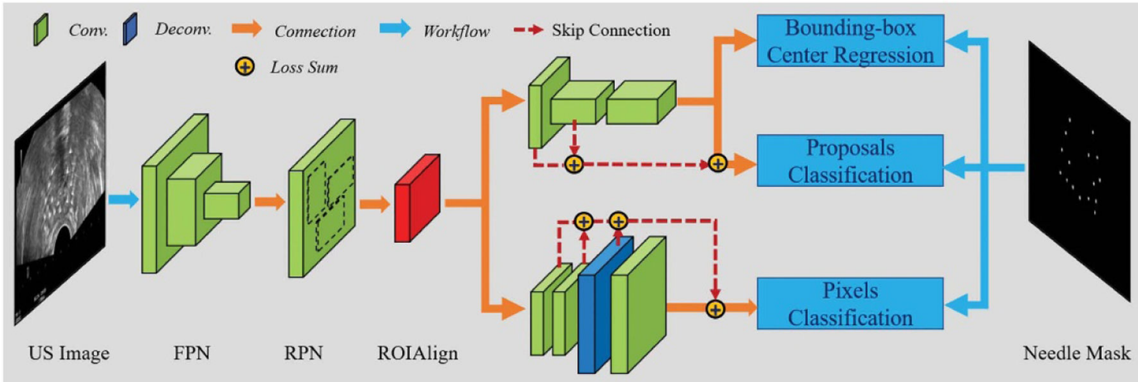

Figure 1.

The schematic workflow of training the proposed deep network, including feature pyramid networks (FPN), region proposal networks (RPN), ROI aligning layer.

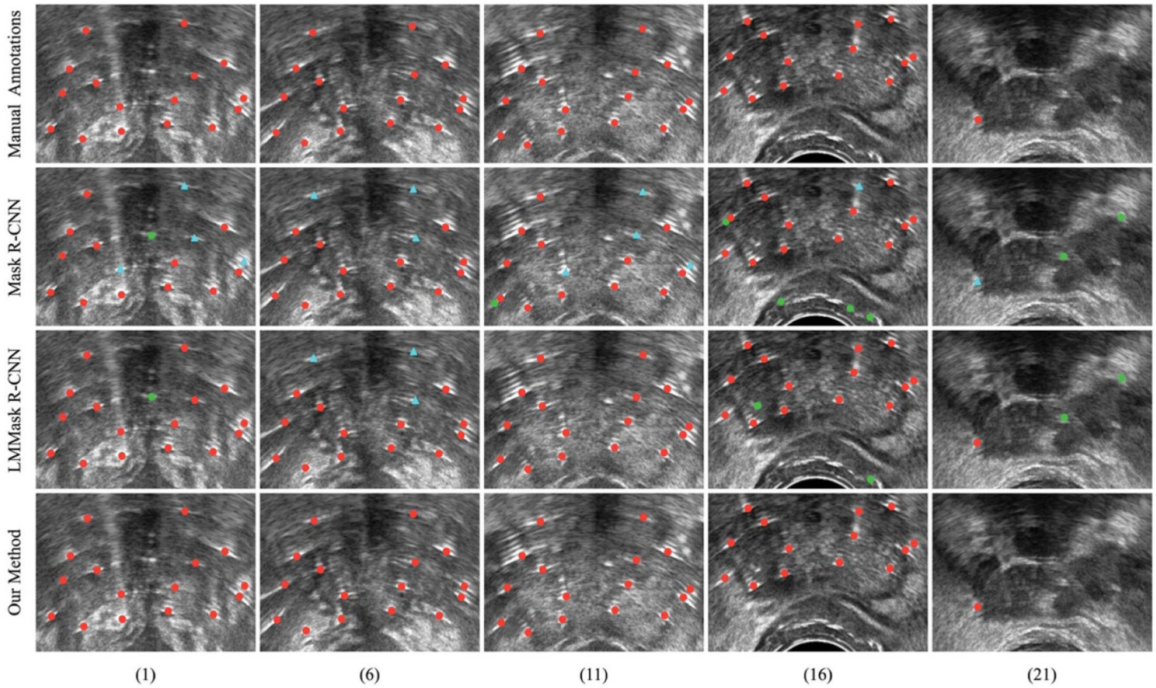

Figure 2.
2D visualizations on 5 slices of a patient's US images. The five columns correspond to the 1th, 6th, 11th, 16th, 21th slices of the US images, respectively. The green points are the missed needle localizations, while the cyan triangles denote the false localization points.

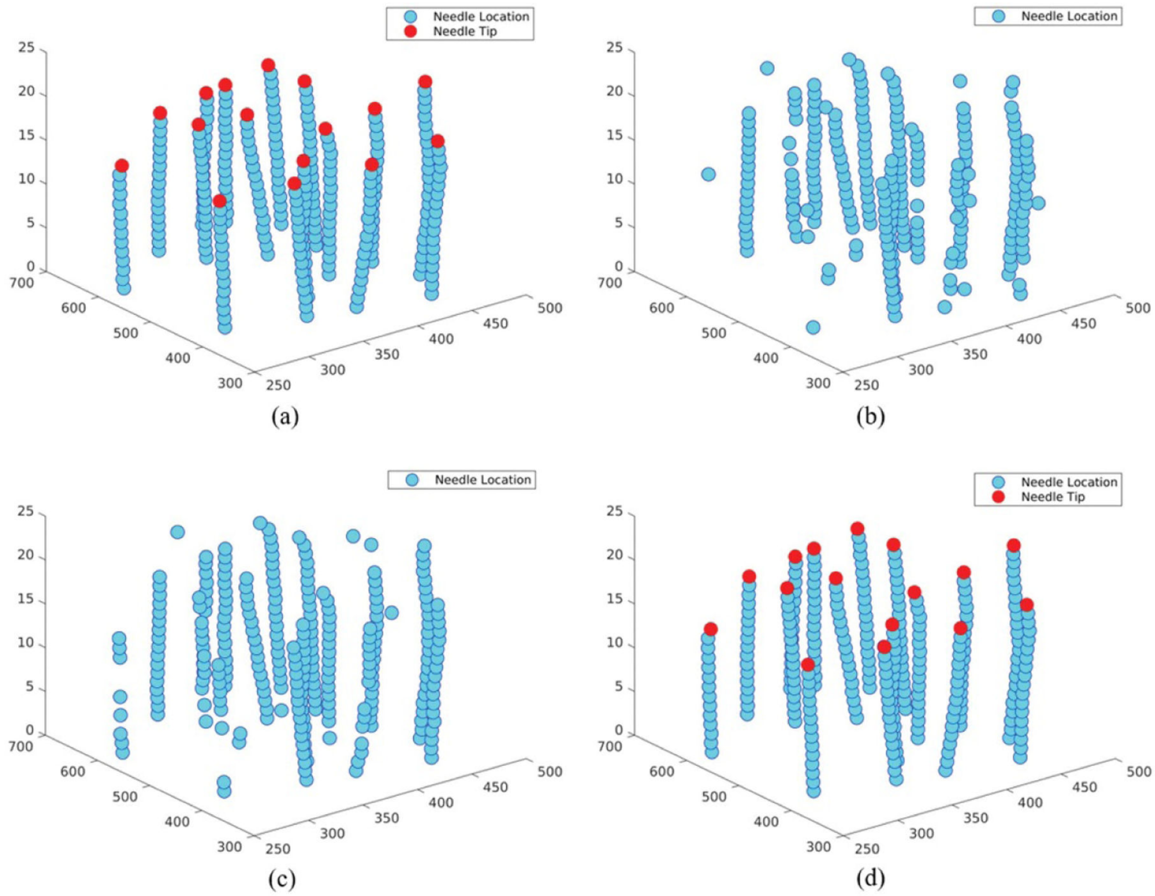

Figure 3.
3D visualizations of the identified needles on a patient's US images with (b) mask R-CNN, (c) LMMask R-CNN and (d) our workflow. The needles and tips manually identified by our physicist are also shown in (a) as ground truth.

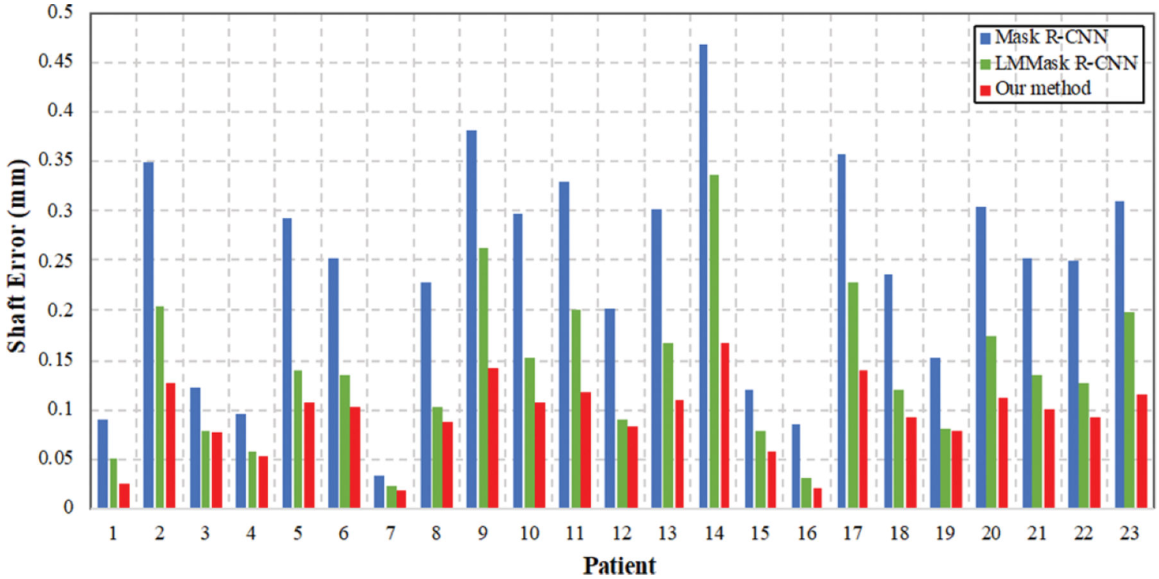

Figure 4.
Average shaft errors on each patient using the three methods.

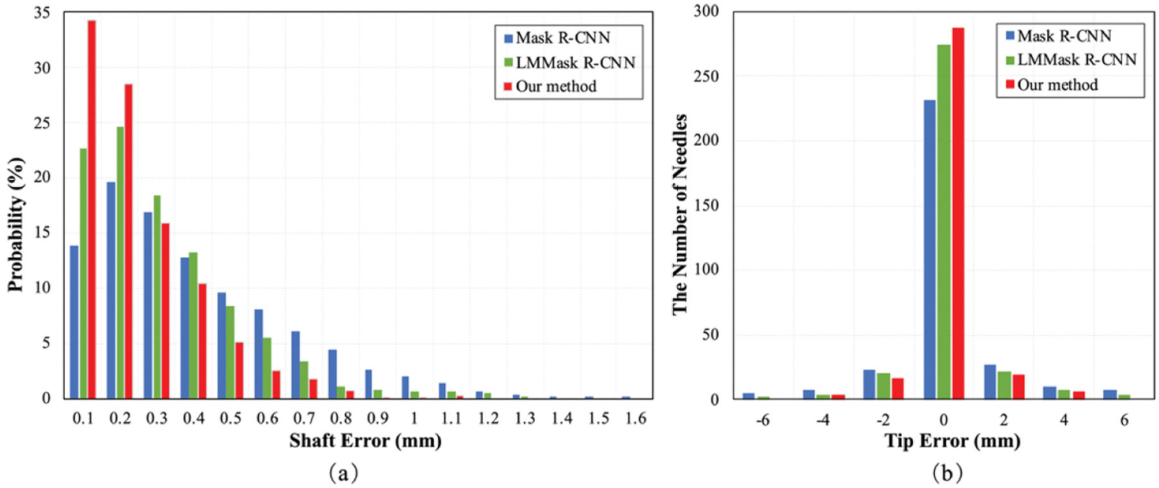

Figure 5.
The distributions of shaft localization errors (a) and tip errors (b) using the three used methods.

Table 1.

The improved DBSCAN algorithm for needle refinement and tip detection.

Needle based−DBSCAN Algorithm
Input: Detected needle point coordinates D; Patient's US images $\mathbf{X}$;
Output: Needle shaft points $\mathbb{C}$ and Needle tips Γ .
1. **While** NOT STOP
2. Performing DBSCAN on D;
3. Selecting the cluster C with the most points; if $\#\{C\} \le 3$, then STOP;
4. Fitting a needle model M in the cluster C;
5. Calculating the intersection points E of M on all slices in $\mathbf{X}$;
6. Finding out the point T with biggest intensity drop; % Tip Detection
7. Fetching points $E1$ before T in E; $C = C \cup E1$; % Filling Missed Points
8. Seeking points $C1 \subset C$ whose distances to M are greater than needle diameter 1.67 mm;
9. $D = D - C + C1, \mathbb{C} \leftarrow C - C1, \Gamma \leftarrow$ T;
10. **End**

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Evaluation results on all needles for 23 patients (mm). τ_i in equation (8) was set as 6 mm or 3 mm.

	Mask R-CNN (6 mm)	Mask R-CNN (3 mm)	LMMask R-CNN (6 mm)	LMMask R-CNN (3 mm)	Our Method (6 mm)	Our Method (3 mm)
E_{shaft}	0.190 ± 0.157	0.177 ± 0.164	0.114 ± 0.081	0.125 ± 0.083	0.091 ± 0.043	0.089 ± 0.040
E_{tip}	0.789 ± 1.172	0.711 ± 0.942	0.493 ± 0.630	0.473 ± 0.512	0.330 ± 0.363	0.316 ± 0.304
Acc	0.920	0.832	0.982	0.935	0.982	0.953

Table 3.

Evaluations on all needle shaft point retrieval.

	Mask R-CNN	LMMask R-CNN	Our Method
Precision	0.736	0.867	0.926
Recall	0.860	0.912	0.941
F1	0.793	0.889	0.933

Table 4.

Evaluation results with different manual annotations (mm). τ_i in equation (8) was set as 6 mm or 3 mm.

	$E_{shaft}(6\ \text{mm})$	$E_{shaft}(3\ \text{mm})$	$E_{tip}(6\ \text{mm})$	$E_{tip}(3\ \text{mm})$	$Acc(6\ \text{mm})$	$Acc(3\ \text{mm})$
Physicist 2	0.088 ± 0.032	0.078 ± 0.038	0.312 ± 0.341	0.294 ± 0.247	0.982	0.953
Physicist 3	0.092 ± 0.040	0.082 ± 0.040	0.336 ± 0.355	0.304 ± 0.286	0.982	0.953
Mean Positions	0.089 ± 0.039	0.086 ± 0.041	0.347 ± 0.369	0.311 ± 0.294	0.982	0.953