Clinical and Translational Radiation Oncology 51 (2025) 100925

Contents lists available at ScienceDirect

Clinical and Translational Radiation Oncology

journal homepage: www.sciencedirect.com/journal/clinical-and-translational-radiation-oncology

Fully automatic reconstruction of prostate high-dose-rate brachytherapy interstitial needles using two-phase deep learning-based segmentation and object tracking algorithms☆

Mohammad Mahdi Moradi [a], Zahra Siavashpour [b,c,*], Soheib Takhtardeshir [a], Eman Showkatian [d], Ramin Jaberi [e,f], Reza Ghaderi [g], Bahram Mofid [b], Farzad Taghizadeh-Hesary [g,h,*]

[a] Faculty of Electrical Engineering Shahid Beheshti University Tehran Iran
[b] Radiotherapy Oncology Department Shohada-e Tajrish Educational Hospital Shahid Beheshti University of Medical Science Tehran Iran
[c] Clinical Research and Development Unit of Shohaday Tajrish Hospital Tehran Iran
[d] Faculty of Medical Sciences Department of Medical Physics Iran University of Medical Science Tehran Iran
[e] Radiation Oncology Department Yas Hospital Tehran University of Medical Sciences Tehran Iran
[f] Department of Physics University of Surrey Guildford UK
[g] ENT and Head and Neck Research Center and Department The Five Senses Health Institute School of Medicine Iran University of Medical Sciences Tehran Iran
[h] Radiation Oncology Department Iran University of Medical Sciences Tehran Iran

ARTICLE INFO

Keywords:
Brachytherapy
Deep Learning
Prostate
Catheters
Neural Networks

ABSTRACT

The critical aspect of successful brachytherapy (BT) is accurate detection of applicator/needle trajectories, which is an ongoing challenge. This study proposes a two-phase deep learning-based method to automate localization of high-dose-rate (HDR) prostate BT catheters through the patient's CT images. The whole process is divided into two phases using two different deep neural networks. First, BT needles segmentation was accomplished through a pix2pix Generative Adversarial Neural network (pix2pix GAN). Second, a Generic Object Tracking Using Regression Networks (GOTURN) was used to predict the needle trajectories. These models were trained and tested on a clinical prostate BT dataset. Among the total 25 patients, 5 patients that consist of 592 slices was dedicated to testing sets, and the rest were used as train/validation set. The total number of needles in these slices of CT images was 8764, of which the employed pix2pix network was able to segment 98.72 % (8652 of total). Dice Similarity Coefficient (DSC) and IoU (Intersection over Union) between the network output and the ground truth were 0.95 and 0.90, respectively. Moreover, the F1-score, recall, and precision results were 0.95, 0.93, and 0.97, respectively. Regarding the location of the shafts, the proposed model has an error of 0.41 mm. The current study proposed a novel methodology to automatically localize and reconstruct the prostate HDR-BT interstitial needles through the 3D CT images. The presented method can be utilized as a computer-aided module in clinical applications to automatically detect and delineate the multi-catheters, potentially enhancing the treatment quality.

Introduction

Prostate cancer is the second most frequently occurring cancer in men, accounting for nearly 7.3 % of new cancer diagnoses yearly based on the 2020 global cancer statistical report [1]. Surgery, systematic therapy, and radiation therapy are the treatment options in terms of prostate cancer management. External beam radiation therapy (EBRT) has been accepted as an effective treatment option for low-risk prostate cancer. However, optimum biochemical control will not be achieved using conventional EBRT alone to treat patients with intermediate to

☆ This article is part of a special issue entitled: 'Brachytherapy' published in Clinical and Translational Radiation Oncology.

* Corresponding authors at: Radiotherapy Oncology Department, Shohada-e Tajrish Educational Hospital, Shahid Beheshti University of Medical Science, Tehran, Iran (Z. Siavashpour). ENT and Head and Neck Research Center and Department, The Five Senses Health Institute, School of Medicine, Iran University of Medical Sciences, Tehran, Iran (F. Taghizadeh-Hesary).

E-mail addresses: Z_siavashpour@sbmu.ac.ir (Z. Siavashpour), farzadth89@gmail.com (F. Taghizadeh-Hesary).

https://doi.org/10.1016/j.ctro.2025.100925
Received 30 April 2024; Received in revised form 7 January 2025; Accepted 16 January 2025
Available online 19 January 2025

high-risk prostate malignancies [2]. Therefore, dose escalation is needed to improve biochemical relapse-free survival and prostate cancer-specific survival [3].

Brachytherapy (BT) is an effective alternative or complementary treatment to EBRT for reaching higher target doses and lower organs at risk (OARs) toxicity for the low-risk to locally advanced malignancies [4,5]. The radioactive source/sources is/are placed into the prostate directly or through about 15–20 plastic [6] or stainless steel [7] needles during prostate BT. Therefore, these treatment techniques potentially improve radiotherapy's conformity and the therapeutic ratio [2].

Typically, there are two different types of prostate BT based on the source's duration of implantation and dose rate: permanent low-dose-rate brachytherapy (LDR-BT) and temporary high-dose-rate brachytherapy (HDR-BT). Some benefits were reported for each of these two BT modalities. During LDR-BT, tiny radioactive seeds are inserted into the prostate permanently. However, in HDR-BT, a single high dose rate source (e.g., ^{192}Ir or ^{60}Co) dwells inside the array of transperineally interstitial catheters after treatment planning evaluation and confirmation are inserted into the prostate [5].

The significant advantages of HDR-BT over LDR-BT include the ease of the process and lower dependency on operator skill, higher cost-effectiveness due to the feasibility of using one single radioactive source for many treatment courses, lower risk of operator error, a higher chance of treatment optimization, and feasibility of needle insertion outside the prostate gland to improve prescribed dose coverage of extracapsular extension or seminal vesicle invasion, and elimination of source migration risk [8,9].

Computed tomography (CT), magnetic resonance (MR), and transrectal ultrasound (TRUS) are the three main imaging modalities that have been used for treatment planning after applicator insertion. There are some reported pros and cons for each of these modalities: the power of soft tissue differentiation, cost, physician and staff efficiency, patient mobilization and reposition, and compatibility of treatment planning systems [2]. Overall, CT is one of the pioneer imaging techniques used in 3D treatment planning of prostate cancer BT worldwide [10].

Applicator digitization (also known as reconstruction) and defining source track relative to the target and OAR is recognized as the second cause of uncertainty in prostate BT dosimetry [11]. There are some predefined applicators' libraries in the treatment planning systems database from which a user can load and allocate the inserted applicators to the image series of patients. However, using these libraries is more valuable and practical for the rigid applicator set. Plastic needle tracking and defining all their trajectories is still a mandatory task that is time-consuming, labor-intensive, and prone to inter/intra-observer variability as users should do through the patient's image slices. Therefore, it is highly subjective and dependent on the user's skill and experience. This task becomes more challenging, especially when needles touch and cross each other in some of the reconstructed image slices during their path in the patient's body. Automatization of needles reconstruction was investigated previously. An electromagnetic (EM) sensor-based navigation system was proposed for fast and accurate HDR catheter reconstruction. However, the proposed method needs an EM tracking system installed beneath the BT table and some DOF EM sensors to be attached at the stylet tip and inserted into the needles for helping their detection [12]. All these technical and mechanical requirements can be used as a logical excuse for refusing the routine use of this method.

Machine learning and deep learning have been attractive tools for solving (proved as a state-of-the-art solution for) optimization, image segmentation, radiomics, and prediction problems during a reasonable time in medical science [13,14]. Some valuable publications developed deep learning algorithms for BT treatment planning, automatic organ segmentation, and applicator digitization [15,16]. Artificial intelligence and deep learning-based algorithms are proposed to automatically reconstruct different BT applicators through CT, MRI, and TRUS images [15]. Some artificial intelligence-based algorithms were proposed for 3D CT- and MRI-based gynecological HDR-BT rigid applicator automatic reconstruction [17,18]. Deep learning was also successfully used for automatic needle segmentation of TRUS-guided prostate BT [19], in which a deep learning-based model is proposed to segment and reconstruct the applicators automatically.

An automatic HDR-BT needle reconstruction algorithm is still needed due to the abundant CT-based treatment planning application. This research study proposes two new phases, deep learning-based models, for auto-localization of interstitial HDR prostate BT needle trajectories through the patients' CT image set. The proposed method consists of two phases: first, CT images are segmented using the generative adversarial networks (Pix2Pix) based on U-Net generator; next, the segmented images are passed to a tracking network (GOTURN) in a sequence of slides to extract the path of each needle.

Methods

Database

HDR-BT of prostate cancer using ^{60}Co

Clinical data was collected from Yas Hospital, Tehran, Iran, where HDR-BT was performed under regional anesthesia with endorectal ultrasound guidance. Treatments included definitive or salvage monotherapy and EBRT boost, delivered in four or two fractions. Single-needle implants ensured precise positioning, verified by measuring the exposed needle length. Dedicated CT markers of BT were inserted into interstitial plastic needles (Eckert & Ziegler BEBIG, Berlin, Germany). Pelvic CT images (HiSpeed Dual Scanner, GE Medical Systems, USA) with a 2 mm slice thickness were acquired and imported into the HDR PLUS BT planning system, which utilizes a ^{60}Co source database from the same company.

OARs, including the rectum, urethra, and bladder, were delineated by an experienced radiation oncologist. Clinical target volumes (CTV), encompassing low- and high-risk areas, were contoured using insights from patient MRI (T2w, diffusion-weighted, and dynamic contrast-enhanced images). An experienced medical physicist performed applicator reconstruction by mapping needle trajectories from their tips to the perineal exit points using the TPS applicator module.

Treatment planning aimed for CTV V100 $>$ 95 % and D90 $>$ 100 % of the prescribed dose (PD), with a PD of approximately 9.5 Gy per fraction over 2–4 fractions, depending on treatment goals. DVH parameters of BT included D^3_{2cm}, $D^3_{0.1cm}$, D_{10}, and D_{30}, representing the minimum dose to maximum irradiated tissue volumes of 2 cm^3, 0.1 cm^3, 10 %, and 30 %, respectively. Dose constraints included rectum $D^3_{2cm} \leq 75$ Gy and urethra $D^3_{0.1cm} \leq 120$ Gy, $D_{10} \leq 120$ Gy, and $D_{30} \leq 105$ Gy, calculated using the EQD2 formula for the entire radiotherapy regimen [20].

Dataset processing

The datasets used in this study were selected from the twenty-five patients' 3D CT images for whom interstitial HDR-BT plastic needles were inserted. Patient's images and corresponding radiotherapy structures files (RT structures files), which comprise ground truth labels, were exported from the treatment planning system. An experienced medical physicist checked and manually verified these labels to ensure that all the needles were appropriately segmented and there are no missing contours among different slices. Thus, the segmentation reflects a single operator's input, which reduces inter-observer variability. However, since manual needle reconstruction is inherently subjective, minor variability may exist. The same expert conducted all verifications and consistent labeling practices to mitigate potential uncertainties.

Needle masks were extracted from RT structure files and converted into binary masks. The dataset includes 25 patients (5467 slices) with a slice thickness of 2 mm and pixel spacing of 0.84 mm. The average number of needles in the training and validation sets is 17.35, and 16.4 for the test set. The number of slices per test case patient is 18, 16, 15, 17, and 16. After filtering out non-informative slices, 17 patients (2041

slices) were used for training, 3 patients (391 slices) for validation, and 5 patients (592 slices) for testing.

To address overfitting due to the limited dataset, various data augmentation techniques were applied. These included random rotations (+7° to + 15°, +20° to + 30°, −7° to −15°, −20° to −30°) around the x- and y-axes, vertical and horizontal flips, random translations (+7 to + 15, +20 to + 30, −7 to −15, −20 to −30) along the x- and y-axes, and two rounds of cropping. These augmentations expanded the training dataset 22-fold, resulting in 44,902 augmented slices for training.

Proposed model

The segmentation algorithm was an image-to-image translation algorithm that learns the mapping from input to output domain [21]. Pix2Pix conditional GANs (cGAN) consists of two networks called Generator and Discriminator, in which the Generator and Discriminator are conditioned on additional information [22]. A Deep Tracker Network called GOTURN is presented to estimate the catheters' trajectories from the 3D CT images (Appendix 1). An overview of the BT needles segmentation and the general framework of the proposed method is given in Fig. 1. The proposed method comprises preprocessing, segmentation, and tracking, which will be discussed thoroughly in the following sections.

Preprocessing

Before feeding the dataset into the model, a series of preprocessing has been applied to the CT images. The intensities of the images were normalized between −1 and 1. A region of interest (ROI) selection has been applied to the images to crop the location of needles in CT images. It means that the original 512 × 512 image slices were cropped to focus specifically on the central region, as this area typically contained the most relevant information. By reducing the image dimensions to 256 × 256, the ROI was effectively isolated, which helped to minimize irrelevant background data. This refinement did not only reduce computational complexity but also enhanced the model's focus on critical anatomical structures, improving segmentation accuracy and overall model performance. Finally, this cropped area fed to the model with a 256 × 256 matrix size.

Needle segmentation

The Generator maps CT images to their corresponding segmentation masks, focusing on needle masks in each slice. Its architecture is based on a modified U-Net (mU-Net), a widely used model for segmentation tasks, consisting of an Encoder and a Decoder with skip connections between corresponding blocks [23]. The Encoder extracts multi-level features, while the Decoder reconstructs target images, such as segmentation masks, from these features.

To improve performance, the original U-Net was modified in several ways:

Elimination of Pooling Layers: Pooling layers, which can lose spatial information, were replaced with convolutional layers for downsampling.
Normalization Layers: Instance Normalization was introduced to enhance stability and performance.
Activation Functions: LeakyReLU replaced ReLU in the Encoder to address the dying ReLU problem, as it maintains a non-zero slope for negative inputs, improving feature learning. The Decoder retains ReLU for its computational efficiency.

The network architecture is illustrated in Fig. 2.a, with Downsample and Upsample components detailed in Fig. 2.b.

- *Downsampling:* Includes convolutional layers with a stride of two, Instance Normalization, and LeakyReLU activation. Unlike Batch Normalization, Instance Normalization normalizes each channel independently, reducing computational overhead and increasing speed.
- *Upsampling:* Comprises transpose convolution, Instance Normalization, Dropout (0.5 probability), and ReLU activation.

A combination of adversarial loss and L1 distance was employed as the Generator's loss function, inspired by the Pix2Pix framework. The Generator (G) minimizes the gap between generated and original data, while the Discriminator (D) maximizes this gap in a minGmaxD framework, as shown in Equation (1). Here, zzz represents random noise, xxx is the Generator input, and yyy is the desired output. Additionally, L1 distance (Equation (2) ensures the Generator produces outputs closely resembling the desired results while deceiving the Discriminator [24].

$$L(G, D) = E_{x,y}[\log D(x, y)] + E_{x,z}[\log (1 - D(G(x, y)))] \quad (1)$$

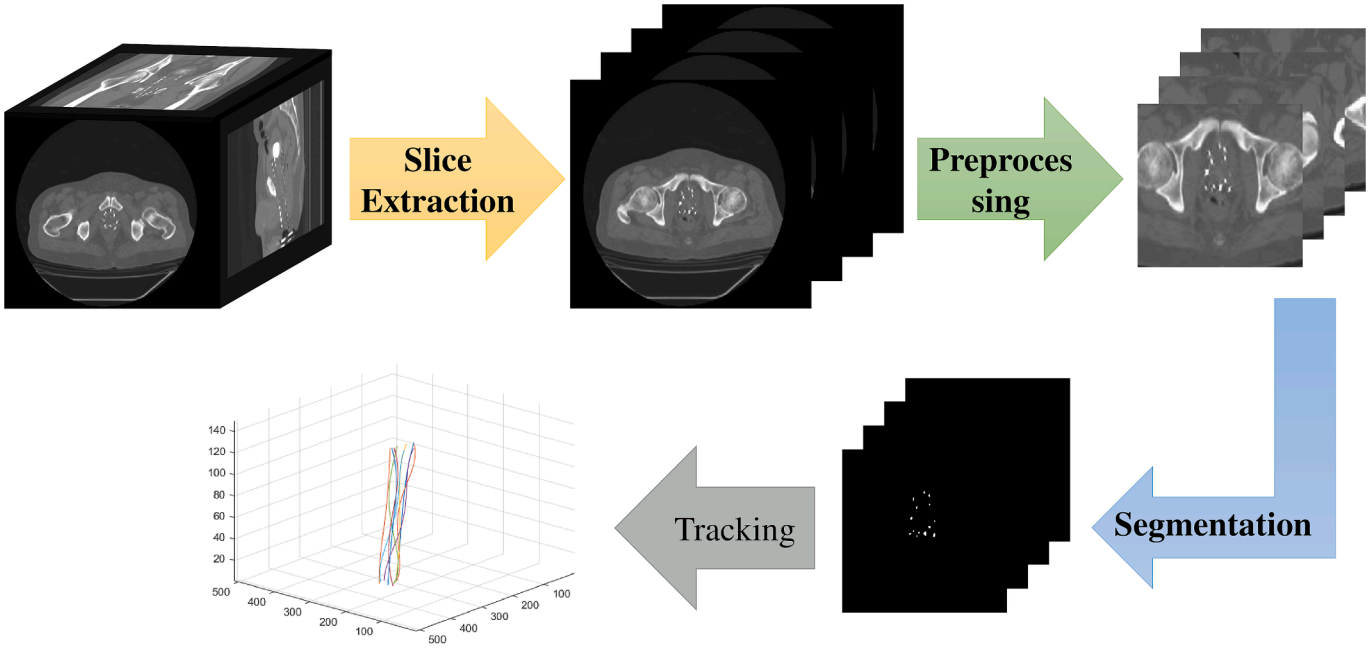

Fig. 1. The general framework of the proposed method.

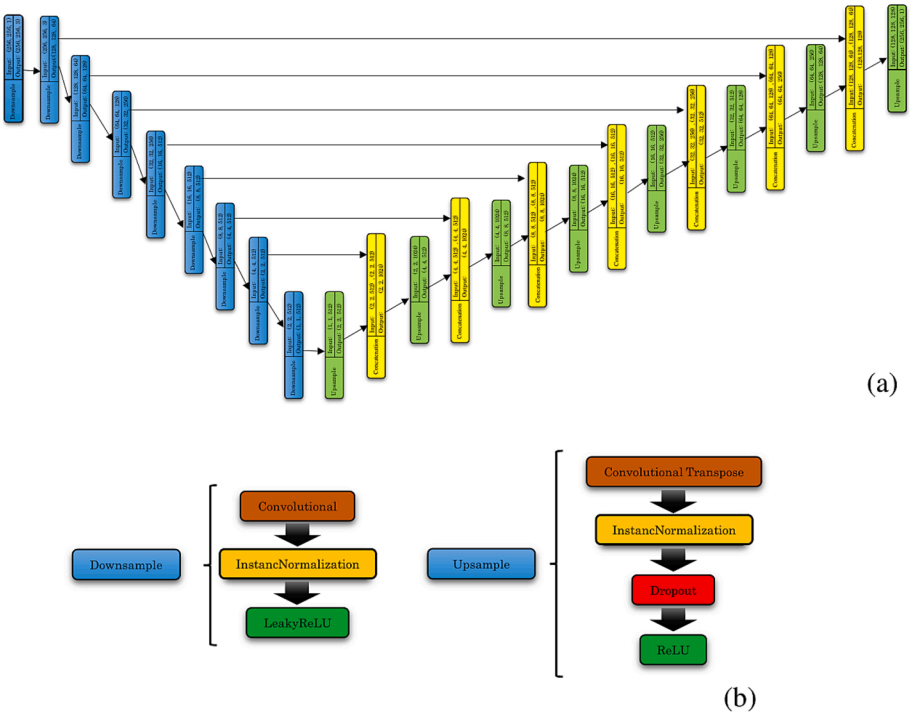

Fig. 2. Generator structure, (a) Block Diagram of the Generator, (b) The context of the Downsample and Upsample in the Generator of the Pix2Pix.

$$L_{L1}(G) = \mathrm{E}_{x,y,z}[\|y - G(x,z)\|] \tag{2}$$

The Discriminator employs a PatchGAN architecture, which evaluates images in localized sections rather than as a whole. By dividing an image into 70 × 70 blocks, the Discriminator focuses on local structures, producing high-frequency results and requiring fewer parameters than traditional methods. Each block is analyzed through convolutional layers, and the output is a 30 × 30 array. The final decision on the entire image is made by averaging these arrays. This localized approach enhances the detection of fine details.

The Discriminator's architecture, illustrated in Fig. 3, consists of three Downsample blocks, two Zero-Padding layers, LeakyReLU activation functions, and Convolutional layers. Its cost function is Binary Cross-Entropy, which differentiates between generated and ground-truth images.

The proposed model was trained for 300 epochs using the Adam optimizer, which adapts the estimation of first and second-order moments for efficient gradient descent. Early stopping was implemented to prevent overfitting, with training halting if validation loss failed to improve by at least 0.001 over ten epochs. The initial learning rate was set to 0.001, and the Plateau algorithm further optimized training by adjusting learning rates dynamically to enhance convergence.

Trajectory prediction

Following needle segmentation in CT slices, the GOTURN tracker was employed to predict the path of each BT needle. GOTURN is a fast, feed-forward tracker based on image comparison. It does not require online training and learns a general relationship between an object's motion and appearance. Using two successive frames, GOTURN determines the target object's position in the current frame, as shown in Fig. 4.

The network comprises two identical convolutional subnetworks, each processing a frame to extract features. The search region is restricted to a specific area around the object's previous location,

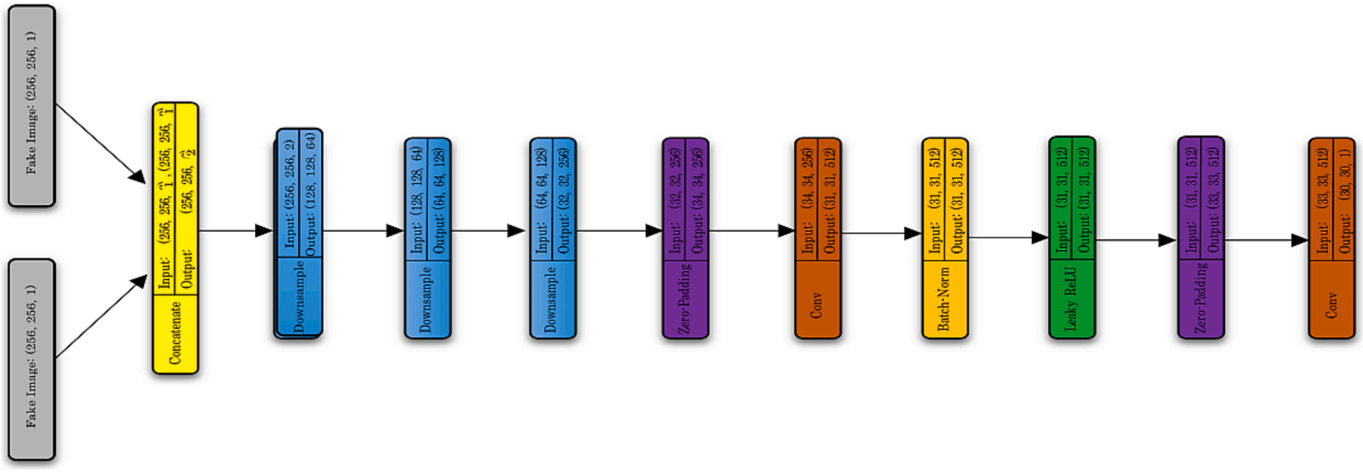

Fig. 3. The Pix2Pix Discriminator architecture.

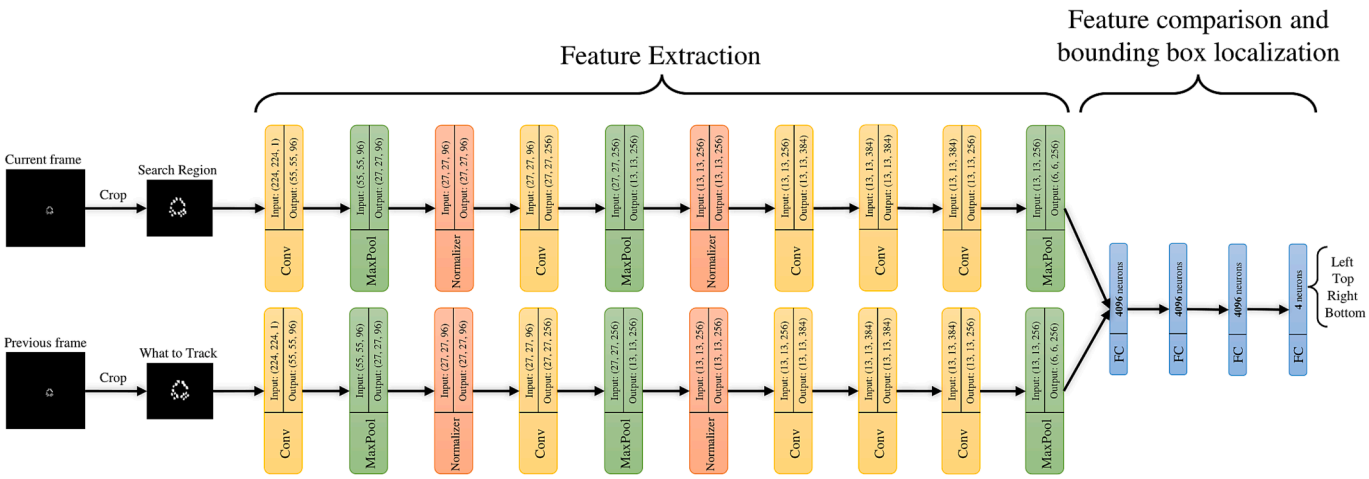

Fig. 4. The GOTURN architecture.

defined as twice the width and height of the object's prior bounding box. This approach leverages the object's smooth motion, accelerates the tracking process, and enhances performance.

GOTURN's architecture includes five convolutional layers and three pooling layers, based on the first five layers of CaffeNet [25]. The extracted features from the current and previous frames are combined and processed through fully connected layers. Each fully connected layer contains 4096 neurons, except the final layer, which has four neurons representing the bounding box coordinates of the target object [26].

Evaluation metrics

The proposed deep-learning-based approach is evaluated using a series of Metrics. These evaluation metrics included Dice Similarity Coefficient (DSC), Intersection over Union (IoU), F1-score, recall, precision, and shaft localization error which are used to assess the proposed method. These metrics are examined in the following.

DSC is a statistical metric that evaluates the similarity between the segmented image and ground truth. DSC is twice the overlap area of the images divided by the total number of pixels in each target and segmented surface (Equation (3)):

$$DSC = \frac{2 \times |target \cap predict|}{2 \times |target \cap predict| + |target - predict| + |predict - target|} \quad (3)$$

IoU, also known as the Jaccard Index, measures the percentage of the segmented image's overlap and its equivalent ground truth. This metric calculates the number of common pixels in segmented and ground truth images divided by common and uncommon pixels. IoU is presented below (i.e., Equation (4):

$$IoU = \frac{|target \cap predict|}{|target \cap predict| + |target - predict| + |predict - target|} \quad (4)$$

Precision, recall, and F1 score are the other quantitative metrics used to evaluate the algorithm's performance regarding incorrect localizations and missed localizations. Their calculated as follow:

$$precision = \frac{\|\{Generated\ image\} \cap \{Real\ image\}\|}{|\{Generated\ image\}|} \quad (5)$$

$$recall = \frac{\|\{Generated\ image\} \cap \{Real\ image\}\|}{|\{Real\ image\}|} \quad (6)$$

$$F1-score = 2 \times \frac{precision \times recall}{precision + recall} \quad (7)$$

Shaft Localization Error evaluates the predicted position of the needles in each cross-sectional CT image. In the needle shaft localization error, the contour for each segmented needle is elicited, and in the next step, their centers were calculated. This metric is defined in Equation (8). In this equation, M is the total number of segmented needle centers, o_i is the actual center, and t_i is the predicted center of the desired needle.

$$Error_{shaft} = \frac{1}{M}\sum_{i=1}^{M}\|o_i - t_i\| \quad (8)$$

In order to have a better illustration of the performance of the tracker model and visual comparison with the actual routes, the results were presented in the polar r-θ instead of the Cartesian x-y system. Data are offered in two forms: r-z and θ-z curves in which z is slice number. The relation between the pixel coordinates in these two systems was given through Equations (9) and (10) as follow:

$$r = \sqrt{x^2 + y^2} \quad (9)$$

$$\theta = \arctan\frac{x}{y} \quad (10)$$

Results

Segmentation performance

As mentioned earlier, 5 patients (592 slices) have been selected for testing to evaluate the performance of the proposed method. The average DSC and IoU of our model for these cases are equal to 0.95 and 0.90, respectively. Furthermore, the mean F1-score, Recall, and Precision were 0.95, 0.93, and 0.97, respectively. Additionally, among the 8764 needle locations in the CT slices, only 112 of them have been missed in segmentation.

The results of four test data samples have shown in Fig. 5.a. in axial view (images have zoomed in for better visualization). In this figure, the first, second, and third rows are the model's input, its corresponding ground truth, and the output of the proposed model, respectively. The difference between the ground truth and the model's segmentation results has been presented in the last row. Moreover, in Fig. 5.b., one of the few missing catheter's locations has been presented. The DSC of the test samples (2D slices) in Fig. 5.a. are 0.91, 0.97, 0.96, and 0.91, respectively. Additionally, the DSC of Fig. 5.b. that one miss occurs in is equal to 0.90.

Tracking performance

The results of the GOTURN model are shown in Fig. 6. Among the 592 test images, 20 consecutive slices through the z-axis (i.e.,

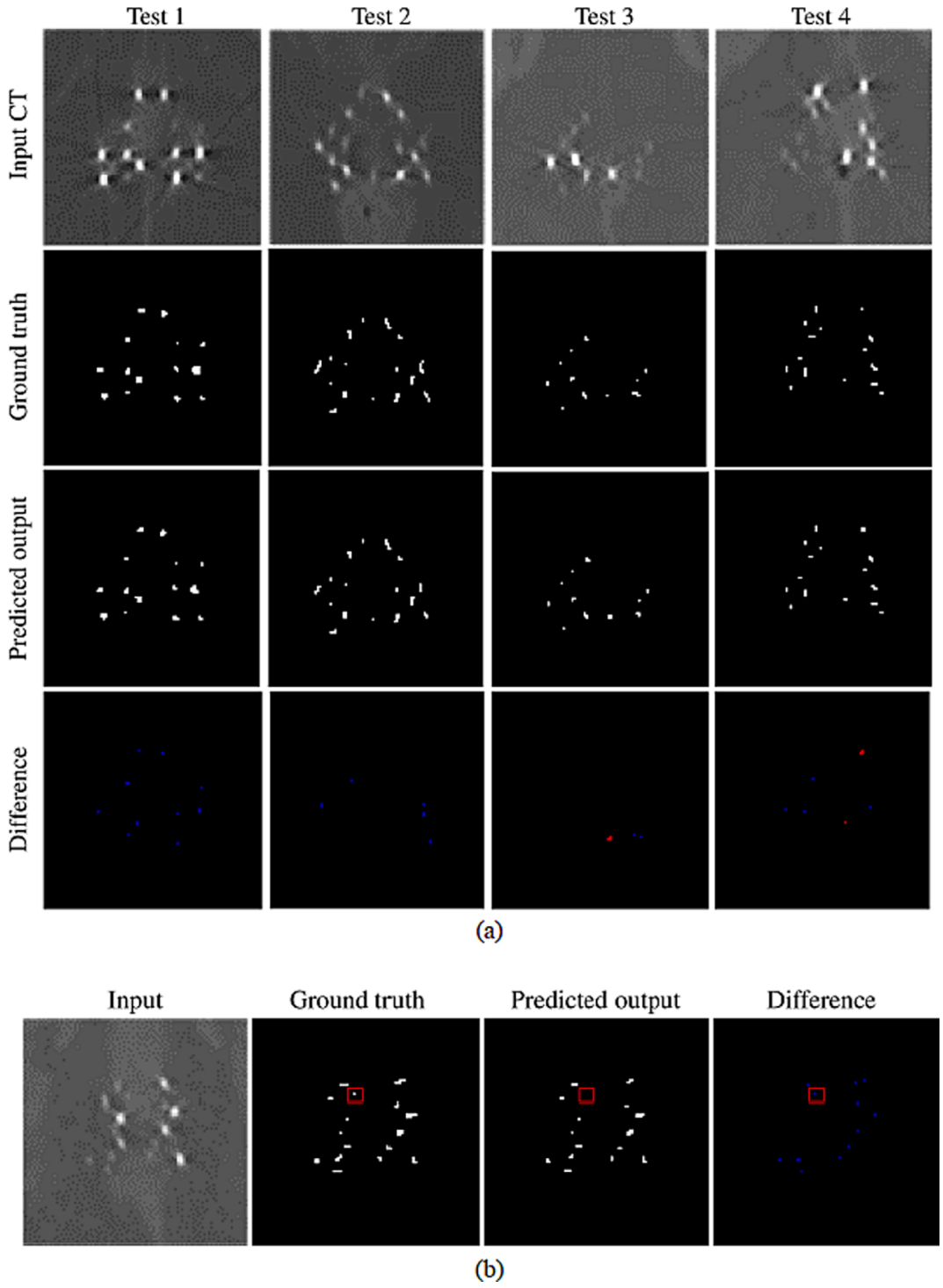

Fig. 5. (a) Needles localization estimation results of a single patient's four different test CT slices. The red and blue colors indicate the regions in a needle that are more or less segmented relative to its ground truths segmentation, respectively. (b) One sample slice in which miss localization has occurred.

craniocaudal direction) are selected for visualization of the tracker network output. There are nine catheters in this series of slices, and the tracker must be capable of following the path of each separately. This figure shows the predicted path for these catheters and the true path for those 20 consecutive slices. In this figure, the stars represent the center of the catheter number i, in the slice number j, and the continuous lines indicate the predicted path of that catheter by a tracker. The shaft localization error of the catheters equals 0.41 mm by this model, which is acceptable considering the amount of available data. Our model completes the entire process in just 16.16 s on the Nvidia K80/T4, broken down as follows: 0.09 s for preprocessing, 12.95 s for segmentation, and 3.12 s for tracking.

Comparative evaluation

In order to evaluate the performance of our proposed model in contouring the catheters of a 3D CT image, the results of a specific patient have shown in the 3D view in Fig. 7. In this figure, the ground truth reconstructed catheters of the image, and their fusion has presented, that as can be seen, they are well matched. Eventually, obtained results

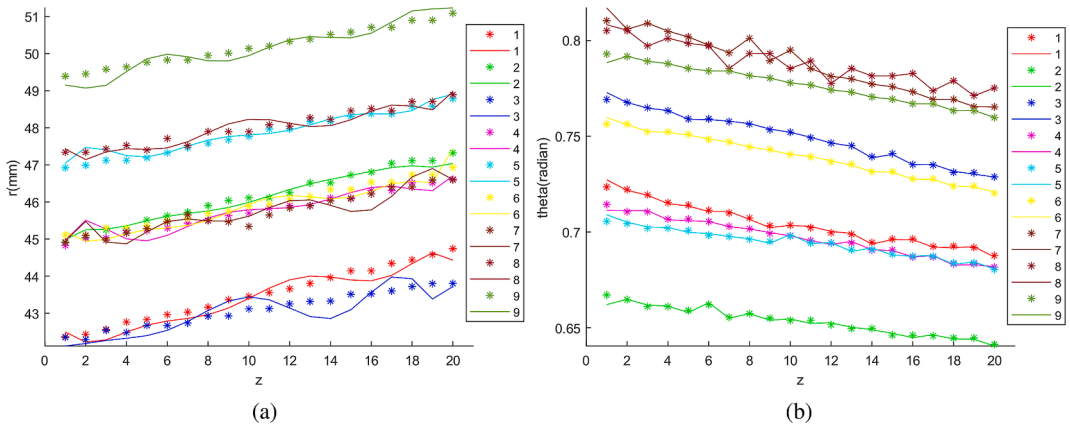

Fig. 6. GOTURN results in which continuous lines are the needle's predicted route and the stars represent the true locations of each needle in each slice, (a) r-z plane, (b) θ-z plane.

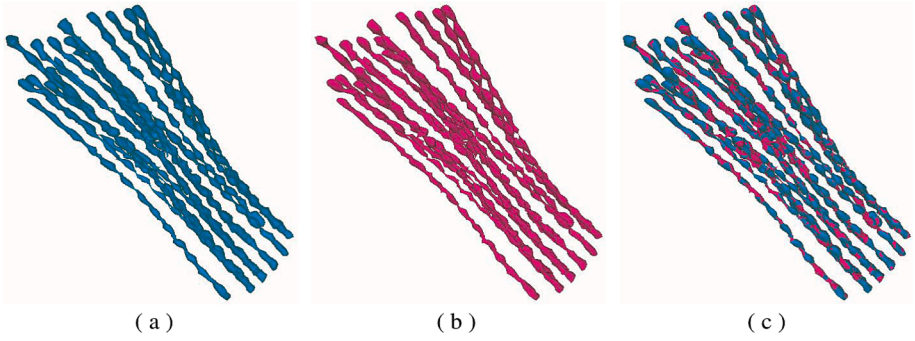

Fig. 7. Multiple catheters in 3D view. (a) Ground Truth, (b) reconstructed needle, (c) Their fusion.

for the shaft error, tip error, and DSC compared to the [17], which are the two most recent studies in this field on CT images, are presented in Table 1. Meanwhile, for better assessment, the model results for each patient in the test set in terms of DSC, IoU, Precision, Recall, and F1-score have been presented in Fig. 8. The results of our model for shaft error and tip error of each case have been presented in Fig. 9.

Discussion

To have an optimal HDR prostate BT procedure, precise and accurate delineation of the position of the needles is crucial. Generally, interstitial needle localization can be accomplished using ultrasound, MRI, and CT images. Due to lower SNR, speckles, and artifacts, localization by transrectal ultrasound (TRUS) images is challenging. MRI images are supreme in terms of soft-tissue contrast, which boosts the visualization of the prostate. We proposed a method based on catheter reconstruction using CT images which leads to optimal geometrical precision in needle positions reconstruction [28]. Providing electron density information for model-based treatment planning systems, higher availability, and lower cost of CT compared to MRI are the main reasons for its popularity for imaging to post-implant treatment planning. Therefore, automatization of applicator localization through CT images is reasonable. This issue is more beneficial for developing countries that mostly face higher workloads BT departments due to a lack of high-tech radiotherapy techniques such as IMRT/VMAT or SBRT. Commonly catheter digitization procedure manually is both overwhelming and error-prone.

Table 1
Obtained results from two recent similar along with our model.

Publications	DSC	Shaft error (mm)	Tip error (mm)
Study [27]	0.89	0.50	0.80
Study [17]	0.93	–	0.63
Our model	0.94	0.41	0.72

An automatic approach for the segmentation and localization of the CT images can be significantly beneficial. This can be affordable since it takes approximately 10–20 min for an experienced medical physicist to manually carry out these tasks in HDR prostate BT, while our model completes the entire process in just 16.16 s. The current study aimed to introduce and test a two-phase deep-learning-based approach for the automatic interstitial plastic needle of CT-based HDR prostate BT.

Typically, clinicians put fiducial markers around the tumor targets to employ them as a reference for target description. The intensity of the needles and these markers are similar in the CT images. Therefore, straightforward segmentation methods like thresholding cannot discern between those markers and desired needles, leading to inaccurate needle localization and wrong trajectories.

In this study, a new two-phase deep learning-based model has been proposed to automatically segment and reconstruct BT needles in patients' CTs, which has achieved remarkable performance on the needle segmentation, shaft error, and tip error. Digitization of the needles is carried out through the Pix2Pix GAN network, in which the segmented images are generated using a U-Net-based model. Another model is used to improve the segmentation process performance alongside the U-Net, which is based on PatchGAN. This network competes with the U-Net for achieving better results in the generated images. Additionally, given that there is a limited number of patients, the robustness of the model is improved by employing Data Augmentation. This model is able to bring down the treatment planning time severely. The computation time can even be more diminished by using a further efficient network or implementing methods like wight pruning. This saving time can lead to accelerating the procedure of clinical workflow. Eventually, the path of

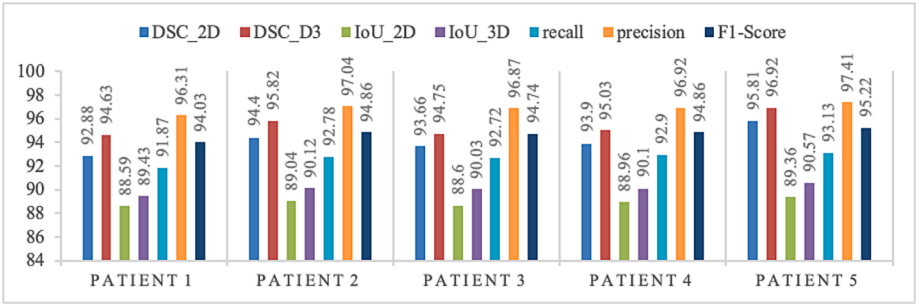

Fig. 8. The results of DSC, and IoU in 2D and 3D space, Recall, Precision, and F1-Score on the test dataset.

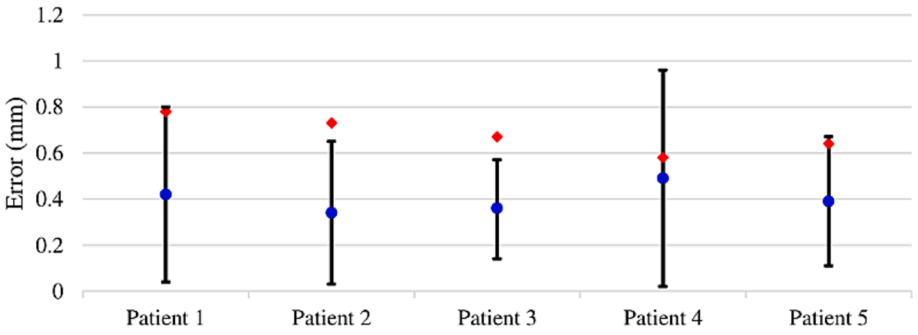

Fig. 9. Mean of tip (Blue circles) and shaft (red diamonds) errors and the test sets standard deviations.

the catheters is obtained using a deep neural network named GOTURN, in which the trajectory is calculated by comparing the current and previous frames.

In contrast to most previous studies that have employed U-Net for segmentation, Pix2Pix GAN has been selected and implemented in this study. One of the most challenging issues in medical applications is the problem of unbalanced pixel categories and small size of segmentation that U-Net cannot tackle with this issue perfectly. Furthermore, Pix2Pix GAN updates its weights through two paths; the first uses skip connections similar to the U-Net, and the other is through the external connection between Generator and Discriminator. On the other hand, the GANs contain an additional network called Discriminator based on the PatchGAN, as mentioned earlier. It divides the input into some portions and determines whether they are actual or counterfeit individually. This gives rise to improving the resolution and quality of the generated data and the performance of the Generator. This approach leads the generator and discriminator's total loss to diminish from 0.31 to 0.11 and 0.23 to 0.14, respectively.

Generally, our model can detect 98.72 % of needle localization in each slice with the shaft error, tip error, and DSC equal to 0.41 mm, 0.68 mm, and 0.95, respectively. The precision and recall metrics are an indicator of the number of false positives and negatives of our model, which are equal to 0.97 and 0.93. Consequently, as can be seen in Fig. 1, the number of false positives is less than the number of false negatives. Moreover, according to Fig. 5.b., it can be figured out that our results show that missed localizations occur mainly for needles with low intensity in the CT image, which may be improved by redefining the CT protocol. As shown in Fig. 6, the tracker network can track each catheter's path well in different slices, and even dense areas have not caused the target to be missed or track the incorrect ones.

There are two recent studies for auto-segmenting of rigid gynecological tandem and ovoid intracavitary applicators [17,27]. Segmentation of these applicators through patients' CT images was performed using the U-Net architecture [17]. The input of this network comprises three successive CT slices to the 3D shape information that would be captured. The segmented voxels that are the U-Net output are clustered using deep learning methods and divided into different categories for each needle. Then a smooth curve is defined to place the voxels in each cluster as the needle trajectory [17]. This model has attained the DSC, HD95%, and tip error of 0.93, 0.71 mm, and 0.63 mm, respectively, under circumstances that its amount of data through Data Augmentation has reached 95000. In another study carried out by [15], segmentation is also done by U-Net architecture, and clustering is performed using a density-based linkage clustering algorithm. For each channel, the average coordinates of all points in a single slice were calculated. Therefore, the needles' central points and trajectory were obtained [15]. This method has attained the DSC, HD95%, shaft error, and tip error of 0.89, 1.66 mm, 0.50 mm, and 0.80 mm, respectively, while it approximately consists of 6100 slices. However, as mentioned earlier, our model comprises only 2041 data for training, and using Data Augmentation, it was converted to 44,902 data. Therefore, having a smaller population size, facing non-rigid flexible needles with smaller diameters, and a higher number of objects to detect the current model performance is promising compared to the previously published models. On the other hand, according to Table 1, our proposed model may potentially achieve better results by accessing additional databases.

In another study, Weishaupt et al. focused on automating the segmentation and localization of treatment applicators for HDR-BT in prostate cancer using different approaches [29]. They introduced a segmentation approach by reformatting CT images into sagittal slices and applying a 2D U-Net architecture, followed by a density-based clustering algorithm to classify individual catheters in 3D. Their study, focusing on titanium needles, achieved submillimeter accuracy with shaft errors of 0.13 ± 0.09 mm and tip errors within -0.1 ± 0.6 mm. Their method also featured a polynomial fitting algorithm to resolve intersecting needles, achieving a segmentation time of approximately one minute. The use of sagittal slices provided geometric continuity along needle paths, potentially improving segmentation, especially for rigid applicators. In comparison, our study utilizes axial CT images, which align with standard clinical workflows and are more widely compatible with treatment planning systems. We employ a Pix2Pix GAN for segmentation, which offers advantages over U-Net in handling small

objects and unbalanced pixel categories, followed by the GOTURN tracker to predict needle trajectories across slices. Despite working with non-rigid plastic needles, which present additional segmentation challenges, our model achieves competitive performance with a shaft error of 0.41 mm and tip error of 0.68 mm. Notably, our approach completes segmentation and tracking in 16.16 s, making it suitable for real-time clinical applications. While both methods address distinct challenges, our approach focuses on speed and robustness, complementing Weishaupt et al.'s emphasis on geometric precision [29].

We acknowledge the importance of benchmarking our proposed architecture against widely used models like nnUNet to further validate its performance and complexity. However, due to constraints in time and computational resources, such direct comparisons could not be conducted as part of this study. Despite this limitation, our proposed method demonstrates significant advantages in handling unbalanced datasets, as reflected in the performance metrics reported. The design of our architecture specifically addresses challenges inherent to our dataset, such as severe class imbalance or noisy annotations, making it well-suited for this application. We anticipate that applying nnUNet or similar baseline models in future studies will provide a more comprehensive evaluation of the effectiveness of our approach and further support our findings.

Automatic BT applicator reconstruction is an exciting area in this treatment domain due to its practical applications and impacts on the accuracy of patient treatment. However, there are some limitations during this pilot study. One of these limitations is the amount of data used despite the satisfying results. Increasing the amount of data will lead to the reliability of the proposed model. By training the model with a variety of cases and testing its performance with more data from several BTs, the validity of the proposed model will increase, and it can be used more reliably in clinical applications. Besides the reliability and robustness, this can make the model able to deal with challenging situations more efficiently. As shown in Fig. 5.b., in our proposed model, some needles are missed through automatic reconstruction. Even though only 1.28 % misses, these pitfalls should be carefully considered and avoided during a real clinical scenario. Meanwhile, this number of misses can demonstrate that our model consists of few false-negative and should have an acceptable recall parameter. Therefore, every automatic reconstruction should be double-checked through the CT slices. However, it is expected that fewer misses occurred by feeding the model with a higher number and verities of training sets, although this number of misses (112) due to the number of our available data is acceptable compared with some previously published research.

Funding details

None.

CRediT authorship contribution statement

Mohammad Mahdi Moradi: Conceptualization, Writing – original draft, Validation. **Zahra Siavashpour:** Conceptualization, Writing – original draft, Validation, Writing – review & editing. **Soheib Takhtardeshir:** Visualization, Investigation. **Eman Showkatian:** Visualization, Investigation. **Ramin Jaberi:** Supervision. **Reza Ghaderi:** Visualization, Investigation. **Bahram Mofid:** Supervision. **Farzad Taghizadeh-Hesary:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

None.

Informed consent

Not applicable.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ctro.2025.100925.

References

[1] Sung H. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, vol. 71, no. 3. in CA: a Cancer Journal for Clinicians, vol. 71. pp. 209–249, 2021.
[2] Dutta SW. Prostate cancer high dose-rate brachytherapy: review of evidence and current perspectives. *Expert Rev Med Devices* 2018;15(1):71–9.
[3] Pasalic D, et al. Dose escalation for prostate adenocarcinoma: a long-term update on the outcomes of a phase 3, single institution randomized clinical trial. Int J Radiat Oncol*Biol*Phys 2019;104(4):790–7. https://doi.org/10.1016/j.ijrobp.2019.02.045.
[4] Vuolukka K. Stereotactic body radiotherapy for localized prostate cancer–5-year efficacy results. *Radiat Oncol* 2020;15(1):1–8.
[5] Yamada Y. American Brachytherapy Society consensus guidelines for high-dose-rate prostate brachytherapy. *Brachytherapy* 2012;11(1):20–32.
[6] G. Corey, A. B. M. Yoosuf, G. Workman, M. Byrne, D. M. Mitchell, and S. Jain, UK & Ireland Prostate Brachytherapy Practice Survey 2014-2016, *jcb*, vol. 10, no. 3, pp. 238–245, 2018, doi: 10.5114/jcb.2018.76839.
[7] Orlando N, et al. A power Doppler ultrasound method for improving intraoperative tip localization for visually obstructed needles in interstitial prostate brachytherapy. *Med Phys* 2023;50(5):2649–61. https://doi.org/10.1002/mp.16336.
[8] Fischer-Valuck BW. A brief review of low-dose rate (LDR) and high-dose rate (HDR) brachytherapy boost for high-risk prostate. *Front Oncol* 2019;9:1378.
[9] Gray C, Campbell K. High dose rate brachytherapy versus low dose rate brachytherapy for the treatment of prostate cancer: a review of clinical effectiveness and cost-effectiveness; 2019.
[10] Peach MS, Trifiletti DM, Libby B. Systematic review of focal prostate brachytherapy and the future implementation of image-guided prostate HDR brachytherapy using MR-ultrasound fusion. *Prostate Cancer* 2016;2016:1–13. https://doi.org/10.1155/2016/4754031.
[11] Kirisits C. Review of clinical brachytherapy uncertainties: analysis guidelines of GEC-ESTRO and the AAPM. *Radiother Oncol* 2014;110(1):199–212.
[12] Liang F. Reconstruction of brachytherapy catheters and needles using EM sensor-based navigation system, in *Frontiers in Biomedical Devices. 2017*, American Society of Mechanical Engineers.
[13] Chan MF, Witztum A, Valdes G. Integration of AI and machine learning in radiotherapy QA. *Front Artif Intell* 2020;3:76.
[14] Currie G. Machine learning and deep learning in medical imaging: intelligent imaging. *J Med Imag Radiat Sci* 2019;50(4):477–87.
[15] Hu H. A review of the application of deep learning in brachytherapy. *Open Access Library J* 2020;7(07):1.
[16] Nicolae A. Evaluation of a machine-learning algorithm for treatment planning in prostate low-dose-rate brachytherapy. *Int J Radiat Oncol* Biol* Phys* 2017;97(4):822–9.
[17] Jung H. Deep-learning–assisted automatic digitization of applicators in 3D CT image-based high-dose-rate brachytherapy of gynecological cancer. *Brachytherapy* 2019;18(6):841–51.
[18] Zaffino P. Fully automatic catheter segmentation in MRI with 3D convolutional neural networks: application to MRI-guided gynecologic brachytherapy. *Phys Med Biol* 2019;64(16):165008.
[19] Orlando N. Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images. *Med Phys* 2020;47(6):2413–26.
[20] Crook J, Marbán M, Batchelar D. HDR prostate brachytherapy. *Seminars in Radiation Oncology*. Elsevier; 2020.
[21] Isola P. Image-to-image translation with conditional adversarial networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2017.
[22] Mirza M, Osindero S. Conditional generative adversarial nets; 2014.
[23] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer; 2015.
[24] Lu L. Dying relu and initialization: Theory and numerical examples; 2019.
[25] Qin Z. How convolutional neural network see the world-A survey of convolutional neural network visualization methods; 2018.
[26] Held D, Thrun S, Savarese S. Learning to track at 100 fps with deep regression networks. *European Conference on Computer Vision*. Springer; 2016.

[27] Hu H, et al. Deep learning applications in automatic segmentation and reconstruction in CT-based cervix brachytherapy. *J Contemp Brachytherapy* 2021;13 (3):325–30. https://doi.org/10.5114/jcb.2021.106118.
[28] Hellebust T. Place of modern imaging in brachytherapy planning. *Cancer/ radiothérapie* 2018;22(4):326–33.
[29] Weishaupt LL, et al. Approaching automated applicator digitization from a new angle: Using sagittal images to improve deep learning accuracy and robustness in high-dose-rate prostate brachytherapy. *Brachytherapy* 2022;21(4):520–31. https:// doi.org/10.1016/j.brachy.2022.02.005.