Fully Automatic Reconstruction of Prostate High Dose Rate Brachytherapy Interstitial Needles by Using Two Phases Deep Learning Based Segmentation and Object Tracking Algorithms

Mohammad Mahdi Moradi
Shahid Beheshti University

Zahra Siavashpour (z_siavashpour@sbmu.ac.ir)
Shohada-eTajrish Educational Hospital, Shahid Beheshti University of Medical Science

Soheib Takhtardeshir
Shahid Beheshti University

Eman Showkatian
Iran University of Medical Science

Isaac Shiri
Geneva University Hospital

Ramin Jeberi
Tehran University of Medical Sciences

Reza Ghaderi
Shahid Beheshti University

Article

Keywords:

Posted Date: May 23rd, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1590410/v1

Abstract

The essential step of successful brachytherapy would be precise applicator/needles trajectory detection, which is an open problem yet. This study proposes a two-phase deep learning-based method to automate the localization of high-dose-rate (HDR) prostate brachytherapy catheters through the patient's CT images. The whole process is divided into two phases using two different deep neural networks. First, brachytherapy needles segmentation is 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 brachytherapy dataset. Among the total 25 patients, 5 patients that consisted 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 is 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 brachytherapy 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 2020, the global cancer statistical report[1]). Surgery, systematic therapy, and radiation therapy are the treatment options in terms of prostate cancer management. External beam radiotherapy (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 high-risk prostate malignancies[2]. Therefore, dose escalation is needed to improve biochemical relapse-free survival (bRFS) and overall survival (OS).

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[3, 4]. The radioactive source is placed into the prostate directly or through about 20 ± 5 plastic needles during prostate brachytherapy. 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[4].

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[5, 6].

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]. However, CT knows as the pioneer imaging technique used in 3D treatment planning and is still the most available modality in many brachytherapy centers worldwide.

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 brachytherapy dosimetry[7]. There is some predefined applicators' library 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 of 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 to help their detection[8]. 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[9, 10]. Some valuable publications developed deep learning algorithms for brachytherapy treatment planning, automatic organ segmentation, and applicator digitization[11, 12]. Artificial intelligence and deep learning-based algorithms are proposed to automatically reconstruct different brachytherapy applicators through CT, MRI, and TRUS images[11]. Some artificial intelligence-based algorithms were proposed for 3D CT- and MRI-based gynecological HDR-BT rigid applicator

automatic reconstruction[13, 14]. Deep learning was also successfully used for automatic needle segmentation of TRUS-guided prostate brachytherapy[15], 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 a new two-phase, deep learning-based models, for auto-localization of interstitial HDR prostate brachytherapy 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.

Results

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. 1 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. 2, one of the few missing catheter locations has been presented. The DSC of the test samples (2D slices) in Fig. 1 is 0.91, 0.97, 0.96, and 0.91, respectively. Additionally, the DSC of Fig. 2 that one miss occurs is equal to 0.90.

The results of the GOTURN model are shown in Fig. 3. Among the 592 test images, 20 consecutive slices 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. Figure 3.a. and Fig. 3.b 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 0.4056mm by this model, which is acceptable considering the amount of available data.
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. 4. 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 for the Shaft Error, Tip Error, and DSC compared to the[13], 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. 5. The results of our model for shaft error and tip error for each have been presented in Fig. 6.

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

	DSMC	shaft error (mm)	Tip error (mm)
[23]	0.89	0.50	0.80
[13]	0.93	-	**0.63**
Our model	**0.94**	**0.41**	0.72

Discussion

To have a perfect HDR prostate brachytherapy 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 position reconstruction[16]. 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 in brachytherapy 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.

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 minutes for an experienced physicist to manually carry out these tasks in HDR prostate brachytherapy, while our model is able to do it only in 16.16 sec on the Nvidia K80 / T4 (0.09 sec for preprocessing, 12.95 sec for segmentation, and 3.12 sec for tracking). 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 brachytherapy.

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 in 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 to achieve 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 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 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 Denerator 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 0.41mm, 0.68mm, 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 Figs. 7 and 8, the number of false positives is less than the number of false negatives. Moreover, according to Fig. 2, 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. 3, 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[11, 13] for auto-segmenting of rigid gynecological tandem and ovoid intracavitary applicators. Segmentation of these applicators through patients' CT images was performed using the U-Net architecture[13]. 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[13]. 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[11], 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[11]. This method has attained the DSC, HD95%, shaft error, and tip error of 0.89, 1.66mm, 0.50mm, 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 44902 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.

Automatic brachytherapy 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 were 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 BT, 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. 2, 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 occur 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.

Methods

Database

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

Clinical data of the current study was obtained from the radiotherapy department of Yas hospital, Tehran, Iran. Applicator insertion has been done under regional anesthesia with the patient in the dorsal lithotomy position. Endorectal ultrasound guidance has been utilized.

HDR-BT has been performed as a definitive or salvage monotherapy treatment or EBRT boost in four or two treatment fractions, respectively. Two treatments were delivered with a single needles implant by 6 hours interval while assuring the labeled needles tip's exact position by checking and measuring their remaining length out of the perineal surface.

Dedicated brachytherapy CT markers have been inserted into the plastic needles. A pelvic CT image (HiSpeed Dual Scanner, General Electric Company, Medical Systems Group, USA) with a maximum 2mm slice thickness has been acquired. These images have been imported to the HDR PLUS brachytherapy treatment planning system (Eckert & Ziegler BEBIG, Berlin, Germany), which uses a ^{60}Co source database of the same company.

OARs, including the rectum, urethra, and bladder, have been delineated by an experienced radiation oncologist. Clinical target volumes (CTV) (both low- and high-risk ones) also have been contoured getting insight from the patient MRI especially T2w, diffusion-weighted (DW), and dynamic contrast-enhanced MRI. A brachytherapy medical physicist has done applicator reconstruction by allocating the needle's trajectory from their tip to the ends out of the patient's perinea using TPS related applicator module.

Treatment planning has been done to reach the CTV V_{100} (percentage of the prostate receiving 100% of the prescribed dose) > 95% and D_{90} (dose to 90% of the prostate gland) > 100% of the prescribed dose (PD) as the optimum treatment goals. BT DVH parameters such as $D_{2\mathrm{cm}^3}$, $D_{0.1\mathrm{cm}^3}$, D_{10}, and D_{30}, representing the minimum dose to maximum irradiated tissue volumes of 2cm^3 and 0.1cm^3 and 10% and 30% of tissue volume, respectively. Therefore, considered rectum dose constraints are $D_{2\mathrm{cm}^3} \leq 75$, and urethral dose constraints are $D_{0.1\mathrm{cm}^3} \leq 120$ G_y, $D_{10} \leq 120$ G_y, and $D_{30} \leq 105$ G_y[17].
This study was conducted in accordance with the Declaration of Helsinki. The study protocol was evaluated and approved by the Ethics Committee of the Clinical Research Development Unit of Shohadaye Tajrish Hospital, Shahid Beheshti University of Medical Science (Ethics Committee ID: IR.SBMU.RETECH.REC.1399.527). The need for consent was waived by Ethics Committee. Patients' information was anonymized.

Dataset processing

The datasets used in this study were selected from the twenty-five patients' 3D CT images for whom interstitial HDR brachytherapy 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. A physicist checked all the ground truth labels to ensure that all the needles have been appropriately segmented and there are no missing contours among different slices. Following, needle masks were extracted from RT structures files and converted to binary masks. This database comprises 25 patients (5467 slices) with a slice thickness of 2 mm and pixel spacing of 0.84mm. The average number of needles in the training and validation set is equal to 17.35, and for the testing, the set is 16.4, while the number of slices for each of the test case patients is 18, 16, 15, 17, and 16. After cleaning datasets, among the 25 patients, 17 patients (2041 slices), 3 patients (391 slices), and 5 patients (592 slices) were selected for training, validation, and testing, respectively. Due to the lack of a sufficient available database, different data augmentation techniques have been used to deal with the over-fitting problem. In order to expand the variation of data, different augmentation methods, including random rotating between + 7 to + 15, +20 to + 30, -7 to -15, and − 20 to -30 degrees around the x-axis and y-axis, vertical and horizontal flip, random translation between + 7 to + 15, +20 to + 30, -7 to -15, and − 20

to -30 along the x-axis and y-axis, and two times cropping, have been employed. Eventually, the overall training dataset will be increased 22 times relative to the initial training dataset (i.e., 44902 augmented data for training).

Proposed Model

The segmentation algorithm was an image-to-image translation algorithm that learns the mapping from input to output domain[18]. Pix2Pix Conditional GANs (cGAN) consists of two networks called Generator and Discriminator, in which the Generator and Discriminator are conditioned on additional information[19]. A Deep Tracker Network called GOTURN is presented to estimate the catheters' trajectories from the 3D CT images. An overview of the brachytherapy needles segmentation and routing process is given in Fig. 7. The proposed method comprises pre-processing, 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 to 1. A region of interest (ROI) selection has been applied to the images to crop the location of needles in CT images, as it has been illustrated in Fig. 9 as an example. Finally, this area has been cropped and fed to the model with a 256×256 matrix size, as represented in Figs. 2 and 3.

Needle Segmentation

The Generator is responsible for mapping CT images into their corresponding segmentation masks, containing the needles' mask in each slice. The architecture of the Generator is based on a modified U-Net (mU-Net). U-Net is a well-known model used broadly in segmentation tasks and consists of an Encoder and Decoder in which each block of the Encoder is connected to its corresponding block in Decoder[20]. The Encoder encodes images to the different levels of feature, and the Decoder maps the lower resolutions learned features by the Encoder into the target images such as translated images. Firstly, considering that Pooling layers are lossy and do not preserve all the spatial information, these layers were eliminated, and the Downsampling is done just by convolutional layers. Secondly, normalization layers have been added to the network architecture to improve the model's performance and stability. Contrary to the original U-Net, which used only the ReLU activation function, we used LeakyReLU to deal with the dying ReLU problems, so all neurons become effective in predicting the final segmentation mask[21]. LeakyReLU can solve this problem to some extent since its function does not have a zero slope for negative inputs. Accordingly, both have been used to hold speeds of the ReLU and the ability of the LeakyReLU. Moreover, since the Encoding part extracts the input features, it has considerable importance since the output will be built according to these features. Accordingly, LeakyReLU has been used only in the Encoder. The network architecture has been illustrated in Fig. 10.a. This model consists of two

components named Downsample and Upsample, as mentioned earlier. Contents of each block have been illustrated in Fig. 10.b.
Downsampling includes Convolutional layers with a stride of two, Instance Normalization, and a LeakyReLU activation function. Unlike the Batch Normalization, which normalizes all existing data in a batch, Instance Normalization normalizes each channel in a data in a batch, separately and resulting in fewer computations and more speed. Upsampling consists of Transpose Convolution, Instance Normalization, Dropout with probability ratio equal to 0.5, and ReLU activation function. A combination of adversarial and L1 distance has been used as a loss function for the Generator network in Pix2Pix. Generator G tries to minimize the existing gap between the generated and original data, and Discriminator D tries to maximize this distance so as not to be defeated. This minGmaxD is presented in Eq. 1. In this equation, z is random noise, x is the input of the G, and y is the desired output. Another loss function is the L1 distance, which aims to enable the Generator to produce outputs as close as possible to the desired output besides deceiving the Discriminator. This loss function is presented in Eq. 2.

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

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

The Discriminator is a PatchGAN which instead of determining whether the whole image is real or fake, an image is divided into different blocks and sections, and the Discriminator specifies whether each block is real or fake. Thus, since the Discriminator considers only the local structures of the image, it will produce high-frequency results and require far fewer parameters than concluding on the whole image. In the PatchGAN, any individual image is divided into 70×70 blocks and is passed through a series of convolutional layers, and an array with the size of 30×30 is generated in the output. The Discriminator made its final decision about the whole image by averaging these arrays. The architecture of the Discriminator is shown in Fig. 11. Discriminator comprises three Downsample blocks, two Zero-padding, Leaky ReLU activation function, and Convolutional layers. The Discriminator cost function is a Binary Cross Entropy whose task was to distinguish between generated images and Ground truth.
The proposed model is trained on 300 epochs. A stochastic gradient descent method called 'Adam' is used based on the adaptive estimation of first and second-order moments. Furthermore, an early stopping algorithm is used to prevent the model from overfitting. The initial learning rate was set to 0.001 and the Plateau algorithm is employed to facilitate and accelerate the training phase. According to this algorithm, model training would be stopped if the losses in the experimental set did not improve by at least 0.001 over ten epochs.

Trajectory prediction

Following needles segmentation in CT slices, the GOTURN tracker was used to predict each brachytherapy needle's path. This network is a fast tracker, a simple feed-forward that does not require any online training and learns a general relationship between the object's body motion and appearance. The GOTURN is a tracker algorithm that is based on image comparison. As depicted in Fig. 12, GOTURN contains two equal convolutional networks that use two successive frames to find the target object's

position in the current frame, and each of these frames is fed into a series of Convolutional layers to extract their features. The search region for finding the target object is restricted to a certain area around its previous status since the position of this object in the previous image is determined. It can lead to speed up the tracking process and enhance the performance of the model. Therefore, an area twice the width and height of the object's bounding box in the prior frame will be cropped relative to its center in the previous and current image. This is practical since the objects smoothly move, and it could be around their prior status. As shown in Fig. 12, this model contains five Convolutional and three Pooling layers in two groups which are based on the five initial layers of the CaffeNet[22]. The output of these layers is combined and passed on a series of fully connected layers to compare the obtained features of the current and previous frames. Each of these layers consists of 4096 neurons except for the last one, which comprises four neurons representing the coordinates of the favorable object bounding box[23].

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 (Eq. 3):

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

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., Equations 4):

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

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:

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

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

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

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 Eq. 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.

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

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-θinstead 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:

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

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

Declarations

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of competing interest

The authors declare that they have no known competing financial or personal relationships that could be viewed as influencing the work reported in this paper.

Funding

This work did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

Informed consent and patient details

This study was conducted in accordance with the Declaration of Helsinki. The study protocol was evaluated and approved by the Ethics Committee of the Clinical Research Development Unit of Shohadaye Tajrish Hospital, Shahid Beheshti University of Medical Science (Ethics Committee ID: IR.SBMU.RETECH.REC.1399.527). The need for consent was waived by Ethics Committee. Patients' information was anonymized.

References

1. Sung, H., *et al., Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians*, 2021. *71(3): p. 209–249.*
2. Dutta, S.W., *et al., Prostate cancer high dose-rate brachytherapy: review of evidence and current perspectives. Expert review of medical devices*, 2018. *15(1): p. 71–79.*
3. Vuolukka, K., *et al., Stereotactic body radiotherapy for localized prostate cancer–5-year efficacy results. Radiation Oncology*, 2020. *15(1): p. 1–8.*
4. Yamada, Y., *et al., American Brachytherapy Society consensus guidelines for high-dose-rate prostate brachytherapy. Brachytherapy*, 2012. *11(1): p. 20–32.*
5. Fischer-Valuck, B.W., *et al., A brief review of low-dose rate (LDR) and high-dose rate (HDR) brachytherapy boost for high-risk prostate. Frontiers in oncology*, 2019. *9: p. 1378.*
6. Gray, C. *and K. Campbell, High Dose Rate Brachytherapy versus Low Dose Rate Brachytherapy for the Treatment of Prostate Cancer: A Review of Clinical Effectiveness and Cost-Effectiveness.* 2019.
7. Kirisits, C., *et al., Review of clinical brachytherapy uncertainties: analysis guidelines of GEC-ESTRO and the AAPM. Radiotherapy and oncology*, 2014. *110(1): p. 199–212.*
8. Liang, F., *et al. Reconstruction of Brachytherapy Catheters and Needles Using EM Sensor-Based Navigation System. in Frontiers in Biomedical Devices.* 2017. *American Society of Mechanical Engineers.*
9. Chan, M.F., A. Witztum, *and* G. Valdes, *Integration of AI and machine learning in radiotherapy QA. Frontiers in artificial intelligence, 2020. 3: p. 76.*
10. Currie, G., *et al., Machine learning and deep learning in medical imaging: intelligent imaging. Journal of medical imaging and radiation sciences*, 2019. *50(4): p. 477–487.*
11. Hu, H., *A Review of the Application of Deep Learning in Brachytherapy. Open Access Library Journal*, 2020. *7(07): p. 1.*
12. Nicolae, A., *et al., Evaluation of a machine-learning algorithm for treatment planning in prostate low-dose-rate brachytherapy. International Journal of Radiation Oncology* Biology* Physics*, 2017. *97(4):*

p. 822–829.

13. Jung, H., *et al., Deep-learning–assisted automatic digitization of applicators in 3D CT image-based high-dose-rate brachytherapy of gynecological cancer. Brachytherapy, 2019. 18(6): p. 841–851.*
14. Zaffino, P., *et al., Fully automatic catheter segmentation in MRI with 3D convolutional neural networks: application to MRI-guided gynecologic brachytherapy. Physics in Medicine & Biology*, 2019. *64(16): p. 165008.*
15. Orlando, N., *et al., Automatic prostate segmentation using deep learning on clinically diverse 3D transrectal ultrasound images. Medical physics*, 2020. *47(6): p. 2413–2426.*
16. Hellebust, T., *Place of modern imaging in brachytherapy planning. Cancer/Radiothérapie*, 2018. *22(4): p. 326–333.*
17. *Crook, J., M. Marbán, and D. Batchelar. HDR prostate brachytherapy. in Seminars in radiation oncology. 2020. Elsevier.*
18. Isola, P., *et al. Image-to-image translation with conditional adversarial networks. in Proceedings of the IEEE conference on computer vision and pattern recognition.* 2017.
19. *Mirza, M. and S. Osindero, Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784, 2014.*
20. Ronneberger, O., P. Fischer, *and T. Brox. U-net: Convolutional networks for biomedical image segmentation. in International Conference on Medical image computing and computer-assisted intervention.* 2015. *Springer.*
21. Lu, L., *et al., Dying relu and initialization: Theory and numerical examples. arXiv preprint arXiv:1903.06733, 2019.*
22. Qin, Z., *et al., How convolutional neural network see the world-A survey of convolutional neural network visualization methods. arXiv preprint arXiv:1804.11191*, 2018.
23. Held, D., S. Thrun, *and S. Savarese. Learning to track at 100 fps with deep regression networks. in European conference on computer vision.* 2016. *Springer.*

Figures

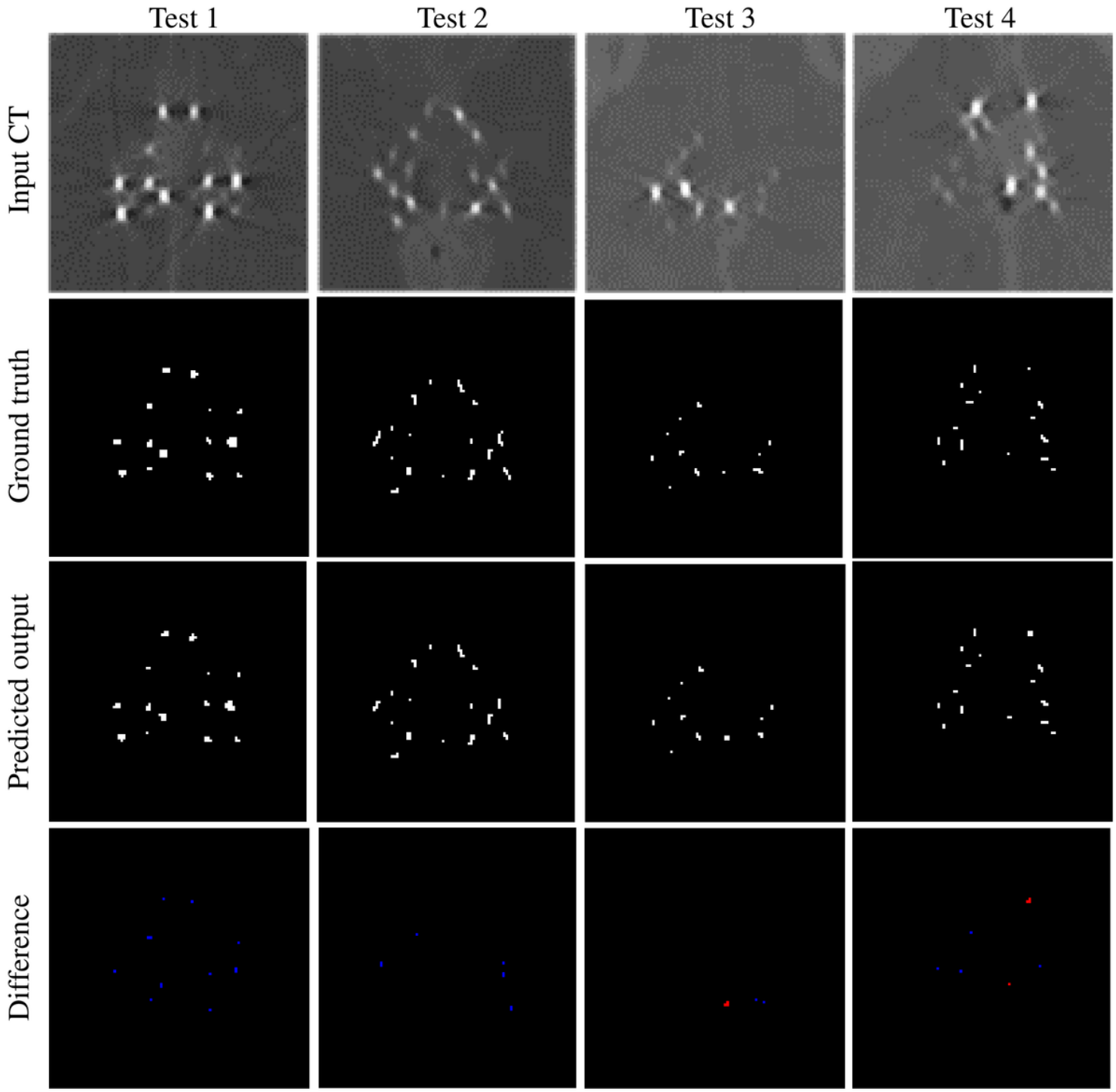

Figure 1

Needles localization estimation results of a single patient's five different test CT slices. In this figure, the red and blue color indicates the regions in a needle that are more or less segmented relative to its ground truths segmentation, respectively.

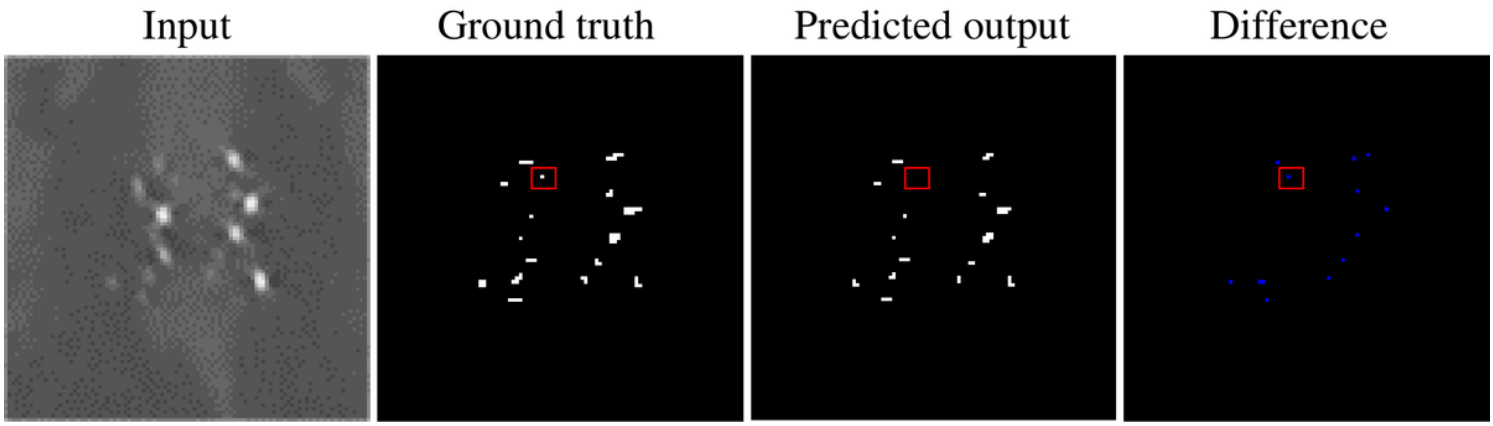

Figure 2

One sample slice in which miss localization has occurred.

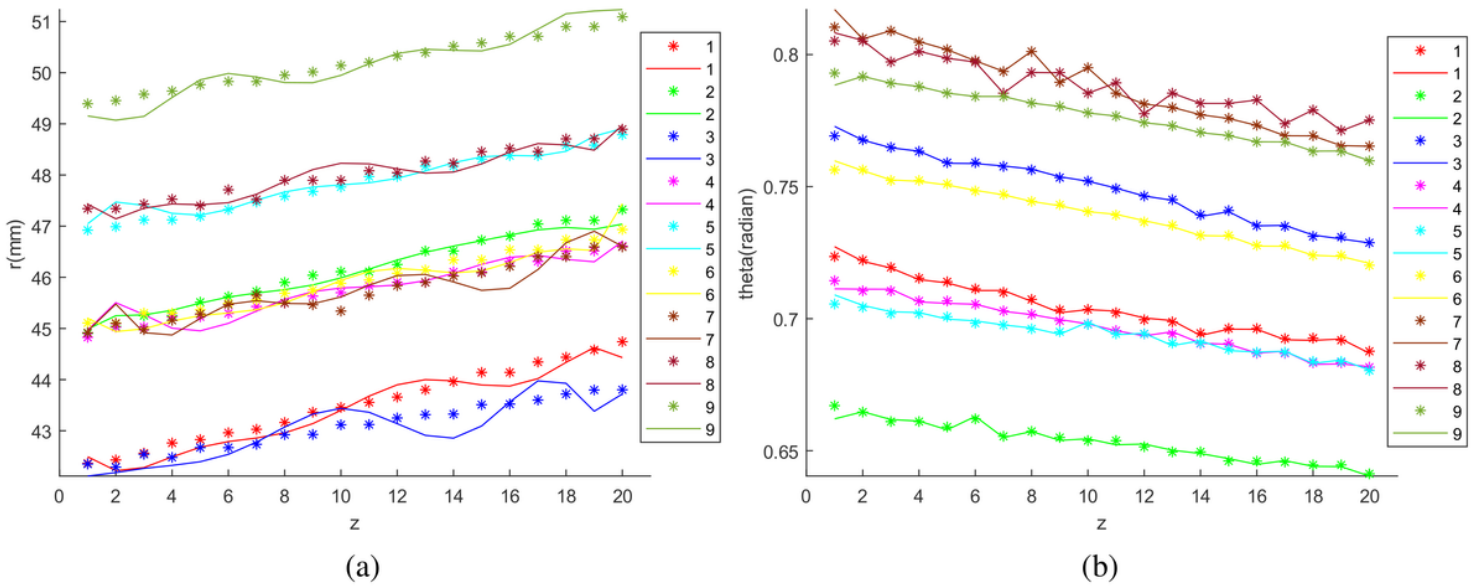

Figure 3

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) -z plane.

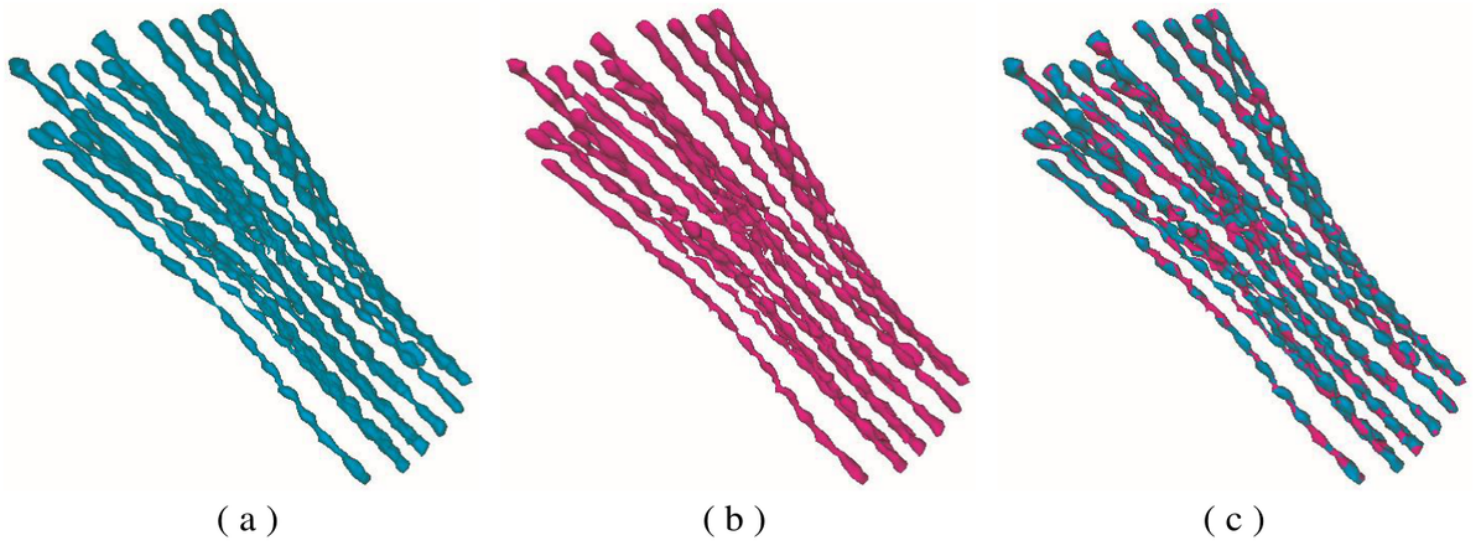

Figure 4

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

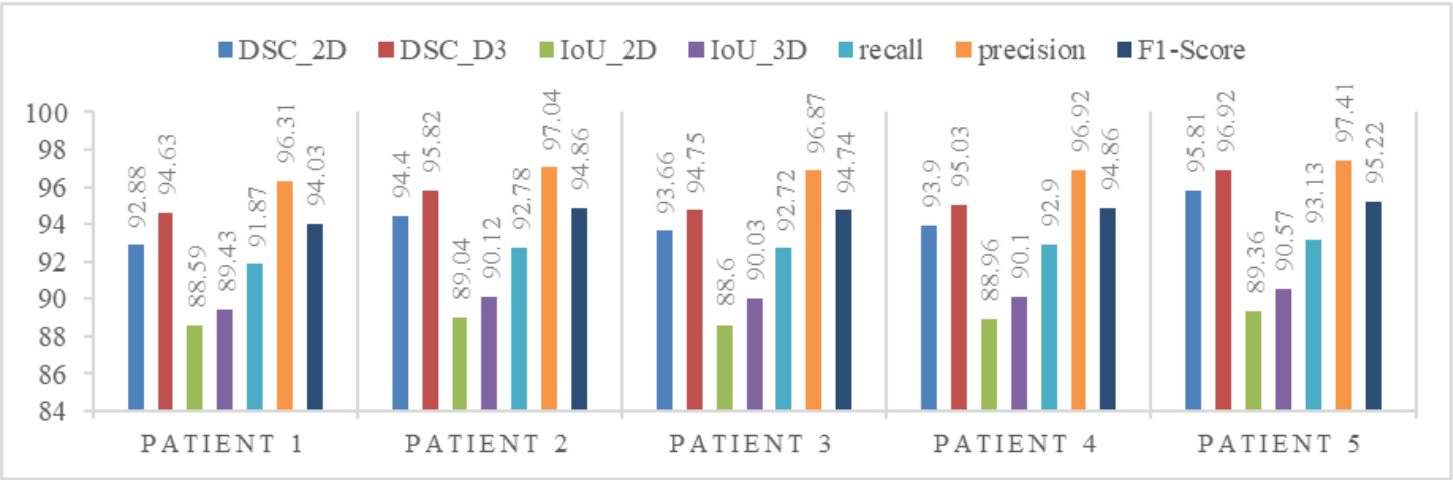

Figure 5

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

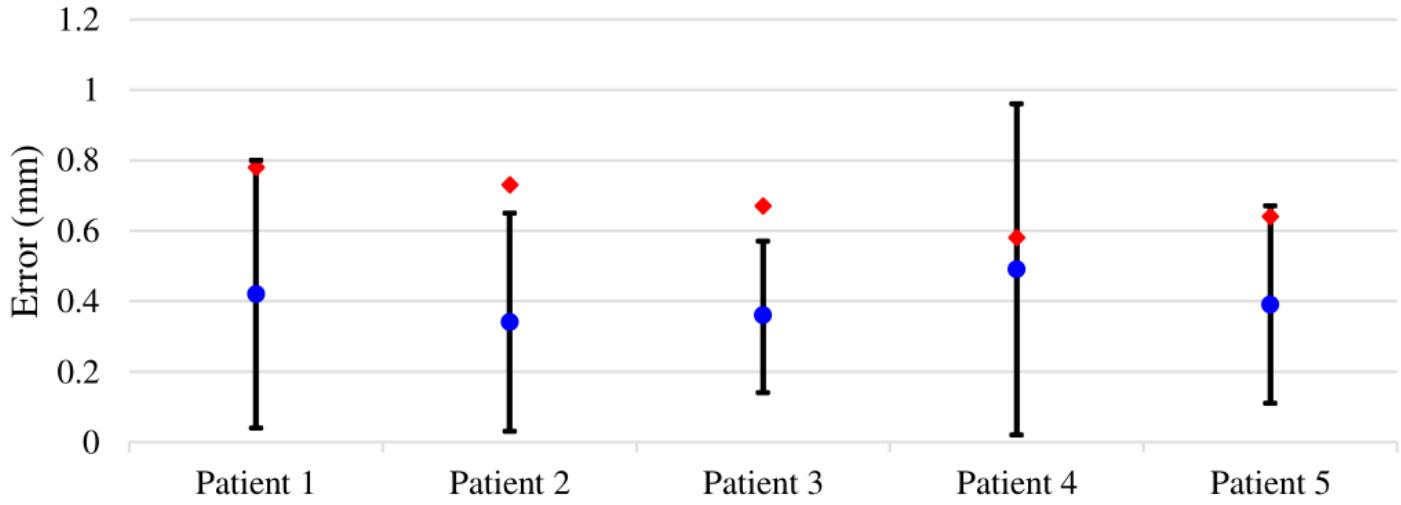

Figure 6

Tip position (Blue circles), mean of shaft errors (red diamonds), and the standard deviations of the test sets.

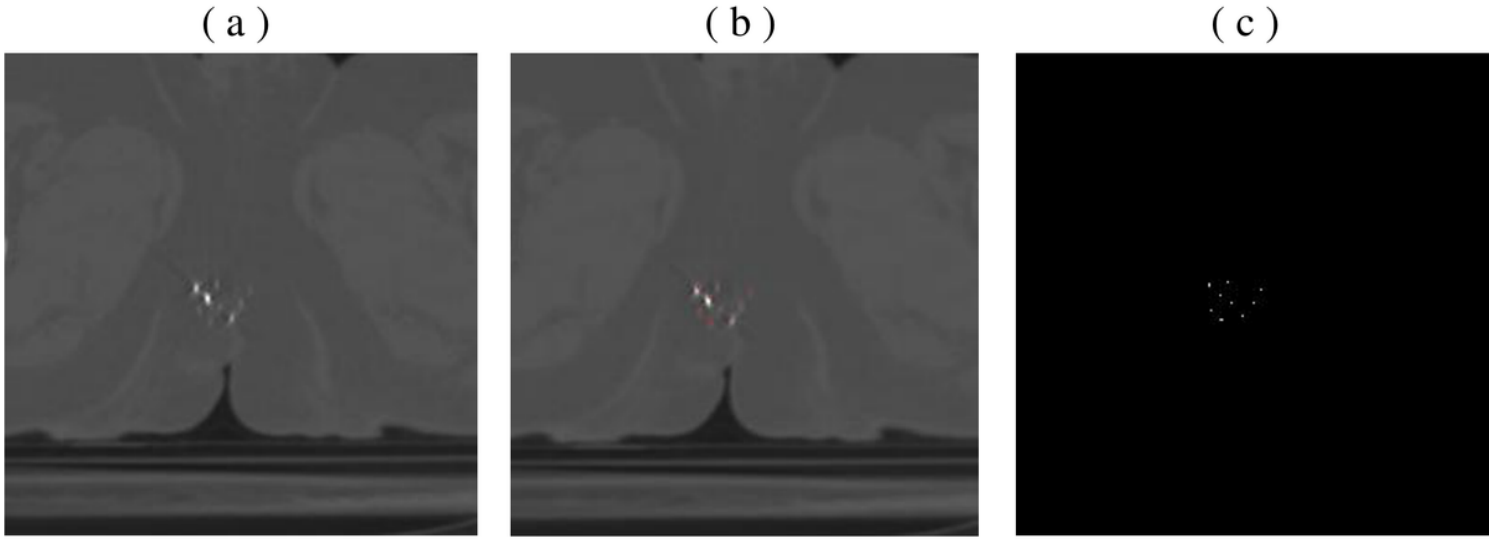

Figure 7

One two-dimensional sample of test set, (a) axial image with catheters, (b) its annotations, (c) catheter's segmentation image.

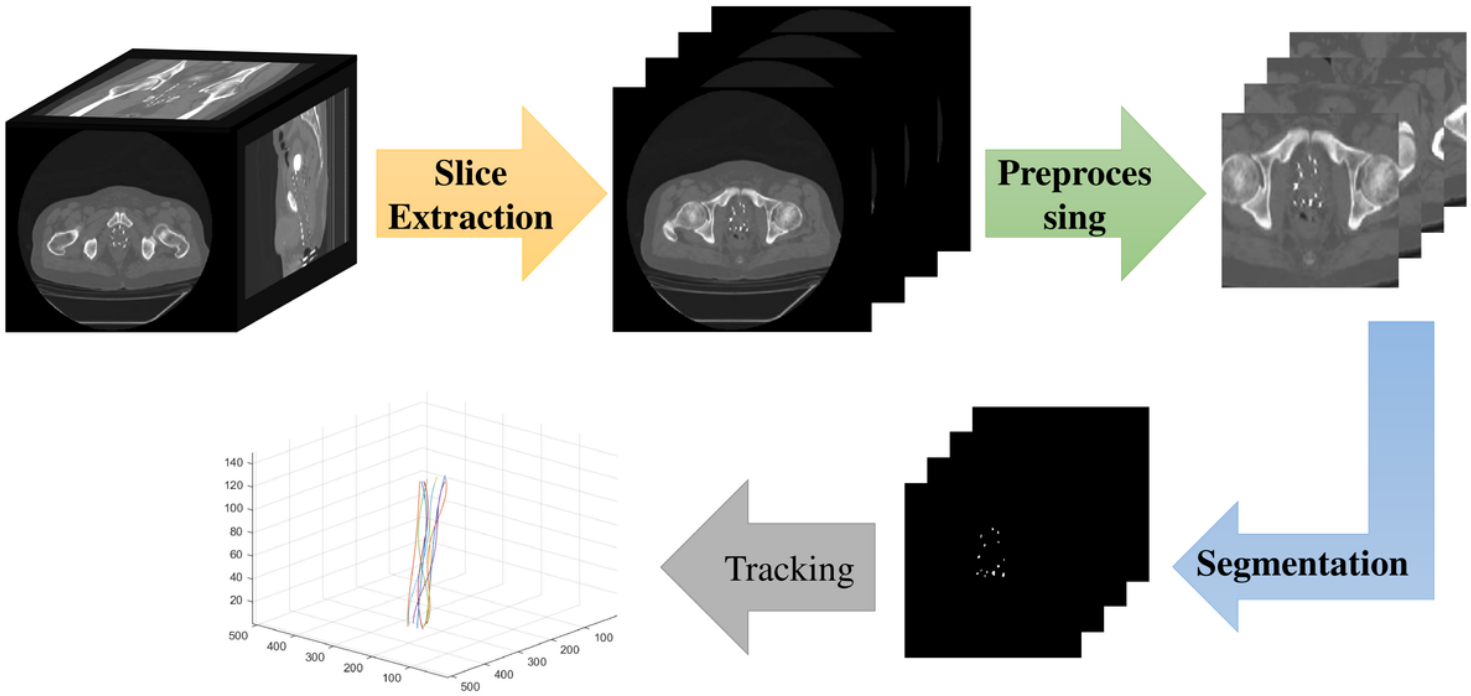

Figure 8

The general framework of the proposed method.

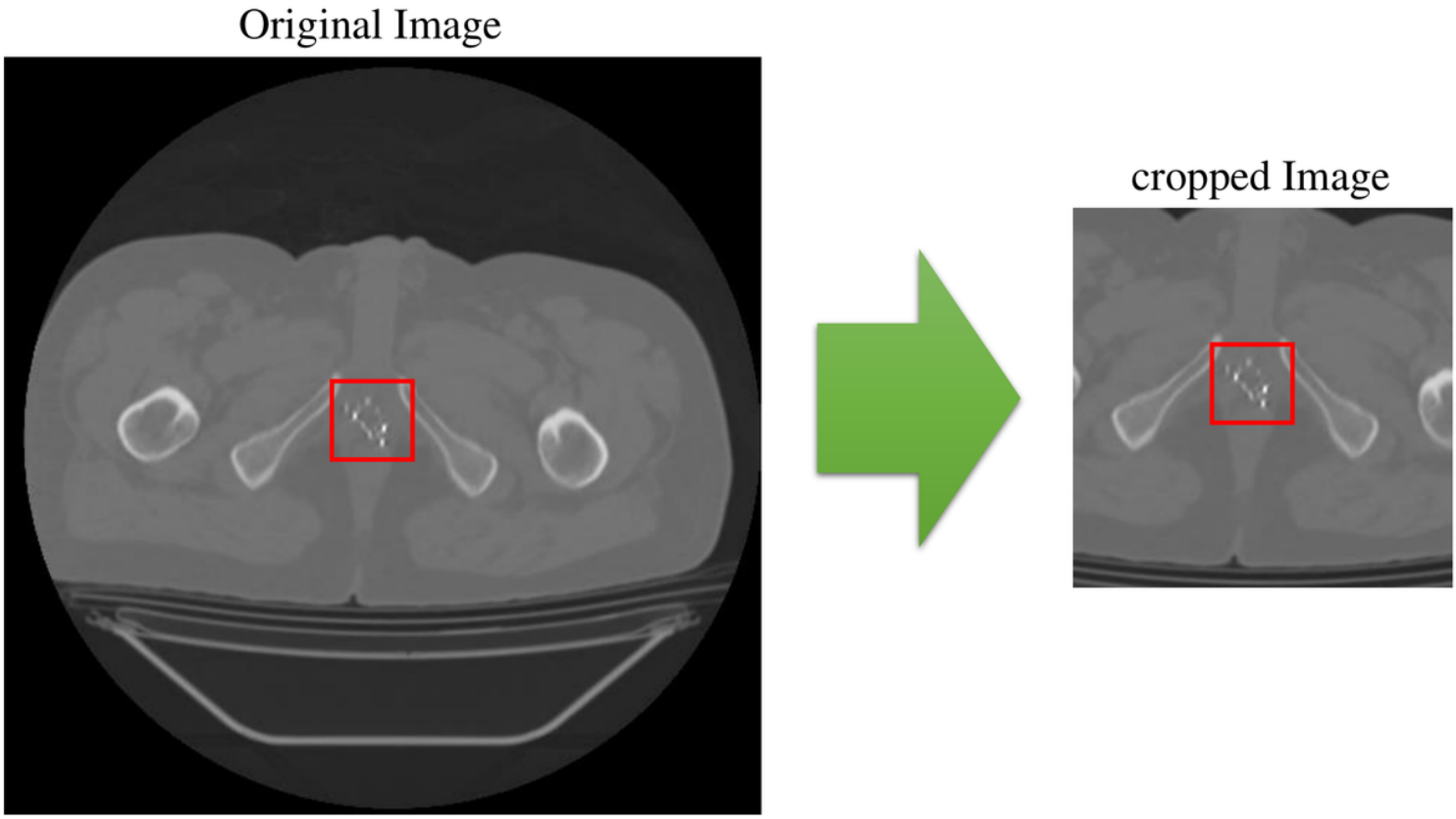

Figure 9

A sample of data cropping.

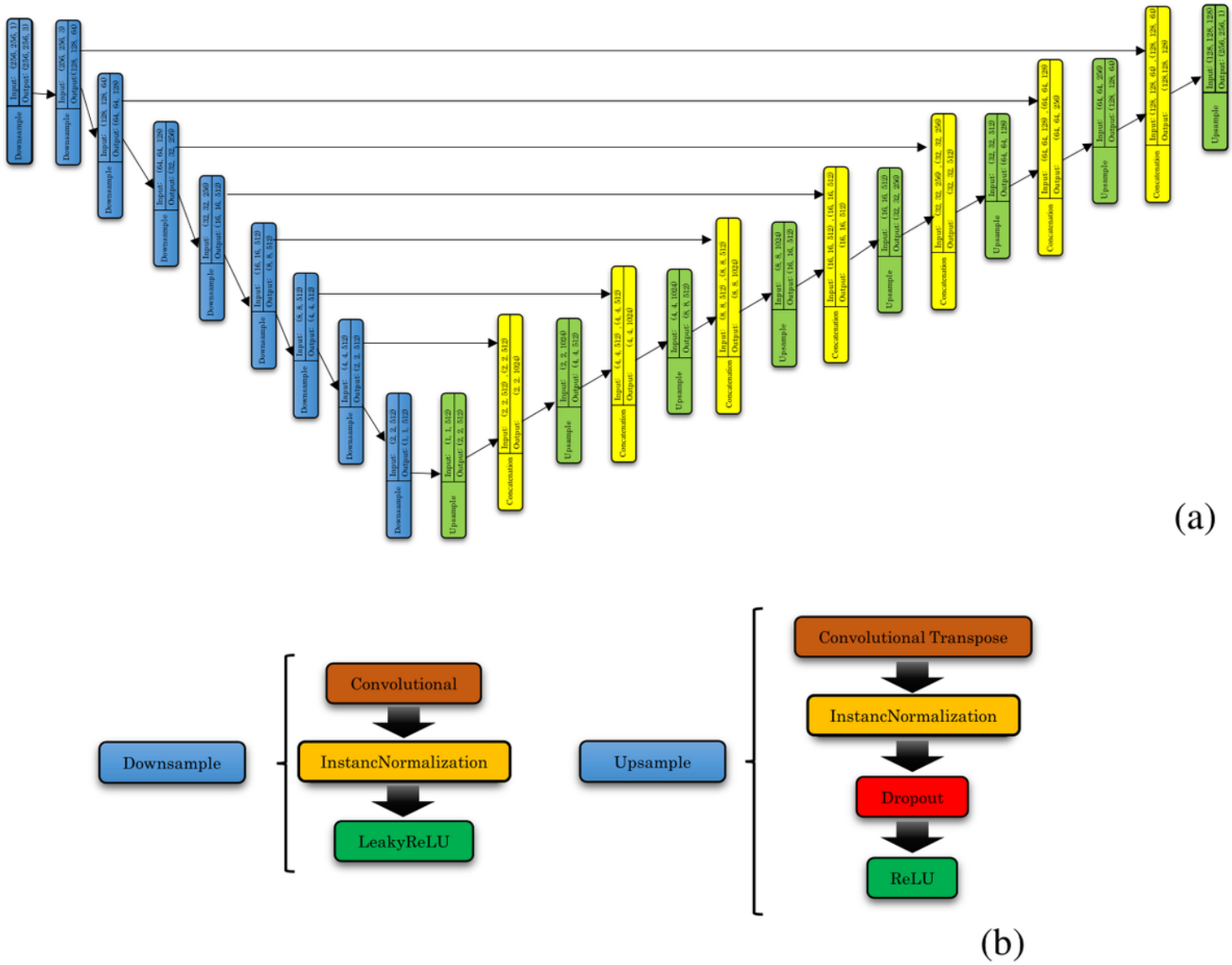

Figure 10

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

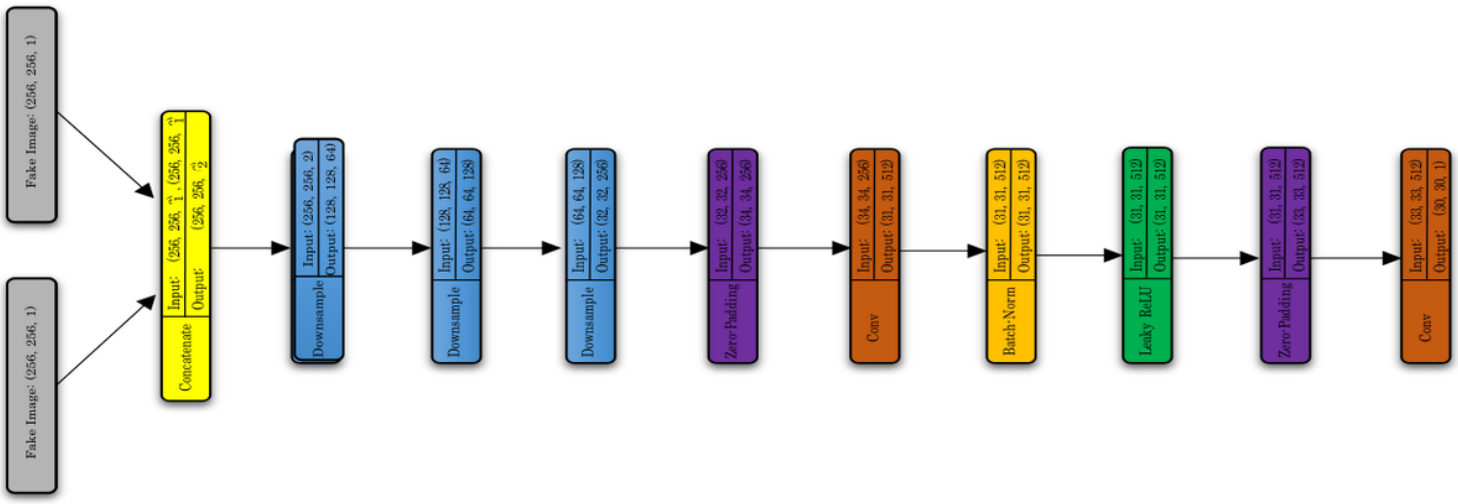

Figure 11

The Pix2Pix Discriminator architecture.

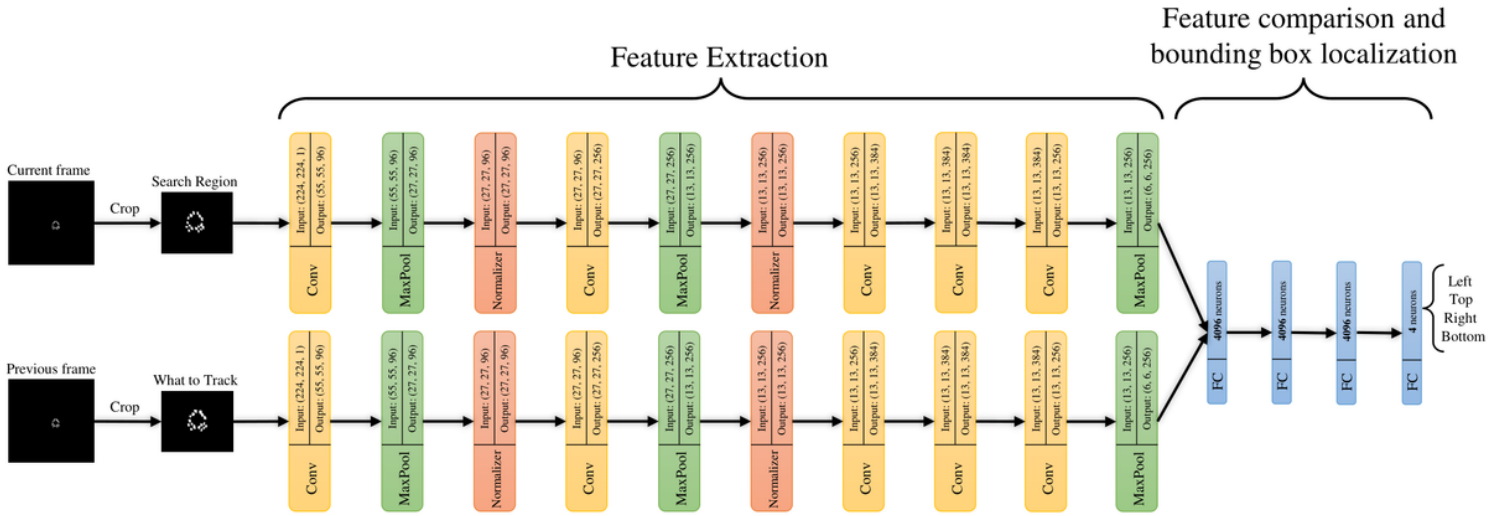

Figure 12

The GOTURN architecture.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Graphicalabstract.png