Radiotherapy and Oncology 140 (2019) 167–174

Contents lists available at ScienceDirect

Radiotherapy and Oncology

journal homepage: www.thegreenjournal.com

Original Article

Incorporating imaging information from deep neural network layers into image guided radiation therapy (IGRT)

Wei Zhao, Bin Han, Yong Yang, Mark Buyyounouski, Steven L. Hancock, Hilary Bagshaw, Lei Xing *

Stanford University, Department of Radiation Oncology, Stanford, USA

ARTICLE INFO

Article history:
Received 7 December 2018
Received in revised form 6 May 2019
Accepted 17 June 2019
Available online 11 July 2019

Keywords:
Prostate
Radiotherapy
Image-guided radiation therapy
Deep learning
Motion management
Convolutional neural network

ABSTRACT

Background and purpose: To investigate a novel markerless prostate localization strategy using a pre-trained deep learning model to interpret routine projection kilovoltage (kV) X-ray images in image-guided radiation therapy (IGRT).

Materials and methods: We developed a personalized region-based convolutional neural network to localize the prostate treatment target without implanted fiducials. To train the deep neural network (DNN), we used the patient's planning computed tomography (pCT) images with pre-delineated prostate target to generate a large amount of synthetic kV projection X-ray images in the geometry of onboard imager (OBI) system. The DNN model was evaluated by retrospectively studying 10 patients who underwent prostate IGRT. Three out of the ten patients who had implanted fiducials and the fiducials' positions in the OBI images acquired for treatment setup were examined to show the potential of the proposed method for prostate IGRT. Statistical analysis using Lin's concordance correlation coefficient was calculated to assess the results along with the difference between the digitally reconstructed radiographs (DRR) derived and DNN predicted locations of the prostate.

Results: Differences between the predicted target positions using DNN and their actual positions are (mean ± standard deviation) 1.58 ± 0.43 mm, 1.64 ± 0.43 mm, and 1.67 ± 0.36 mm in anterior–posterior, lateral, and oblique directions, respectively. Prostate position identified on the OBI kV images is also found to be consistent with that derived from the implanted fiducials.

Conclusions: Highly accurate, markerless prostate localization based on deep learning is achievable. The proposed method is useful for daily patient positioning and real-time target tracking during prostate radiotherapy.

 Radiotherapy and Oncology 140 (2019) 167–174

In the past two decades, we have witnessed tremendous advances in conformal radiation therapy (RT) and new modalities such as volumetric modulated arc therapy (VMAT), which has greatly augmented our ability to shape the isodose distribution to better conform radiation dose to the tumor target while sparing the organs at risks (OARs) [1,2]. In practice, however, just being able to produce conformal dose distributions is not enough as it only fulfills part of the requirements of precision RT. For patients to truly benefit from the advanced planning and dose delivery techniques, we must also ensure that the planned dose is delivered to the right location and, in the case of treatment of a moving tumor, at the right time. For example, the position of prostate may change from fraction to fraction (inter-fraction prostate motion) as well as during the dose delivery (intra-fraction prostate motion) [3–6]. An effective method for tumor localization, preferably in real-time, is thus of great clinical significance for the success of precision RT.

In practice, much efforts have been devoted to developing various image guidance strategies to ensure the accuracy of beam targeting [7–10], which include the use of stereoscopic or monoscopic kV X-ray imaging [11–15], hybrid kV and MV imaging [16,17], cone-beam computed tomography (CBCT) [18–21], onboard MR imaging [22], ultrasound imaging [23,24], etc. In the case of kV image guided RT, which requires less daily dose and less time to setup the patient as compared to CBCT [21], the low soft tissue contrast makes it difficult to see the tumor on the projection X-ray images. Thus, metallic fiducials are often implanted into the tumor volume or adjacent normal tissue to facilitate the patient setup and real-time tumor tracking [25–29]. With the use of fiducials, the planning target volume (PTV) margin for prostate RT can be reduced to 3 mm [30]. Radiofrequency (RF) tumor tracking system with implanted RF transponders has also developed to localize

* Corresponding author at: Department of Radiation Oncology, Stanford University, 875 Blake Wilbur Drive Room G233, Stanford, CA 94305-5847, USA.
E-mail addresses: zhaow85@stanford.edu (W. Zhao), hanbin@stanford.edu (B. Han), yongy66@stanford.edu (Y. Yang), mbuyyou@stanford.edu (M. Buyyounouski), shancock@stanford.edu (S.L. Hancock), hbagshaw@stanford.edu (H. Bagshaw), lei@stanford.edu (L. Xing).

https://doi.org/10.1016/j.radonc.2019.06.027
0167-8140/

the prostate and lung tumors [31,32]. However, the implantation of the fiducial is an invasive procedure which can introduce bleeding, infection and discomfort to the patient. It also requires the service of an interventional radiologist or other specialist and prolongs the treatment procedure. Besides, studies have shown fiducial can migrate within the patient and the prostate exhibited random deformation leading to uncertainty to patient setup and target localization [33]. Hence, non-invasive or markerless tumor motion monitoring is of great clinical relevant. Prior efforts have shown online markerless target monitoring is feasible for spine and lung tumor using several elegant methods, such as 2D/3D registration [34–36], template matching [37–40], machine learning techniques [41], mean-shift algorithm [42], and registration between daily setup CBCT projection and MV/kV images [43]. Although markerless tracking is possible for some sites, it is highly challenging in the abdominal–pelvic region [44–46].

Recently, deep learning has attracted much attention for various medical applications, such as positron emission tomography (PET)/magnetic resonance imaging (MRI) attenuation correction [47], and segmentation of organ at risk for radiotherapy [48]. The purpose of this study was to investigate a novel markerless prostate localization strategy using a pre-trained deep learning model to interpret routine projection kV X-ray images.

Methods and materials

In this institutional review board (IRB)-approved study, the trained prostate localization model is validated by retrospective analysis of 10 patients treated with VMAT from 2015 to 2017. In addition to the pCT images, all patients underwent either kV CBCT scans or orthogonal kV projection using the OBI system (Varian Medical System, Palo Alto, CA) for positioning before treatment.

Deep learning for tumor target localization

In deep learning, a computer model learns to perform prediction or classification from images or other forms of data. For the stated target localization problem, the following are the major tasks to accomplish in order to obtain a reliable deep learning model: construction of hierarchical neural network, collection and annotation of training datasets, training of deep learning model, and validation. Fig. 1 shows the steps and workflow of the proposed deep learning-based target bounding box localization process. The first step is to generate training datasets of kV projection X-ray images reflecting various situations of the anatomy, including different levels of translation, rotation and organ deformation of the patient. For this purpose, we (i) placed the patient's pCT images in the OBI geometry; (ii) introduced a series of changes (translations/rotations/deformations) in pCT to mimic different clinical scenarios; and (iii) for each of the changes, generated a DRR with the corresponding prostate contour and bounding box as annotated to train a deep learning model for subsequent localization of the prostate target. The deformation was done using a robust deformable model described by motion vector fields (MVFs). Validation tests using both DRR and monoscopic X-ray projection images obtained from kV on-board imager (OBI) system

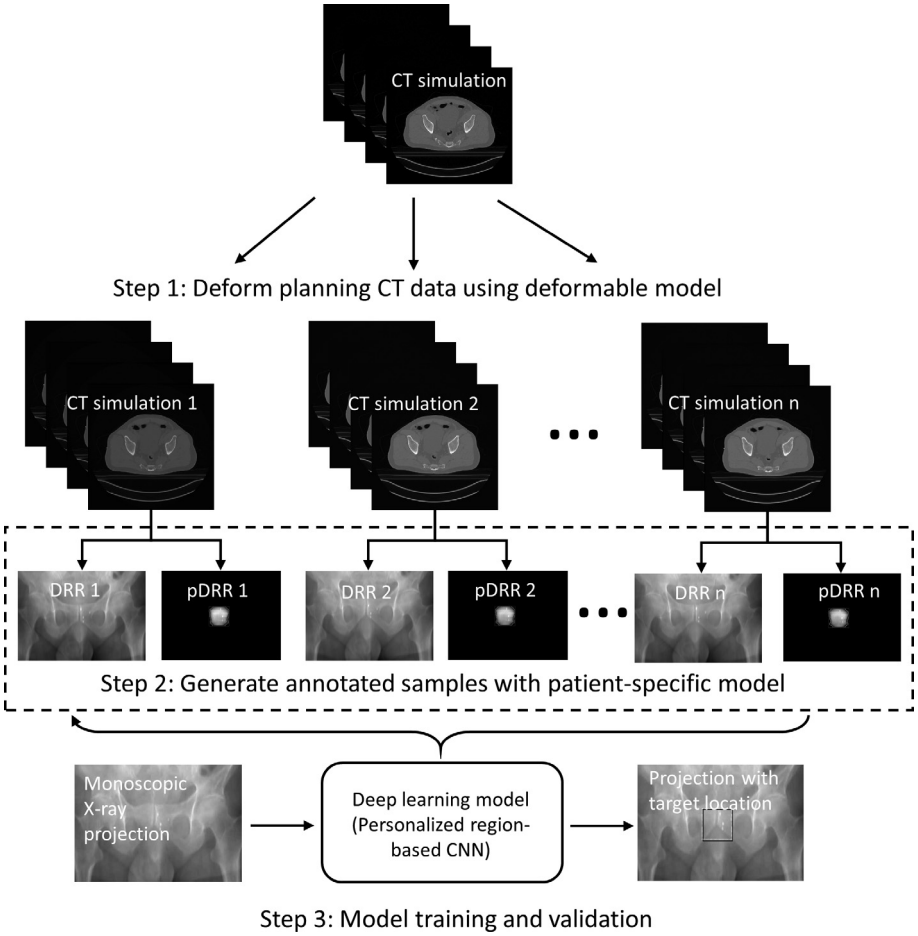

Fig. 1. Overall flowchart of the proposed deep learning-based treatment target localization method. *Abbreviation*: CT = computed tomography, DRR = digitally reconstructed radiograph, pDRR = prostate digitally reconstructed radiograph, CNN = convolutional neural network.

were performed. These steps are described below in some more detail.

Deep learning model

In our deep learning model, the input is a monoscopic X-ray projection image from a given direction, say anterior–posterior (AP), or left-lateral (L-Lat), or an oblique direction obtained by the OBI system. Conventional object detection model using region-based convolutional neural networks (CNN) can achieve nearly real-time detection for given region proposals [49]. However, the calculation of the proposals is computationally intensive and presents a bottleneck for real-time detection of target region. In this study, we employ a region proposal network (RPN) to provide proposals for the region-based CNN, and the two networks share all the image convolutional features with each other [50], as illustrated in Fig. 2. The framework jointly generates region proposal and refines its spatial locations. The online kV X-ray image is first passed through VGG16 [51] feature extractor which has 13 convolutional layers, 13 ReLU layers and 4 pooling payers to generate its convolutional feature maps. By doing so, the CNN exploits the spatial correlation between the investigated prostate target and its surrounding area and encodes it into the feature maps. The VGG16 was pre-trained to take advantage of transfer learning and its top 4 convolutional layers were frozen to reduce the number of training weights and to accelerate training. Based on the feature maps, the RPN is constructed by adding two additional layers. The first one is a 3×3 convolutional layer and the second one is two sibling 1×1 convolutional layers, a box-classification layer (cls) and a box-regression layer (reg). By providing the training DRR samples and their corresponding labels, the RPN outputs a set of region proposals, each with a score. For each of the proposals, a position sensitive region of interest pooling layer extracts a fixed length feature vector from the convolutional feature maps, which are then fed into two fully connected ReLU layers and two sibling fully connected layers for a multi-task loss calculation. With this framework, the RPN and region-based CNN share the 13 convolutional layers and ultimately enable real-time and accurate target detection.

Generation of labeled training datasets for deep learning

Training of deep learning model requires a large number of annotated datasets and this often presents a bottleneck problem in the realization of the potential of deep learning methods. Instead of collecting a large amount of projection kV images from the clinic, which is labor intensive and impractical in realistic clinical settings, we propose to generate projection kV images from augmented pCT images for a specific patient. To this end, for each patient, we first introduce a series of changes including translation, rotations and deformations in pCT to mimic different clinical scenarios. For each of the changes, we generated a DRR with the corresponding prostate contour and bounding box as the annotation (more details in Appendix A).

The DRR data for any given direction are generated using an accurate forward X-ray projection model from the change incorporated pCT. In the forward projection calculation, we use the OBI geometry with the source to detector distance of 1500 mm and the source to patient distance of 1000 mm. The projection pixel size and array size are the same as the OBI, i.e., 0.388 mm and 1024×768, respectively. In this way, the generated DRR is geometrically consistent with the kV projection image acquired using the OBI system. The DRR calculation is implemented using CUDA C with graphics processing unit (Nvidia GeForce GTX Titan X, Santa Clara, CA) acceleration. We calculate the bounding boxes of the prostate on DRRs and the information (top-left corner, width, height) together with the corresponding DRRs serving as annotation for deep learning model training (more details in Appendix B).

Training and validation of the prostate localization model

For each patient, we retrieved pCT images, target structure contours and patient setup images (CBCT images or orthogonal kV projection for different courses of treatment). We then created a set of 900 synthetic DRRs for each of the three directions (AP, lateral, and an oblique direction) and used the 900 pairs of DRRs and their corresponding annotations to train the DNN model. Validation was performed using 100 additional pairs of DRRs by comparing the predicted prostate bounding box position to that annotated bounding box (ground truth) on the DRRs. The training and testing sets were generated using difference MVFs, and they are independent from each other.

The mean absolute difference (MAD) and standard deviation of the observed individual error were used to quantify the accuracy of the predicted prostate position. Lin's concordance correlation coefficient (ρ_c) [52] was also calculated to assess the results along with the difference between the annotated position and DNN predicted position of the prostate. ρ_c is defined as:

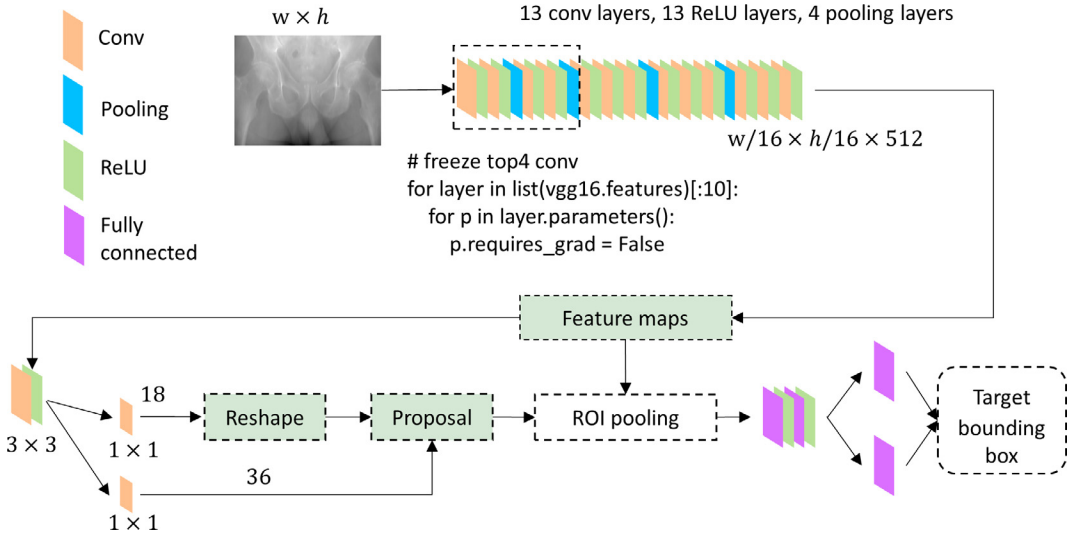

Fig. 2. Schematic illustration of the deep learning model for localizing the prostate. *Abbreviation*: ROI = region of interest.

$$\rho_c = \frac{2\rho\sigma_x\sigma_y}{\sigma_x^2 + \sigma_y^2 + \left(\mu_x - \mu_y\right)^2}$$

where μ_x and μ_y are the mean values for the predicted positions and the annotated positions of the 100 independent testing samples, respectively. σ_x and σ_y are the corresponding standard deviations, and ρ is the correlation coefficient between the predicted and annotated positions.

In addition to the validation studies using the 100 synthetic DRRs for each patient and each projection direction, we also performed validation using OBI-measured kV projections images for three out the ten patients who were setup using fiducial imaging. Due to the image quality difference between DRRs and measured OBI images, spatial invariant transformations including histogram matching and Gaussian filtration were applied to the OBI images. To quantify the accuracy of the deep learning-predicted target position, the distances between the fiducials and the sides of model-predicted bounding box were calculated and compared to the planned values from the patient's treatment plan (calculated based on the DRR with the corresponding prostate bounding box generated from the pCT). Here the fiducials are used as surrogates of the prostate target to quantify the accuracy of the deep learning-predicted target position. Ideally, the bounding box follows the prostate movement and, in the absence of prostate deformation or fiducials migration, the distances of the fiducials to the sides of the bounding box remain unchanged.

Results

Fig. 3 shows some examples of prostate bounding boxes (dashed yellow lines) derived from the proposed approach together with the known bounding box positions (blue lines) by projecting the prostate target onto the corresponding DRR planes for three patients in AP, oblique, and L-Lat directions. To facilitate visualization, the patient's DRRs after MVF deformations are presented in the figure as the background. From Fig. 3, it is seen clearly that the predicted bounding box positions match the known positions very well in all three directions.

Lin's concordance correlation coefficients and MAD between the DNN predictions and the annotated positions of the top-left corner of prostate bounding boxes for all 10 patients are summarized in Table 1. As can be seen, the predicted and annotated positions agree each other well with observed ρ_c values greater than 0.88 for all three directions and all ten patients. Meanwhile, in all directions tested here, the MADs are less than 3 mm in all cases, which is important for patient setup and prostate tracking during IGRT based on the use of the kV OBI imaging.

The deep learning-based prediction model is computationally efficient once the training of the model is done. In a PC platform (Intel Core i7-6700 K, RAM 32 GB) with GPU (Nvidia GeForce GTX Titan X, Memory 12 GB), the training took about two hours for a given direction. The execution of the prediction model took less than 200 ms, which is well suited for real-time tracking of the tumor target position during the dose delivery of radiation therapy or during other interventional procedures.

Fig. 4 shows the predicted target position (bounding box drawn with dashed yellow line) overlaid on top of the patient's DRRs, kV project images for two different treatment sessions of the three patients who were setup under the guidance of kV fiducial imaging and quantitative evaluation of the predicted position with respect to the fiducials. The kV project images used as input to the DNN model were obtained using OBI before the patient's RT treatment. In the reference simulated projections, the deviations between the

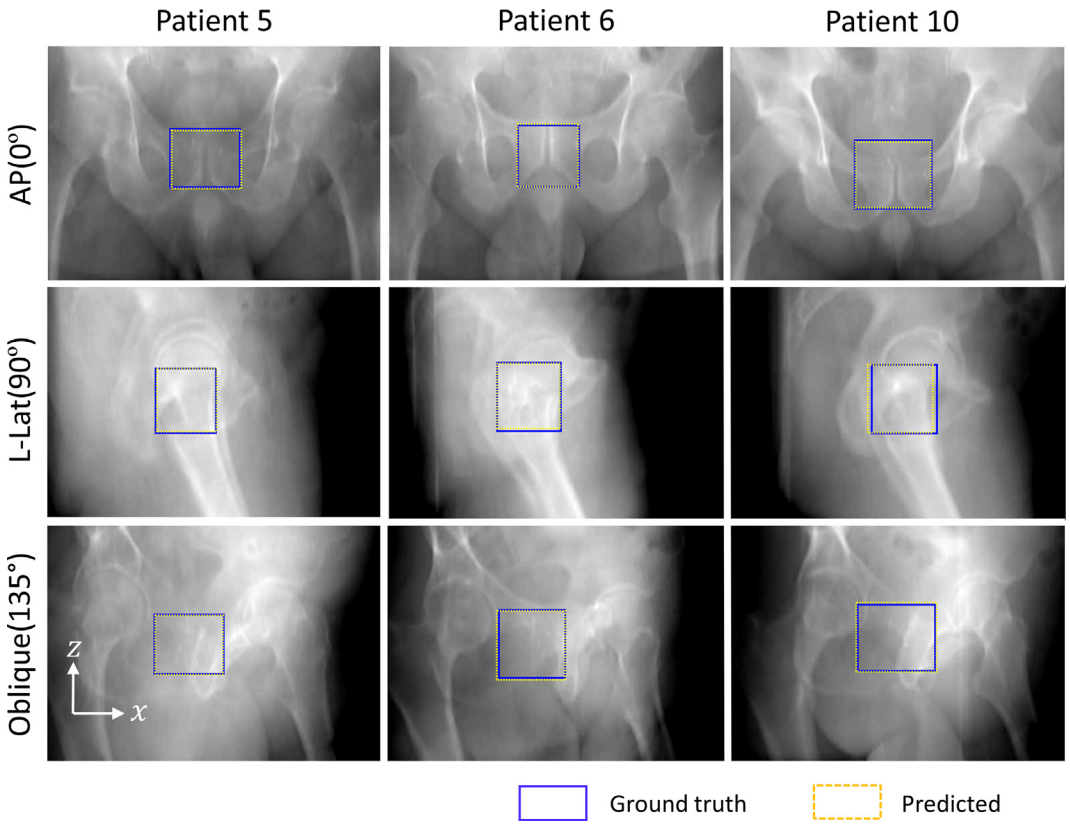

Fig. 3. Examples of the prostate boundary boxes derived from the deep learning model (yellow dashed box) and their corresponding annotations (blue box), overlaid on top of the patients' DRRs. The first, second and third columns show the results in AP, oblique and L-Lat directions, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Mean absolute difference and Lin's concordance correlation coefficients [52] between the predicted and annotated prostate positions in AP, L-Lat, and oblique directions.

Patient index	AP				L-Lat				Oblique			
	Δx (mm)	ρ_c	Δz (mm)	ρ_c	Δx (mm)	ρ_c	Δz (mm)	ρ_c	Δx (mm)	ρ_c	Δz (mm)	ρ_c
1	1.58±1.10	0.97	1.03±0.69	0.93	2.51±1.83	0.94	1.49±1.10	0.91	1.93±1.65	0.96	1.33±0.93	0.93
2	1.78±1.49	0.97	1.45±1.02	0.98	1.73±1.28	0.96	1.08±0.78	0.94	2.16±1.64	0.94	1.81±1.24	0.98
3	1.69±1.51	0.96	1.76±1.58	0.94	1.48±1.15	0.96	1.27±0.95	0.91	1.87±1.86	0.97	1.41±1.31	0.94
4	2.27±1.93	0.96	1.43±1.13	0.94	1.98±1.46	0.95	1.85±1.38	0.95	2.37±1.81	0.91	1.57±1.29	0.90
5	1.78±1.77	0.95	0.98±0.64	0.97	2.16±1.43	0.93	1.15±0.78	0.92	1.63±1.26	0.96	1.28±0.97	0.93
6	2.07±1.22	0.96	2.48±2.32	0.91	1.50±1.20	0.96	1.19±0.80	0.89	1.79±1.22	0.95	1.69±1.05	0.91
7	1.66±1.33	0.95	0.91±0.66	0.94	1.89±1.37	0.95	0.91±0.75	0.91	1.80±1.16	0.95	1.34±0.92	0.93
8	1.81±1.62	0.93	1.37±1.01	0.91	1.97±1.39	0.94	1.49±1.08	0.90	2.45±1.70	0.92	1.11±0.93	0.92
9	1.28±1.03	0.97	0.90±0.71	0.93	1.60±1.07	0.91	1.40±0.84	0.89	1.48±1.18	0.97	1.57±0.98	0.88
10	1.50±1.08	0.95	1.82±1.28	0.90	2.37±1.70	0.93	1.70±0.71	0.91	1.50±1.06	0.96	1.36±1.13	0.89
Mean(Std)	1.58 ± 0.43				1.64 ± 0.43				1.67 ± 0.36			

Note: AP = anterior–posterior, L-Lat = left-lateral, std = standard deviation.
Data are shown as mean ± standard deviations.

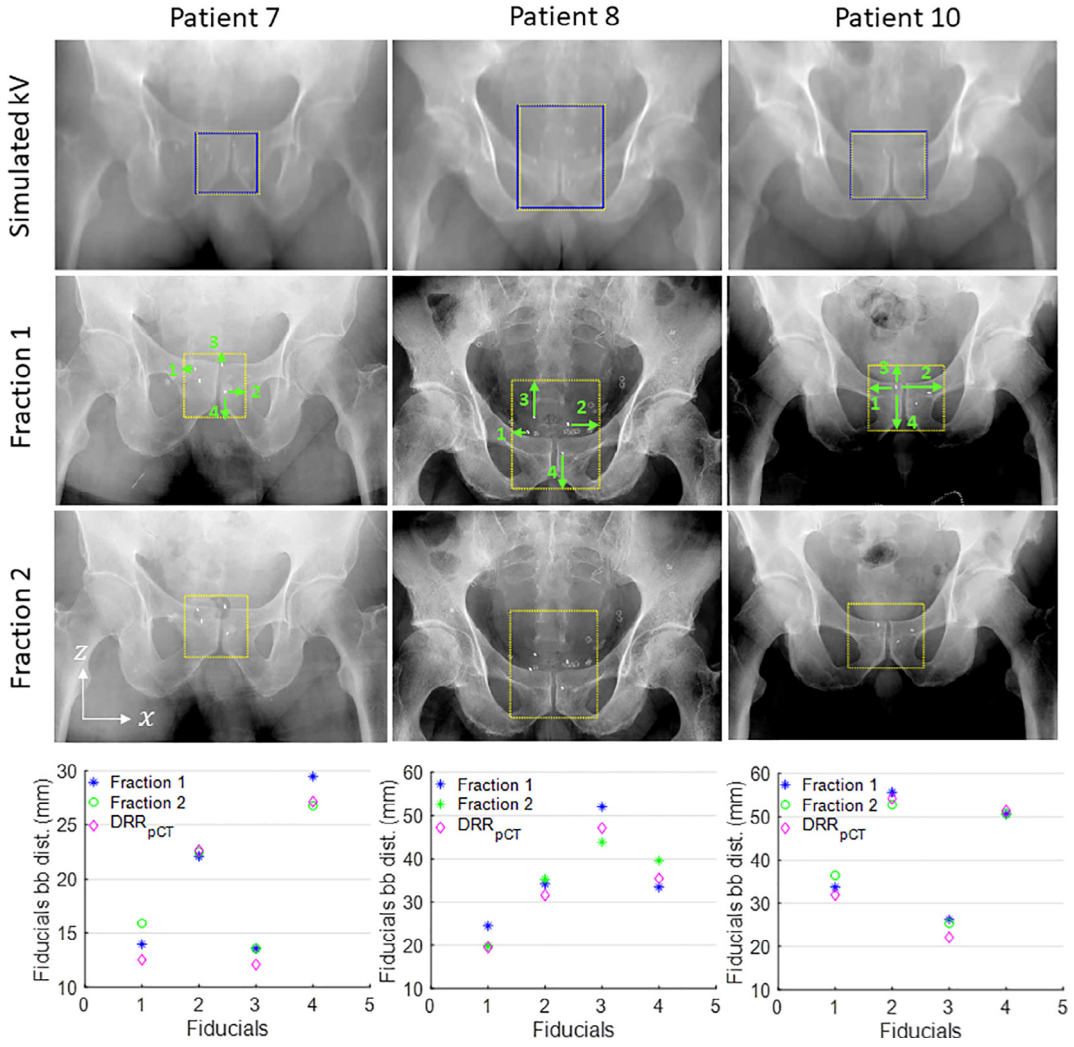

Fig. 4. Predicted and actual positions of the prostate target overlaid on top of the AP simulated projections (1st row) as well as the OBI images (2nd and 3rd rows) for two different treatment sessions of three patients who were setup using kV fiducial imaging. DRR$_{pCT}$ denotes DRR generated using original planning CT images. The fiducial markers seen in OBI images provide additional assurance about the accuracy of the proposed deep learning-based prediction model. *Abbreviation*: OBI = on-board imager.

predicted and the annotated positions are quite small. These patients had implanted fiducials and analysis of the positions of these fiducials relative to the corresponding bounding boxes affords additional assurance of the correctness of our deep learning model. The fiducials relative to the bounding boxes were specified by their distances to the sides of the bounding box (Fig. 4, 2nd row). The fourth row of Fig. 4 shows the calculated distances and the reference values obtained from the patient's treatment plans. It is seen that the change in the distances of the fiducials to the sides of the bounding box is small. For the three cases with kV

fiducial images (2nd, 3rd, 4th rows of Fig. 4), the MAD between the measured distances and that derived from DRRs using original pCT images was 2.29 mm, indicating that the deep learning model is capable of identify the position of the prostate target on measured kV projection images.

Discussion

Management of inter- and intra-fraction organ motion is one of the important issues in RT and other interventional procedures [25,53]. In current practice, the advantage of highly conformal RT techniques is not fully utilized because of the uncertainties in beam targeting. Indeed, current RT suffers from the inability of to see the tumor volume or anatomical structures in images (such as OBI X-ray projection images) acquired online or offline. As a compromised solution, implanted fiducials are often used to localize the tumor target, which is invasive, costly and risky to certain patients. Markerless image guidance based on the correlation between the on-treatment projection image and pCT image without machine learning has also been implemented clinically [39,54]. But these approaches only work when the contrast of the tumor target is sufficiently high, which is often not the case in reality. In this study, we have proposed a target positioning method with an ingenious use of a deep learning model. We have shown that the integration of deep layers of image information greatly augments the traditional X-ray imaging modality and enables us to visualize the prostate target that would otherwise be invisible to humans' eyes. The proposed approach mitigates the need for implanted fiducials without compromising the target localization accuracy. It enables us to localize the prostate in the absence of visible image contrast by effectively utilizing deep layers of image information.

Another salient feature of the proposed technique is that the training of the deep learning model does not rely on the collection of a large amount of data from the routine clinic practice, which has been recognized as a bottleneck in the applications of deep learning techniques [55]. Instead, we developed a strategy of generating synthetic yet practical training datasets covering different clinical scenarios by hypothetically introducing a large number of physically realizable changes in the internal anatomy and patient positioning. This significantly simplifies the process of building a predictive model and makes it practical for a variety of clinical applications. Since the model trained under a specific projection direction cannot be used by another direction, an angle-specific model needs to be trained for real-time target tracking. Additionally, this study is focused on demonstrating the feasibility of deep learning-based prostate localization using monoscopic X-ray imaging, which is incapable of detecting small rotations. However, the proposed framework can be extended to the case of stereoscopic X-ray based prostate localization, which can detect a prostate rotation reliably because of the presence of additional information from an orthogonal angle.

While the current study is focused on patient with radiation therapy of intact prostate, the proposed approach is quite general and is applicable to track other types of organs as long as the volumes are contoured in the pCT [56]. For patients who underwent surgical interventions, we can track the planning target volume (PTV) in the OBI images. We also note that this study is focused on geometric verification, which has been the standard of practice in patient setup. Clinically, a large population of patients are treated with orthogonal projection images-based setup and these patients will benefit directly from the current work as the improved target localization accuracy will provide high confidence for PTV margin reduction, leading to lower toxicity and improved IGRT. Furthermore, because of the high accuracy and computational speed, the proposed technique is readily applicable for real-time target tracking with either on-board monoscopic kV imaging system or other stereoscopic imaging systems.

In conclusion, we have proposed a deep learning model for localization of prostate bounding box based on projection images acquired prior or during therapy. This represents the first attempt of applying deep learning to IGRT. Two novelties of the proposed approach are: (1) a model training procedure based on the use of synthetic DRRs derived from the pCT images; and (2) a combined use of CNN and RPN for nearly real-time localization of the prostate. The use of synthetic DRRs mitigates the need for a vast amount of X-ray images for the model training, which makes the proposed approach a clinically practical solution. Applications of the proposed technique to clinical prostate IGRT cases strongly suggest that highly accurate prostate tracking in X-ray images are readily achievable using the proposed deep learning model. The approach allows us to see otherwise invisible structures on the X-ray images and alleviates the need for implanted fiducials. We emphasize that the proposed method is quite broad and general. The approach can be readily generalized to improve image guidance in many other disease sites, such as the lung, liver, brain, spine, and head and neck IGRT. Finally, the proposed strategy should also be directly applicable to other types of image-guided interventions.

Declaration of Competing Interest

None

Acknowledgements

This work was partially supported by NIH (1R01 CA176553 and R01E0116777), a Faculty Research Award from Google Inc. and a gift fund from Huiyihuiying Med Co. We also gratefully acknowledge the support of NVIDIA Corporation for the GPU donation.

Appendix A

To obtain the labeled DRR images corresponding to numerous possible positions and shapes of the involved anatomical structures for training the deep learning model for a specific patient, we employ a well validated deformable image registration to deform the planning CT data in addition to conventional data augmentation operations such as rotation and random translation of the images. For each planning CT image set, we first replace the fiducials with soft tissue to generate fiducial-free CT image set, which is rotated around the superior-interior direction from -4° to 4° with 2° interval. Then, to deform the high-quality planning CT data, we extract MVFs by registering pre-treatment CBCT data from different courses of treatment. For a CT data set with size of $512 \times 512 \times 93$, the corresponding MVF for our deformable model is a 4D matrix with size of $512 \times 512 \times 93 \times 3$, where the three $512 \times 512 \times 93$ in the 4D matrix contain displacements along the x-axis, y-axis and z-axis, respectively. For each patient, a set of 20 CBCT datasets is extracted and 20 MVFs are generated by registering. In order to increase the sample size, for each of the two MVFs, we randomly generate a new MVF' as follow:

$$MVF' = rand * MVF_1 + (1 - rand) * MVF_2$$

where $MVF_{1,2}$ are two MVFs in the 20 MVFs, and $rand$ is a uniformly distributed random number in the interval $(0, 1)$. With this method, a total of 50 MVFs is generated. After applying the MVFs to the rotation incorporated pCT images, a series of 250 CT data that cover different clinical scenarios are generated and these data are used to mimic the specific patient in different positions and status. This

procedure is performed using a Matlab (MathWorks, Natick, MA) function *imwarp*. The deformed CT datasets are then divided into two parts, where 225 deformed CT datasets are used for model training and the other 25 deformed CT datasets are used for model testing.

The corresponding DRRs along the directions of 0, 90 and 135 degrees are then generated for each deformed CT data using accurate forward X-ray projection model. Each of the DRR is then randomly shifted 4 times in the LR (left–right) and SI (superior-interior) directions with limitations of 10 mm and 20 mm, respectively. Hence, a total of 900 DRRs are generated for each projection direction for the training of deep learning model, and 100 DRRs are used for testing.

To annotate the DRRs, for each patient, we extract the delineated prostate contour used for treatment planning. The prostate is applied to the same changes (translations, rotations, deformations) as part of the pCT data. The prostate after change is projected in OBI geometry to produce prostate projection on the corresponding DRR. The bounding box annotation of the prostate is then generated in the corresponding DRR image.

Appendix B

We use 10 epochs to train the network and the learning rate is set to 0.001. For efficient training of the deep learning model, the annotated samples are cropped into the size 700×1000. Before training, all training samples are randomly permuted. During the training process, we employ an alternating optimization technique to enable the sharing of the convolutional layers between RPN and region-based CNN. We first train RPN to generate reasonable target proposals, and then refine the location of the proposals by training the region-based CNN. After the first training, the region-based CNN is in turn used to initialize RPN and then the parameters of the 13 convolutional layers are fixed. A second training is performed thereafter by fine-tuning layers unique to RPN. Finally, the fully connected layers (FCs) are fine-tuned while keeping other layers fixed.

References

[1] Veldeman L et al. Evidence behind use of intensity-modulated radiotherapy: a systematic review of comparative clinical studies. Lancet Oncol 2008;9:367–75.
[2] Staffurth J. A review of the clinical evidence for intensity-modulated radiotherapy. Clin Oncol 2010;22:643–57.
[3] Crook J et al. Prostate motion during standard radiotherapy as assessed by fiducial markers. Radiother Oncol 1995;37:35–42.
[4] Ghilezan MJ et al. Prostate gland motion assessed with cine-magnetic resonance imaging (cine-MRI). Int J Radiat Oncol Biol Phys 2005;62:406–17.
[5] Kotte AN et al. Intrafraction motion of the prostate during external-beam radiation therapy: analysis of 427 patients with implanted fiducial markers. Int J Radiat Oncol Biol Phys 2007;69:419–25.
[6] Adamson J, Wu Q, Yan D. Dosimetric effect of intrafraction motion and residual setup error for hypofractionated prostate intensity-modulated radiotherapy with online cone beam computed tomography image guidance. Int J Radiat Oncol Biol Phy 2011;80:453–61.
[7] Xing L et al. Overview of image-guided radiation therapy. Med Dosim 2006;31:91–112.
[8] Dawson LA, Jaffray DA. Advances in image-guided radiation therapy. J Clin Oncol 2007;25:938–46.
[9] Verellen D et al. Innovations in image-guided radiotherapy. Nat Rev Cancer 2007;7:949.
[10] Jaffray DA. Image-guided radiotherapy: from current concept to future perspectives. Nat Rev Clin Oncol 2012;9:688.
[11] Soete G et al. Clinical use of stereoscopic X-ray positioning of patients treated with conformal radiotherapy for prostate cancer. Int J Radiat Oncol Biol Phys 2002;54:948–52.
[12] Xie Y et al. Intrafractional motion of the prostate during hypofractionated radiotherapy. Int J Radiat Oncol Biol Phys 2008;72:236–46.
[13] Guckenberger M et al. Potential of image-guidance, gating and real-time tracking to improve accuracy in pulmonary stereotactic body radiotherapy. Radiother Oncol 2009;91:288–95.
[14] Ng JA et al. Kilovoltage intrafraction monitoring for prostate intensity modulated arc therapy: first clinical results. Int J Radiat Oncol Biol Phy 2012;84:e655–61.
[15] Kim J-H et al. The accuracy and precision of Kilovoltage Intrafraction Monitoring (KIM) six degree-of-freedom prostate motion measurements during patient treatments. Radiother Oncol 2018;126:236–43.
[16] Liu W et al. Real-time 3D internal marker tracking during arc radiotherapy by the use of combined MV–kV imaging. Phys Med Biol 2008;53:7197.
[17] Wiersma R, Mao W, Xing L. Combined kV and MV imaging for real-time tracking of implanted fiducial markers a. Med Phys 2008;35:1191–8.
[18] Jaffray DA et al. Flat-panel cone-beam computed tomography for image-guided radiation therapy. Int J Radiat Oncol Biol Phy 2002;53:1337–49.
[19] Han K et al. A comparison of two immobilization systems for stereotactic body radiation therapy of lung tumors. Radiother Oncol 2010;95:103–8.
[20] Korreman S et al. The European Society of Therapeutic Radiology and Oncology-European Institute of Radiotherapy (ESTRO–EIR) report on 3D CT-based in-room image guidance systems: a practical and technical review and guide. Radiother Oncol 2010;94:129–44.
[21] Barney BM et al. Image-guided radiotherapy (IGRT) for prostate cancer comparing kV imaging of fiducial markers with cone beam computed tomography (CBCT). Int J Radiat Oncol Biol Phys 2011;80:301–5.
[22] Acharya S et al. Online magnetic resonance image guided adaptive radiation therapy: first clinical applications. Int J Radiat Oncol Biol Phys 2016;94:394–403.
[23] Fontanarosa D et al. Review of ultrasound image guidance in external beam radiotherapy: I. Treatment planning and inter-fraction motion management. Phys Med Biol 2015;60:R77.
[24] Kaar M et al. Automatic patient alignment system using 3D ultrasound. Med Phys 2013;40.
[25] Shirato H et al. Real-time tumour-tracking radiotherapy. Lancet 1999;353:1331–2.
[26] Schweikard A et al. Robotic motion compensation for respiratory movement during radiosurgery. Comput Aided Surg 2000;5:263–77.
[27] Budiharto T et al. Intrafractional prostate motion during online image guided intensity-modulated radiotherapy for prostate cancer. Radiother Oncol 2011;98:181–6.
[28] Zelefsky MJ et al. Improved clinical outcomes with high-dose image guided radiotherapy compared with non-IGRT for the treatment of clinically localized prostate cancer. Int J Radiat Oncol Biol Phys 2012;84:125–9.
[29] Campbell WG, Miften M, Jones BL. Automated target tracking in kilovoltage images using dynamic templates of fiducial marker clusters. Med Phys 2017;44:364–74.
[30] Moseley DJ et al. Comparison of localization performance with implanted fiducial markers and cone-beam computed tomography for on-line image-guided radiotherapy of the prostate. Int J Radiat Oncol Biol Phys 2007;67:942–53.
[31] Zhu X et al. Tradeoffs of integrating real-time tracking into IGRT for prostate cancer treatment. Phys Med Biol 2009;54:N393.
[32] Shah AP et al. An evaluation of intrafraction motion of the prostate in the prone and supine positions using electromagnetic tracking. Radiother Oncol 2011;99:37–43.
[33] Nichol AM et al. A magnetic resonance imaging study of prostate deformation relative to implanted gold fiducial markers. Int J Radiat Oncol Biol Phys 2007;67:48–56.
[34] Birkfellner W et al. Stochastic rank correlation: a robust merit function for 2D/3D registration of image data obtained at different energies. Med Phys 2009;36:3420–8.
[35] Gendrin C et al. Monitoring tumor motion by real time 2D/3D registration during radiotherapy. Radiother Oncol 2012;102:274–80.
[36] Furtado H et al. Real-time 2D/3D registration using kV-MV image pairs for tumor motion tracking in image guided radiotherapy. Acta Oncologica 2013;52:1464–71.
[37] Berbeco RI et al. Towards fluoroscopic respiratory gating for lung tumours without radiopaque markers. Phys Med Biol 2005;50:4481.
[38] Cui Y et al. Robust fluoroscopic respiratory gating for lung cancer radiotherapy without implanted fiducial markers. Phys Med Biol 2007;52:741.
[39] Hazelaar C et al. First experience with markerless online 3D spine position monitoring during SBRT delivery using a conventional linac. Int J Radiat Oncol Biol Phys 2018;101:1253–8.
[40] Hazelaar C et al. Feasibility of markerless 3D position monitoring of the central airways using kilovoltage projection images: managing the risks of central lung stereotactic radiotherapy. Radiother Oncol 2018;129:234–41.
[41] Lin T et al. Markerless gating for lung cancer radiotherapy based on machine learning techniques. Phys Med Biol 2009;54:1555.
[42] Zhang X et al. A kernel-based method for markerless tumor tracking in kV fluoroscopic images. Phys Med Biol 2014;59:4897.
[43] Zhang P et al. Design and validation of a MV/kV imaging-based markerless tracking system for assessing real-time lung tumor motion. Med Phys 2018;45:5555–63.
[44] Keall PJ et al. Review of real-time 3-dimensional image guided radiation therapy on standard-equipped cancer radiation therapy systems: Are we at the tipping point for the era of real-time radiation therapy? Int J Radiat Oncol Biol Phys 2018;102:922–31.
[45] Dahele M, Verbakel W. In regard to Keall et al.. Int J Radiat Oncol Biol Phys 2019;103:282–3.
[46] Keall PJ et al. In reply to Dahele and Verbakel. Int J Radiat Oncol Biol Phys 2019;103:283–4.

[47] Liu F et al. Deep learning MR imaging–based attenuation correction for PET/MR imaging. Radiology 2017;286:676–84.
[48] Ibragimov B, Xing L. Segmentation of organs-at-risks in head and neck CT images using convolutional neural networks. Med Phys 2017;44:547–57.
[49] Girshick R, Fast r-cnn. arXiv preprint arXiv:1504.08083, 2015.
[50] Ren S et al. Faster r-cnn: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 2017;39:1137–49.
[51] Simonyan K, Zisserman A, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.
[52] Lawrence I, Lin K. A concordance correlation coefficient to evaluate reproducibility. Biometrics 1989:255–68.
[53] Jaffray DA et al. Flat-panel cone-beam computed tomography for image-guided radiation therapy. Int J Radiat Oncol Biol Phys 2002;53:1337–49.
[54] Cui Y et al. Multiple template-based fluoroscopic tracking of lung tumor mass without implanted fiducial markers. Phys Med Biol 2007;52:6229.
[55] Shin H-C et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans Med Imaging 2016;35:1285–98.
[56] Wei Zhao, Shen Liyue, Han Bin, Yang Yong, Cheng Kai, Toesca Diego AS, Koong Albert C, Chang Daniel T, Xing Lei. Markerless pancreatic tumor target localization enabled by deep learning. Int J Radiat Oncol Biol Phys 2019. , https://www.redjournal.org/article/S0360-3016(19)30836-3/abstract.