HHS Public Access

Author manuscript

Proc SPIE Int Soc Opt Eng. Author manuscript; available in PMC 2022 September 21.

Published in final edited form as:

Proc SPIE Int Soc Opt Eng. 2022 ; 12034: . doi:10.1117/12.2611769.

Predicting 3D dose distribution with scale attention network for prostate cancer radiotherapy

Saba Adabi[*,a], **Tzu-Chi Tsen**[a], **Yading Yuan**[*,a]

[a]Department of Radiation Oncology, Icahn School of Medicine at Mount Sinai, NY, NYC, USA, 10029

Abstract

The growing demand for radiation therapy to treat cancer has been directed to focus on improving treatment planning flow for patients. Accurate dose prediction, therefore, plays a prominent role in this regard. In this study, we propose a framework based on our newly developed scale attention networks (SA-Net) to attain voxel-wise dose prediction. Our network ‘s dynamic scale attention model incorporates low-level details with high-level semantics from feature maps at different scales. To achieve more accurate results, we used distance data between each local voxel and the organ surfaces instead of binary masks of organs at risk as well as CT image as input of the network. The proposed method is tested on prostate cancer treated with Volumetric Modulated Arc Therapy (VMAT), where the model was training with 120 cases and tested on 20 cases. The average dose difference between the predicted dose and the clinical planned dose was 0.94 Gy, which is equivalent to 2.1% as compared to the prescription dose of 45 Gy. We also compared the performance of SA-Net dose prediction framework with different input format, the signed distance map vs. binary mask and showed the signed distance map was a better format as input to the model training. These findings show that our deep learning-based strategy of dose prediction is effectively feasible for automating the treatment planning in prostate cancer radiography.

Keywords

Deep learning; dose prediction; scale attention; prostate radiotherapy

1. INTRODUCTION

Radiotherapy treatment planning has been widely used for treating many cancers. The ability to accurately predict dose distributions where the planning target volume (PTV) receives the highest and organ at risk (OAR) receives the lowest amount of dose, bringing the opportunity to improve the quality of the treatment plans. Current Automated treatment planning solutions are included rule implementation and reasoning (RIR) [1] and knowledge-based planning modules (KBP) [2]. In RIR approach, a template is built based on OAR prescription and the plan is being optimized. To enhance OAR, a few auxiliary structures is being generated. While KBP Fits a regression model like PCA or SVM to predict new patients’ one-dimensional dose-volume histogram DVH and to create new

* saba.adabi@mountsinai.org and yading.yuan@mountsinai.org.

treatment planning using a large amount of patient data. Although the abovementioned method proven to be effective, one of the main problems with them was lacking spatial distribution information. To address the shortcoming of the above dose prediction approach on lacking spatial distribution information, the three-dimensional dose prediction methods were developed. Recently machine learning paradigms drew the researchers' attention to automate radiotherapy treatment planning. In particular, for KBP approach, prediction methods were developed to achieve volumetric dose prediction using Neural Network models [3–4]. Furthermore, researchers explored the application of Deep learning method for automated dose map prediction by developing different convolutional Neural network architecture such as Alexnet and transfer learning [5], variations of U-Net and /Dense-Net [6–8] and variant of Res-Net [9].

In this paper we proposed a three-dimensional dose prediction model utilizing a dynamic scale attention concept who integrates low-level details with high-level semantics from feature maps at different scales and distance maps information.

2. METHODS

Our method consists of our proposed network (SA-Net) architecture [12, 13] which in turn include the explanation of its encoding, decoding pathways and the scale attention block. The data preparation and preprocessing for extraction of geometrical features for the structures are given. The remainder are the implementation and evaluation criteria for predict dose image after training the model with fed data.

2.1 Network architecture

The network we used in this study is a 3D encoder-decoder network architecture named scale attention network (SA-Net). Our dose prediction network's inputs are OARs' distance maps, CT images and PTV distance map. The output is a 3d dose prediction. The network comprises of an encoder block, a scale attention block (SA-block) for learning and feature selection using full-scale information, and a decoder block. The backbone of our network is based on ResNet [10]. We used instance normalization and added a squeeze-and-excitation module into each residual block with a reduction ratio of 4 to improve the functionality of the proposed model. In this way we formed a ResSE block. The preliminary scale consists of one ResSE block with 24 features and then progressively divide equally the feature map dimension while doubling the feature width using a convolution (with stride = 2) at the first convolution layer of the first ResSE block in the adjacent scale level. For the rest of the remaining scales, two ResSE blocks are used except the endpoint of the encoding pathway. In the decoding pathway, the reverse pattern of encoding is used, instead with only one ResSE block in each spatial scale. After using a transpose convolution with a stride of 2 to double the feature map dimension and reduce the feature width by 2, The upsampled feature maps are then added to the output of the SA-block. For information fusion between the encoding and decoding pathways, the summation is used instead of concatenation in order to facilitate the information flowing and reduce memory consumption [11, 12]. Figure 1 outlines the schematic flow chart of SA-Net. The architecture of SA-Net. ResSE stands for a squeeze-and-excitation block embedded in a residual module [13]. the endpoint of the

encoding pathway has a dimension of 384 × 16 × 9 × 4. The output of the decoding pathway has one channel with the same spatial size as the input, i.e., 1×256 × 144 × 64.

2.2 Patient dataset and preprocessing

The database comprised of images retrospectively collected from 140 patients with prostate cancer treated with Volumetric Modulated Arc Therapy (VMAT) at Mount Sinai hospital under the supervision of specialists. A commercial treatment planning system (Eclipse, Varian Medical Systems, Palo Alto, CA) was used to provide the clinically delivered dose and OAR contours for each patient. The target dose prescription for all patients was 45 Gy. Simulation CT scans and organ contours of PTV, rectum, bladder, bladder wall, rectum wall, seminal vesicles and femoral heads were extracted from the clinical data set (Figure 2.) The structures were converted to label maps. The signed distance map of an organ is derived from the organ region that is contoured by a radiation oncologist, in which each voxel value represents the Euclidean distance of this voxel to the closest organ surface in 3D space. In order to prepare patient data for deep learning models, several preprocessing steps, including padding, expansion and cropping, were performed to frame each patient data in a 256×144×64 voxel tensor. Moreover, images were resampled from 2.5×2.5×3 mm to 3×3×3 mm grid. An in-house automated preprocessing framework developed for preprocessing purpose.

120 patients were randomly selected in the training phase, where we further divided it into a training set with 100 patients and a validation set with 20 patients. Given a set of patient clinical dose images and their corresponding predicted dose images, we use the Mean Absolute Error (MAE) as the loss function for our regression problem minimization through backpropagation.

2.3 Implementation

The proposed framework was implemented utilizing Python and PyTorch using CT scans and organ contours on a workstation with two NVIDIA GeForce 1080 TI graphics processing unit (GPU) with 11 GB memory. We used Adam optimizer [14] to update the weights with initial learning rate of 0.0003, and weight decay of 0.0001. The batch size was two and the optimization process ended at 60,000 iterations.

2.4 Evaluation

We evaluated the performance prediction model using distance map information and using only binary labels based on dose error. Dose error is voxel based mean absolute difference between ground truth and the predicted dose. The dose score is the mean dose error over all patients in validation or test set as given in equation.1. The lower is the dose score value, the better is the modeling.

$$A_h = \frac{1}{\mathcal{O}^h} \sum_{p \in \mathcal{O}^h} \frac{\left\| \mathrm{d}_p - \hat{d}_p \right\|_1}{|v^p|}, \forall h \in \left\{ \mathrm{val},\ \mathrm{test} \right\}, \forall p \in \mathcal{O}^{\mathrm{val}} \cup \mathcal{O}^{\mathrm{test}} \quad (1)$$

We also employed Gamma analysis [12], a quantitative similarity evaluation technique as another evaluation performance metric for our model. The Gamma between a dose-to-voxel $\tilde{d}_{\alpha,r}$ and a reference dose distribution d_α is defined as

$$\gamma\left(\tilde{d}_{\alpha,r}, d_\alpha\right) = \min_{r' \in M} \sqrt{\frac{|r - r'|^2}{\alpha^2} + \frac{\left|\tilde{d}_{\alpha,r} - d_{\alpha,r'}\right|^2}{\beta^2}} \quad (2)$$

Where $r' \in M$ is a search over a neighborhood of voxels in the reference dose d_α, α is the spatial distance threshold criterion and β is the dose difference threshold criterion. The higher is the gamma value the better is the modeling.

3. RESULTS

We applied our train model on the independent testing dataset including 20 prostate patients to further evaluate the performance of the SA-net prediction dose framework. Figure 3 shows an example of the predicted dose distribution by our proposed vs. clinical dose distributions for one patients.

Figure 4 shows some examples for PTV (yellow region), the prescribed 45 Gy isodose line from clinical plans (green dash line) and the predicted 45Gy isodose line (red dash line). For the testing dataset, the average predicted dose on PTV was found to be 46.85±0.74 Gy, while the planned dose was 46.81±0.78 Gy. Figure 5 shows an example of DVH comparisons for PTV and two OARs (Rectum Wall and Bladder Wall).

Table 1 summaries the dose constraints on the two OARs Rectum Wall and Bladder Wall. These results show that our proposed 3D method is similar to the clinical plan.

We also compared the performance of SA-Net dose prediction framework with different input format, i.e., signed distance map vs. binary mask. The calculated dose score was 0.94 when the signed distance maps were used as input channels, while it was 0.99 for binary masks.

For Gamma analysis, we set Gamma at 80% of prescription dose with a tolerance of dose difference threshold criterion of 5% and a spatial distance threshold criterion neighborhood of 5mm. the results indicate that proposed framework generated an average gamma pass rate of 97.48±2.12 in comparison where the binary labeled structure input channel used by average of 96.99±1.872. The results show the signed distance map was a better format as input to the model training. Note that for dose score, the lower the better; for gamma analysis, the higher the better.

4. CONCLUSIONS

In this study, we developed a 3D dose prediction framework based on SA-Net and signed distance maps. SA-Net replaces the long-range skip connections between the same scale in the U-Net with full-scale skip connections to utilize feature maps in full scales. Attention mechanism fine-tunes the weights of each scale feature to lay emphasis on the key scales

while suppressing that are not as important in an adaptive manner. SA-Net reduced the number of trainable parameters from 17.8M (U-Net) to 16.5M (SA-Net) providing a top performance with limited GPU resources. SA-Net has shown strong voxel-wise prediction capability in various image segmentation tasks including the 3rd place in MICCAI 2020 Brain Tumor Segmentation (BraTS) Challenge [12] using multi-parametric MRI images, and the 4th place in primary Head&Neck tumor segmentation using PET/CT image in MICCAI 2020 HECTOR Challenge [13]. The application of the proposed method on prostate radiotherapy showed that the average absolute dose difference between the predicted dose and the clinically delivered dose is small and acceptable. The findings of this study demonstrated that our deep learning-based strategy of dose prediction is successfully feasible for automating treatment planning in prostate cancer radiotherapy.

ACKNOWLEDGMENT

This work is partially supported by a research grant from Varian Medical Systems (Palo Alto, CA, USA), UL1TR001433 from NIH/NCATS, and R21EB030209 from NIH/NIBIB. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

REFERENCES

[1]. Purdie TG, Dinniwell RE, Fyles A, Sharpe MB “Automation and intensity modulated radiation therapy for individualized high-quality tangent breast treatment plans,” Int J Radiat Oncol Biol Phys. 90(3), 688–695 (2014). [PubMed: 25160607]

[2]. Wu B, Ricchetti F, Sanguineti G, Kazhdan M, Simari P, Jacques R, Taylor R and McNutt T “Data-driven approach to generating achievable dose-volume histogram objectives in intensity-modulated radiotherapy planning” Int J Radiat Oncol Biol Phys;79(4), 1241–1247 (2011). [PubMed: 20800382]

[3]. Shiraishi S and Moore KL 2015 Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy,

[4]. Campbell WG, Miften M, Olsen L, Stumpf P, Schefter T, Goodman KA and Jones BL “Neural network dose models for knowledge-based planning in pancreatic SBRT,” Med. Phys 44(12), 6148–6158 (2017). [PubMed: 28994459]

[5]. Kajikawa T, Kadoya N, Ito K, Takayama Y, Chiba T, Tomori S, et al. Automated prediction of dosimetric eligibility of patients with prostate cancer undergoing intensity-modulated radiation therapy using a convolutional neural network.

[6]. Nguyen D, Jia X, Sher D, Lin MH, Iqbal Z, Liu H and Jiang S, “3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture,” Phys. Med. Biol, 64(6), 65020 (2019).

[7]. Yuan Y, Tseng T, Chao M, Stock R, and Lo Y., “Predicting three dimensional dose distribution with deep convolutional neural networks,” Med. Phys, 44(6), (2017)

[8]. Yuan Y,T Tseng Lo. Y, “3D deep planning radiotherapy system and method,” U.S. Patent Application No. 16/633,567 (2020).

[9]. Liu Z, Fan J, Li M, Yan H, Hu Z, Huang P, Tian Y, Miao J and Dai J 2019 A deep learning method for prediction of three-dimensional dose distribution of helical tomotherapy Med. Phys 46 1972–83 [PubMed: 30870586]

[10]. Chen X, Men K, Li Y, Yi J and Dai J “A feasibility study on an automated method to generate patient specific dose distributions for radiotherapy using deep learning,” Med. Phys 46(1), 56–64 (2019). [PubMed: 30367492]

[11]. He K, Zhang X, Ren S and Sun J “Deep residual learning for image recognition”. Proc. CVPR 770–778, 16

[12]. Yuan Y, “Automatic brain tumor segmentation with scale attention network. In Crime A, Bakas S (eds) BrainLes 2020: Brainlesion: Glimoa, Multiple Sclerosis, Stroke and Traumatic Brain Injuries,”. LNCS. Springer, Cham., 12658 285–294 (2021).

[13]. Yuan Y “Automatic head and neck tumor segmentation in PET/CT with scale attention network in Head and Neck Tumor Segmentation,”. HECKTOR 2020. LNCS. Spinger, Cham. 12603 44–52 (2021).

[14]. Kingma DP and Ba J 2017 Adam: a method for stochastic optimization (arXiv:1412.6980)

[15]. Hu J, Shen L and Sun G, “Squeeze-and-excitation networks,” Proc. IEEE CVPR, 7132–7141 (2018)

[16]. Low DA and Dempsey JF, “Evaluation of the gamma dose distribution comparison method,”. Med. Phys 30(9), 2455–2464 (2003). [PubMed: 14528967]

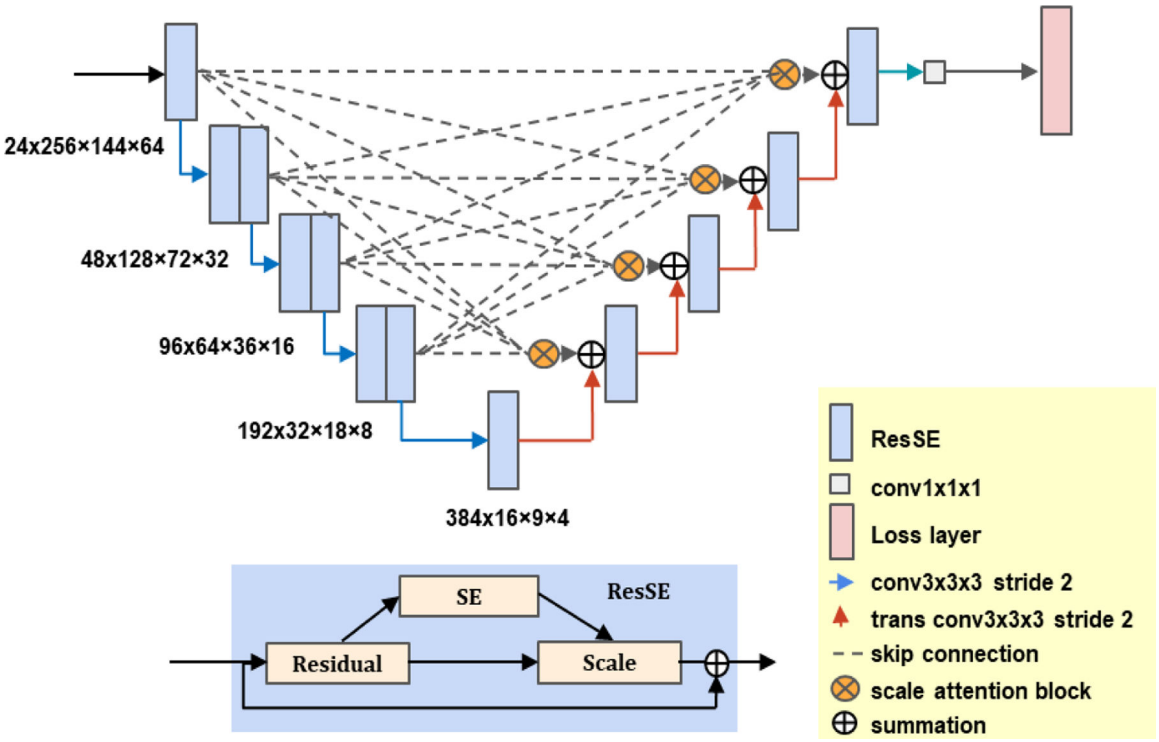


Figure 1.
schematic diagram of proposed network, Input is a 10× 256×144×64 tensor followed by one ResSE block with 24 features

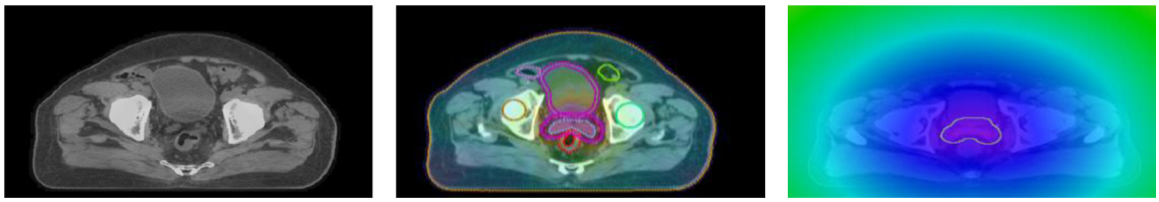


Figure 2.
Example of left: CT image, middle: CT contoured image including OARs and PTV and right: the distance map image of PTV.

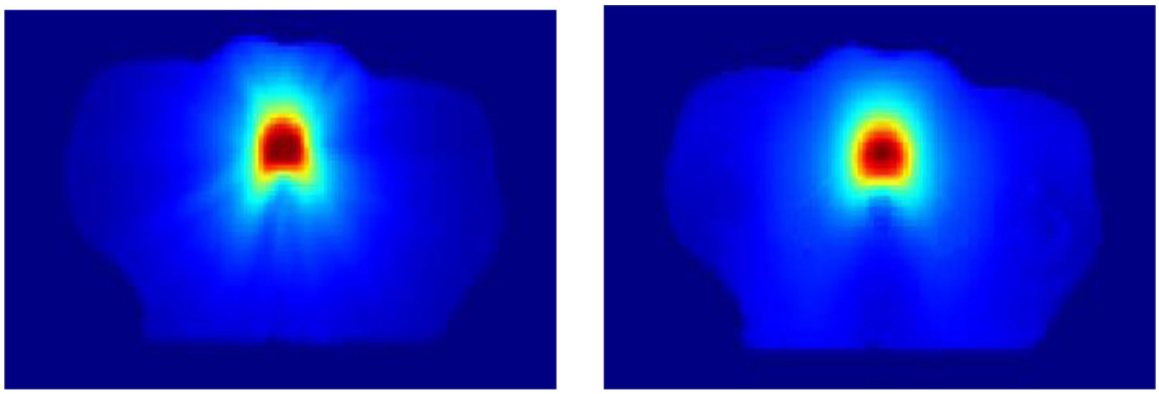


Figure 3.
Example of right predicted dose distribution vs. left clinical dose distributions

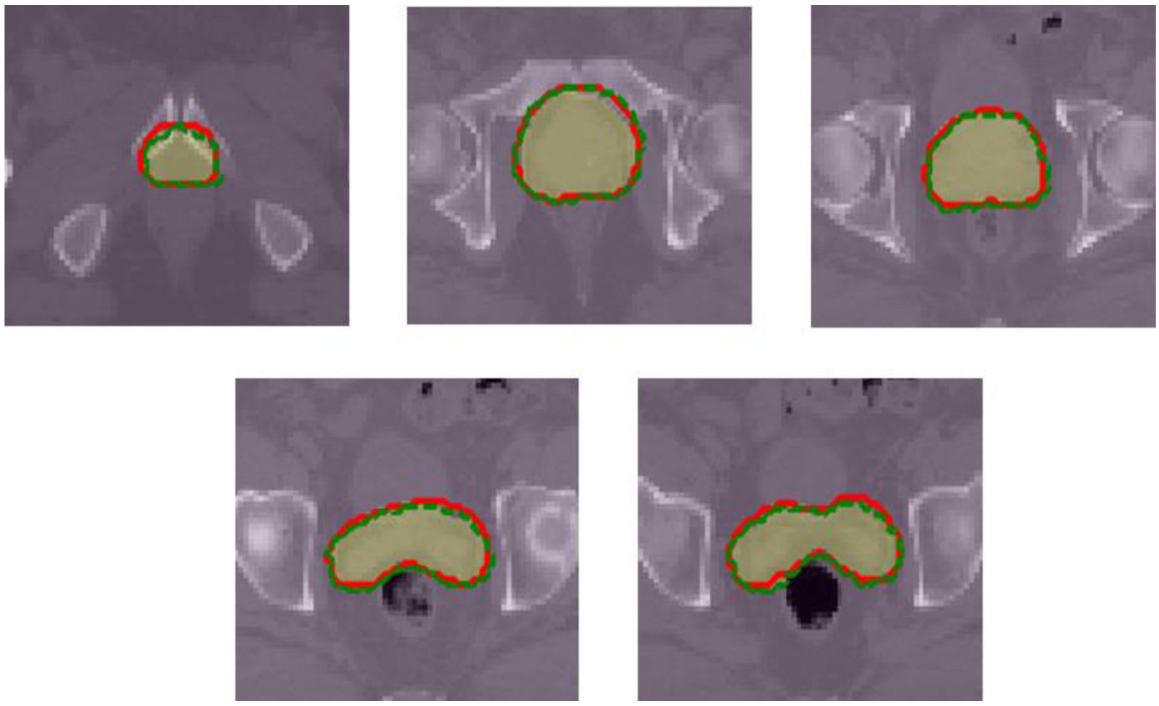
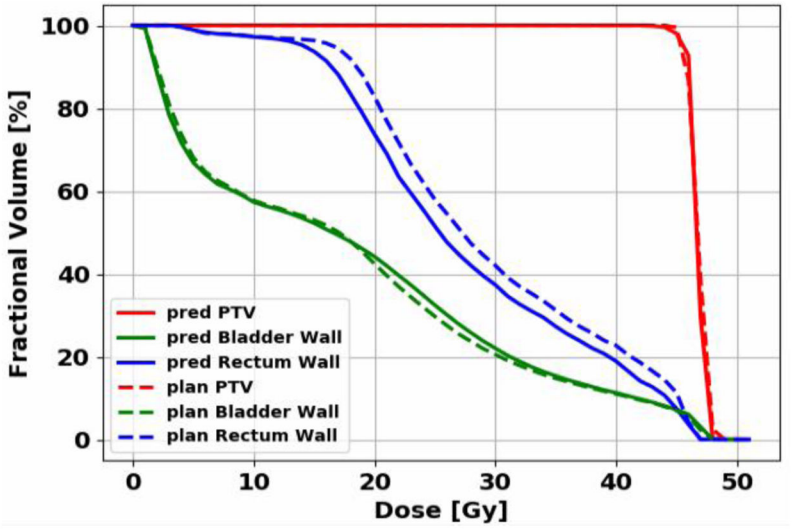
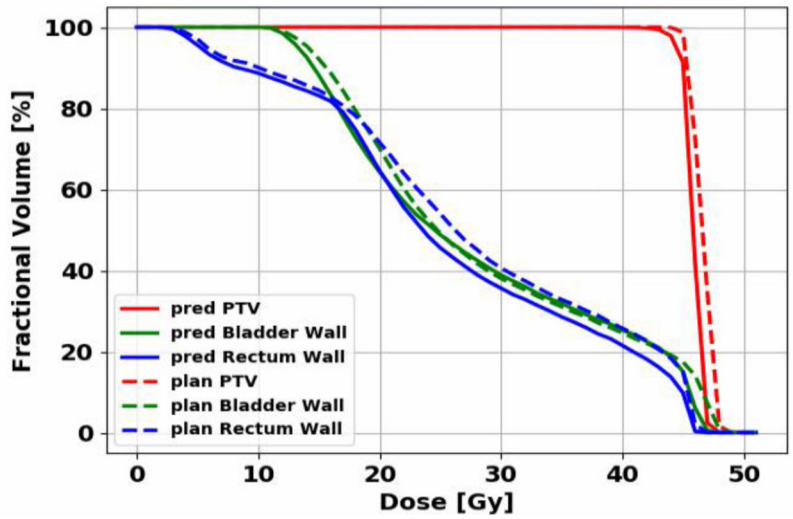


Figure 4.
Example of comparison between clinical plans and predicted dose.



(a)



(b)

Figure 5.

Example of DVH comparisons for two patients, solid lines show DVHs generated from clinical dose distributions, dashed line are generated from our predictions.

Table 1.

Dose Constraint comparison on rectum wall and bladder wall.

OAR Constraints		Clinical Plans	SA Net
Bladder Wall	$V_{30} < 50\%$	36.02	35.05
	$V_{40} < 30\%$	20.65	20.08
Rectum Wall	$V_{30} < 50\%$	36.73	33.01
	$V_{40} < 30\%$	21.84	16.72