

A Deep-Learning-based Multi-segment VMAT Plan Generation from Patient Anatomy for Prostate Simultaneous Integrated Boost (SIB) Cases: A Feasibility Study of Prostate Radiotherapy Application

by

Qingyuan Zhu

Graduate Program in Medical Physics
Duke Kunshan and Duke University

Date: _____

Approved:

Jackie Wu, Advisor

Fang-Fang Yin

James Bowsher

Thesis submitted in partial fulfillment of
the requirements for the degree of Master of Science in the Graduate Program in
Medical Physics in the Graduate School of Duke Kunshan and Duke University

2021

ABSTRACT

A Deep-Learning-based Multi-segment VMAT Plan Generation from Patient Anatomy for Prostate Simultaneous Integrated Boost (SIB) Cases: A Feasibility Study of Prostate Radiotherapy Application

by

Qingyuan Zhu

Graduate Program in Medical Physics
Duke Kunshan and Duke University

Date: _____

Approved:

Jackie Wu, Advisor

Fang-Fang Yin

James Bowsher

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Graduate Program in Medical Physics in the Graduate School of Duke Kunshan and Duke University

2021

Copyright by
Qingyuan Zhu
2021

Abstract

Introduction: Several studies have implemented deep learning (DL) based fluence-map-predictions for Intensity Modulated Radiation Therapy (IMRT) treatment planning. However, DL-based Volumetric Modulated Arc Therapy (VMAT) planning remains a challenging problem. One of the main difficulties in DL-based VMAT planning is the conversion/generation of leaf sequences from the predicted fluence intensity maps. Leaf sequences are required for a large number of control points, and they also need to meet physical restrictions of the multi-leave collimator (MLC). This study aims to improve the performance of a DL-based automated VMAT planning algorithm by generating MLC leaf sequences from predicted intensity maps using differentiated threshold levels. Based on this study, we developed an algorithm to convert DL-predicted intensity maps to multi-segment VMAT plans to improve the performance of these one-arc plans.

Methods: Our deep learning model utilizes a series of 2D projections of a patient's anatomy and contour structures to generate dynamic MLC sequences with 7-8 segments in a VMAT plan. The backbone of this model is a novel U-net implementation that has a 4-resolution-step analysis path and a 4-resolution-step synthesis path. In the DL model, a total of 131 patients were involved, with 120 patients in the training and 11 patients in the testing groups, respectively. These patients were prescribed with 70Gy/58.8 Gy to the primary/boost PTVs in 28 fractions in a simultaneous integrated boost (SIB) regime. In this study, the predicted

intensity maps are separated into 7-8 segments along the collimator angle. Hence, the arcs could separately simulate predicted intensity maps with independent weight factors. This separation also potentially allows MLC leaves to simulate more dose gradient in the predicted intensity maps.

Results: After dose normalization ($PTV_{70} V70Gy=95\%$), all 11 multi-segment test plans met institutional clinic guidelines of dose distribution outside PTV. Bladder ($V70Gy=5.3\pm 3.3\%$, $V40Gy=16.1\pm 8.6\%$) and rectum ($V70Gy=4.5\pm 2.3\%$, $V40Gy=33.4\pm 8.1\%$) results in multi-segment plans were encouraging compared with one-arc plans generated by U-Net prediction. 3D max dose results in plans ($D1cc=112.6\pm 1.9\%$) were comparable to one-arc plans ($D1cc=106.7\pm 0.8\%$). On average, our algorithm used 600 seconds for a plan generation in contrast to the current clinical practice (>20 minutes).

Conclusion: Results suggest that multi-segment plans can generate a prostate VMAT plan with clinically acceptable dosimetric quality. The proposed multi-segment plan generation algorithm has the capability to achieve higher modulation and lower maximum dose. With its high efficiency, this technique may hold great potentials for a real-time planning application in future implementation.

Keywords: real-time planning, deep learning, treatment planning

Contents

Abstract	iv
List of Tables	viii
List of Figures	ix
Acknowledgments	x
1. Introduction	1
1.1 Volumetric Modulated Arc Therapy	2
1.2 Simultaneous Integrated boost technique and its application on VMAT	4
1.3 Methods for realizing automated planning	5
1.4 Innovation and Impact	7
1.5 Research Objectives	7
2. Materials and Methods	9
2.1 Materials	9
2.2 General Workflow	9
2.3 Ground Truth VMAT plans generation	10
2.4 U-Net Dose map Prediction	12
2.5 MLC Position and Weighting Algorithm	13
2.6 Plan settings	18
2.7 Model Validation	18
3. Results	20
3.1 Isodose Distributions	20
3.2 DVH statistics	25
4. Discussion	28

4.1 Tasks and goals of Automated VMAT planning	28
4.2 Limitations.....	29
4.3 Future research	30
5. Conclusion	31
References	32

List of Tables

Table 1: Ground truth generation, Step 1: Optimization objectives (PTV-only) for pilot planning	11
Table 2: Ground truth generation, Step 2: Optimization objectives (OAR included) for pilot planning	11
Table 3: Plan validation indices	19
Table 4: Major dosimetric parameters comparison	25

List of Figures

Figure 1: Progressive sampling for VMAT planning.....	4
Figure 2: Overall workflow	10
Figure 3: U-Net architecture	12
Figure 4: Prediction and Ground truth segmentation	14
Figure 5: Thresholds of different segments.....	15
Figure 6: Loss change over trials.....	17
Figure 7: Aperture shape comparison between multi-segment plan, one arc plan and ground truth plan. (Blue contour represents PTV58.8)	20
Figure 8: Aperture shape comparison between one-arc plan and ground truth plan of four different control points.	21
Figure 9: Isodose line comparison between Multi-segment plan (left) and one-arc plan (right)	22
Figure 10: Comparison of Isodose line (sagittal view) of other 2 plans	23
Figure 11: Hot spot difference between plans with different jaw openings (0.1cm vs. 0.4cm).....	24
Figure 12: Dose-volume histogram comparison between Multi-segment and one-arc plans.....	26
Figure 13: Major Dose statistics comparison (box plot).....	27
Figure 14: Loss of details in binary masks when applying CC method.....	30

Acknowledgments

I would like to thank my supervisor, Dr. Q. Jackie Wu, for her guidance through each stage of the process. I am so much appreciated her patience and insightfulness during my research process. I would also like to acknowledge Dr. James Bowsher, Dr. David Huang, and Dr. Xin Chen for inspiring my interest in the development of innovative technologies and gave me a solid foundation for my research work. Ph. D student Xinyi Li and senior student Yimin Ni, gave me so much help during my research process in such a COVID-19 pandemic. It is really a special experience for me to work remotely. During this period, I learned a lot, and I got lots of help from the professors, students, staff, Duke medical physics program, and Duke Kunshan medical physics program. Finally, I would like to thank my family, without their support, I cannot go this far and explore such a beautiful world.

1. Introduction

2020 WHO annual report¹ states that cancer has become the second leading cause of death in the world, accounting for approximately 1/6 of death globally. Various treatment technologies, such as surgery, radiation therapy, chemotherapy, immunotherapy, gene therapy, were invented and applied since the 20th century. Many researchers have been making explorations continuously in this field. At present, among all these treatment modalities, external beam radiation therapy (EBRT) is one of the most commonly adopted modalities for cancer treatment. According to a report in 2017, EBRT takes up to more than 50% of all the cancer treatment in developed countries and is playing a more and more important role in developing countries². The major task of EBRT is to lower the dose delivered to the organ at risk and ensure there is enough dose delivered to the tumor volume. In order to satisfy this goal, several treatment techniques have been emerging during the past few decades, including Three-Dimensional Conformal Radiation Therapy (3D-CRT), Intensity-Modulated Radiation Therapy (IMRT), and Volumetric Modulated Arc Therapy (VMAT)³. Due to the essence of dynamic leaf sequence and continuous beam delivery of the VMAT, the design of patient-specific treatment in the current clinic workflow requires a lot of time and computational power. The design of the plans requires extensive effort from the planners and the oncologists, and the quality of the resultant plans depends on planners' and oncologists' personal experience to a great extent, and it may take a long time to modify and re-calculate the VMAT plan. Therefore, despite the high efficiency of dose delivery and good

dose distribution, the overall efficiency of VMAT still needs to be improved. Many research and efforts have been carried out to solve the vital problem. In recent years, with the fast development of artificial intelligence, 3D dose predictions for radiotherapy^{4,5} were considered as a promising technique for carrying out inverse planning with higher efficiency. Additionally, several recent studies tried to generate MLC sequences directly from the input dose distributions or DVH statistics and have got really encouraging outcomes^{6,7}. However, inverse optimization and manual fine-tuning of the plan are still needed as post-processing to reach a deliverable plan. Therefore, a highly automated treatment planning algorithm has not yet been realized. In this study, we proposed an innovative work that separates the predicted fluence intensity map into multiple segments that cover different PTV regions of SIB treatment. In this prostate VMAT planning application study, the developed multi-segment VMAT plans meet the clinical standards with fully automated implementation and high execution efficiency. The proposed multi-segment VMAT plan generation algorithm has the potentials to achieve higher modulation and lower maximum dose.

1.1 Volumetric Modulated Arc Therapy

VMAT is an advanced radiation therapy technique that is capable of generating plans with highly conformal dose distributions. The completion of this concept also went through an evolution process. It is Yu⁸ that first developed an intensity-modulated arc therapy (IMAT) technique that uses the MLC to dynamically shape the fields and rotate the gantry in the arc therapy mode. However, due to its

constant linac dose rate during the delivery, it would take several arcs to deliver the dose if the dose distribution is complex at one gantry angle. Therefore, many efforts have been made to find optional dose rate rotational delivery methods. To meet the requirements of the dose delivery technique above, The TPS need to compute the dose by sampling the delivery at a large number of gantry angles, which also brings a big problem for the TPS to optimize the MLC leaf motion with many parameters and the physical constraint of MLC leaves. In 2007, Otto³ first put forward an optimizing method called progressive sampling. This improvement in the optimization method has been made possible by using only one arc for VMAT plans, which significantly reduced the time required for VMAT delivery.

The "progressive sampling technique" is as shown in Otto's introductory paper³. Figure 1 displays an example of how progressive sampling is used to optimize a VMAT delivery. The gantry angle, dose distribution, and the physical constraint of leaf motion are the parameters and restrictions for optimization. In this approach, the arc is subdivided into a small number of gantry angle samples, whose beam shapes or intensities are varied during the first iterations of the optimization. There are few restrictions on MLC leaf position for these fields since the initial samples are far apart. After several iterations, a new arc sample is added with a field shape interpolated between the first two samples (e.g., Figure 1b), and optimization continues. After more iterations, another new arc sample is added in the same fashion midway between samples 2 and 3. This process continues until the gantry

angle sampling meets the desired standard.

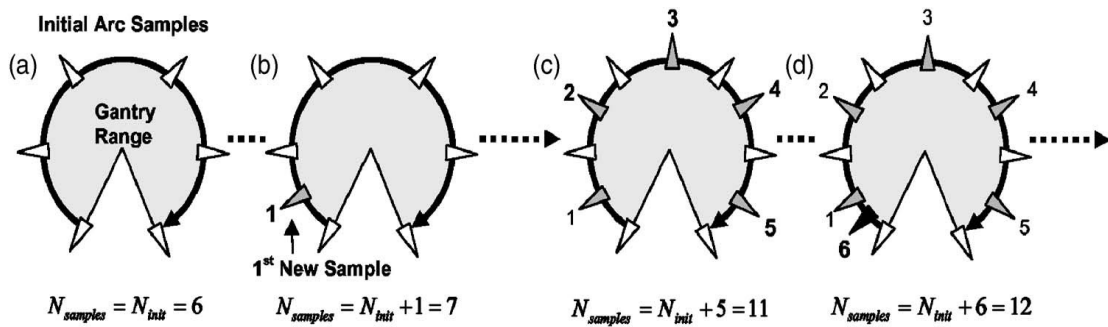


Figure 1: Progressive sampling for VMAT planning⁹

Several preliminary studies have demonstrated equivalent dosimetric quality is adequately sampled (every 2 degrees).¹⁰⁻¹² In treatment sites with a relatively simple geometry like the prostate, the VMAT outperformed the other two techniques.¹¹ With equivalent treatment quality and higher delivery efficiency, VMAT soon gained popularity in clinical practices.

1.2 Simultaneous Integrated boost technique and its application on VMAT

The simultaneous boost technique¹³ in radiotherapy consists of delivering the boost treatment (additional doses to reduced volumes) simultaneously with the primary (large-field) treatment (primary PTV) for all treatment sessions. In the early days, when the traditional EBRT technique was the most popular, boost fields are combined with the large fields over the whole treatment time. The boost dose is delivered simultaneously (i.e., or within the same few minutes) with the large-field dose for all fractions. Suppose in the original technique, the large-field technique consists of two opposing open fields, and the subsequent boost dose is given with

two angled beams with wedges. In that case, four fields will be used for each treatment session in the simultaneous boost technique by combining the two open fields and the two wedged fields with appropriate weighting factors. By applying this technique, the total dose delivered by the large fields will be spread over all fractions, and a smaller dose per fraction can be used.

The Simultaneous Integrated Boost (SIB) technique becomes more and more popular when modern radiation therapy techniques were involved. The ever-growing computational power also promoted its development in the radiation therapy field. Several studies explored the SIB technique's application on different regions and treatment techniques¹⁴⁻¹⁶.

The delivery efficiency of VMAT makes it especially appealing for clinical practices. In particular, the total delivery time of a one-arc VMAT plan is less than five minutes³. This fast implementation could alleviate the timing stress caused by stringent requirements for patient positioning and image guidance of the current EBRT practice. Several trials have been made to evaluate the performance of VMAT on delivering the SIB technique, and most of them got positive and promising outcomes¹⁷⁻¹⁹.

1.3 Methods for realizing automated planning

Though the VMAT-SIB technique has a great potential to be a highly effective method for fast dose delivery, it cannot be dismissed that it would take a long time for physicians and physicists to design a highly conformal treatment plan. A large number of control points and the time-consuming inverse-planning process would

take up around 20-30 minutes to make a clinical plan that meets all the standards.

Knowledge-based planning²⁰⁻²¹ (KBP) and artificial intelligence (AI) technologies are introduced into RT treatment planning to accelerate the process.

As we can see from its name, knowledge-based planning involves extracting planning knowledge from prior plans to facilitate the design of new plans. Key anatomical and dosimetric features such as OAR volumes and their dose-volume statistics are computed as references, and a machine learning model is used to predict the optimization objectives of the new patient's plan.

Due to the burst of development in Artificial Intelligence, researchers are finding solutions for problems that seem impossible to solve in the past, including Natural Linguistic Processing (NLP) and the automation of many manual tasks. AI has quickly found its application in the medical field. Several deep learning models for segmentation or diagnosis have shown good efficiency and accuracy.²²⁻²⁴ For example, AI-based algorithms have outperformed experienced physicians on finding and recognizing lung nodules²⁵. And this emerging technology has also been incorporated in the daily quality assurance (QA) procedures in radiation oncology²⁶⁻²⁷. It is also playing a more and more important role in generating RT plans. Several examples of the application of deep learning in treatment planning are summarized below.

Nguyen *et al.* reported a U-Net-based model for dose distribution prediction of prostate IMRT.⁴ The resultant dose distributions are shown to be able to guide the optimization process. Another recent work has demonstrated the efficacy of novel

deep learning architecture, DoseNet, for prostate SBRT dose prediction.²⁶ Besides dose prediction, IMRT fluence map prediction is another research area that has gained increasing popularity. Lee *et al.* developed a deep neural network to generate fluence maps directly from the organ contour and dose distribution, eliminating the time-consuming inverse planning process.²⁷ In addition, there are new proposed integrated models with high automation to perform treatment planning automatically²⁸⁻³⁴.

1.4 Innovation and Impact

This study proposes a novel multi-segmented leaf sequencing algorithm to contribute to AI-based automatic prostate treatment planning. The algorithm has demonstrated encouraging dosimetric outcomes in test cases. Future studies will further explore more leaf sequencing possibilities to improve VMAT plan quality.

1.5 Research Objectives

In the previous study⁶, a one-arc delivery method was introduced. Due to the stochastics in DL network training, the predicted intensity maps do not necessarily meet all physical limitations of MLC, which was implied in the training ground truth. Additionally, dose gradients, an important feature in the SIB cases, are often missed or misplaced in the generated one-arc plans, which means the dose fall-off between the boosted PTV and primary PTV cannot be reflected accurately. In clinical cases, if a patient's seminal vesicle is large, directly applying the MLC aperture of the one-arc plans would lead to a lack of dose in PTV_{58.8} areas. In this study, multiple segments with the same collimator angle simulate the predicted intensity maps by

assigning different beam weight values to each segment. This study may facilitate the balance of target coverage and OAR sparing, thus improving plan quality.

2. Materials and Methods

2.1 Materials

The dataset consisted of 131 anonymized patient plans, with 120 patients in the training group and 11 patients in the testing groups, respectively. These patients were prescribed 70Gy/58.8Gy to the primary/boost PTVs in 28 fractions in a simultaneous integrated boost (SIB) regime. In this study, PTV₇₀ was defined as the prostate volume plus 5mm margin, and PTV_{58.8} was defined as the sum of prostate and seminal vesicles with a 5mm margin. In the plan quality analysis, we took rectum and bladder as our OARs.

In this study, treatment planning was performed using Eclipse™ v13.7 software (Varian Medical System, Palo Alto, CA). Dose calculations were carried out with the AAA (Anisotropic Analytical Algorithm) with 1mm grid resolution.

2.2 General Workflow

This study aims to improve the dosimetric quality of the plan generated from the U-Net predicted intensity map. The study workflow can be divided into three sequential steps. (a) Generate the MLC sequences from U-Net predicted intensity map. (b) separate the whole sequence into 7 or 8 segments (c). Modify the arc weight of these segments using a simulated annealing algorithm. An overview of the workflow is presented in Figure 2. This workflow is the same as Ni et al.'s work⁶ until MLC leaf sequencing, which is the focus of this study.

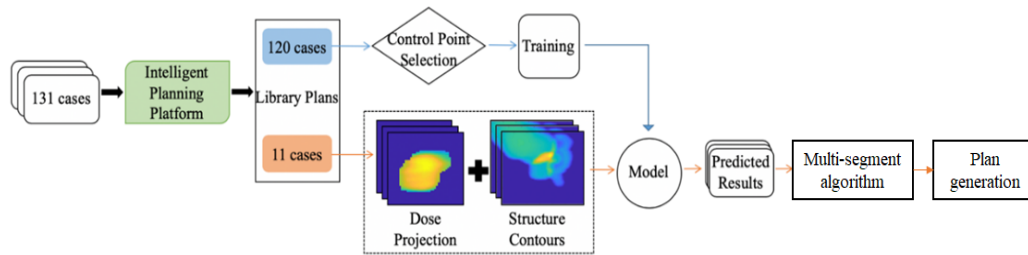


Figure 2: Overall workflow

The customized U-Net input was 2D projections generated from patient anatomy for the predefined 100 control points at specific gantry angles. The ground truth was generated from MU-weighted MLC apertures in the library one-arc VMAT plans. The network predictions were then interpolated to the intensity maps at 178 control points, which were subsequently resampled and sent to TPS using an automated scripting interface for plan finalization. Finally, the generated plans were compared with Ni et al.'s original DL plans⁶ for plan quality validation.

2.3 Ground Truth VMAT plans generation

In Ni et al.'s work⁶, the ground truth plans were generated using systematic adjustments of optimization objectives. Specifically, the process includes two steps. The first step is to generate dose distributions without considering OAR constraints. All the plans are normalized to that 100% of the prescription dose (70Gy) covers 95% of the PTV₇₀ volume. Structures that are not mentioned before are listed below:

- i. PTV_{58.8exp}: Represents PTV_{58.8} with one extra slice at the top. The extra slides were shrunk by 5 mm from the edge of the original top slice
- ii. Z_P1minusP2: PTV_{58.8exp} – Slices that contains PTV₇₀ component (only the upper portion).

The optimization objectives are listed in Table 1⁶:

Structure	Upper or lower	Dose (cGy)	Volume (%)	Weight
PTV ₇₀	Lower	7000*0.96	99.9	110
PTV ₇₀	Lower	7000	97.5	115
PTV ₇₀	Upper	7000*1.04	1.5	110
PTV ₇₀	Upper	7000*1.05	0	115
BODY	Upper	7000*1.05	0	350
PTV _{58.8}	Lower	5880*0.97	99.9	110
PTV _{58.8}	Lower	5880	98	115
PTV _{58.8}	Upper	7000*1.05	0	110
Z_P1minusP2	Upper	7000*92%	0.0	55
Z_P1minusP2	Upper	7000*86%	50	55

Table 1: Ground truth generation, Step 1: Optimization objectives (PTV-only) for pilot planning

The second step is to take OAR into considerations. (Here, FHR represents Right Femoral Head, and FHL represents Left Femoral Head). More objectives based on DVH statistics were introduced to refine the OAR dose constraints. The objectives needed are listed in Table 2⁶:

Structure	Upper or lower	Dose (Gy)	Volume (%)	Weight
Bladder	Upper	71.0	0	80
	Upper	70.3	0.1	75
	Upper	$D_{50\%}^{pilot} * 0.85$	50	50
	Upper	$D_{70\%}^{pilot} * 0.85$	70	50
Rectum	Upper	70.8	0	80
	Upper	70.7	0.1	80
	Upper	$\text{Min}(D_{50\%}^{pilot} * 0.70, D_{50\%}^{pilot} - 100)$	50	80
	Upper	50	$\text{Min}(V_{50Gy}^{pilot} * 0.6, 20\%)$	70
	Upper	21	$V_{21Gy}^{pilot} * 0.6$	70
FHR	Upper	35	0	55
FHL	Upper	35	0	55

Table 2: Ground truth generation, Step 2: Optimization objectives (OAR included) for pilot planning

2.4 U-Net Dose map Prediction

In this study, we used the same U-Net architecture⁶ to capture the patient anatomic features and generate the predicted intensity maps and thus generating MLC sequences for the VMAT plans. U-Net was chosen as our deep learning model because it has proven to be an efficient deep learning network in the field of medical image processing that can achieve a balance between accuracy and efficiency. Additionally, U-Net has the capability of performing successful image segmentation with limited training samples³⁶. Since our data set is not that large, U-Net is suitable for this study from the perspective of the data scale.

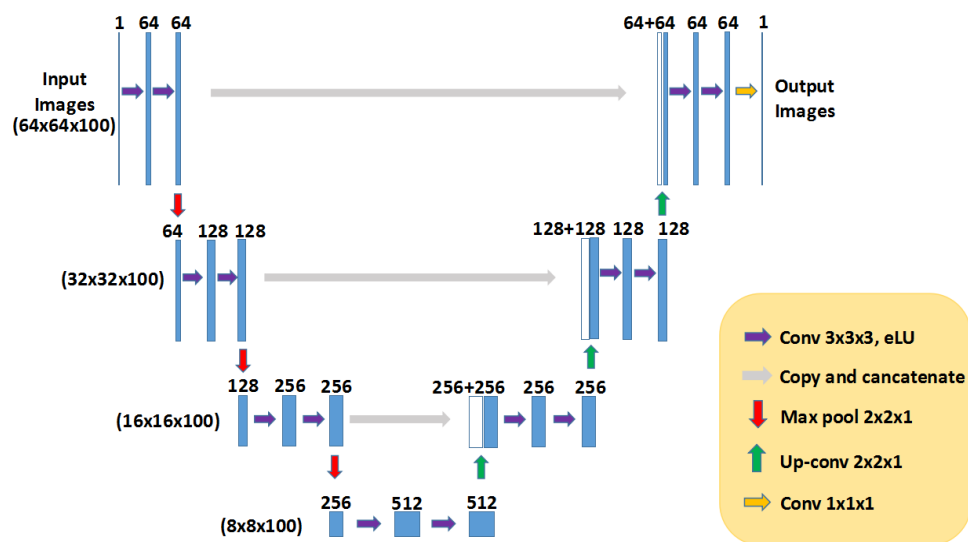


Figure 3: U-Net architecture⁶

Figure 3 demonstrates the U-Net architecture used in this study. The network consists of a 4-resolution-step contracting path (left side) and a 4-resolution-step expansive path (right side). The contracting path follows the typical architecture of a convolutional network. It consists of the repeated application of two $3 \times 3 \times 3$ convolutions, each followed by an exponential linear unit (eLU) and a $2 \times 2 \times 1$ max

pooling operation with stride 2 for down-sampling. Every step in the expansive path consists of an up-sampling of the feature map. Concatenations are performed between three pairs of units with the same resolution as indicated by the gray arrow in the figure. Then two $3 \times 3 \times 3$ convolutions were performed, each followed by an eLU. In the final layer, a 1×1 convolution is used to combine the 64-channel layer into a one-channel output. In total, the network has 23 convolutional layers.

The input of the U-Net architecture is a 4-D matrix with six channels. The raw data is generated from performing Radon transform on organs and PTVs of patients' CT slices. The channels are 1. Bladder, 2. Rectum, 3. PTV 70, 4. PTV 58.8, 5. Seminal vesicle, 6. Digital Reconstructed Radiographs (DRR) of the pelvic region. The outputs are binarized MLC aperture shapes at the 100 selected control points with a dimension of $64 \times 64 \times 100$.

2.5 MLC Position and Weighting Algorithm

The previous study has shown the feasibility of using U-Net to generate VMAT plans that can meet most of the clinical standards⁶. However, some dosimetric indices could still be improved, such as maximum dose, $PTV_{58.8}$ coverage, dose distribution in the seminal vesicles, etc. In this study, we plan to divide the U-Net-predicted intensity maps into 7-8 segments and then improve the max dose, DVH statistics by optimizing each segment's weight.

Specifically, the size of the prediction matrixes is $64 \times 64 \times 100$ (2.5 mm/pixel). After loading the predicted intensity maps into MATLAB, the MLC sequences generation follows these steps:

(a). Interpolate the $64 \times 64 \times 100$ input into $75 \times 75 \times 100$ (2.5 mm/pixel both).

Divide the predicted intensity map in the PTV into eight segments. Each segment corresponds to two pairs of MLC leaves.

(b). Read the ground truth MLC sequences and separate the PTV into eight segments in the same way as the prediction dose map. (Figure 4)

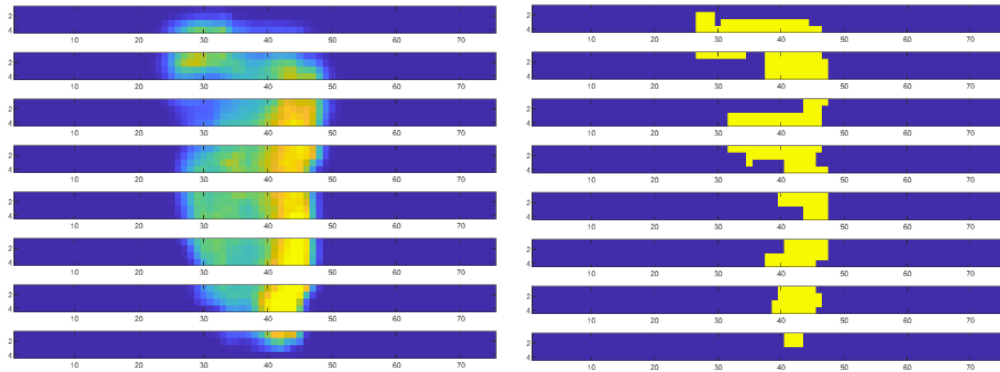


Figure 4: Prediction and Ground truth segmentation

In this step, the common method for performing segmentation is to separate the images from the 20th row to the 51st row (75 rows in total); each segment contains four rows. In some cases, it's possible that the patient's prostate is extremely large and exceeds the range of the segmentation. If it happens, we would add 1-2 rows in the first and/or last segments and make sure the PTV is fully covered.

(c). For each control point, we set 450 different constants (from 0.005 to 0.500, 0.001 for each step) as thresholds. Then we apply the threshold x to the formula:

$$\text{Segmented pixel map} > x \quad (1.1)$$

The original pixel dose map would become a binary mask. Cross-correlation would be calculated between the ground truth and the binary map. The thresholds that correlate to the max cross-correlation values would be recorded as the control

point's optimized threshold values x' . The minimum threshold of 0.005 is set in case that the aperture opening that is extremely large.

(d). After finishing all the control point threshold values' calculation, we apply the optimized threshold values to the predicted intensity maps that have the same size as the CT images (512×512×178, 1 mm/pixel, linearly interpolated from the 75×75×100 predicted intensity maps) and get a new binary mask. This mask would be converted to MLC leaf positions directly. The mask is generated by:

$$Mask(i,j) = \begin{cases} 1, & P(i,j) > x \\ 0, & elsewhere \end{cases} \quad (1.2)$$

Here the $P(i,j)$ is the segmented pixel map with a resolution of 1mm/pixel. In the binary mask $Mask(i,j)$, we search the first non-zero pixel from left to right and record this pixel location as the left MLC position, then search the first non-zero pixel

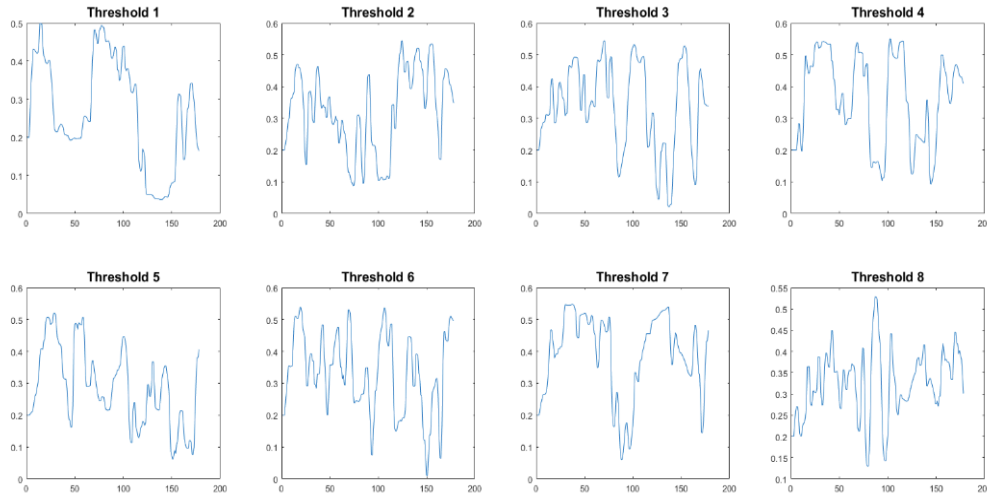


Figure 5: Thresholds of different segments

from right to left and record the pixel location as the right MLC position. Then the MLC position is written into Dicom files and imported into the Eclipse™ TPS system.

The default field weight factor is set as 1.

(e). In order to fine-tune the beam weight factor and improve the dose conformity, lower the hot spot, and fully cover the PTV_{58.8}, we would introduce a simulated annealing algorithm to modify the beam weight. The simulated annealing algorithm is an effective way to find the global optimum of a given function. The process is shown as follow:

1). Define a loss function L

L is the definition that includes objectives to be optimized. In this study, we define the loss function as:

$$L = \text{ReLU}(PTV_{58.8} - 0.84) * 25 + (SV - 0.88) * 20 + (body1cc - 1) * 5 + (body1cc - 1.08) * 25 + (rectum1cc - 1) * 5 + (body1cc - 1.08) * 25 + (bladder1cc - 1) * 5 + (bladder1cc - 1.08) * 15$$

Here, ReLU represents Rectified Linear Unit, SV represents the seminal vesicle. All the organ and OAR dose values are relative to the prescription. We add an SV penalty to prevent SV from receiving an excessive dose. Also, a two-stage penalty is added to organs and OARs in order to avoid hot spots and extremely large high-dose areas.

2). Select current state w

w is the initial state of the weight factor, whose numerical value is 1.0 for each segment.

3). Add a random perturbation to state w and calculate the new loss

To ensure convergence, exponential decay is applied to the amplitude of the perturbation (i is the step-index from 1 to 400, $rand$ means a random number between 0 to 1):

$$(rand - 0.5) * (0.0001 + \exp(-0.03 * i) * 0.1) \quad (1.3)$$

4). Renew the state w

If the loss is smaller than that in the previous state, accept the move; otherwise, we also accept the perturbation when an exponentially decreasing probability (given by prior experience) is smaller than the value of $rand$ in the current step:

$$P = \exp(-0.03 * i) * 0.3 \quad (1.4)$$

The algorithm stops after 400 iterations.

5). Start next trial

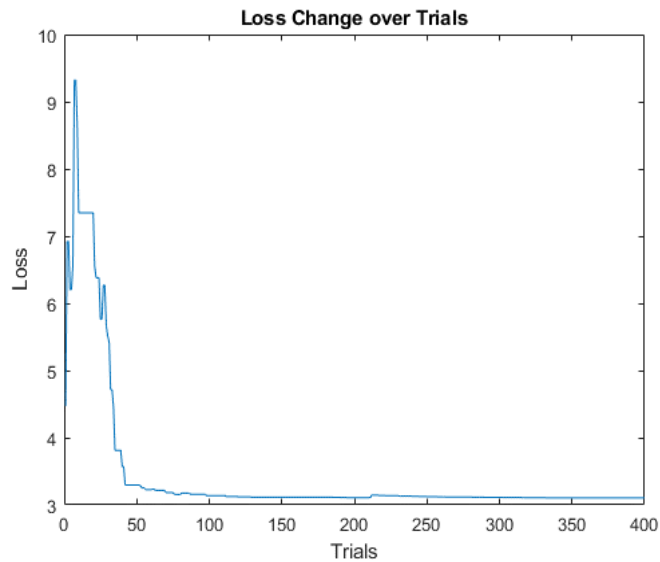


Figure 6: Loss change over trials

For each plan, the beam weight fine-tuning algorithm tuned the beam weight factor. The modified RT plan is the final plan and then calculated by TPS. The

simulating result of one case is shown in Fig 6. We can see that the optimization converged.

2.6 Plan settings

This set of plans use 10X as the beam energy and 600 MU/minute as the dose rate. The gantry rotation angle starts from 179° to 181° counter-clockwise. The collimator rotation angle is set as 5°.

The jaw opening in the Y direction for each segment is 1.2 cm. (two pairs of MLC take up 1cm, and 0.1cm margin is added on both sides of the MLC leaves).

2.7 Model Validation

To validate the efficacy and accuracy of the current model, ten indices are chosen to evaluate the dosimetric quality of the VMAT plan, as listed in Table III.

Table III.

Indices	Unit	Description
$V_{58.8Gy}^{PTV58.8}$	(%)	The percent volume of PTV _{58.8} covered by 58.8Gy isodose line
D_{1cc}^{max}	(%)	The maximum percent dose received by 1cc of patient body
D_{1cc}^{rectum}	(Gy)	The maximum absolute dose received by 1cc of the rectum
V_{70Gy}^{rectum}	(cc)	The absolute volume of rectum receiving more than 70Gy dose
V_{65Gy}^{rectum}	(%)	The relative volume of rectum receiving more than 65Gy dose
V_{40Gy}^{rectum}	(%)	The relative volume of rectum receiving more than 40Gy dose
$D_{1cc}^{bladder}$	(Gy)	The maximum absolute dose received by 1cc of bladder

$V_{70Gy}^{bladder}$	(cc)	The absolute volume of bladder receiving more than 70Gy dose
$V_{65Gy}^{bladder}$	(%)	The relative volume of bladder receiving more than 65Gy dose
$V_{40Gy}^{bladder}$	(%)	The relative volume of bladder receiving more than 40Gy dose

Table 3: Plan validation indices

3. Results

3.1 Isodose Distributions

The outputs of the U-Net model and MLC optimization algorithm are binarized MLC aperture shapes at 178 control points in 7-8 full arcs, which are sent to a commercial treatment planning system (TPS) for dose calculation and plan finalization. On average, this procedure would cost 10 min for a plan generation in contrast to the current clinical practice (>30 minutes). Figure 7 shows the MLC aperture shapes of our multi-segment plan and the ground truth plan and one-arc U-Net predicted plan in one example case. Figure 7.a-7.i are examples of the seven segments in the multi-segment plan. Figure 7.h is the MLC aperture of the one-arc U-Net predicted plan, and Figure 7.i is the MLC aperture of the ground truth plan.

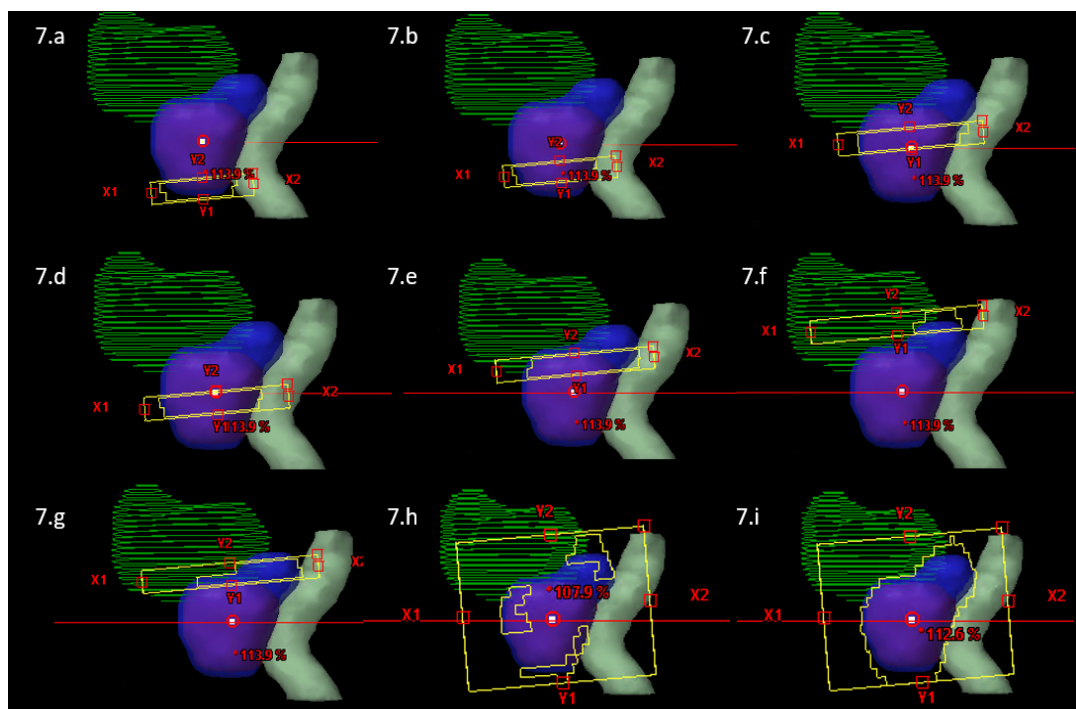


Figure 7: Aperture shape comparison between multi-segment plan, one arc plan and ground truth plan. (Blue contour represents PTV58.8)

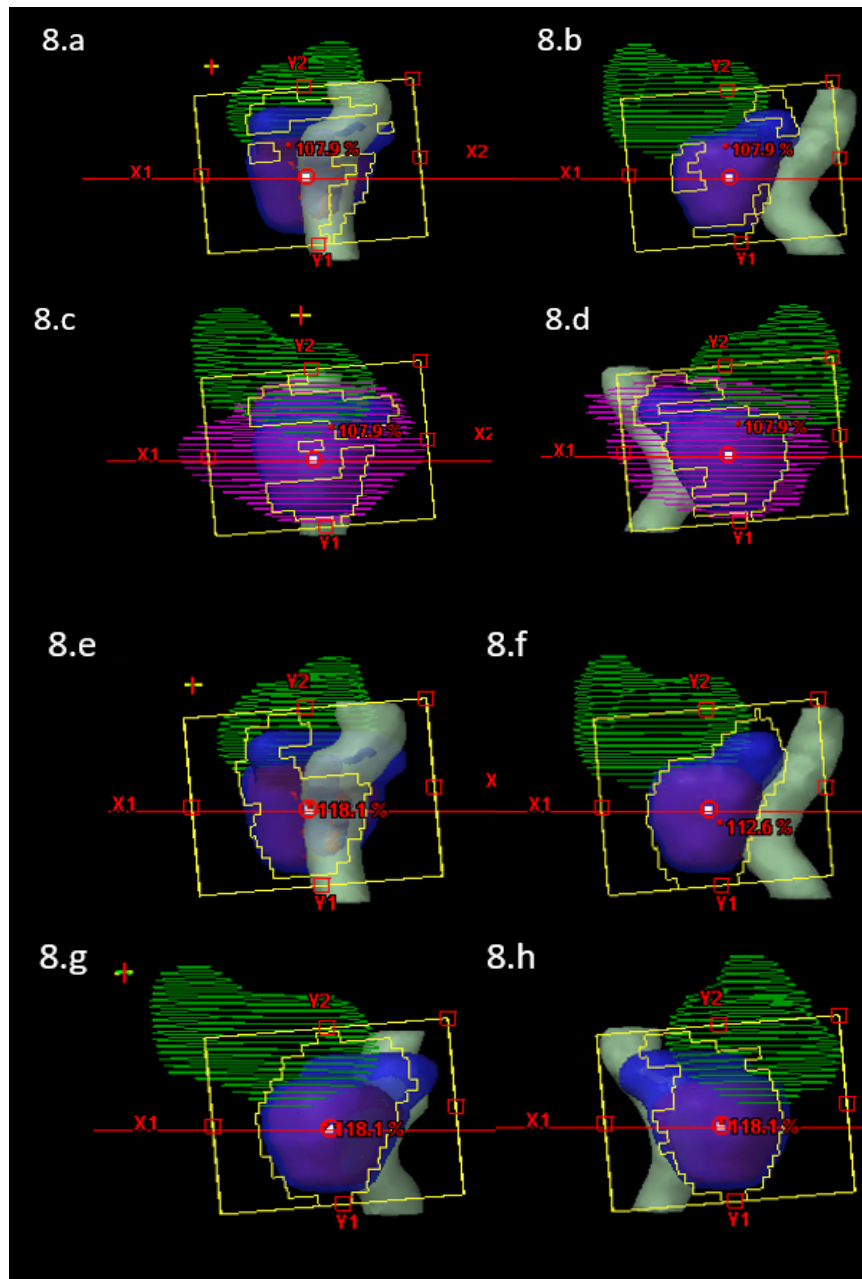


Figure 8: Aperture shape comparison between one-arc plan and ground truth plan of four different control points.

Since the MLC sequences of one-arc plans and multi-segment plans are similar, we take a one-arc plan in comparison with the ground truth plan to show the whole aperture shape in a simple way in Figure 8.

As can be seen from Figure 7 and Figure 8, the open-or-close pattern of the aperture shape of multi-segment plan and one-arc plan can be regarded as an

approximate delineation of the patient's PTV_{58.8}. (Figure 8.f) However, only a few OAR regions could be protected if there are conflicts between the OARs and the PTV (Figure 8.e, 8.h). However, the aperture shape of the ground truth plan seems to address different areas inside the PTV at different control points, and the OAR protection is better than the one-arc plan and ground truth plan. Figure 9 shows the

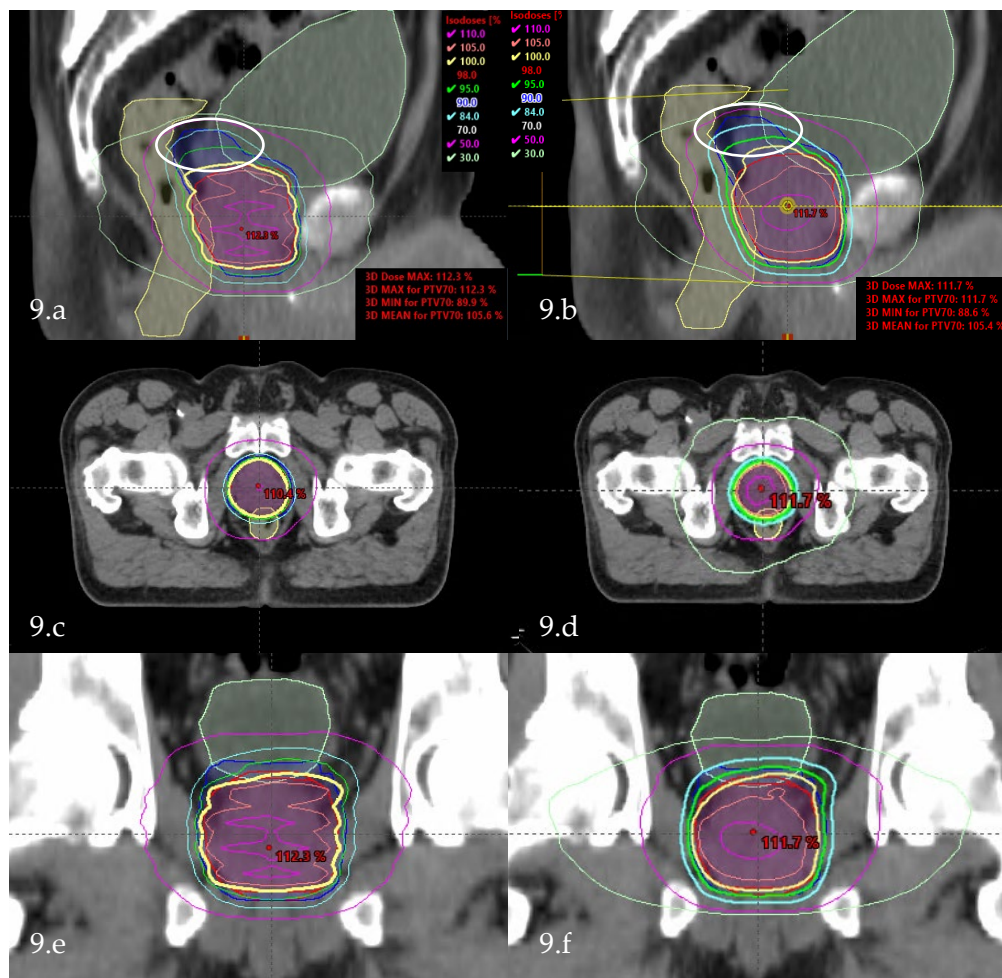


Figure 9: Isodose line comparison between Multi-segment plan (left) and one-arc plan (right)

isodose distribution comparison between the Multi-segment plan and the one-arc plan for one patient case.

Overall, the bladder and rectum sparing between the two plans are similar, and both meet institutional guidelines. The hot spots of the two plans do not show much difference (112% and 111%). In the one-arc plan, the upper part of the PTV_{58.8} (marked with white circles) could not get enough dose (Fig 9a - 9b, also shown in Fig 10.c and 10.d). The multi-segment plan showed improvements in the dose conformity on the coverage of PTV_{58.8} and SV dose fall-off. (Fig 10.e and 10.f) Both plans have a good 100% dose coverage (yellow isodose line) to PTV₇₀ (inner red contour) with good conformity as the plans are normalized to $V_{70Gy} = 95\%$ for PTV₇₀. In addition, the dose sparing to left and right femoral heads between the two plans is comparable. Figure 10 shows two more cases on the comparison of the sagittal view of the plans.

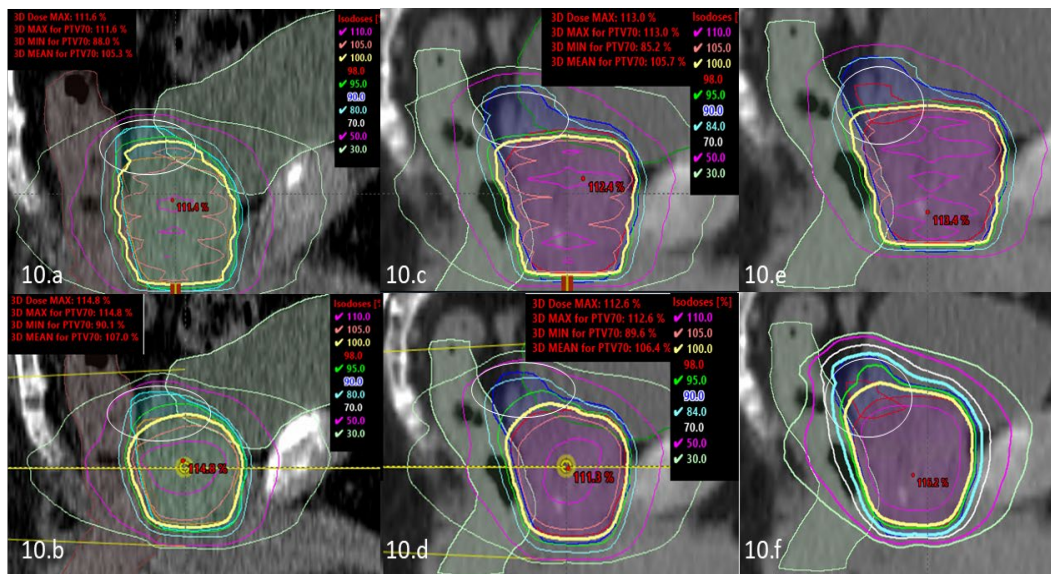


Figure 10: Comparison of Isodose line (sagittal view) of other 2 plans

We need to mention that the dose statistics, especially hot spots, have a very close relationship with the jaw opening in multi-segmented plans. The difference of a large margin (0.4 cm) between the jaw and the MLC leaf would make a maximum dose difference around 1% of the prescription dose (Figure 11.a - 11.b) in one-arc plans, but more than 10% in multi-segment plans (Figure 11.c – 11.d). Therefore, we managed to minimize the jaw opening in every field (0.1cm margin) to get a lower hotspot.

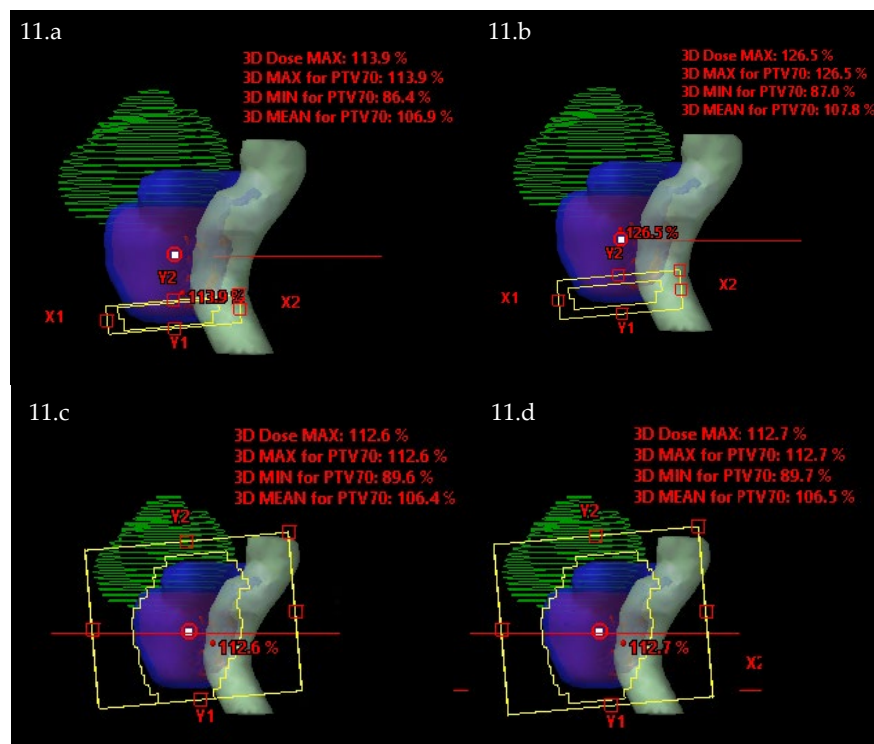


Figure 11: Hot spot difference between plans with different jaw openings (0.1 cm vs. 0.4 cm)

3.2 DVH statistics

In Figure 12, DVHs of the multi-segment plan and one-arc plan are compared for the same example case shown in Figure 9 above. Two PTVs and major OARs are included in the figure, each labeled by a unique color. The DVHs for the PTV₇₀, rectum, and bladder of the multi-segment plan are similar to those of the one-arc plan. The DVH for the rectum shows a moderate difference at middle to medium and high dose intervals (between 25Gy to 70Gy). As for the DVH for bladder, the multi-segment plan have a slightly better performance than the one-arc plan. The DVH curve and statistics can verify the finding that the multi-segment plan shows

<i>Structure</i>	<i>Parameter</i>	<i>Multi-segment</i>	<i>One-arc</i>	<i>TPS Plan</i>
<i>Body</i>	$D_{1cc}(\%)$	112.64±1.9	118.9±4.1	106.7±0.8
<i>PTV_{58.8}</i>	$V_{58.8Gy}(\%)$	94.21±2.9	92.4±2.2	96.8±0.7
<i>Rectum</i>	$D_{1cc}(Gy)$	73.3±1.1	74.0±3.0	70.5±1.1
	$V_{70Gy}(cc)$	4.5±2.3	2.8±1.8	1.4±0.7
	$V_{65Gy}(\%)$	8.6±2.8	4.7±1.5	3.6±1.1
	$V_{40Gy}(\%)$	33.4±8.1	26.3±5.9	24.0±5.0
<i>Bladder</i>	$D_{1cc}(Gy)$	74.0±1.9	75.8±2.0	72.6±0.8
	$V_{70Gy}(cc)$	5.3±3.3	6.8±3.6	4.1±2.0
	$V_{65Gy}(\%)$	4.4±3.0	5.5±2.3	3.7±2.0
	$V_{40Gy}(\%)$	16.1±8.6	19.4±9.2	17.7±8.9

Table 4: Major dosimetric parameters comparison

better dose coverage to PTV_{58.8} (V_{58.8Gy} = 95.6%) than the one-arc plan (V_{58.8Gy} = 91.6%).

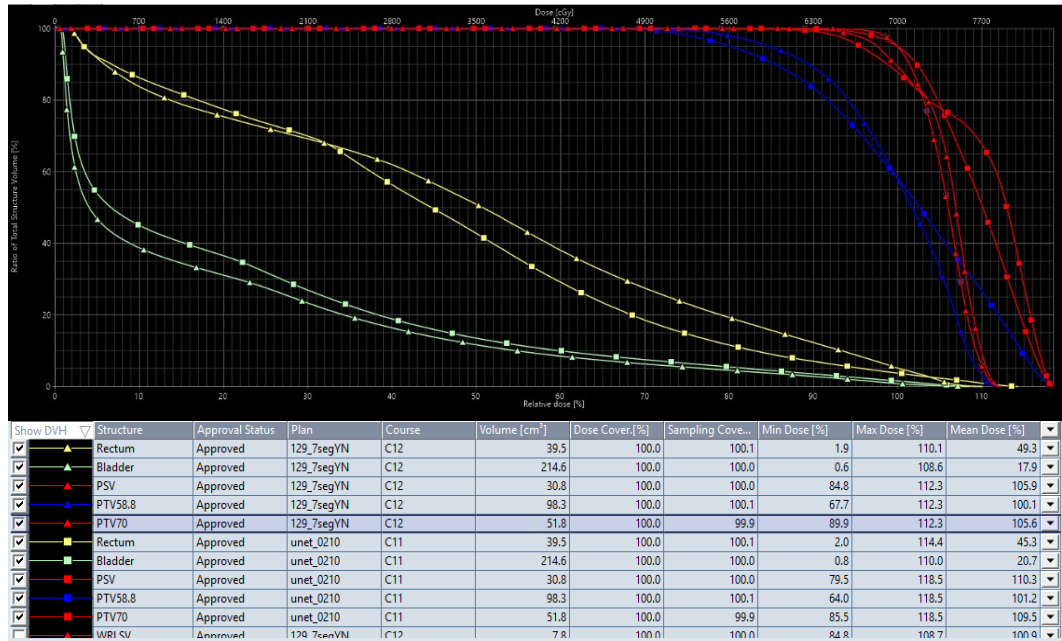


Figure 12: Dose volume histogram comparison between Multi-segment and one-arc plans.

We introduce the commercial TPS plan results (abbreviated as TPS plans below) for comparison to further evaluate the dose statistics. Table IV summarizes the quantitative dosimetric evaluations of Multi-segment plans, One-arc plans, and TPS plans for the 11 test cases. After dose normalization (PTV₇₀ V_{70Gy}=95%), all 11 Multi-segment plans achieved clinically acceptable spatial dose distribution. As can be derived from the table, although all the dose statistics seem not to be as good as TPS plans, the multi-segment plans show improvement on the hot spot, PTV_{58.8} coverage, and bladder dose than one-arc plans. In terms of low dose sparing, V_{40Gy} results of bladder and rectum in multi-segment plans were even better than one-arc plan results. However, rectum dose in multi-segment plans still needs to be improved.

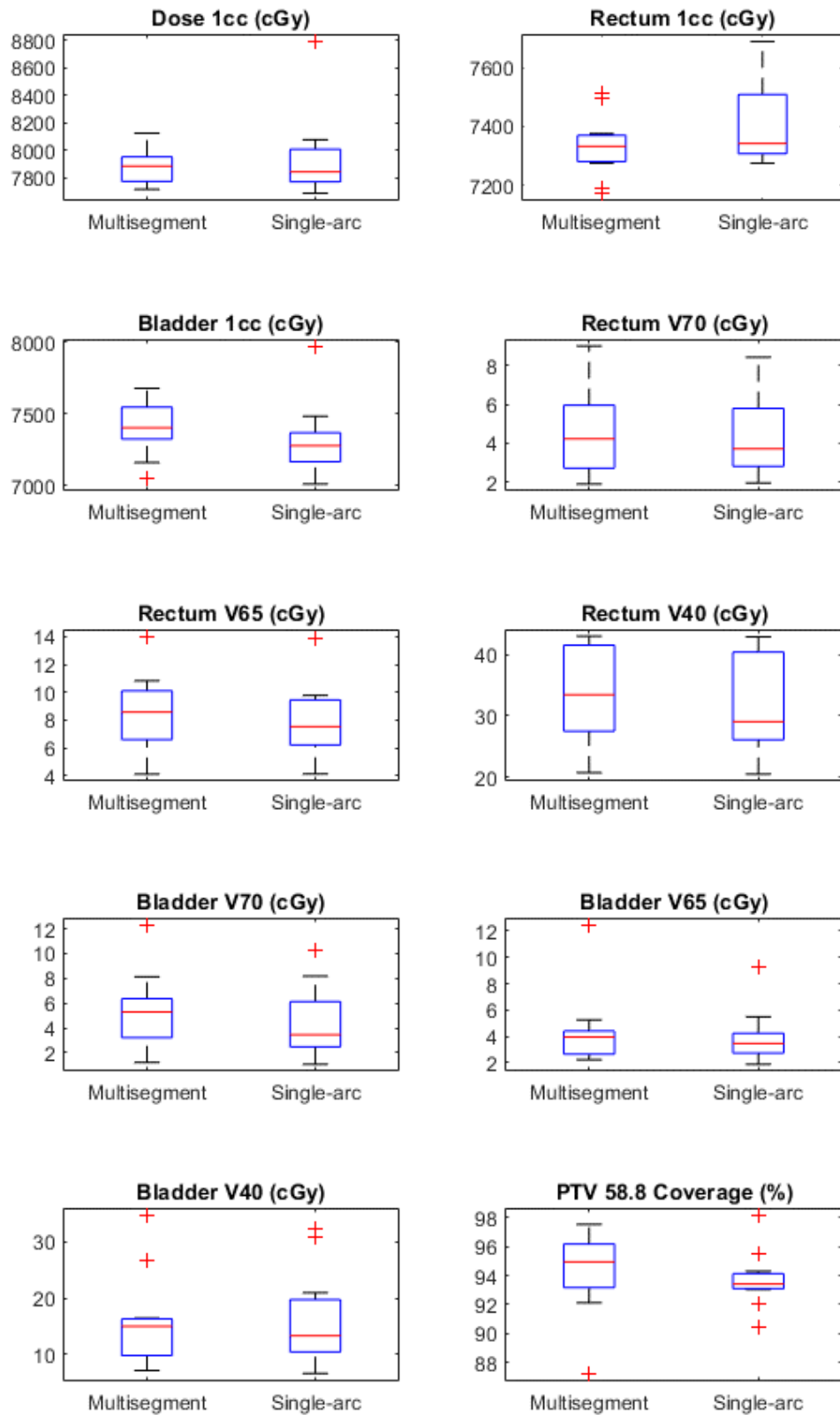


Figure 13: Major Dose statistics comparison (box plot)

As we can see from the box plot (Figure 13), there are several outliers on the multi-segment plans. The reason for the outliers of the box plot is the extremely large prostate cases. Though additional MLC pairs were added into the arcs, further optimization methods still need to be introduced regarding cases with large anatomical structures.

4. Discussion

4.1 Tasks and goals of Automated VMAT planning

Over the past several decades, treatment planning technologies have been rapidly growing and evolving. From 3D-CRT to IMRT to VMAT, the required computation and human intervention become heavier and heavier to simultaneously achieve better target dose conformity and high delivery efficiency. Partially or fully automated treatment planning can greatly improve the efficiency of the planning process and contribute to the clinical field. With state-of-the-art AI algorithms, automation tools and prediction models become available for many elements in treatment planning, ranging from dose prediction to automated MLC algorithm. Undoubtedly, the ultimate goal is to implement a fully automated treatment planning process with a self-learning capability. However, numerous preliminary studies and clinical trials are required before the realization of this ultimate goal. In this work, we developed an automatic multi-segment VMAT plan generation algorithm for prostate SIB cases. This study is a trial to enhance the dosimetric performance of an MLC dynamic sequence prediction-based automatic treatment

planning algorithm. The proposed study design might provide new insights for future research in the field of VMAT treatment planning.

4.2 Limitations

The proposed method has the following limitations: First, the model was only tested on 11 patient cases, which might not be enough to validate the model's efficacy. Second, adding too many arcs can lower the clinical efficiency and add to the computational work, which seems impractical in clinical applications. Third, the process of plan generation can also be optimized. Currently, we use the cross-correlation method to choose a constant threshold for each aperture, which means we have to compute the CC hundreds of times (450×7 times at one control point) for one patient). This affects the efficiency of the algorithm to a great extent. As shown in Figure 14, in one-arc plans, applying the CC method would lead to the missing of details (marked by white circles) in binary masks, whatever which threshold value is applied. Even if we make several segments, the loss of details in binary masks still happens among different pairs of MLC in one segment. Fourth, this proposed method was designed and validated for prostate cases. The prostate has a relatively simple geometric structure, while this procedure may come across many difficulties upon being applied to other structures. As we can see from the result, the MLC aperture of multi-segment plans cannot handle the conflict between the OARs and the PTV well, which may cause more problems regarding OAR dose sparing in other anatomical structures with more OAR restrictions. In other words, its potential application for other treatment sides with a higher complexity remained unclear.

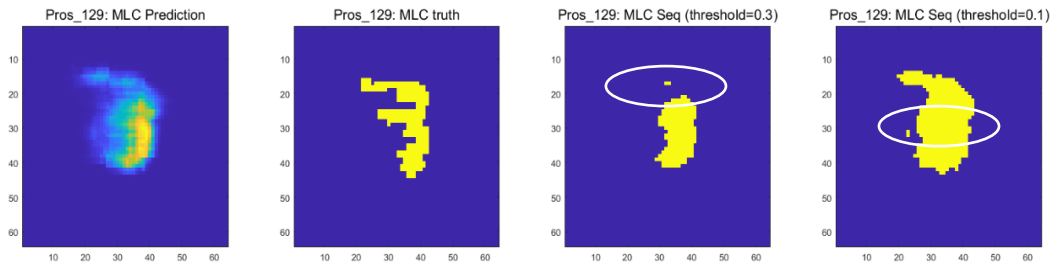


Figure 14: Loss of details in binary masks when applying CC method

4.3 Future research

The first future research direction is to potentially improve the robustness of this algorithm through: 1) including more patient cases in both the training and validation of the model and 2) refining DL network design and training. Compared to IMRT plans, VMAT plans have much more gantry angles to be taken into consideration. Therefore, the complexity of modeling is greatly increased. The most straightforward solution to this increased complexity is to include more training cases for the modeling. We have observed that there are some common problems that all of our U-Net predicted plans have, such as the distortion of the predicted intensity maps at the first and the last few control points. Then we could look back on our training process and try to optimize and modify the hyperparameters in the DL network. Additionally, we could investigate the performance of more recent deep network architectures like nnU-Net³⁵. It shows a strong capability in the field of image segmentation.

The second direction is to further improve the MLC leaf sequencing algorithm. We plan to use other algorithms to replace the computation of cross-correlation. We have also tried to calculate the MLC position by analyzing the

predicted intensity map line by line and thus selecting the optimized MLC leaf pair by pair. We have finished an experimental plan by this method and got encouraging dose statistics with higher efficiency. The multi-segment algorithm can improve the DVH statistics, yet multi-segment delivery may add to the treatment time. We plan to combine several arcs into one arc after the optimization process.

In the future, we plan to integrate our procedure into a GUI-based application. Developing such a fully automatic treatment planning system can benefit the frontline medical staff in radiation oncology in many aspects. For clinics that are short of human resources or overloaded with patient visits, this system can approximate treatment plans for the planners to modify and improve, saving time for the planners on repetitive work to handle more complicated tasks. Then the system can get more cases for self-learning and updating, thus improving the system's overall performance.

5. Conclusion

In conclusion, we proposed a multi-segment VMAT plan generation method to automate VMAT radiotherapy planning and a multi-segment MLC leaf sequencing algorithm. Compared to one-arc plans, the proposed multi-segment VMAT plan generation algorithm has a higher degree of freedom, better coverage of PTV_{58.8}, and lower hot spots. The generated multi-segment VMAT plans have the potentials for future development to realize predicted radiation intensity maps with high complexity and achieve desirable plan quality.

References

1. CP, W., E, W., & BW, S. (2020). World Cancer Report: Cancer Research for Cancer Prevention. Retrieved from <http://publications.iarc.fr/Non-Series-Publications/World-Cancer-Reports/World-Cancer-Report-Cancer-Research-For-Cancer-Prevention-2020>
2. Baskar, R., & Itahana, K. (2017). Radiation therapy and cancer control in developing countries: Can we save more lives? *International Journal of Medical Sciences*, 14(1), 13 - 17.
3. Otto, K. (2008). Volumetric modulated arc therapy: IMRT in a single gantry arc. *Medical Physics*, 35(1), 310 - 317.
4. Nguyen, D., Long, T., Jia, X., Lu, W., Gu, X., Iqbal, Z., Jiang, S. (2019). A feasibility study for predicting optimal radiation therapy dose distributions of prostate cancer patients from patient anatomy using deep learning. *Scientific Reports*, 9(1), 1–10.
5. Fan, J., Wang, J., Chen, Z., Hu, C., Zhang, Z., & Hu, W. (2019). Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique. *Medical Physics*, 46(1), 370–381.
6. Ni et al. (2020). A Deep-learning Method of Automatic VMAT Planning Via MLC Dynamic Sequence Prediction (AVP-DSP) Using 3D Dose Prediction: a Feasibility Study of Prostate Radiotherapy Application. Duke University. ProQuest Dissertations Publishing, 2020. 27743011
7. Ma, M., Kovalchuk, N., Buyyounouski, M. K., Xing, L., & Yang, Y. (2019). Dosimetric features-driven machine learning model for DVH prediction in VMAT treatment planning. *Medical Physics*, 46(2), 857 - 867.
8. Yu C., Li X., Ma L., Chen D., Naqvi S., Shepard D., et al.. (2002). Clinical implementation of intensity-modulated arc therapy. *Int J Radiat Oncol Biol Phys*; 53:453–63.
9. F., Khan, J., Gibbons (2014) Physics of radiation therapy 5th edition. Wolters Kluwer
10. Verbakel WF, Cuijpers JP, Hoffmans D, Bieker M, Slotman BJ, Senan S. (2009). Volumetric intensity-modulated arc therapy vs. conventional IMRT in head-and-neck cancer: a comparative planning and dosimetric study. *Int J Radiat Oncol Biol*, 74, 252–9.
11. Quan, E. M., Li, X., Li, Y., Wang, X., Kudchadker, R. J., Johnson, J. L., Kuban, D. A., Lee, A. K., & Zhang, X. (2012). A comprehensive comparison of IMRT and

VMAT plan quality for prostate cancer treatment. *International journal of radiation oncology, biology, physics*, 83(4), 1169–1178.

12. Ong CL, Verbakel WF, Cuijpers JP, Slotman BJ, Lagerwaard FJ, Senan S. (2010) Stereotactic radiotherapy for peripheral lung tumors: a comparison of volumetric modulated arc therapy with 3 other delivery techniques. *Radiother Oncol* 2010; 97: 437–42.
13. J., Lebesque, R., Keus, (1991). The simultaneous boost technique: the concept of relative normalized total dose. *Radiotherapy and Oncology*; 22, 45-55
14. X. Allen Li, Jian Z. Wang, Paul A. Jursinic, Colleen A. Lawton, Dian Wang, (2005). Dosimetric advantages of IMRT simultaneous integrated boost for high-risk prostate cancer, *International Journal of Radiation Oncology Biology Physics*, Volume 61, 1251-1257
15. F., Wong, A., Ng, V., Lee, C., Lui, K., Yuen, S., Tung, (2010). Whole-Field Simultaneous Integrated-Boost Intensity-Modulated Radiotherapy for Patients With Nasopharyngeal Carcinoma, *International Journal of Radiation Oncology Biology Physics*; 76, Issue 1, 138-145
16. Q., Wu, R., Mohan, M., Morris, A., Lauve, R., Ullrich (2003). Simultaneous integrated boost intensity-modulated radiotherapy for locally advanced head-and-neck squamous cell carcinomas. I: dosimetric results. *International Journal of Radiation Oncology Biology Physics*; 56, Issue 2, 573-585
17. J., Cheng, E., Huang, S., Hsu, C., Wang. (2016). Simultaneous integrated boost (SIB) of the parametrium and cervix in radiotherapy for uterine cervical carcinoma: a dosimetric study using a new alternative approach. *Br J Radiol*.
18. S., Cilla, F., Deodato, C., Digesù, G, Macchia, V., Picardi, M., Ferro, G., Sallustio, M., De Spirito, A., Piermattei, A., Morganti. (2014). Assessing the feasibility of volumetric-modulated arc therapy using simultaneous integrated boost (SIB-VMAT): An analysis for complex head-neck, high-risk prostate and rectal cancer cases. *Med Dosim*.
19. D., Wang, J., Chen, X., Zhang , T., Zhang, L., Wang, Q., Feng, Z., Zhou, J., Dai, N., Bi. (2020). Sparing Organs at Risk with Simultaneous Integrated Boost Volumetric Modulated Arc Therapy for Locally Advanced Non-Small Cell Lung Cancer: An Automatic Treatment Planning Study. *Cancer Manag Res*.
20. Yang, Y., & Xing, L. (2004). Clinical knowledge-based inverse treatment planning. *Physics in Medicine and Biology*, 49(22), 5101 – 5117.

21. Zhu X, Ge Y, Li T, Thongphiew D, Yin F, Wu Q. (2011) A planning quality evaluation tool for prostate adaptive IMRT based on machine learning. *Med Phys.* 2011;38(2):719-726.
22. Dou, Q., Yu, L., Chen, H., Jin, Y., Yang, X., Qin, J., & Heng, P.-A. (2017). 3D deeply supervised network for automated segmentation of volumetric medical images. *Medical Image Analysis*, 41, 40 - 54.
23. Çiçek, Ö., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, & W. Wells (Eds.), *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016* (pp. 424 - 432). Springer International Publishing.
24. Y., Xie, Y., Xia, J., Zhang, Y., Song, D., Feng (2019) Knowledge-based Collaborative Deep Learning for Benign-Malignant Lung Nodule Classification on Chest CT, *IEEE Transactions on Medical Imaging*, vol. 38, no. 4, pp. 991-1004.
25. Walker, G. V., Awan, M., Tao, R., Koay, E. J., Boehling, N. S., Grant, J. D., Sittig, D. F., Gunn, G. B., Garden, A. S., Phan, J., Morrison, W. H., Rosenthal, D. I., Mohamed, A. S. R., & Fuller, C. D. (2014). Prospective randomized double-blind study of atlas-based organ-at-risk autosegmentation-assisted radiation planning in head and neck cancer. *Radiotherapy and Oncology: Journal of the European Society for Therapeutic Radiology and Oncology*, 112(3), 321 - 325.
26. Valdes, G., Chan, M. F., Lim, S. B., Scheuermann, R., Deasy, J. O., & Solberg, T. D. (2017). IMRT QA using machine learning: A multi-institutional validation. *Journal of Applied Clinical Medical Physics*, 18(5), 279 - 284.
27. Valdes, G., Scheuermann, R., Hung, C. Y., Olszanski, A., Bellerive, M., & Solberg, T. D. (2016). A mathematical framework for virtual IMRT QA using machine learning. *Medical Physics*, 43(7), 4323 - 4334.
28. Kearney, V., Chan, J. W., Haaf, S., Descovich, M., & Solberg, T. D. (2018). DoseNet: A volumetric dose prediction algorithm using 3D fully-convolutional neural networks. *Physics in Medicine and Biology*, 63(23), 235022.
29. Lee, H., Kim, H., Kwak, J., Kim, Y. S., Lee, S. W., Cho, S., & Cho, B. (2019). Fluence-map generation for prostate intensity-modulated radiotherapy planning using a deep-neural-network. *Scientific Reports*, 9(1), 1 - 11.
30. Sheng Y, Li T, Yoo S, Yin F-F, Blitzblau R, Horton JK, Ge Y, Wu QJ. Automatic Planning of Whole Breast Radiation Therapy Using Machine Learning Models. *Frontiers in Oncology*. 2019 Aug 7;9:1-8.

31. Li X, Zhang J, Sheng Y, Chang Y, Yin FF, Ge Y, Wu QJ, Wang C. (2020) Automatic IMRT Planning via Static Field Fluence Prediction (AIP-SFFP): A Deep Learning Algorithm for Real-time Prostate Treatment Planning . *Phys Med Biol.* 10.1088/1361-6560/aba5eb.
32. Wang W, Sheng Y, Wang C, Zhang J, Li X, Palta M, Czito B, Willett C, Wu Q, Ge Y, Yin FF, Wu QJ, (2020) Fluence Map Prediction Using Deep Learning Models-Direct Plan Generation for Pancreas Stereotactic Body Radiation Therapy. *Frontiers in Artificial Intelligence.* 2020;3:68.
33. Wang W, Sheng Y, Palta M, Czito B, Willett C, Hito M, Yin FF, Wu Q, Ge Y, Wu QJ, (2021) Deep Learning-based Fluence Map Prediction for Pancreas Stereotactic Body Radiation Therapy with Simultaneous Integrated Boost. *Advances in Radiation Oncology.* 2021;6:1-9.
34. Jensen PJ, Zhang J, Koontz BF, Wu QJ. (2021) A Novel Machine Learning Model for Dose Prediction in Prostate Volumetric Modulated Arc Therapy Using Output Initialization and Optimization Priorities. *Front Artif Intell.* 2021;4:624038.
35. Li X, Wang C, Sheng Y, Zhang J, Wang W, Yin FF, Wu Q, Wu QJ, Ge Y (2021): An artificial intelligence-driven agent for real-time head-and-neck IMRT plan generation using conditional generative adversarial network (cGAN). *Med Phys* 2021, 48:2714-2723.
36. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*, 234 - 241.
37. Fabian Isensee, Paul F. Jäger, Simon A. A. Kohl, Jens Petersen, Klaus H. Maier-Hein (2020). Automated Design of Deep Learning Methods for Biomedical Image Segmentation. *Nature Methods* (2020)