Flexible-C^m GAN: Towards Precise 3D Dose Prediction in Radiotherapy

Riqiang Gao Bin Lou Zhoubing Xu Dorin Comaniciu Ali Kamen
Digital Technology and Innovation, Siemens Healthineers, Princeton NJ 08540
riqiang.gao@siemens-healthineers.com

Abstract

Deep learning has been utilized in knowledge-based radiotherapy planning in which a system trained with a set of clinically approved plans is employed to infer a three-dimensional dose map for a given new patient. However, previous deep methods are primarily limited to simple scenarios, e.g., a fixed planning type or a consistent beam angle configuration. This in fact limits the usability of such approaches and makes them not generalizable over a larger set of clinical scenarios. Herein, we propose a novel conditional generative model, Flexible-C^m GAN, utilizing additional information regarding planning types and various beam geometries. A miss-consistency loss is proposed to deal with the challenge of having a limited set of conditions on the input data, e.g., incomplete training samples. To address the challenges of including clinical preferences, we derive a differentiable shift-dose-volume loss to incorporate the well-known dose-volume histogram constraints. During inference, users can flexibly choose a specific planning type and a set of beam angles to meet the clinical requirements. We conduct experiments on an illustrative face dataset to show the motivation of Flexible-C^m GAN and further validate our model's potential clinical values with two radiotherapy datasets. The results demonstrate the superior performance of the proposed method in a practical heterogeneous radiotherapy planning application compared to existing deep learning-based approaches.

1. Introduction

Radiation therapy (RT) is an essential modality for cancer treatment and is applicable to about 50% of patients [12, 25]. However, many studies demonstrate that millions of patients currently do not have access to radiotherapy due to limited infrastructures and trained experts to handle complex planning procedures [15, 20, 59].

RT treatment planning is a process that involves a multidisciplinary team (e.g., oncologists, therapists, physicists) to figure out the treatment beam configurations and intensity for cancer patients [25]. The modern RT treatments can be divided into two broad categories using static- or dynamic-beams. The intensity modulated radiotherapy (IMRT) [56] and volumetric modulated arc therapy (VMAT) [48, 57] are the most common static- and dynamic- beam types, respectively [10]. IMRT uses several personalized but fixed beam angles, delivering radiation precisely to the tumor while sparing the surrounding normal tissues according to the location of the tumor and anatomical organs at risk (OARs). During VMAT, the treatment beam is on while its treatment head is moving on an arc trajectory [10]. As shown in Figure 1, the dose maps of static- and dynamic- beam RT plans can look significantly different, which results from different delivery nature energy fluences in those two modes. Moreover, even using the same planning mode, different configurations (e.g., beam angles, isocenter) are needed for different patients due to different tumor locations/shapes, anatomy structures, and other clinical parameters.

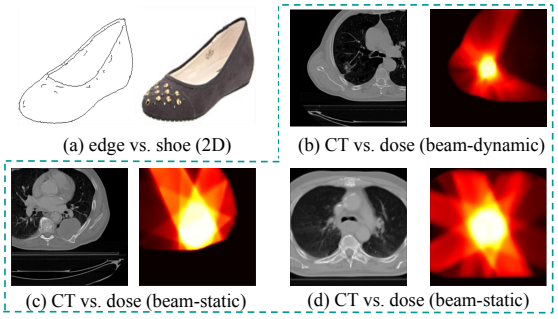


Figure 1. Vanilla image-to-image translation (a) and dose prediction (b)-(d). (a) has a clear shape match between the source and target domains. (b)-(d) illustrate our challenges, including heterogeneous patterns, no clear match between source and target, and 3D data (showing 2D for simplicity) is harder than 2Ds.

Knowledge-based planning (KBP) aims to use computer technologies to reduce the time for individualized treatment plans [6, 43]. Historically, KBP technologies relied on statistical models or handcrafted features [46, 54]. While providing promising results, these methods are hard to generalize beyond an inherently targeted limited set of scenarios [33]. Advanced artificial intelligence (e.g., deep learning [37]) has shown great potential to alter the way oncology therapies are administered [25, 53]. An integral part of

KBP methods is to predict the dose distribution that should be delivered to a patient [6,32,59]. Three-dimensional (3D) deep learning models have been applied in dose map prediction across different cancers [7, 32, 33, 41, 45, 55] where the inputs are generally computed tomography (CT) image/volume and masks of organs at risk (OARs) / planning target volume (PTV). However, the existing automatic dose prediction models mainly focus on relatively simple scenarios, e.g., single planning mode or/and consistent angle configuration, which significantly limits the model flexibility.

In this paper, we propose a novel conditional generative model, flexible-multiple-condition GAN, short for Flexible-C^m GAN or FCGAN, for precise 3D dose prediction in heterogeneous RT contexts. In addition to the conditions (CT, PTV/OAR masks) that other methods in literature have used, we further integrate two conditions: the planning mode (i.e., static- or dynamic- beam) and the angle configuration. Furthermore, we show that our model is robust in those scenarios where angle configuration may not be available. Briefly, our contributions include

- We proposed a novel GAN variant, FCGAN, that considers multi-level conditions and handles missing condition values with a new miss-consistency loss.
- We derived a *differentiable* and *spatially-unbiased* loss function from a widely applied dose-volume histogram and show its effectiveness within the deep learning training for 3D dose prediction.
- We introduced the deep 3D dose prediction for practical heterogeneous treatment scenarios (i.e., multi-type, multi-beam configuration), which enables easy and fast user-interaction by changing input conditions and checking results interactively during inference.
- We conducted experiments on two clinical radiotherapy datasets and a face dataset, validating that our approach is superior to state-of-the-art deep models.

2. Related Work

RT Dose Planning. There are mainly two branches of dose prediction models: (1) handcrafted dose-volume (DV) feature-based and (2) deep learning-based. In the first branch, the dose-volume statistical models are mainly based on conventional machine learning or statistics at low-dimension feature bases [2, 3, 14, 29]. The cumulative dose-volume frequency distributions [13], simply as dose-volume histograms (DVHs), have been used for evaluation in both traditional models and recently deep learning models [14]. RapidPlan™ is a commercial KBP tool developed by Varian Medical Systems (Palo-Alto, CA), which is using DVHs as the feature set, and allows clinical objectives to be automatically inferred from previous plans [1]. Furthermore, the user is enabled to modify and integrate clinical objectives into the iterative optimization pipeline to create a plan for a new patient [1, 16, 35].

Recently, 3D-based image models (especially UNet-based networks) have been adapted in dose map prediction for different cancers, such as in prostate [32, 33, 45], head-and-neck [41, 55, 58], esophageal [4, 64], and lung [7, 27]. Most of them use reconstruction (e.g., L1 or L2) or/and DVH-based losses to train the 3D dose prediction model in a single-mode context (e.g., only IMRT). For example, Kearney et al. [32] apply a fully convolutional neural network to predict the dose map. A cascaded 3D U-Net [41] has achieved the best performance in a KBP challenge of IMRT on head-and-neck cancer [6]. Vanilla adversarial losses [19, 26] are used with reconstruction loss to make the predicted dose more realistic [4, 33, 45]. Soomro et al. [55] extend [45] with a dilated DenseNet for head-and-neck cancer but in a narrow scenario of equidistant static beams.

Conditional GAN and I2I Translation. Generative adversarial nets (GANs) [19] have been successful in generating realistic images [8,9,21,36]. However, there is no control on generated samples for vanilla GAN. Conditional GAN [44] provides a solution to guide the generation by feeding additional information, which has been widely explored in general computer vision such as image-to-image translation [11, 18, 26, 50, 65], text-to-image generation [49], and image imputation [38, 40]. Conditional GANs have also been extended to medical imaging addressing specific healthcare problems with some necessary modifications [63], such as CT-to-MRI translation [24], disease diagnosis or image segmentation after data imputation [17, 28].

Dose prediction can be formulated broadly as image-to-image translation (I2I), while it is significantly different from the well-known ones, as illustrated in Figure 1. Simple I2I translation usually shares similar structures between the source and target domains, e.g., edges to color image [26, 65] (as in Figure 1a), and MRI to CT [24]. The source (CT/masks) has a relatively weak connection and no unique mapping to the target (i.e., dose map) in the practical heterogeneous contexts (as in Figures 1b, 1c, and 1d).

3. Methodology

3.1. Problem Description and Motivation

In order to obtain precise individualized RT plans, three types of data are needed: (1) planning mode (static- or dynamic- beams), (2) beam angle configuration, and (3) fluence map for each beam, which determines the beam shape and intensity. (1) and (2) are relatively easy to tune in a human-computer interaction system, while obtaining (3) is labor-intensive and time-consuming. A typical target of the KBP pipeline, as in Figure 2, is to obtain the fluence map for each beam (or subfield), given the planning mode and beam angle configuration. There are two main steps to ob-

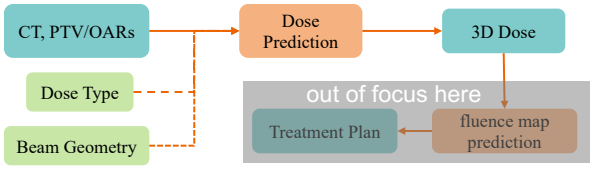


Figure 2. Overview of a typical KBP pipeline. The teal blocks are inputs or outputs, the orange blocks are algorithm modules, and the green ones are the introduced flexible conditions.

tain final fluence maps [5,6,61,62]: (1) dose prediction, and (2) fluence map prediction when given the predicted dose map. As the first and essential step, the dose prediction is the main focus of this paper.

In most dose prediction models via deep learning (e.g, [4,32,33,45,64]), the inputs (i.e., source) are the computed tomography (CT) image/volume, the organ at risk (OAR) masks, and planning target volume (PTV) mask. However, the dose map (i.e., target) lacks clear structure or texture information from the source. Furthermore, the reference dose, used as ground truth during training, is subjective and can biased as it is designed by a group of expert humans. It is less challenging if 1) the planning context is simple such as using the same dose type or/and the same beam angles; 2) the relative locations of PTV and OARs are nearly consistent for some cancers, such as in prostate cancer and head-and-neck. However, in clinical practice, specialists may need to consider which dose types should be delivered (e.g., beam-static or beam-dynamic) and the suitable set of beam angles. Tumors in some cancers (e.g., lung cancer, liver cancer) can vary in a large range of locations within the host organ, resulting in the relative positions of PTV and OARs being largely heterogeneous.

We argue that two key points need to be addressed for precise dose prediction: individualism and realism. *individualism* is required for the plan to be precise specifically in heterogeneous set of conditions. *Realism* makes the subsequent tasks (e.g., fluence map prediction [39, 42, 62], deliverable dose [5]) more manageable. GANs have been specifically useful in generating realistic samples, and their conditional versions allow additional restrictions and constraints which could be beneficial in our task. Conditional GAN [44] is defined as min-max game of discriminator D with generator G using following loss:

$$\begin{aligned} V(D,G) =& \mathbb{E}_{x\sim p_{data}}[\log D(x|y)]+ \\ & \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z|y)))], \end{aligned} \tag{1}$$

where G generates samples based on the condition y and random noise z. A simple conditional GAN [44] can not satisfy our use-case as we need to handle (1) multi-level conditions with heterogeneous types and (2) missing conditions during training and testing.

3.2. Flexible-C^m GAN Mechanism

We describe the general methodology of Flexible-C^m GAN in this section, and the instantiation for 3D dose prediction is shown in Sec. 3.3. Math symbol explanations of our model are in Supplement A.

Given M conditions $\{C^i\}_{i=1}^M$ (i.e., $\mathbf{C}$) and their missing indicators $\mathbf{m}$ ($m^i = 0$ if i-th condition is missing, otherwise $m^i = 1$), our adversarial loss becomes:

$$\begin{aligned} V(D,G) =& \mathbb{E}_{x\sim p_{data}}[\log D(x|\mathbf{C},\mathbf{m})]+ \\ & \mathbb{E}_{z\sim p_z(z)}[\log(1-D(G(z|\mathbf{C},\mathbf{m})))]. \end{aligned} \tag{2}$$

The i-th missing condition is imputed with default or random values (termed as $\bar{C}^i$). To let the model be robust to the missing condition, we introduce a miss-consistency loss L_{mc} based on condition regularization loss L_{cr}:

$$L_{cr} = \sum_{i,m^i>0} L^i(E^i(G(z|\mathbf{C},\mathbf{m})),\cdot), \tag{3}$$

$$L_{mc} = \sum_{j\neq i, m^j>0} |E^j(G(\cdot|m^i=0,\cdot)) - E^j(G(\cdot))|. \tag{4}$$

where $E^i(\cdot)$ extracts feature from the prediction $G(\cdot)$ for the condition C^i. $L^i(\cdot,\cdot)$ measures the discrepancy between the prediction and the reference corresponding to C^i, and L_{mc} reflects how predictions related to observed condition j are consistent when another condition i is given versus the scenario in which it is missing.

3.3. FCGAN Instantiation for 3D Dose Prediction

3.3.1 3D Dose Prediction Framework

Our overall framework for 3D dose prediction is illustrated in Figure 3. The input includes CT, PTV/OAR masks, planning mode (beam-static or beam-dynamic), angle/beam plates, and a condition mask indicating if any condition is missing. The angle/beam plates computation and detailed generator G structures are described in Supplement D.

For three-dimensional conditions (CT, PTV/OARs masks, angle/beam plates), the condition regularization terms of Eq. 3 (related E^i is *Identity*) are jointly covered by a reconstruction loss L_{rec} and shift-dose-volume (SDV) loss L_{sdv}. The L_{rec} of N samples is the mean absolute error (MAE) of the reference dose $\mathbf{Y}_i$ and its prediction $\hat{\mathbf{Y}}_i$:

$$L_{rec} = \frac{1}{N}\sum_{i=1}^{N} ||\mathbf{Y}_i - \hat{\mathbf{Y}}_i||_1. \tag{5}$$

The details of L_{sdv} are shown in Sec. 3.3.2.

Inspired by [17, 47] with categorical condition, we introduce the cross-entropy loss (CEL) to instantiate Eq. 3 (related E^i is the pre-trained encoder in Figure 3) for the planning mode condition C^m:

$$L_{cls} = \frac{1}{N}\sum_{i=1}^{N} -C_i^m \log(\hat{p}_i) - (1-C_i^m)\log(1-\hat{p}_i). \tag{6}$$

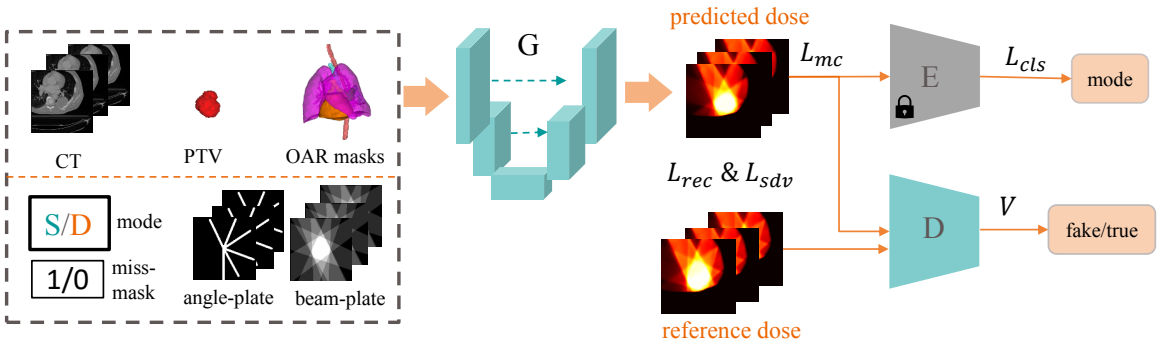


Figure 3. Flexible-C^m GAN for 3D dose prediction. CT, PTV/OAR masks, mode, angle/beam plates, and miss-mask are fed into a U-shape-based Generator (G) to predict dose maps. Note that the condition can be missing, indicated by the missing indicator. Reconstruction loss (L_{rec}) and SDV loss (L_{sdv}) minimize the discrepancy between predicted and reference doses from different perspectives. With a pre-trianed encoder (E), L_{cls} regularizes the mode of predicted dose. L_{mc} enforces prediction consistency of one condition whether another condition is missing or not, and V is the GAN loss to achieve realistic prediction.

Thus, the total FCGAN loss for 3D dose prediction is:

$$L_{total} = V + 3L_{rec} + L_{cls} + L_{sdv} + L_{mc}, \quad (7)$$

where the adversarial loss V follows Eq. 2, and $\mathbf{C}$ includes three types of condition: (1) CT and PTV/OAR masks, which are also used by previous methods, e.g., [6, 41, 45], (2) planning mode, and (3) angle/beam plates. The ratio of L_{rec} is set relatively larger to avoid the error is overwhelm by small/zero dose voxels. The first three items are common in various conditional GANs [17, 19, 41, 47]. The last two losses (L_{sdv} and L_{mc}) are our contributions, and we have a set of ablation studies to validate their effectiveness.

3.3.2 Shift-Dose-Volume Loss

Original DVH [13] is not differentiable (for error back-propagation) and can be sub-optimal from a spatial view. Several DVH-based loss functions have been proposed for dose prediction training [27, 45, 58]. The majority of those methods are empirically designed based on the computation of DVH, e.g., approximate *Sign* function with *Sigmoid* function. Differently, we derive a new DVH-based loss from DVH definition by shifting histogram errors to their voxel space, which holds a mathematical connection between the differentiable loss function and original DVH.

Assume the max dose is D_T and there are T bins with unique widths for the DVH. So, the bin width is $w = \frac{D_T}{T}$. Given a region of interest (ROI, i.e., PTV or OAR) mask M (as the teal region in Figure 4b), according to DVH definition: the fractional volume $f(D_t)$ (y-axis) at t-th threshold D_t (x-axis) is $f(D_t) = \frac{\sum_j \mathbb{1}(Y_j - D_t) M_j}{\sum_j M_j}$, where j is voxel index and $t = \frac{TD_t}{D_T}$ is bin index of D_t, as in Figure 4a.

Let us denote voxel indexes of t-th bin as $\mathbf{M}_t$ (as the orange in Figure 4a). To minimize the error from a DVH perspective, we want errors of each bin to be minimized, and we shift the error measurement of each bin to the error of voxels that contribute to the fractional volume $f(D_t)$.

Definition 1. *The shift-dose-volume loss L^t_{sdv} of each bin $[D_t, D_t + \frac{D_T}{T})$ is the absolute error $\varepsilon(\cdot)$ of all voxels that contribute to $f(D_t)$ multiply the bin width w, i.e., expected error of $f(D_t)$ from voxel perspective. The L_{sdv} of each ROI is the sum of each bin L^t_{sdv}, i.e., $\sum_{t=1}^{T} L^t_{sdv}$.*

Given referenced dose $\mathbf{Y}$, we have:

$$\begin{aligned}
&\sum_j M_j \sum_{t=1}^{T} L^t_{sdv} = \sum_{t=1}^{T} \sum_j \mathbb{1}(Y_j - D_t) \cdot \varepsilon(Y_j \cdot M_j) \cdot w \\
&= \sum_{t=1}^{T} \varepsilon(\mathbf{Y} \odot (\mathbf{M}_t + ... + \mathbf{M}_T)) \cdot w \\
&= \sum_{t=1}^{T} t \cdot \varepsilon(\mathbf{Y} \odot \mathbf{M}_t) \cdot w = \sum_{t=1}^{T} D_t \cdot \varepsilon(\mathbf{Y} \odot \mathbf{M}_t),
\end{aligned} \quad (8)$$

where $\odot$ is the Hadamard product. When T is large enough to make every $Y_j \in [D_t, D_t + \frac{D_T}{T})$ indexed by $\mathbf{M}_t$ is close to D_t, so $\lim_{T \to \infty} D_t \cdot \varepsilon(\mathbf{Y} \odot \mathbf{M}_t) = \forall_{j \in \mathbf{M}_t} Y_j \cdot \varepsilon(\mathbf{Y} \odot \mathbf{M}_t)$. Thus, the SDV loss for a single organ becomes

$$\sum_{t=1}^{T} L^t_{sdv} = ||\mathbf{Y} \odot (\mathbf{Y} - \hat{\mathbf{Y}}) \odot \mathbf{M}||_1. \quad (9)$$

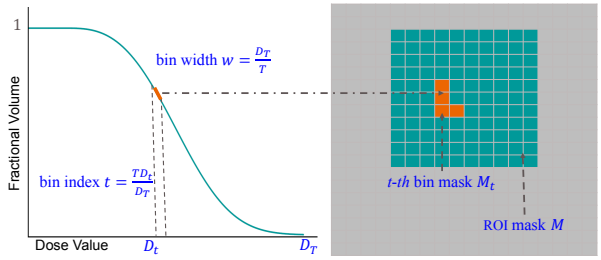
Generally, given S ROIs masked by $\{\mathbf{M}^s_i\}$ for i-th patient, we have L_{sdv} for N patients as:

$$L_{sdv} = \frac{1}{N} \sum_{i=1}^{N} \sum_{s=1}^{S} \lambda_s ||\mathbf{Y}_i \odot (\mathbf{Y}_i - \hat{\mathbf{Y}}_i) \odot \mathbf{M}^s_i||_1, \quad (10)$$

where $\lambda_s = 1$ if $\mathbf{M}^s_i$ mask PTV else $\lambda_s = 0.5$. In summary, the proposed L_{sdv} has the following properties (detailed derivation from Eq. 8 to Eq. 10, justification and math symbol clarification are in Supplement A and E).

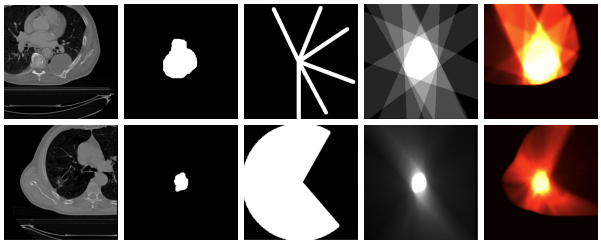
Property 1: *Higher dose voxels contribute more to both DVH computation and our SDV loss.*

Property 2: *Optimal L_{sdv} leads to an exact match of 3D dose maps in its ROI, while zero DVH gaps (i.e., optimum) theoretically can come from sub-optimal spatial mismatch.*



(a) dose volume histogram (b) 3D dose map

Figure 4. Shifting original (frequency-level) DVH definition to voxel-level loss. (a) shows a DVH and (b) is its relative volume. Teal is the ROI mask, the orange is corresponding to t-th bin.



(a) CT (b) PTV (c) angle plt. (d) beam plt. (e) ref. dose

Figure 5. The illustration of beam-static (upper) and beam-dynamic (lower) samples (3D). The CT and PTV/OAR masks are source data. Angle-plate and beam-plate are created using our pre-processing pipeline. The last column is reference dose.

3.3.3 RT Data Pre-processing

We introduce an effective way to utilize RT geometry by creating spatial matrices (termed as *plates*). As in Figure 5, angle plates are binary masks indicating angles (dynamic beam plates like sectors due to covering a range of angles). The beam plates are created based on angles and PTV masks incorporating commonly-known geometry information. The details of creating angle/beam plates and data augmentation are in Supplement B and C.

Considering that closer regions to the `isocenter` (i.e., planning center) are more interesting for the planner, we centralize the `isocenter` as the volume center (Figure 5) with two physical ranges in our experiments: 96 mm $\times$ 225 mm $\times$ 225 mm and 144 mm $\times$ 256 mm $\times$ 256 mm. The axial range is smaller since treatment beam pass orthogonally to the axial direction. Our major experiments are conducted with 96 mm $\times$ 224 mm $\times$ 224 mm with data size 32 $\times$ 192 $\times$ 192. To separate the RT planning prediction from the clinically-dependent dose prescription, we normalize the reference dose to 0-5 during the training. Seven organs at risk (OARs) have been included in the multiple channel input: brachial, chest wall and rib, esophagus, heart, proximal bronchial, spinal cord, and lung. We follow [41] to impute zeros for missing organ masks.

(a) M&G (b) no-glass (c) female (d) unknown

Figure 6. Illustration with face images demonstrates our method can be generalizable. A single generation model can accept multiple conditions (male and glass) as (a), single condition as in (b) and (c), and even no condition (d) with desired output.

4. Experiments

To evaluation our proposed model, we provide illustration examples with the FFHQ dataset [31] and in-depth validations in 3D dose prediction with two clinical datasets.

4.1. Illustration Example with Face Synthesis

We use a state-of-the-art GAN model (vision-aided mechanism [36] with StyleGAN3 [30]) as the backbone, which was initially developed for image generation without conditions. We extend the official source code [1] with our Flexible-C^m GAN mechanism. FFHQ is a representative face dataset for evaluating GAN [31]. We include two conditions (wearing glasses and gender) trained on 2,000 128$\times$128 images and utilize the public annotation [2].

Figure 6 shows the generated images with a single model from four different situations: (a) multiple conditions, (b,c) partial conditions, and (d) no condition. When all conditions are specified, the generated images follow those conditions (e.g., all faces are male and with glasses in Figure 6a). When one or more conditions are missing, the generated images are with a reasonable diversity in those conditions (e.g., male and female are shown in Figure 6b, 6d).

[1] https://github.com/nupurkmr9/vision-aided-gan/tree/main/stylegan3

[2] https://github.com/DCGM/ffhq-features-dataset

4.2. Dose Prediction for Lung Cancer

We have two datasets for clinical validation: D1WA (Dataset 1 with angle configuration available) and D2NA (Dataset 2 with no angle configuration). After quality assurance, D1WA and D2NA have 365 and 793 patients, respectively. Dataset details are in Supplement F. All the experiments are trained in PyTorch 1.11 [51], and the optimizer is Adam [34] with a weight decay 1e-4. The initial learning rate is set to 1e-4 and is multiplied by 0.3 at the 200-th and 300-th epochs. The max training epoch is 350. The train, validation, and test splits for D1WA and D2NA datasets are 256, 30, 70 and 613, 60,120, respectively.

4.2.1 Comparison to Dose Prediction Models

We compare our method with state-of-the-art models (as below) in 3D dose prediction. All the models has been adapted in lung cancer and for 3D contexts.

Barragan et al. [7] developed a new deep learning architecture for dose prediction combining DenseNet [23] and UNet [52] structures, which was originally proposed to handle heterogeneous beam configuration in IMRT context.

DoseGAN [33] utilize an attention-gated mechanism in the adversarial training contexts to predict dose map for prostate cancer patients.

DeepDoseNet [55] is based on ResNet [22] / Dilated DenseNet structures and with a discrete dose-volume histogram loss [45], which achieves a new state-of-the-art in the 2020 AAPM OpenKBP challenge [6] for head-and-neck cancer patients with IMRT treatment.

Wang et al. [58] utilized IMRT beam masks for to decompose the dose map to sub-fractions, and introduced two value-based and criteria-based DVH losses [6].

Jhanwar et al. [27] proposed a new moment-based loss function, which incorporate DVH metrics in a differentiable manner, to predict 3D dose in IMRT contexts.

The backbone of our **FCGAN**, termed as Dose ResUNet (DRUNet), changes the classical CNN layers in UNet [52] to residual blocks [22]. We add random noise when concatenating features from the left path to the right path while adding multi-conditions when integrated with FCGAN mechanism. DRUNet details are in Supplement D.

4.2.2 Evaluation Metrics

We provide four quantitative metrics for comparison. More discussion of metrics is in Supplement G.

Shift-DV Error (SDE). SDE is derived from our SDV loss (*Definition* 1), which shifts the expected errors of DVH space to voxel space (a holistic view of DVH errors).

Discrete-DVH Error (DDE). DDE measures error between some discrete values of DVH. Motivated by [6], we calculate the absolute difference with following: PTV_1, PTV_{95}, PTV_{99} (doses received by 1%, 95%, 99% of voxels in PTV), and OAR^i_{mean} (the mean dose of i-th OAR).

Models	SDE (↓)	DDE(↓)	MAE(↓)	CEL(↓)
Barragan et al. [7]	7.30	1.95	3.14	2.10
DoseGAN [33]	7.43	1.81	3.33	<u>0.91</u>
DeepDoseNet [55]	<u>6.26</u>	1.60	3.11	2.07
Wang et al. [58]	6.41	<u>1.58</u>	<u>2.71</u>	1.30
Jhanwar et al. [27]	6.78	1.62	3.09	2.02
FCGAN (ours)	**5.80**	**1.48**	**2.64**	**0.05**

Table 1. Results on D1WA dataset with baselines. The **bold** and <u>underline</u> show the best and second best performances.

Models	SDE (↓)	DDE(↓)	MAE(↓)	CEL(↓)
Barragan et al. [7]	5.99	1.69	1.42	2.56
DoseGAN [33]	6.20	1.70	1.52	<u>0.92</u>
DeepDoseNet [55]	<u>5.13</u>	1.53	<u>1.40</u>	2.63
Wang et al. [58]	5.40	<u>1.50</u>	**1.39**	2.61
Jhanwar et al. [27]	5.67	1.64	1.43	2.56
FCGAN$^-$ (ours)	**4.71**	**1.47**	1.49	**0.03**

Table 2. Results on D2NA dataset with baselines. The **bold** and <u>underline</u> show the best and second best performances.

Mean Absolute Error (MAE). MAE measures the mean absolute difference of all voxels between the predicted and the reference dose maps.

To reduce bias from the clinical-dependent dose scale for different patients and make different metrics for a consistent range of doses, we linearly normalized dose maps to 0-10 when computing SDE and 0-50 for MAE and DDE.

Cross-Entropy Loss (CEL). CEL is used for evaluating the planning mode of the predicted dose. It is noteworthy to mention that the pre-trained model used to compute CEL uses the same training set as the pre-trained model in Figure 3 (can lead to some overfit). This metric reflects how consistent of planning mode is between predicted and reference doses to some degree.

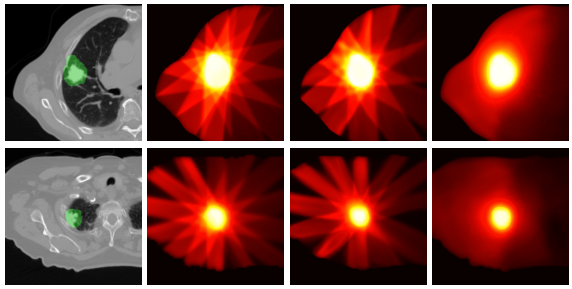
4.2.3 Experimental Results

Table 1 shows that our model outperforms all the representative baselines on all four metrics on D1WA (with angles). It indicates that our FCGAN mechanism utilizes multi-condition (e.g., mode, angles) effectively (e.g., SDE reduces 7.3%, ours vs. the second best [55]).

Table 2 shows our downgraded model (FCGAN$^-$, no angle condition is used in D2NA) has overall superior performance (e.g., best SDE, DDE, and CEL). The dose prediction from baselines can be over-fitted to some Gaussian-like distributions to achieve lower MAE in this heterogeneous context, as in Figure 7. Although our FCGAN$^-$ achieves higher MAE, it is reasonable due to realistic regularization and no angle specified. Also, both the predicted and the ref-

Models	SDE (↓)	DDE(↓)	MAE(↓)	CEL (↓)	SDE (↓)	DDE(↓)	MAE(↓)	CEL (↓)
	test on D2NA, no angle condition				test on D1WA			
FCGAN (train w/ D1WA)	6.62	2.06	1.91	0.14	5.80	1.48	2.64	0.05
FCGAN^{-} (train w/ D2NA)	**4.71**	**1.47**	**1.49**	**0.03**	23.30	7.49	5.61	0.13
FCGAN (train w/ D1WA+D2NA)	5.41	1.50	1.54	0.04	**5.39**	**1.39**	**2.61**	**0.03**

Table 3. Comparison with training combining two datasets. Training or/and testing on D2NA can not use angle condition.



(a) CT/PTV (b) reference (c) ours (d) baseline

Figure 7. Examples show that our model can generate realistic dose maps, which will help the overall KBP pipeline. The baseline (DRUNet) predictions overfit to gaussian-like distributions when the model is trained in this heterogeneous context.

Models	SDE (↓)	DDE(↓)	MAE(↓)	CEL (↓)
train and inference with only beam-static samples				
DRUNet	7.47	1.74	3.65	0.71
FCGAN*	**6.60**	**1.62**	**3.09**	**0.01**
train and inference with only beam-dynamic samples				
DRUNet	6.91	1.61	**1.73**	0.36
FCGAN*	**5.77**	**1.31**	1.91	**0.09**

Table 4. Results on D1WA with single mode. FCGAN* denotes FCGAN only handling single planning mode here.

erence doses can be acceptable due to the subjective nature of RT. More details are in Supplement G.

Table 3 shows the performance when combining those two clinical datasets. Due to the large population gap, model training on one dataset is hard to generalize to another. The performance especially worsens when it is trained on D2NA and tested on D1WA, because D1WA is more heterogeneous, and its angle information cannot be used. Our FCGAN can integrate those two datasets into a single model even if they have different patient populations and conditions, which shows overall (or approximately) superior performance than training with a single dataset. Data specifics and more discussions are in Supplement F.

4.3. Single Mode Models

Our FCGAN is also superior when trained separately on static-beam and dynamic-beam modes, as in Table 4. Interestingly, even for the baseline (DRUNet), the CEL loss is smaller than training two modes together (as in Table 1). The downsides of single-mode models are 1) the requirements for more compute resources for two separate models and 2) potentially limited training samples making deep learning models underutilized.

Models	SDE (↓)	DDE(↓)	MAE(↓)	CEL (↓)
DRUNet	7.01	1.75	3.12	1.12
FCGAN$^{(-2)}$	6.81	1.84	2.72	0.06
FCGAN$^{(-1)}$	6.10	1.57	2.71	**0.04**
FCGAN	**5.80**	**1.48**	**2.64**	0.05

Table 5. Ablation study of different components. FCGAN$^{(-2)}$ has no L_{sdv} and no L_{mc}. FCGAN$^{(-1)}$ has no L_{mc}.

4.4. Ablation Studies

In this section, we have ablation studies on L_{sdv} and L_{mc}. More ablations, including data augmentation, different physical ranges, and different backbones are shown in Supplement C and H.

As shown in Table 5, the FCGAN$^{(-2)}$ (using the FCGAN mechanism but no L_{sdv} and no L_{mc}) utilize the additional conditions effectively with smaller errors, e.g., MAE reduces 13% (from 3.12 to 2.72). Furthermore, L_{sdv} and L_{mc} bring improvements separately, e.g., SDE reduces 10% (from 6.81 to 6.10) by adding L_{sdv} and further reduces 5% (from 6.10 to 5.80) by adding L_{mc}.

4.5. User Intervention Inference

Our model allows users to flexibly change planning mode and angle configurations and instantly make additional inferences of 3D dose map. As in Figure 8, our prediction is close to the reference when the mode/angle conditions are the same as the reference. Further, Our model can have corresponding output by modifying angles and the mode at inference stage and have reasonable outcomes even when user leaves one condition (e.g., angles) empty.

5. Discussion and Limitation

3D dose prediction plays a vital role in knowledge-based planning framework, where it accuracy directly affects the following fluence map prediction [39, 60–62] or other optimization-based models (e.g., dose mimicking [5]). Moreover, the predicted 3D dose can be helpful in a conventional optimization-based pipeline, e.g., for plan validation and for decision support. We proposed a new conditional generative model to predict *individualized* and *real-*

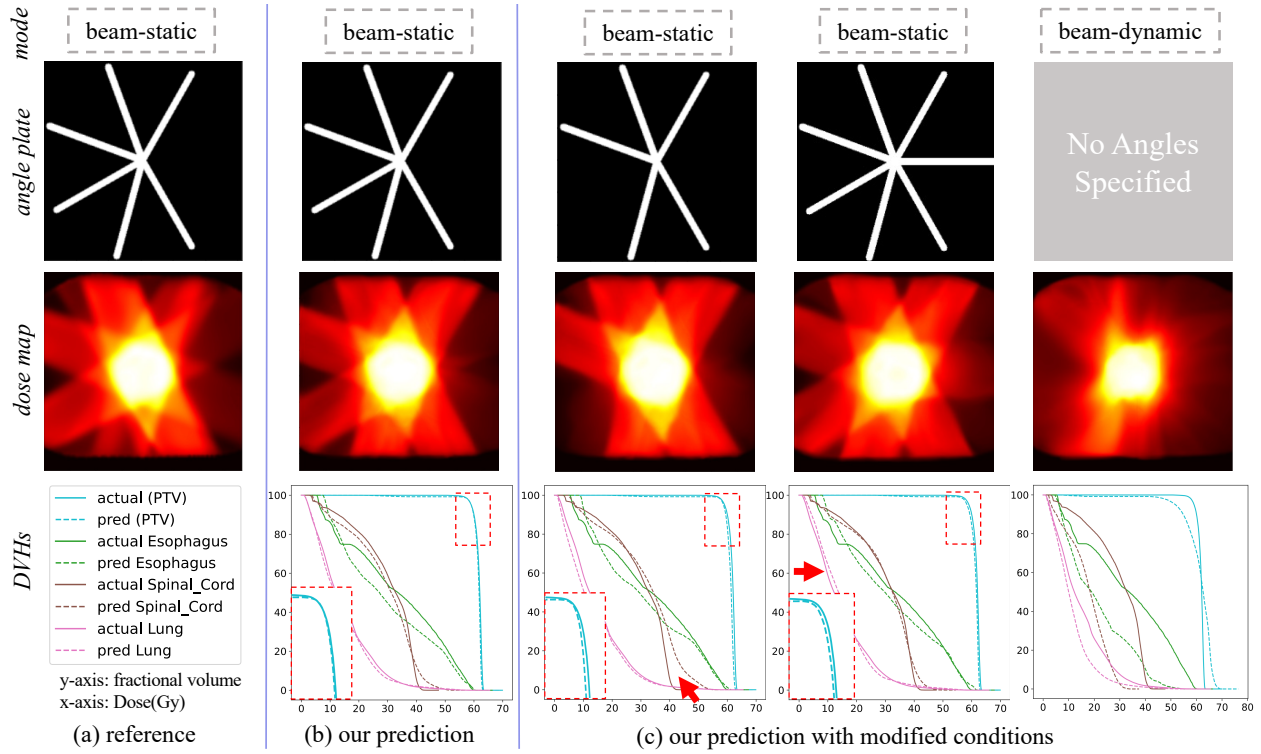


Figure 8. Dose prediction examples allowing user interface. The mode and angle plate are two types of condition that can be modified. When given the same conditions as reference (a), our prediction (b) has a reasonable match with the delivered dose. (c) shows three conditional modifications with the desired prediction, even if a condition is missing. The condition modification will be reflected on predicted dose maps and DVHs. The closer dash line and solid line in DVHs, the better matching between prediction and reference.

istic dose utilizing multi-condition and handle the potential missing condition in practical heterogeneous contexts. During inference, we allow the users to manipulate conditions or provide partial conditions to generate 3D dose (Figure 8). This is essential in a RT planning pipeline since the treatment is subjective to clinicians/experts, and different institutions may have different preferences. Previous dose prediction deep models have been explored mainly in a simpler context, e.g., single mode or single beam configuration. Considering a more practical and challenging context (e.g., lung cancer with multiple modes and different angle configurations), handling different situations with separate models is impractical due to implementation complexity. We enable the usage of additional geometry information and human-machine interaction with a single model.

Limitations and Future Work. Currently, our model has only been tested on lung cancer datasets. Lung cancer is one of the most heterogeneous compared to widely studied cancers (e.g., prostate) with RT planning. We have shown that our model is promising in this challenging context and will yet to show how it generalizes to other cancer sites in the near future. Second, although we use pretraining mechanism [36] (a state-of-the-art GAN) for the face dataset, the exact same approach is not applicable in 3D dose contexts due to a shortage of external pretrained 3D dose model and memory constraints. For future work, we will explore other techniques, for example utilizing self-supervised learning, to obtain appropriate pretrained models and to boost performance. Third, another future work is to build an entire deep learning RT planning pipeline (adding fluence map prediction as in Figure 2 and other necessary parts).

Conclusion. We proposed a new GAN variant Flexible-C^m GAN with two novel loss functions motivated by the practical challenges in RT dose planning. We are the first to advance the precision of 3D dose prediction by adding flexible conditions such as beam geometry and treatment mode within a deep learning framework. Our miss-consistency loss upgrades our model to be able to handle flexible conditions. Shift-dose-volume loss makes widely used DVH metric differentiable and allows us to use it in a deep learning context reducing the possibility of sub-optimal solutions. We conducted extensive experiments on two clinical datasets to validate our model and its sub-components. We also demonstrated that Flexible-C^m GAN is generalizable to other tasks (e.g., face synthesis).

Disclaimer. The information in this paper is based on research results that are not commercially available. Future commercial availability cannot be guaranteed.

References

[1] Varian rapidplan knowledge-based planning. https://www.varian.com/products/radiotherapy/treatment-planning/rapidplan-knowledge-based-planning. Accessed: 2022-09-30. 2

[2] Lindsey M. Appenzoller, Jeff M. Michalski, Wade L. Thorstad, Sasa Mutic, and Kevin L. Moore. Predicting dose-volume histograms for organs-at-risk in IMRT planning. *Medical Physics*, 2012. 2

[3] Aaron Babier, Justin J. Boutilier, Andrea L. McNiven, and Timothy C.Y. Chan. Knowledge-based automated planning for oropharyngeal cancer. *Medical Physics*, 2018. 2

[4] Aaron Babier, Rafid Mahmood, Andrea L. McNiven, Adam Diamant, and Timothy C.Y. Chan. Knowledge-based automated planning with three-dimensional generative adversarial networks. *Medical Physics*, 2020. 2, 3

[5] Aaron Babier, Rafid Mahmood, Binghao Zhang, Victor G L Alves, Ana Maria Barragán, Barrag´ Barragán-Montero, Joel Beaudry, Carlos E Cardenas, Yankui Chang, Zijie Chen, Jaehee Chun, Kelly Diaz, Harold David Eraso, Erik Faustmann, Sibaji Gaj, Skylar Gay, Mary Gronberg, Bingqi Guo, Junjun He, Gerd Heilemann, Sanchit Hira, Yuliang Huang, Fuxin Ji, Dashan Jiang, Jean Carlo, Jimenez Giraldo, Hoyeon Lee, Jun Lian, Shuolin Liu, Keng-Chi Liu, José Jos´, José Marrugo, Kentaro Miki, Kunio Nakamura, Tucker Netherton, Dan Nguyen, Hamidreza Nourzadeh, Alexander F I Osman, Zhao Peng, José Darío, Quinto Muñoz, Mu˜ Muñoz, Christian Ramsl, Dong Joo Rhee, Juan David Rodriguez, Hongming Shan, Jeffrey V Siebers, Mumtaz H Soomro, Kay Sun, Andrés Andr´ Andrés, Usuga Hoyos, Carlos Valderrama, Rob Verbeek, Enpei Wang, Siri Willems, Qi Wu, Xuanang Xu, Sen Yang, Lulin Yuan, Simeng Zhu, Lukas Zimmermann, Kevin L Moore, Thomas G Purdie, Andrea L Mcniven, and Timothy C Y Chan. OpenKBP-Opt: An international and reproducible evaluation of 76 knowledge-based planning pipelines. *Physics in Medicine & Biology*, 2022. 3, 7

[6] Aaron Babier, Binghao Zhang, Rafid Mahmood, Kevin L. Moore, Thomas G. Purdie, Andrea L. McNiven, and Timothy C. Y. Chan. OpenKBP: The open-access knowledge-based planning grand challenge. *Medical Physics*, 2020. 1, 2, 3, 4, 6

[7] Ana María Barragán-Montero, Dan Nguyen, Weiguo Lu, Mu Han Lin, Roya Norouzi-Kandalan, Xavier Geets, Edmond Sterpin, and Steve Jiang. Three-dimensional dose prediction for lung IMRT patients with deep neural networks: robust learning from heterogeneous beam configurations. *Medical Physics*, 2019. 2, 6

[8] Sam Bond-Taylor, Adam Leach, Yang Long, and Chris G. Willcocks. Deep Generative Modelling: A Comparative Review of VAEs, GANs, Normalizing Flows, Energy-Based and Autoregressive Models. *arXiv:2103.04922*, 2021. 2

[9] Andrew Brock, Jeff Donahue, and Karen Simonyan. Large Scale GAN Training for High Fidelity Natural Image Synthesis. In *ICLR*, 2019. 2

[10] S Chiavassa, I Bessieres, M Edouard, M Mathot, and A Moignier. Complexity metrics for IMRT and VMAT plans: a review of current literature and applications. Technical report, 2019. 1

[11] Yunjey Choi, Minje Choi, Munyoung Kim, Jung Woo Ha, Sunghun Kim, and Jaegul Choo. StarGAN: Unified Generative Adversarial Networks for Multi-domain Image-to-Image Translation. In *IEEE CVPR*, 2018. 2

[12] Geoff Delaney, Susannah Jacob, Carolyn Featherstone, and Michael Barton. The role of radiotherapy in cancer treatment: Estimating optimal utilization from a review of evidence-based clinical guidelines. *Cancer*, 2005. 1

[13] R E Drzymala, ' R Mohan, L Brewster, J Chu, M Goitein, W Harms, B S ' And, and M Urie. DOSE-VOLUME HISTOGRAMS. *Inr. J. Radrarion Oncology Biol. Phys*, 1991. 2, 4

[14] M. A. Ebert, A. Haworth, R. Kearvell, B. Hooton, B. Hug, N. A. Spry, S. A. Bydder, and D. J. Joseph. Comparison of DVH data from multiple radiotherapy treatment planning systems. *Physics in medicine and biology*, 2010. 2

[15] Jacques Ferlay, Isabelle Soerjomataram, Rajesh Dikshit, Sultan Eser, Colin Mathers, Marise Rebelo, Donald Maxwell Parkin, David Forman, and Freddie Bray. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *International Journal of Cancer*, 2015. 1

[16] A. Fogliata, L. Cozzi, G. Reggiori, A. Stravato, F. Lobefalo, C. Franzese, D. Franceschini, S. Tomatis, and M. Scorsetti. RapidPlan knowledge based planning: Iterative learning process and model ability to steer planning strategies. *Radiation Oncology*, 2019. 2

[17] Riqiang Gao, Yucheng Tang, Kaiwen Xu, Ho Hin Lee, Steve Deppen, Kim Sandler, Pierre Massion, Thomas A. Lasko, Yuankai Huo, and Bennett A. Landman. Lung Cancer Risk Estimation with Incomplete Data: A Joint Missing Imputation Perspective. In *MICCAI*, 2021. 2, 3, 4

[18] Arnab Ghosh, Richard Zhang, Puneet K Dokania, Oliver Wang, Alexei A Efros, Philip H S Torr, and Eli Shechtman. Interactive Sketch & Fill: Multiclass Sketch-to-Image Translation. In *ICCV*, 2019. 2

[19] Ian J Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *NIPS*, 2014. 2, 4

[20] Surbhi Grover, Melody J. Xu, Alyssa Yeager, Lori Rosman, Reinou S. Groen, Smita Chackungal, Danielle Rodin, Margaret Mangaali, Sommer Nurkic, Annemarie Fernandes, Lilie L. Lin, Gillian Thomas, and Ana I. Tergas. A Systematic Review of Radiotherapy Capacity in Low- and Middle-Income Countries. *Frontiers in Oncology*, 2015. 1

[21] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron Courville. Improved Training of Wasserstein GANs. In *NeurIPS*, 2017. 2

[22] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *IEEE CVPR*, 2016. 6

[23] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely Connected Convolutional Networks. In *IEEE CVPR*, 2017. 6

[24] Yuankai Huo, Zhoubing Xu, Hyeonsoo Moon, Shunxing Bao, Albert Assad, Tamara K. Moyo, Michael R. Savona, Richard G. Abramson, and Bennett A. Landman. SynSeg-Net: Synthetic Segmentation Without Target Modality Ground Truth. *IEEE Transactions on Medical Imaging*, 2019. 2

[25] Elizabeth Huynh, Ahmed Hosny, Christian Guthier, Danielle S. Bitterman, Steven F. Petit, Daphne A. Haas-Kogan, Benjamin Kann, Hugo J.W.L. Aerts, and Raymond H. Mak. Artificial intelligence in radiation oncology. *Nature Reviews Clinical Oncology*, 2020. 1

[26] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-Image Translation with Conditional Adversarial Networks. In *IEEE CVPR*, 2017. 2

[27] Gourav Jhanwar, Navdeep Dahiya, Parmida Ghahremani, Masoud Zarepisheh, and Saad Nadeem. Domain knowledge driven 3D dose prediction using moment-based loss function. *Physics in Medicine and Biology*, 2022. 2, 4, 6

[28] Dakai Jin, Ziyue Xu, Youbao Tang, Adam P. Harrison, and Daniel J. Mollura. CT-realistic lung nodule simulation from 3D conditional generative adversarial networks for robust lung segmentation. In *MICCAI*, volume 11071 LNCS, 2018. 2

[29] Michael K.M. Kam, Ricky M.C. Chau, Joyce Suen, Peter H.K. Choi, and Peter M.L. Teo. Intensity-modulated radiotherapy in nasopharyngeal carcinoma: dosimetric advantage over conventional plans and feasibility of dose escalation. *International Journal of Radiation Oncology Biology Physics*, 2003. 2

[30] Tero Karras, Miika Aittala, Samuli Laine, Erik Härkönen, Janne Hellsten, Jaakko Lehtinen, and Timo Aila. Alias-Free Generative Adversarial Networks. In *NeurIPS*, 2021. 5

[31] Tero Karras, Samuli Laine, and Timo Aila. A Style-Based Generator Architecture for Generative Adversarial Networks. In *IEEE CVPR*, 2019. 5

[32] Vasant Kearney, Jason W. Chan, Samuel Haaf, Martina Descovich, and Timothy D. Solberg. DoseNet: a volumetric dose prediction algorithm using 3D fully-convolutional neural networks. *Physics in Medicine & Biology*, 2018. 2, 3

[33] Vasant Kearney, Jason W. Chan, Tianqi Wang, Alan Perry, Martina Descovich, Olivier Morin, Sue S. Yom, and Timothy D. Solberg. DoseGAN: a generative adversarial network for synthetic dose prediction using attention-gated discrimination and generation. *Scientific Reports*, 2020. 1, 2, 3, 6

[34] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *ICLR*, 2015. 6

[35] Kazuki Kubo, Hajime Monzen, Kentaro Ishii, Mikoto Tamura, Ryu Kawamorita, Iori Sumida, Hirokazu Mizuno, and Yasumasa Nishimura. Dosimetric comparison of RapidPlan and manually optimized plans in volumetric modulated arc therapy for prostate cancer. *Physica Medica*, 2017. 2

[36] Nupur Kumari, Richard Zhang, Eli Shechtman, and Jun-Yan Zhu. Ensembling Off-the-shelf Models for GAN Training. In *IEEE CVPR*, 2022. 2, 5, 8

[37] Yann Lecun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 2015. 1

[38] Dongwook Lee, Junyoung Kim, Won Jin Moon, and Jong Chul Ye. Collagan: Collaborative gan for missing image data imputation. In *IEEE CVPR*, 2019. 2

[39] Hoyeon Lee, Hojin Kim, Jungwon Kwak, Young Seok Kim, Sang Wook Lee, Seungryong Cho, and Byungchul Cho. Fluence-map generation for prostate intensity-modulated radiotherapy planning using a deep-neural-network. *Scientific Reports*, 2019. 3, 7

[40] Steven Cheng Xian Li and Benjamin M Marlin. Learning from irregularly-sampled time series: A missing data perspective. In *ICML*, 2020. 2

[41] Shuolin Liu, Jingjing Zhang, Teng Li, Hui Yan, and Jianfei Liu. Technical Note: A cascade 3D U-Net for dose prediction in radiotherapy. *Medical physics*, 2021. 2, 4, 5

[42] Lin Ma, Mingli Chen, Xuejun Gu, and Weiguo Lu. Generalizability of deep learning based fluence map prediction as an inverse planning approach. *arXiv:2104.15032*, 4 2021. 3

[43] Chris McIntosh, Leigh Conroy, Michael C. Tjong, Tim Craig, Andrew Bayley, Charles Catton, Mary Gospodarowicz, Joelle Helou, Naghmeh Isfahanian, Vickie Kong, Tony Lam, Srinivas Raman, Padraig Warde, Peter Chung, Alejandro Berlin, and Thomas G. Purdie. Clinical integration of machine learning for curative-intent radiation treatment of patients with prostate cancer. *Nature Medicine*, 2021. 1

[44] Mehdi Mirza and Simon Osindero. Conditional Generative Adversarial Nets. *arXiv:1411.1784*, 2014. 2, 3

[45] Dan Nguyen, Rafe Mcbeth, Azar Sadeghnejad Barkousaraie, Gyanendra Bohara, Chenyang Shen, Xun Jia, and Steve Jiang. Incorporating human and learned domain knowledge into training deep neural networks: A differentiable dose volume histogram and adversarial inspired framework for generating Pareto optimal dose distributions in radiation therapy HHS Public Access. *Medical Physics*, 2020. 2, 3, 4, 6

[46] Obioma Nwankwo, Hana Mekdash, Dwi Seno Kuncoro Sihono, Frederik Wenz, and Gerhard Glatting. Knowledge-based radiation therapy (KBRT) treatment planning versus planning by experts: Validation of a KBRT algorithm for prostate cancer treatment planning. *Radiation Oncology*, 2015. 1

[47] Augustus Odena, Christopher Olah, and Jonathon Shlens. Conditional Image Synthesis With Auxiliary Classifier GANs. In *ICML*, 2016. 3, 4

[48] Karl Otto. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Medical Physics*, 2008. 1

[49] Hyojin Park, YoungJoon Yoo, and Nojun Kwak. MC-GAN: Multi-conditional Generative Adversarial Network for Image Synthesis. In *BMVC*, 2018. 2

[50] Taesung Park, Alexei A Efros, Richard Zhang, and Jun-Yan Zhu. Contrastive Learning for Unpaired Image-to-Image Translation. In *ECCV*, 2020. 2

[51] Adam et al. Paszke. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *NeurIPS*, 2019. 6

[52] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, 2015. 6

[53] Berkman Sahiner, Aria Pezeshk, Lubomir M. Hadjiiski, Xiaosong Wang, Karen Drukker, Kenny H. Cha, Ronald M.

Summers, and Maryellen L. Giger. Deep learning in medical imaging and radiation therapy. *Medical Physics*, 2019. 1

[54] Satomi Shiraishi, Jun Tan, Lindsey A. Olsen, and Kevin L. Moore. Knowledge-based prediction of plan quality metrics in intracranial stereotactic radiosurgery. *Medical Physics*, 2015. 1

[55] Mumtaz Hussain Soomro, Victor Gabriel, Leandro Alves, Hamidreza Nourzadeh, and Jeffrey V Siebers. DeepDoseNet: A Deep Learning model for 3D Dose Prediction in Radiation Therapy. Technical report, arxiv, 2021. 2, 6

[56] Bin S. Teh, Shiao Y. Woo, and E. Brian Butler. Intensity Modulated Radiation Therapy (IMRT): A New Promising Technology in Radiation Oncology. *The Oncologist*, 1999. 1

[57] May Teoh, C. H. Clark, K. Wood, S. Whitaker, and A. Nisbet. Volumetric modulated arc therapy: A review of current literature and clinical use in practice. *British Journal of Radiology*, 2011. 1

[58] Bin Wang, Lin Teng, Lanzhuju Mei, Zhiming Cui, Xuanang Xu, Qianjin Feng, and Dinggang Shen. Deep Learning-Based Head and Neck Radiotherapy Planning Dose Prediction via Beam-Wise Dose Decomposition. In *MICCAI*, 2022. 2, 4, 6

[59] Chunhao Wang, Xiaofeng Zhu, Julian C. Hong, and Dandan Zheng. Artificial Intelligence in Radiotherapy Treatment Planning: Present and Future. *Technology in cancer research & treatment*, 2019. 1, 2

[60] Mingqing Wang, Qilin Zhang, Saikit Lam, Jing Cai, and Ruijie Yang. A Review on Application of Deep Learning Algorithms in External Beam Radiotherapy Automated Treatment Planning. *Frontiers in Oncology*, 2020. 7

[61] Wentao Wang, Yang Sheng, Manisha Palta, Brian Czito, Christopher Willett, Martin Hito, Fang-Fang Yin, Qiuwen Wu, Yaorong Ge, and Q. Jackie Wu. Deep Learning-Based Fluence Map Prediction for Pancreas Stereotactic Body Radiation Therapy With Simultaneous Integrated Boost. *Advances in Radiation Oncology*, 2021. 3, 7

[62] Wentao Wang, Yang Sheng, Chunhao Wang, Jiahan Zhang, Xinyi Li, Manisha Palta, Brian Czito, Christopher G. Willett, Qiuwen Wu, Yaorong Ge, Fang Fang Yin, and Q. Jackie Wu. Fluence Map Prediction Using Deep Learning Models – Direct Plan Generation for Pancreas Stereotactic Body Radiation Therapy. *Frontiers in Artificial Intelligence*, 2020. 3, 7

[63] Xin Yi, Ekta Walia, and Paul Babyn. Generative adversarial network in medical imaging: A review. *Medical image analysis*, 2019. 2

[64] Jingjing Zhang, Shuolin Liu, Hui Yan, Teng Li, Ronghu Mao, and Jianfei Liu. Predicting voxel-level dose distributions for esophageal radiotherapy using densely connected network with dilated convolutions. *Physics in medicine and biology*, 2020. 2, 3

[65] Jun Yan Zhu, Taesung Park, Phillip Isola, and Alexei A. Efros. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. In *ICCV*, 2017. 2