Medical Image Analysis 92 (2024) 103045

Contents lists available at ScienceDirect

Medical Image Analysis

journal homepage: www.elsevier.com/locate/media

Beam-wise dose composition learning for head and neck cancer dose prediction in radiotherapy

Lin Teng [a,b,1], Bin Wang [a,1], Xuanang Xu [c,1], Jiadong Zhang [a], Lanzhuju Mei [a], Qianjin Feng [b], Dinggang Shen [a,d,e,*]

[a] School of Biomedical Engineering, ShanghaiTech University, Shanghai 201210, China
[b] School of Biomedical Engineering, Southern Medical University, Guangzhou 510515, China
[c] Department of Biomedical Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180, USA
[d] Department of Research and Development, Shanghai United Imaging Intelligence Co., Ltd., Shanghai 200230, China
[e] Shanghai Clinical Research and Trial Center, Shanghai 201210, China

ARTICLE INFO

Keywords:
Radiation therapy
Head and neck cancer
Dose prediction
Beam-wise dose learning
Dose-volume histogram

ABSTRACT

Automatic and accurate dose distribution prediction plays an important role in radiotherapy plan. Although previous methods can provide promising performance, most methods did not consider beam-shaped radiation of treatment delivery in clinical practice. This leads to inaccurate prediction, especially on beam paths. To solve this problem, we propose a beam-wise dose composition learning (BDCL) method for dose prediction in the context of head and neck (H&N) radiotherapy plan. Specifically, a global dose network is first utilized to predict coarse dose values in the whole-image space. Then, we propose to generate individual beam masks to decompose the coarse dose distribution into multiple field doses, called beam voters, which are further refined by a subsequent beam dose network and reassembled to form the final dose distribution. In particular, we design an overlap consistency module to keep the similarity of high-level features in overlapping regions between different beam voters. To make the predicted dose distribution more consistent with the real radiotherapy plan, we also propose a dose-volume histogram (DVH) calibration process to facilitate feature learning in some clinically concerned regions. We further apply an edge enhancement procedure to enhance the learning of the extracted feature from the dose falloff regions. Experimental results on a public H&N cancer dataset from the AAPM OpenKBP challenge show that our method achieves superior performance over other state-of-the-art approaches by significant margins. Source code is released at https://github.com/TL9792/BDCLDosePrediction.

1. Introduction

Head and neck (H&N) cancer is a broad category of diverse cancer types, originating from various soft tissue, glands, and bones (Pai and Westra, 2009). To kill cancerous cells while avoiding normal tissue damage, external radiation therapy (RT) is regarded as the preferred treatment, aiming to deliver a high radiation dose (*i.e.*, prescription dose) to the planning target volume (PTV) while minimizing the dose to organs-at-risk (OARs) via multiple focused radiation beams (Khan, 2010). The radiation delivery of RT is followed by RT plan. Currently, one of the common treatments for external RT is intensity-modulated radiation therapy (IMRT), in which delivered beams are highly conformal to the PTV, and the radiation intensity for each beam can be modulated individually (Webb, 2003). Thus, the RT plan in IMRT is more acceptable physically in the clinical workflow.

Dose distribution design in RT plan is a complex process, including CT image acquisition, ROI contouring on the acquired CT image by radiation oncologist manually (Fig. 1(b) and (e)), treatment parameter arrangement (*e.g.*, geometry of beams), and plan parameter optimization (*e.g.*, dose-volume objectives). The result is a spatial dose distribution, which is called a dose distribution map. The intensity value in each voxel represents the amount of radiation dose accepted by the body in the unit of Gray (Gy) (Khan, 2010).

To obtain clinically acceptable RT plan, dosimetrists need to manually adjust treatment parameters in a trial-and-error manner, this process costs hours so that delaying the best treatment period for each patient (Kearney et al., 2018b). In addition, the quality of RT plan has high variability between inter- and intra-institutions due to differences in technological parameters (including treatment planning

* Corresponding author.
E-mail address: dinggang.shen@gmail.com (D. Shen).
[1] The authors contribute equally to this work.

https://doi.org/10.1016/j.media.2023.103045
Received 29 October 2022; Received in revised form 12 October 2023; Accepted 27 November 2023
Available online 9 December 2023

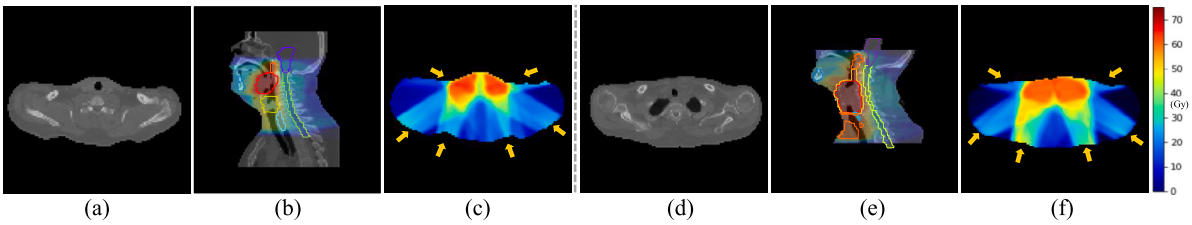

Fig. 1. Examples of two H&N RT plans, with each side of the line corresponding to one subject. Each subject includes the CT images shown as (a) and (d) in axial view; ROI contours shown as (b) and (e) superimposed on the dose distribution map and plan CT image in sagittal view; the dose distribution map shown as (c) and (f) in axial view, where yellow arrows indicate different beam paths.

system and modality) as well as planner's skill level (*e.g.*, years of experience and education) (Nelms et al., 2012). The above-mentioned reasons would result in sub-optimal RT plan and thus affect the result of the plan (Peters et al., 2010). Considering the challenge and time-consuming nature of manual RT plan, developing automated methods (e.g., knowledge-based planning Momin et al., 2021; Shiraishi and Moore, 2016; Ge and Wu, 2019; Babier et al., 2020) is of great clinical value.

Recently, due to the fast development of deep learning (DL), especially the convolutional neural network (CNN) as well as its variants (Ronneberger et al., 2015; Milletari et al., 2016), great success has been achieved in solving a broad array of computer vision problems (Liu et al., 2019; Shan et al., 2021; Luan et al., 2008; Jia et al., 2012; Zacharaki et al., 2008). To automatically predict dose distribution map, many DL-based methods have been proposed, which can be generally classified into three categories: 1) designing advanced network architectures, such as C3D (Liu et al., 2021), DoseNet (Kearney et al., 2018a), HD U-net (Nguyen et al., 2019) and DCNN (Gronberg et al., 2021); 2) introducing additional prior knowledge, such as distance map (Zhang et al., 2019) and gradient map (Tan et al., 2021); 3) proposing domain-specific loss functions, such as dose-volume histogram (DVH) loss (Nguyen et al., 2020). However, these DL-based methods are still facing limitations. For instance, the first group of methods often focuses on improving the model's global performance while losing accuracy for some local regions related to hard-to-learn features. The second group of methods considers physical prior of dose distribution to facilitate the learning of discriminative features while ignoring the geometric prior of beam-shaped radiation in RT, thus causing poor performance along beam paths. The last group of methods utilizes elaborated loss functions to regularize the key indices while imposing high computational overhead and GPU memory consumption on the model training.

In this work, to deal with the aforementioned limitations and achieve high-performance automatic RT plan, we propose a beam-wise dose composition learning (BDCL) method to gradually estimate the dose distribution map in a three-stage (global-beam-global) manner. Specifically, we first employ a global dose network (GDN) to predict a coarse dose distribution over the whole-image space. Then, the coarse dose distribution is decomposed into a series of field doses (beam voters) and further refined by a beam dose network (BDN) according to the geometric prior of the radiation beams. Finally, all the refined beam voters are reassembled into a new global dose distribution, which is further refined by our proposed edge enhancement and DVH calibration processes to meet clinical criteria. We conduct extensive experiments on a public H&N cancer RT dataset, the experimental results show that our method outperforms other state-of-the-art methods by a significant margin and the predicted dose distribution is much closer to the physically deliverable one by using the machine parameters and beam fluence that deliver it. In summary, our main contributions are four-fold:

- We propose to generate beam masks as the prior knowledge of beam-wise radiation delivery by a novel beam mask generator, which guide and decompose the coarse dose distribution map into multiple field doses. This process exploits dose distribution on the beam paths in a beam-wise way, which decomposes the difficult task into a few easy-to-learn sub-tasks.
- We propose an overlap consistency module to make the predictions of overlapped regions between different beam voters consistent, which improves the accuracy of the prediction and accelerates the convergence speed of the model.
- We present a novel multi-beam voting mechanism to reassemble the global dose distribution map from the multiple beam voters, which lays the foundation for global-wise dose refinement.
- We integrate DVH metrics into DL model training by the proposed DVH calibration process, which makes the prediction in ROIs more accurate and efficient. Besides, we also apply edge enhancement to enhance boundary learning, making the prediction sharper.

This work is a substantial extension of our conference paper published on MICCAI 2022 (Wang et al., 2022) in the following highlighted aspects. First, we further improve the performance of our method by proposing the overlap consistency module and the edge enhancement process. Second, we conduct comprehensive ablation studies on the proposed method to justify our designs in a more systematic way. Third, we introduce more radiotherapy-specific DVH metrics to evaluate experimental results, demonstrating that our prediction is closer to the real clinical RT plan than predictions of the state-of-the-art methods in terms of clinical criteria. Last, we have thorough discussions on this study, regarding experimental results, strengths, and limitations of the proposed method.

2. Related works

Automatic and efficient dose prediction in RT plan is highly desired in clinical practice (Kearney et al., 2018b; Zheng et al., 2019; Wang et al., 2020). To achieve deliverable dose prediction, many methods leverage prior knowledge from delivered high-quality RT plan, namely knowledge-based planning (KBP), to produce an unseen patient plan (Wu et al., 2009). According to recent clinical research statistics, KBP holds great promise and application in RT practice, since it can accelerate process of RT plan and reduce manual interventions (McIntosh et al., 2021; Scaggion et al., 2023; Fogliata et al., 2019; El Naqa, 2021). The KBP methods can be divided into two major categories: 1) traditional KBP methods and 2) DL-based KBP methods.

2.1. Traditional KBP methods

Traditional KBP methods utilize various anatomical and geometrical information to build a mathematical or statistical model to predict dosimetry features for a new patient. For example, Chanyavanich et al.

(2011) developed a case-similarity algorithm to identify a matched case for a given query case and adjusted various treatment parameters from the matched case to transfer new plan. Boutilier et al. (2015) utilized overlap volume histogram (OVH) information to predict weight values of dose constraints of OAR by K-nearest neighbors and multinomial logistic regression, which can guide the generation process of new plan. Amit et al. (2015) employed a random forest regression algorithm to predict high-quality dose distribution by learning the relationships between beam angles and anatomical features. Moreover, Campbell et al. (2017) proposed an artificial neural network dose model (ANN-DM) to calculate dose distribution according to a set of geometric parameters (distance-to-PTV, distance-to-OAR, etc.) and plan parameters (PTV volume, photon beam energy, etc.). These methods can achieve decent results when dealing with easy cases or specific datasets. However, due to the limited representation ability of the handcrafted features, their performance generalized to the hard cases or multi-institute data is still under expectation.

2.2. DL-based KBP methods

In contrast to traditional KBP methods, DL-based methods can directly learn discriminative features from the raw data driven by the target. For instance, Kearney et al. (2018a) designed a novel neural network (DoseNet) based on V-Net (Milletari et al., 2016) for prostatic dose prediction. Zhang et al. (2020) applied U-Net-like architecture and dense feature aggregation block to capture features of multiple scales for dose prediction in esophageal cancer patients, known as a densely-connected neural network (DCNN). To further improve the prediction performance and training efficiency, Nguyen et al. (2019) proposed a hierarchically densely connected U-net (HD U-net) architecture based on U-Net (Çiçek et al., 2016) and DenseNet (Huang et al., 2017) for dose distribution prediction in H&N cancer. Moreover, Xu et al. (2021) applied a triple-stage cascaded U-Net to predict H&N cancer dose in a coarse-to-fine manner using the auto-context mechanism. During the OpenKBP-2020 AAPM Grand Challenge, Liu et al. (2021) proposed a U-Net-like cascaded three-dimensional model (C3D) with data augmentation and knowledge distillation technique for dose prediction in H&N cancer, achieving the first place in the Challenge. Besides, Gronberg et al. (2021) designed a 3D U-Net-like network with a densely connected sequence of dilated convolutions as the bottleneck level for dose distribution prediction, achieving second place in the Challenge. To propagate low-level features into the deep layers of the network, Zimmermann et al. (2021) employed ResNet blocks after each down- and up-sampling convolution block and trained the model with the one-cycle scheduler for dose distribution prediction. By employing novel network architectures or advanced training strategies, all the top-ranking methods consider the performance of the global dose distribution while ignoring local information that can be learned from some key regions, which leads to sub-optimal solutions in the prediction. Therefore, in this paper, we propose the BDCL method to consider more local information for facilitating the dose distribution prediction.

3. Method

The workflow of our proposed method is illustrated in Fig. 3. We first estimate a coarse dose distribution over the whole image in the global-wise dose learning stage (Section 3.1). Then, the coarse dose distribution map is fine-tuned along different beam paths individually through the beam-wise dose refinement stage (Section 3.2). Finally, the beam-wise refined dose distribution map is further improved by the edge enhancement and DVH calibration process in the global-wise dose refinement stage (Section 3.3). The implementation details of the proposed method are provided in Section 3.4.

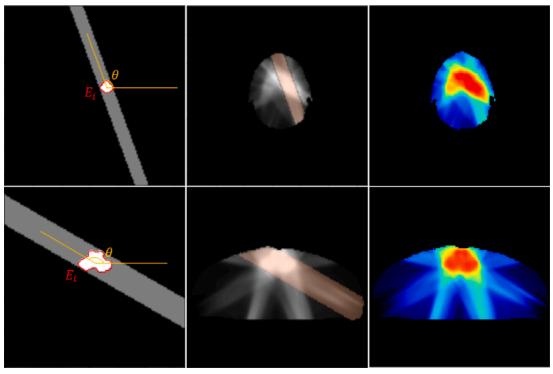

Fig. 2. Illustration of beam mask generation. It shows that the generated beam masks can reasonably characterize the beam paths in radiotherapy. Each row represents a different slice location in the plan CT image.

3.1. Global-wise dose learning

As the first step of our method, we employ GDN to roughly estimate the dose distribution as an initialization for the subsequent steps. The GDN has a 3D U-Net (Ronneberger et al., 2015) architecture with minor modification (as detailed in Section 3.4) and takes as input the CT image concatenated with the segmentation masks of the PTV&OAR. The output of the GDN is a 3D coarse dose distribution map (in the same size as the CT image), which is supervised by the corresponding ground-truth dose distribution map via mean absolute error (MAE) loss during training.

Through this stage, we aim to get a coarse dose distribution map that can generally represent the dose distribution over the whole image. However, due to less consideration of the geometric relationship between beams and ROIs, the predicted dose values are often inaccurate in some local regions (*e.g.*, around regions of paths and ROIs). Therefore, we propose the following beam-wise dose refinement step, as well as the global-wise dose refinement step, to solve the problem.

3.2. Beam-wise dose refinement

In this stage, we refine the coarse dose distribution map by looking into multiple field dose distribution maps individually contributed by different beams. The critical techniques are listed below.

3.2.1. Beam mask generation

In IMRT for H&N cancer, there are typically seven to nine co-planar beams with predefined angular positions, focusing together toward the PTV. Delivering the high-dose radiation in this way will lead to a spoke-shaped dose distribution. Following this geometric prior knowledge, the beam paths can be simulated with the location of PTV regions and the pre-defined angles in the treatment plan, which are represented as a set of binary masks, called beam masks. Since the PTV location varies across the CT slices, the beam masks are built slice-by-slice to make them closer to the real beam paths. Specifically, given the PTV contour E_i in a CT slice, we draw two parallel tangent lines along the predefined beam angle θ. The band region between the two parallel tangent lines is marked as the beam mask, which is illustrated in Fig. 2.

3.2.2. Beam-wise dose learning

In clinical RT plan, the coarse dose distribution map is individually optimized along different beam directions, resulting in differences in the field dose distribution. Such variation in beam-wise dose distribution leads to difficulty in directly learning a dose distribution map to conform all beam paths simultaneously. To overcome this challenge, we propose a beam-wise decomposition strategy. Specifically, we utilize the beam masks to decompose the coarse dose distribution map into

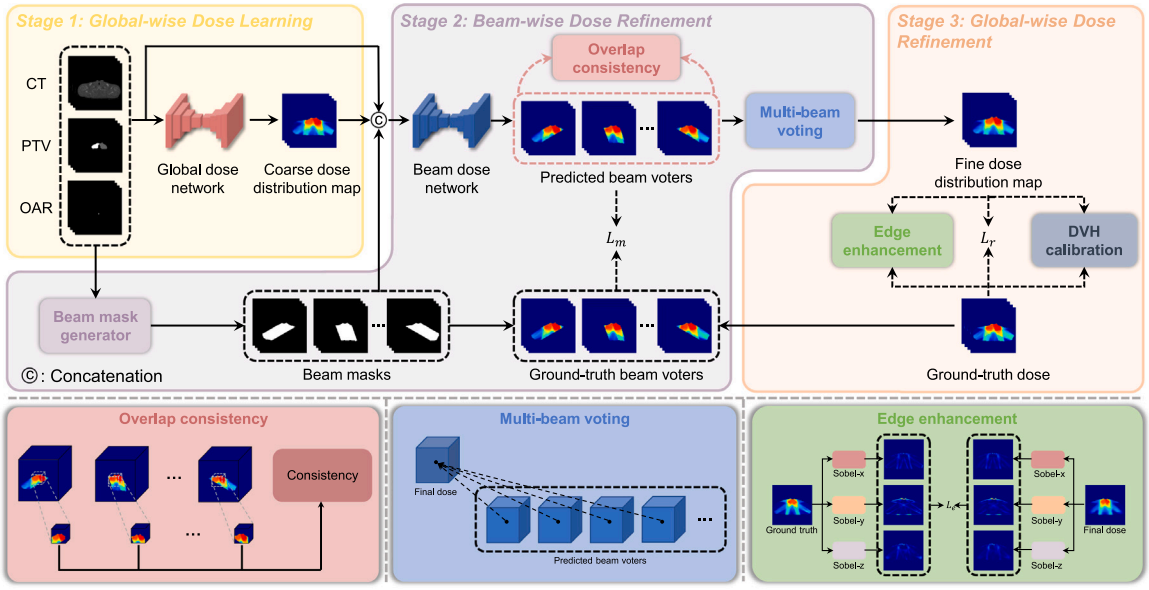

Fig. 3. Overview of the proposed beam-wise dose composition learning (BDCL) method for dose prediction.

multiple field dose distributions, called beam voters. Each beam voter represents the dose distribution prediction of one certain beam path. In this manner, the global dose prediction task is converted into a set of field dose prediction tasks, which can learn the features of each different beam path independently. Given the coarse global dose distribution map produced by the preceding GDN, we concatenate it with the CT image, the PTV&OAR masks, and the beam masks to composite a multi-channel input for the BDN to predict multiple beam voters. The BDN can be any fully convolutional networks (FCNs) whose output has the same spatial size as the input. In this study, we adopt two representative FCN architectures (3D U-Net and ResUnet) as the BDN, whose output is supervised by the ground-truth beam voters through a beam-wise mean MAE loss as below:

$$L_m(P,G) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N_i}\left|P_{ij}-G_{ij}\right|}{\sum_{i=1}^{N} N_i}. \tag{1}$$

Here, N is the number of beam voters, and N_i is the maximum number of voxels in the ith beam voter. P_{ij} and G_{ij} refer to the dose value of the jth voxel in the ith beam voter for the prediction and the ground truth, respectively.

In addition, since multiple beams are focusing on the common PTV regions, beams with different directions have overlapping regions between each other, including but not limited to the PTV regions. These regions should be consistent and predicted to have the same dose values. However, after beam-wise decomposition, each beam voter is supervised independently, causing the risk of inconsistency among the overlapping regions. To solve this problem, we impose the overlap consistency constraint on all the beam voters. Specifically, supposing one voxel in the coarse dose distribution map is passed by k beam voters, in which the predicted dose values are $\alpha_1, \alpha_2, \ldots, \alpha_k$, respectively. We regularize these k values to be consistent with their mean value $\overline{\alpha} = \frac{1}{k}\sum_{j=1}^{k}\alpha_j$. The consistency loss is defined as:

$$L_{cs}(\hat{Y},\overline{\alpha}) = \frac{\sum_{i=1}^{M}\sum_{j=1}^{K_i}\left|\hat{Y}_{ij}-\overline{\alpha}_i\right|}{\sum_{i=1}^{M} K_i}, \tag{2}$$

where M and K_i represent the number of voxels in overlapping regions and beam voters through the ith voxel, respectively. $\hat{Y}_{ij}$ denotes the predicted dose value in the ith voxel of overlapping regions and the jth beam voter passes that voxel, and $\overline{\alpha}_i$ is a mean dose value in the ith voxel of overlapping regions.

3.2.3. Multi-beam dose aggregation

In order to reassemble the refined field doses into one global dose distribution map, we propose a novel multi-beam voting strategy. In this voting strategy, the dose value of one voxel in the final dose distribution map is voted by the beam voters that contain this voxel. Specifically, if one voxel is passed by multiple beam voters, all these voters will vote for this voxel and contribute to the final dose value by averaging operation. On the other hand, if the voxel is only passed by one beam voter, the value of this voxel will be directly assigned as the voxel value on that beam voter. We apply a MAE loss L_r to make the supervision on the global-image space, which is defined as:

$$L_r(Y,\hat{Y}) = \frac{1}{N}\sum_{i=1}^{N}\left|Y_i-\hat{Y}_i\right|, \tag{3}$$

where N is the number of voxels in the dose mask, and $\hat{Y}_i$ and Y_i represent dose values in the ith voxel of the prediction and the ground truth, respectively.

3.3. Global-wise dose refinement

To further improve the predicted quality, we conduct the following edge enhancement procedure and DVH calibration process during the training phase of the BDN.

3.3.1. Edge enhancement

In the real dose distribution map, sharp edges exist in the dose falloff regions. The main reasons are as follows: 1) the radiation delivered into the patient body generates a clear beam path that contains high-dose values on it while the background regions received by little radiation only carry low-dose values; 2) the dose values decreased rapidly from the PTV to its surroundings from the nature of radiotherapy requirements, in which the PTV reaches prescription dose while minimizing the dose for OAR. However, due to the large size of the receptive field, the CNNs tend to produce over-smooth output, which is undesired in the predicted dose distributions with sharp edges.

Inspired by Tan et al. (2021), we propose a 3D edge enhancement procedure, aiming to preserve the dose falloff regions in the prediction. By applying the commonly used Sobel operator on the 3D final dose distribution map, it extracts the gradient feature map that presents the

value difference between neighboring voxels, as shown in Fig. 3. Then, a 3D gradient loss is defined as follows:

$$L_e(Y,\hat{Y}) = \sum_{S\in\{S_x,S_y,S_z\}} \left| S(Y) - S(\hat{Y}) \right|, \tag{4}$$

where Y and $\hat{Y}$ denote the dose distribution map of the ground truth and the prediction. S_x, S_y, S_z are Sobel operators in x, y, z dimensions, respectively. This loss tries to penalize the over-smooth dose distribution near the dose falloff regions.

3.3.2. DVH calibration

In clinical practice, dosimetrists concern more about the dose distribution inside the ROIs than that of the non-ROIs. The DVH curve indicates the quality of dose distribution inside ROIs, in which the horizontal axis represents the absorbed dose value, and the longitudinal axis represents the relative volume of exposure to the corresponding dose. Inspired by this observation, Nguyen et al. (2020) proposed a DVH loss function to calculate the volume difference (y coordinate of the DVH curve) between the prediction and the ground truth. It requires repeated image-based computations since the processing of the whole 3D image is needed for each dose value threshold (x coordinate of DVH curve). If we reduce the computational consumption by increasing the threshold interval, it would lose accuracy to fit the ground-truth DVH curve. Hence, we proposed a value-based DVH loss to balance efficiency and accuracy. Besides, we also design criteria-based DVH loss functions to further emphasize the prediction accuracy inside the ROIs.

Value-based DVH loss: In order to utilize the guidance of DVH criteria for DL model training, we propose a value-based DVH loss defined as follows:

$$\mathcal{L}_{vDVH} = \frac{\sum_{s=1}^{H}\sum_{n=1}^{N_s} \left| R\left(\hat{Y}\cdot W_s\right)_n - R\left(Y\cdot W_s\right)_n \right|}{\sum_{s=1}^{H} N_s}, \tag{5}$$

where H is the maximum number of ROIs, N_s is the foreground voxel number in the sth ROI mask (W_s) and $R(\cdot)$ denotes the sorting operation. Specifically, we first utilize ROIs masks to extract the dose values in ROIs and apply a sorting operation to rank the voxel with similar dose values and simulate the DVH curve. Then, we calculate the difference between the prediction and the ground-truth ranked dose value to help conduct the supervision. In this way, the volume information of the y coordinate of the DVH curve can be hidden in the sorted dose value as the form of rank. Note that only one round of computation is needed on the whole image, and the processing is directly employed on the dose values, and the processing is directly employed on the dose values, which is more efficient and accurate.

Criteria-based DVH loss: In clinical RT plan, dosimetrists generally determine the feasibility of the plan by observing a set of critical points, such as $C_1^{PTV}, C_{95}^{PTV}, C_{99}^{PTV}, C_{0.1cc}^{OAR}$, and C_{mean}^{OAR}, denoting the dose value received by top-ranked 1%, 95%, 99% volume of PTV regions, top-ranked 0.1cc volume and the mean value of dose value in OAR regions. To specifically take advantage of these important points, we propose a criteria-based DVH loss to further improve the accuracy in ROIs, which is defined as follows:

$$\mathcal{L}_{cDVH} = D_1^{PTV} + D_{95}^{PTV} + D_{99}^{PTV} + D_{0.1cc}^{OAR} + D_{mean}^{OAR}, \tag{6}$$

where we use the dose value of 99th, 95th, and 1st percentile in PTV regions, maximum and mean dose value of OAR regions to represent the criteria $C_1^{PTV}, C_{95}^{PTV}, C_{99}^{PTV}, C_{0.1cc}^{OAR}$, and C_{mean}^{OAR}. In Eq. (6), $D_1^{PTV}, D_{95}^{PTV}, D_{99}^{PTV}, D_{0.1cc}^{OAR}, D_{mean}^{OAR}$ denote the difference of the corresponding criteria between the prediction and the ground truth.

The total loss function will be elaborated as follows:

$$L = L_m + \lambda_1 L_r + \lambda_2 L_{cs} + \lambda_3 L_{cDVH} + \lambda_4 L_{vDVH} + \lambda_5 L_e, \tag{7}$$

where λ_1, λ_2, λ_3, λ_4, λ_5 are hyper-parameters for balancing the contributions of different losses.

3.4. Implementation details

All of our experiments are implemented with Pytorch framework and one NVIDIA TITANV GPU with 12 GB. The models are trained for 1000 epochs with early stopping. The batch size is 2, and the Adam optimizer with an initial learning rate of 10^{-4} is used during the training process. Moreover, we perform online data augmentation to avoid overfitting, including random rotation ranging from $-20°$ to $20°$ and random flip along the Z axis. The three PTV masks and seven OAR masks are concatenated into one mask, respectively, in which different ROI masks are assigned by different labels. The intensity of the CT image is first cropped to the range of [600,1400] since the intensity of the CT image from the public official dataset is HU+1000 scale, then divided by 1000 HU. The GDN is a 3D U-Net (Ronneberger et al., 2015) with the convolution layer replacing the pooling layer. The BDN is configured as two backbone networks, *i.e.*, 3D U-Net and ResUnet. For the ResUnet configuration, we add residual connections to each convolutional block in the 3D U-Net. For better understanding and reproducibility of our method, we release our source code on GitHub.[2]

4. Experiments

4.1. Dataset

We utilize the public H&N cancer dataset[3] from the AAPM OpenKBP challenge (Babier et al., 2021), which contains 200 training cases, 40 validation cases, and 100 testing cases. A few cases include a CT image, PTV masks with three prescription doses (56 Gy, 63 Gy, and 70 Gy), seven OAR masks (brainstem, spinal cord, right parotid, left parotid, larynx, esophagus, and mandible), a feasible dose mask (a mask of where dose can be non-zero), a 3D dose distribution map. The dose distribution map is delivered from nine equispaced coplanar beams at $0°, 40°, \ldots, 320°$, with 6 MV, step and shoot, IMRT. There is a varying voxel spacing between different cases, but the approximate voxel spacing is 3.906 mm × 3.906 mm × 2.5 mm. All the images and masks are resampled to the same size of 128 × 128 × 128.

4.2. Evaluation metrics

We employ two official metrics from the OpenKBP challenge (Babier et al., 2021), *i.e.*, Dose score and DVH score, to quantitatively evaluate the prediction performance of models. Dose score calculates the voxel-wise MAE between the predicted and the reference dose distributions. DVH score measures the absolute difference in DVH criteria between the predicted and its corresponding ground-truth values. The lower the two metrics, the better the prediction results.

To validate the quality of the dose distribution from the perspective of dosimetrists, we introduce a set of clinical criteria, such as D_{99}, D_{95}, D_1, conformity index (CI) (Paddick, 2000), homogeneity index (HI) (Helal and Omar, 2015) for PTV dose coverage, and D_{mean} and $D_{0.1cc}$ for OAR dose coverage. Here, D_v is the minimum dose received by v% of the PTV volume. V_D represents the percentage volume that receives a dose level of at least D. D_{mean} denotes the mean dose absorbed by the OAR, respectively. CI is defined as:

$$\mathrm{CI} = \frac{\left(V_T \cap V_{PI}\right)^2}{V_T V_{PI}}, \tag{8}$$

where V_T is the volume of the target (*i.e.*, PTV). V_{PI} is the volume receiving dose greater than or equal to the prescription isodose. And $V_T \cap V_{PI}$ denotes the intersection of the target volume and the prescription isodose volume. CI describes the fitness of the target volume

[2] https://github.com/TL9792/BDCLDosePrediction
[3] https://github.com/ababier/open-kbp/tree/master/provided-data

Table 1
Hyper-parameter selection of λ_1, λ_2, λ_3, λ_4, and λ_5 (representing the weights for L_r, L_{cs}, L_{cDVH}, L_{vDVH}, and L_e, respectively) using NPE of Dose score and DVH score.

λ_1	NPE	λ_2	NPE	λ_3	NPE	λ_4	NPE	λ_5	NPE
0	0.0222	0	0.0220	0	0.0220	0	0.0233	0	0.0221
0.05	0.0221	0.1	**0.0217**	0.1	**0.0217**	0.5	0.0222	0.5	**0.0217**
0.5	**0.0217**	1	0.0233	1	0.0219	1	**0.0217**	1	0.0229
5	0.0225	5	0.0452	5	0.0221	5	0.0221	5	0.0235
10	0.0229	10	0.0537	10	0.0283	10	0.0242	10	0.0240

and the prescription isodose volume. The higher the CI, the better the prediction results. Meanwhile, the HI formulation is defined as:

$$\mathrm{HI} = \frac{D_2 - D_{98}}{D_{50}}, \tag{9}$$

which denotes dose uniformity in the target. In contrast to CI, the lower the HI, the better the prediction results.

In addition, to quantify the disparity between the prediction and the ground truth, we calculate the normalized prediction error (NPE). For CI and HI, NPE takes Eq. (10). For the rest metrics, NPE is calculated in Eq. (11). We adopt $NPE \pm SD$ to represent the prediction error, where SD is the standard deviation and the NPE is defined as:

$$\mathrm{NPE} = \frac{1}{n}\sum_{i=1}^{n} |X_i^{GT} - X_i^{pred}|, \tag{10}$$

$$\mathrm{NPE} = \frac{\frac{1}{n}\sum_{i=1}^{n} |X_i^{GT} - X_i^{pred}|}{D_{prescription}}, \tag{11}$$

where n represents the number of testing cases, X^{GT} denotes the ground-truth value, and X^{pred} is the corresponding predicted value. $D_{prescription}$ represents a prescription dose value delivered in PTV. Furthermore, to prove the significant improvement of our method compared with the state-of-the-art methods, we conduct paired t-test to confirm that the improvement margin is considered statistically significant when the p-value is lower than 0.05.

4.3. Selection of hyper-parameters

To validate the effect of the hyper-parameters in Eq. (7), i.e., the weights λ_1, λ_2, λ_3, λ_4, λ_5 of different loss terms, to the model performance, we conduct a series of experiments using the ResUnet backbone to analyze their optimal values. Specifically, the process of tuning hyper-parameters is divided into two steps. First, we roughly set a reasonable range between [0, 10] with different intervals for five hyper-parameters according to the magnitude and relative importance of the corresponding loss function. For example, L_r plays an important role in the final prediction result since it supervises the network learning on the global image space. Thus, we set a large value for λ_1. As L_{DVH} can constrain the network to focus on ROIs, we also set a relatively large value for λ_2. Besides, since high-frequency information (*e.g.*, gradient information) is more difficult to learn compared with low-frequency information, we define a relatively small value for λ_3. Second, we further explore an optimal value by fixing other hyper-parameters to tune a certain one. The reference metric is the NPE of Dose score and DVH score. For instance, we first fix $\{\lambda_2, \lambda_3, \lambda_4, \lambda_5\}$ to $\{0.1, 0.1, 0.5, 0.5\}$, respectively, and vary λ_1 from 0 to 10. Experiment results are shown in Table 1. It is clear that better performance is achieved when λ_1 is set to 0.5. Then we explore the optimal values of the remaining four hyper-parameters in the same way as above. In the end, the performance of our method is optimal when setting λ_2, λ_3, λ_4, and λ_5 to be 0.1, 0.1, 1, and 0.5, respectively.

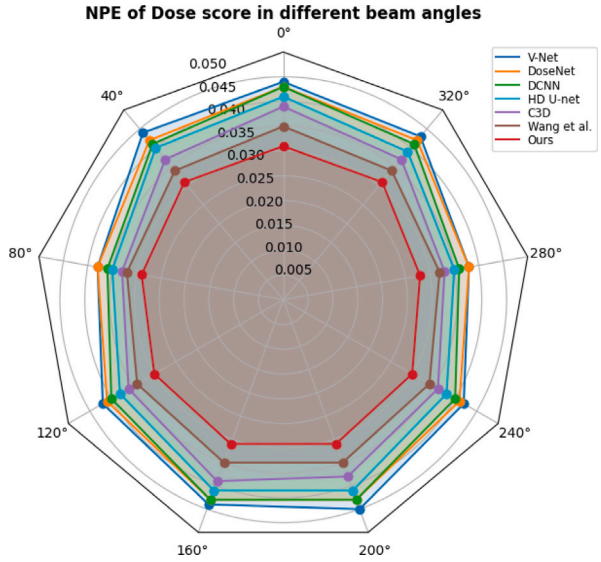

Fig. 4. Radar charts of NPE of Dose score for field doses in all individual beam angles yielded by different SOTA methods. Each line with a different color represents prediction result from different method.

4.4. Comparison with state-of-the-art methods

In this section, we conduct a comparison with state-of-the-art methods for dose prediction to demonstrate the superiority of our method. The competing methods include V-Net (Milletari et al., 2016), DoseNet (Kearney et al., 2018a), HD U-net (Nguyen et al., 2019), DCNN (Zhang et al., 2020), C3D (Liu et al., 2021) and the method proposed of our conference paper (Wang et al., 2022). The first four methods are well-established methods for dose prediction task, and the C3D ranks first in the OpenKBP challenge (Babier et al., 2021, 2022). For a fair comparison, the first three approaches are reproduced by the same implementation strategy as our method. For the last three methods, we directly refer to the results from their papers or released codes. All methods use the same dataset and the same data split from the Challenge.

Table 2 shows the quantitative results with the top approaches from the Challenge. We can see that our method achieves the best performance in terms of Dose score and DVH score. Importantly, Table 3 gives a comprehensive assessment of the dose quality from various clinical criteria. Among the competing methods, we can see that our method obtains superior performance in all NPE of criteria. In addition, we also quantify NPE of Dose score for the field dose distributions in terms of radar chart. Fig. 4 shows that our method exhibits lower error than other methods in all beam angles.

Fig. 5 is the visualization of the predicted dose distribution maps. It shows that our prediction is more consistent with the ground truth, regarding not only the global space but also the local regions, especially for the beam paths, ROIs, and gradient boundaries. This result well-demonstrates the effectiveness of our strategy toward the beam-wise

Table 2
Performance comparison between SOTA methods and our method in terms of Dose score and DVH score.

Method	Dose score [Gy]	DVH score [Gy]
Xu et al. (Xu et al., 2021)	2.753[a]	1.556[a]
Lin et al. (Lin et al., 2021)	2.357[a]	1.465[a]
LukasFetty (Aaron Babier, 2020)	2.650[a]	1.539[a]
Zimmermann et al. (Zimmermann et al., 2021)	2.620 ± 1.100[a]	1.520 ± 1.060[a]
Gronberg et al. (Gronberg et al., 2021)	2.563 ± 1.143*	1.704 ± 1.096*
C3D (Liu et al., 2021)	2.429 ± 1.031*	1.478 ± 1.552*
Wang et al. (Wang et al., 2022)	2.276 ± 1.013*	1.257 ± 1.163*
Ours	**2.066 ± 0.900**	**0.977 ± 1.091**

* p-value < 0.05.
[a] Results reported in original literature.

Table 3
Performance comparison between SOTA methods and our method in terms of PTVs' dose coverage (indicated by NPE ± SD of CI, HI, D_{99}, D_{95}, and D_1) and OARs' dose coverage (indicated by NPE ± SD of D_{mean} and $D_{0.1cc}$).

Method	CI	HI	D_{99}	D_{95}	D_1
V-Net (Milletari et al., 2016)	0.086 ± 0.143*	0.036 ± 0.048*	0.034 ± 0.038*	0.026 ± 0.031*	0.025 ± 0.021*
DoseNet (Kearney et al., 2018a)	0.057 ± 0.115*	0.026 ± 0.041*	0.021 ± 0.034	0.016 ± 0.028*	0.020 ± 0.016*
DCNN (Zhang et al., 2020)	0.057 ± 0.105*	0.030 ± 0.042*	0.025 ± 0.034*	0.018 ± 0.027*	0.021 ± 0.018*
HD U-net (Nguyen et al., 2019)	0.072 ± 0.136*	0.031 ± 0.043*	0.025 ± 0.033*	0.019 ± 0.027*	0.025 ± 0.022*
C3D (Liu et al., 2021)	0.055 ± 0.120*	0.026 ± 0.042*	0.021 ± 0.033	0.016 ± 0.028*	0.019 ± 0.017*
Wang et al. (Wang et al., 2022)	0.063 ± 0.128*	0.022 ± 0.037	0.022 ± 0.032*	0.016 ± 0.027*	0.019 ± 0.015*
Ours	**0.041 ± 0.085**	**0.021 ± 0.041**	**0.020 ± 0.032**	**0.013 ± 0.025**	**0.011 ± 0.012**

Method	D_{mean}				$D_{0.1cc}$		
	Right parotid	Left parotid	Esophagus	Larynx	Brainstem	Spinal cord	Mandible
V-Net (Milletari et al., 2016)	0.048 ± 0.044*	0.053 ± 0.049*	0.035 ± 0.033*	0.049 ± 0.040*	0.053 ± 0.064*	0.049 ± 0.046*	0.022 ± 0.020*
DoseNet (Kearney et al., 2018a)	0.058 ± 0.048*	0.048 ± 0.049*	0.035 ± 0.038*	0.042 ± 0.038*	0.056 ± 0.064*	0.042 ± 0.042*	0.019 ± 0.026*
DCNN (Zhang et al., 2020)	0.068 ± 0.055*	0.065 ± 0.059*	0.053 ± 0.046*	0.041 ± 0.036*	0.052 ± 0.061*	0.046 ± 0.042*	0.020 ± 0.023*
HD U-net (Nguyen et al., 2019)	0.052 ± 0.045*	0.054 ± 0.047*	0.038 ± 0.036*	0.042 ± 0.049*	0.052 ± 0.061*	0.041 ± 0.036*	0.021 ± 0.022*
C3D (Liu et al., 2021)	0.046 ± 0.040*	0.052 ± 0.048*	0.039 ± 0.041*	0.040 ± 0.051*	0.049 ± 0.066*	0.044 ± 0.035*	0.018 ± 0.025*
Wang et al. (Wang et al., 2022)	0.044 ± 0.036*	0.042 ± 0.043*	0.018 ± 0.025	0.029 ± 0.026*	**0.030 ± 0.056**	**0.020 ± 0.019**	0.020 ± 0.028*
Ours	**0.033 ± 0.034**	**0.041 ± 0.040**	**0.019 ± 0.023**	**0.023 ± 0.040**	0.039 ± 0.059	0.029 ± 0.028	**0.012 ± 0.012**

* p-value < 0.05.

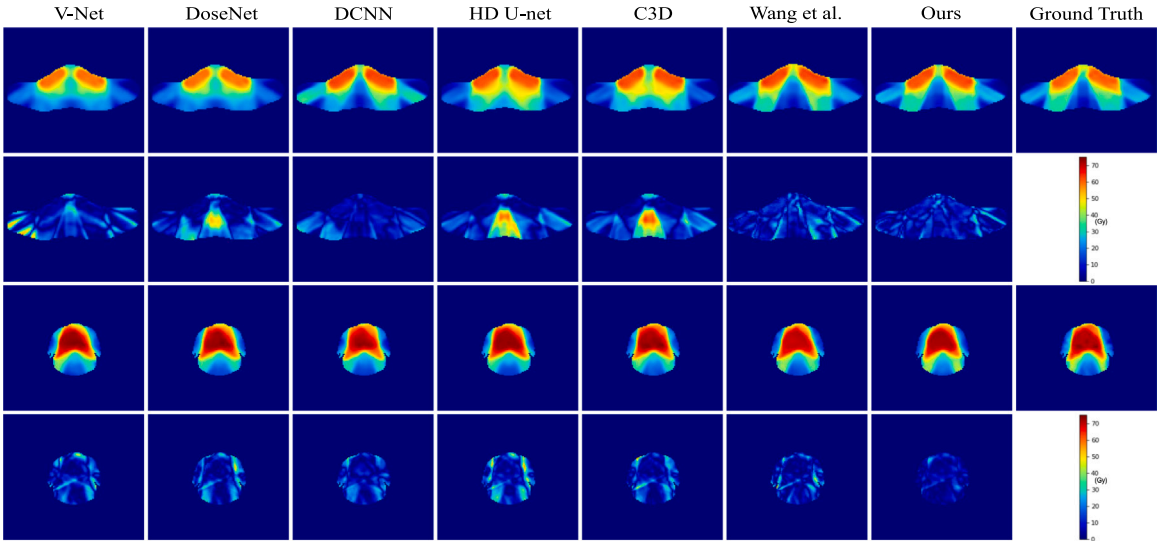

Fig. 5. Visualization of results yielded by different methods from two cases. The first and third rows are the dose distribution maps. The second and fourth rows are the error maps.

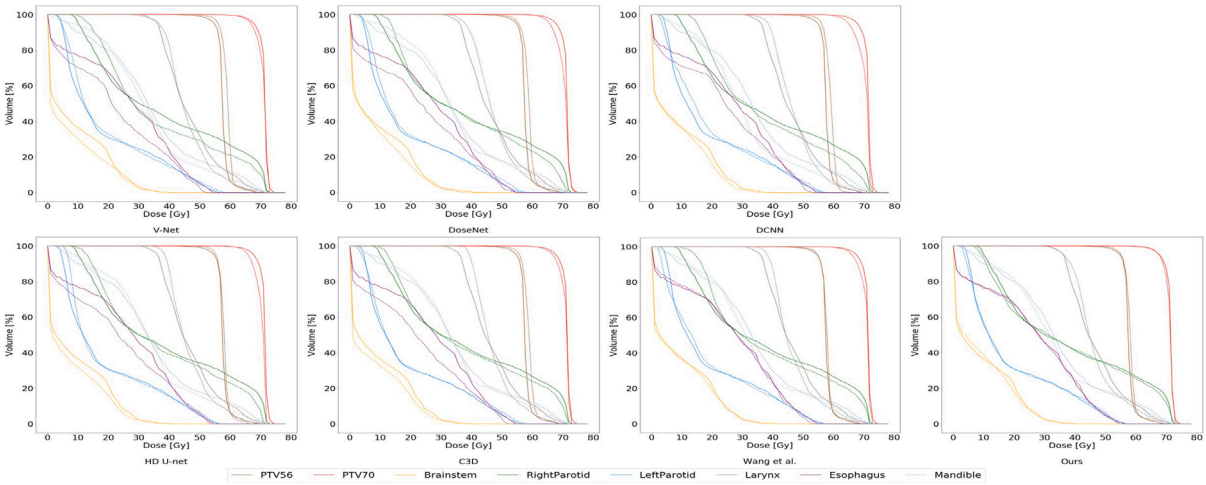

Fig. 6. DVH curve of the ground-truth (solid line) and predicted (dashed line) dose distribution maps yielded by different SOTA methods. Different color curves indicate DVH curves of different ROIs.

dose distributions and dose gradients. In addition, Fig. 6 provides an analysis of the DVH curve. It is apparent that the DVH curves predicted by our method match best with the ground truth, proving that our predictions in ROIs are more accurate and better satisfy the design requirement of treatment plan in clinical workflow.

4.5. Ablation analysis

To verify the effectiveness of key components in our method, we conduct a series of ablation studies under two different backbone configurations, *i.e.*, 3D U-Net and ResUnet. We use a cascaded U-Net as the baseline framework, and then sequentially add key components in an order of (1) beam-wise dose prediction (BDP) including beam-wise decomposition module and multi-beam voting mechanism, (2) DVH calibration (L_{DVH}) including value- and criteria-based DVH loss, (3) overlap consistency module (L_{cs}), and (4) edge enhancement (L_e). The quantitative results are shown in Table 4, and the qualitative results are given in Fig. 7.

From Table 4, we can clearly see that the Dose score and DVH score are decreasing with the addition of key components. Especially for the components of BDP and L_{DVH}, both of the metrics decrease by a large margin under different backbones. For example, under 3D U-Net backbone, BDP can make Dose score reduce from 2.862 to 2.401, demonstrating that it can improve the accuracy of dose prediction. And then adding L_{DVH} brings 31% of improvement (from 1.567 to 1.257) in terms of DVH score, indicating that L_{DVH} can make prediction more precise on ROIs. For the ResUnet backbone, it shows a consistency tendency and has a more excellent performance when the corresponding components are added in sequence.

We further evaluate the L_{cs} and L_e introduced in this work by comparing with our prior conference paper (Wang et al., 2022) (Baseline+BDP $+L_{DVH}$). Fig. 7 shows visual comparison results, indicating that adding L_{cs} can make local details of the dose distribution have significant improvement, and also shows its contribution to learning high-level features so as to achieve better prediction performance. Last, adding L_e makes our prediction best match the ground truth, especially for shapes and boundaries of the dose falloff regions. The error map in the last row confirms our findings.

5. Discussion

Automatic dose prediction is of great clinical significance in RT plan. An accurate dose distribution map provides crucial guidance for dosimetrists and thus can serve as a good initialization for subsequent dose optimization, which solves the problem of the NP-hard to optimize treatment plan in clinical workflows. However, the predicted dose distribution map is not physically deliverable because the prediction model does not consider any machine parameters or beam fluence that deliver the dose. In order to utilize the prediction to generate actual RT plan, many works are emerging (Fan et al., 2019; McIntosh et al., 2017; Sun et al., 2022). For example, Fan et al. (2019) fed the predicted dose distribution map into a treatment planning system (*e.g.*, matRad) to produce clinically acceptable plan. McIntosh et al. (2017) converted the predicted voxel-wise dose distribution into a deliverable RT plan by a voxel-based dose mimicking method.

In order to make the prediction close to the clinical RT plan, we proposed a novel DVH calibration process. Compared with the DVH loss proposed in Nguyen et al. (2020) (detailed in Section 3.3), our proposed method can accelerate the training speed and improve the calculation accuracy. To demonstrate our advantage, we conduct a comparison experiment in terms of convergence speed, GPU memory consumption, and computational time. The experimental results demonstrate that applying our DVH calibration leads to a fast convergence speed compared with the model trained by the DVH loss (Nguyen et al., 2020). Moreover, the total training process with our DVH calibration takes about 12 GB GPU memory and the computational time for an epoch is 2 min, which are significantly lower than those by using the DVH loss (23 GB and 6 min).

Despite the superior performance achieved by our method, there are still some limitations in our work. One limitation is that the beam masks are generated by a parallel-shaped radiation delivery from a line source. However, in clinical practice, beams are generated in the form of a cone-shaped radiation delivery from a point source. How to fit actual beams for delivering radiation is not explored yet, which will be our future work. Another limitation of this work is that we did not utilize the actual individual IMRT field dose (Ehrgott et al., 2010; Losasso et al., 2001) to train the BDN for beam-wise dose refinement. Instead, we used a masked version of the total dose to approximate the individual field dose for each beam. This is because the used public dataset did not contain the individual field dose files. We believe that, by introducing the actual individual IMRT field doses for BDN training,

Table 4
Experimental results of ablation studies on our proposed method using two different backbone networks.

Method	3D U-Net		ResUnet	
	Dose score [Gy]	DVH score [Gy]	Dose score [Gy]	DVH score [Gy]
Baseline	2.862 ± 1.049*	1.586 ± 1.146*	2.803 ± 1.125*	1.533 ± 1.111*
Baseline+BDP	2.401 ± 1.033*	1.567 ± 1.179*	2.306 ± 1.017*	1.395 ± 1.234*
Baseline+BDP+L_{DVH}	2.276 ± 1.014*	1.257 ± 1.163*	2.281 ± 0.962*	1.185 ± 1.161*
Baseline+BDP+L_{DVH} + L_{cs}	2.265 ± 0.976*	1.172 ± 1.182	2.152 ± 0.917*	1.068 ± 1.112*
Baseline+BDP+L_{DVH} + L_{cs} + L_e	**2.113 ± 0.903**	**1.124 ± 1.132**	**2.066 ± 0.900**	**0.977 ± 1.091**

* p-value < 0.05.

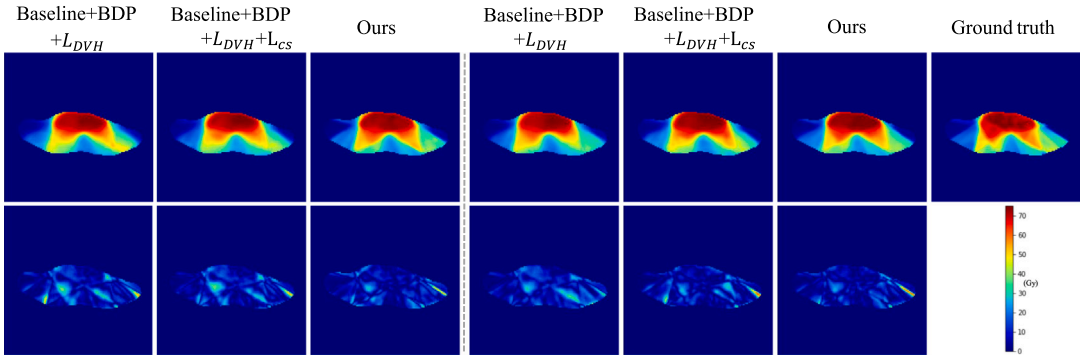

Fig. 7. Visual comparison results of ablation studies using 3D U-Net (left) and ResUnet (right) backbones. The first row shows the dose distribution map. And the second row shows the error map.

the performance of the proposed method can be further improved. Since in this work we mainly focus on the prototype study improving the predicted dose distribution along the decomposed beam paths, we reserve this as one of our future works. The third limitation is that our proposed overlap consistency constraint is only applicable for the traditional IMRT treatment with fixed beams but not for some more advanced treatments such as the IMRT using rotational arc technique or VMAT (Volumetric Modulated Arc Therapy) (Ma et al., 2019), in which the dose ratio of beams with different directions is dynamically changing. Developing a more robust algorithm to predict dose distribution for more types of RT will be one of our future work. The last limitation comes from the use of a single type of dataset, H&N cancer. Thus, in our future work, we will use other datasets with different types of cancer to evaluate the generalization of our proposed method.

6. Conclusion

We have presented a novel beam-wise dose composition learning (BDCL) method for dose prediction in H&N RT plan. In this method, we first utilize the GDN to estimate a coarse dose distribution map over the whole-image space. Then, the coarse dose distribution map is decomposed into multiple field doses, with each individually refined by the BDN. Finally, the refined beam doses are reassembled into a new global dose distribution map by the proposed multi-beam voting mechanism. Moreover, the dose distribution of ROIs and the dose falloff regions are refined by our proposed DVH calibration process (using value- and criteria-based DVH regularization) and edge enhancement procedure. Experimental results show that our method achieves superior performance over other state-of-the-art methods in terms of all metrics. It is worth noting that our prediction is close to the physically deliverable one from the perspective of clinical criteria, and thus can be used as a good starting point for subsequent dose optimization. This will substantially reduce time consumption and plan variation in the clinical workflow.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We have shared the link to our data/code as a footnote in the submitted manuscript.

Acknowledgments

This work was supported in part by National Natural Science Foundation of China [grant number 62131015, 62250710165], Science and Technology Commission of Shanghai Municipality (STCSM) [grant number 21010502600], and The Key R&D Program of Guangdong Province, China [grant number 2021B0101420006].

References

Aaron Babier, B.Z., 2020. Openkbp CodaLab. https://competitions.codalab.org/competitions/23428, (Last Accessed on 30 September 2022).
Amit, G., Purdie, T.G., Levinshtein, A., Hope, A.J., Lindsay, P., Marshall, A., Jaffray, D.A., Pekar, V., 2015. Automatic learning-based beam angle selection for thoracic IMRT. Med. Phys. 42 (4), 1992–2005.
Babier, A., Mahmood, R., McNiven, A.L., Diamant, A., Chan, T.C., 2020. Knowledge-based automated planning with three-dimensional generative adversarial networks. Med. Phys. 47 (2), 297–306.
Babier, A., Mahmood, R., Zhang, B., Alves, V.G., Barragán-Montero, A.M., Beaudry, J., Cardenas, C.E., Chang, Y., Chen, Z., Chun, J., et al., 2022. OpenKBP-Opt: An international and reproducible evaluation of 76 knowledge-based planning pipelines. Phys. Med. Biol. 67 (18), 185012.
Babier, A., Zhang, B., Mahmood, R., Moore, K.L., Purdie, T.G., McNiven, A.L., Chan, T.C., 2021. OpenKBP: The open-access knowledge-based planning grand challenge and dataset. Med. Phys. 48 (9), 5549–5561.
Boutilier, J.J., Lee, T., Craig, T., Sharpe, M.B., Chan, T.C., 2015. Models for predicting objective function weights in prostate cancer IMRT. Med. Phys. 42 (4), 1586–1595.
Campbell, W.G., Miften, M., Olsen, L., Stumpf, P., Schefter, T., Goodman, K.A., Jones, B.L., 2017. Neural network dose models for knowledge-based planning in pancreatic SBRT. Med. Phys. 44 (12), 6148–6158.

Chanyavanich, V., Das, S.K., Lee, W.R., Lo, J.Y., 2011. Knowledge-based IMRT treatment planning for prostate cancer. Med. Phys. 38 (5), 2515–2522.
Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O., 2016. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 424–432.
Ehrgott, M., Güler, Ç., Hamacher, H.W., Shao, L., 2010. Mathematical optimization in intensity modulated radiation therapy. Ann. Oper. Res. 175 (1), 309–365.
El Naqa, I., 2021. Prospective clinical deployment of machine learning in radiation oncology. Nat. Rev. Clin. Oncol. 18 (10), 605–606.
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. Med. Phys. 46 (1), 370–381.
Fogliata, A., Cozzi, L., Reggiori, G., Stravato, A., Lobefalo, F., Franzese, C., Franceschini, D., Tomatis, S., Scorsetti, M., 2019. RapidPlan knowledge based planning: Iterative learning process and model ability to steer planning strategies. Radiat. Oncol. 14, 1–12.
Ge, Y., Wu, Q.J., 2019. Knowledge-based planning for intensity-modulated radiation therapy: A review of data-driven approaches. Med. Phys. 46 (6), 2760–2775.
Gronberg, M.P., Gay, S.S., Netherton, T.J., Rhee, D.J., Court, L.E., Cardenas, C.E., 2021. Dose prediction for head and neck radiotherapy using a three-dimensional dense dilated U-net architecture. Med. Phys. 48 (9), 5567–5573.
Helal, A., Omar, A., 2015. Homogeneity index: Effective tool for evaluation of 3DCRT. Pan. Arab. J. Oncol. 8 (2), 20–23.
Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q., 2017. Densely connected convolutional networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 4700–4708.
Jia, H., Yap, P.-T., Shen, D., 2012. Iterative multi-atlas-based multi-image segmentation with tree-based registration. NeuroImage 59 (1), 422–430.
Kearney, V., Chan, J.W., Haaf, S., Descovich, M., Solberg, T.D., 2018a. DoseNet: A volumetric dose prediction algorithm using 3D fully-convolutional neural networks. Phys. Med. Biol. 63 (23), 235022.
Kearney, V., Chan, J.W., Valdes, G., Solberg, T.D., Yom, S.S., 2018b. The application of artificial intelligence in the IMRT planning process for head and neck cancer. Oral Oncol. 87, 111–116.
Khan, F.M., 2010. The Physics of Radiation Therapy. Lippincott Williams & Wilkins.
Lin, Y., Liu, Y., Liu, J., Liu, G., Ma, K., Zheng, Y., 2021. LE-NAS: Learning-based ensemble with NAS for dose prediction. arXiv preprint arXiv:2106.06733.
Liu, T., Guo, Q., Lian, C., Ren, X., Liang, S., Yu, J., Niu, L., Sun, W., Shen, D., 2019. Automated detection and classification of thyroid nodules in ultrasound images using clinical-knowledge-guided convolutional neural networks. Med. Image Anal. 58, 101555.
Liu, S., Zhang, J., Li, T., Yan, H., Liu, J., 2021. A cascade 3D U-Net for dose prediction in radiotherapy. Med. Phys. 48 (9), 5574–5582.
Losasso, T., Chui, C.-S., Ling, C.C., 2001. Comprehensive quality assurance for the delivery of intensity modulated radiotherapy with a multileaf collimator used in the dynamic mode. Med. Phys. 28 (11).
Luan, H., Qi, F., Xue, Z., Chen, L., Shen, D., 2008. Multimodality image registration by maximization of quantitative–qualitative measure of mutual information. Pattern Recognit. 41 (1), 285–298.
Ma, M., Buyyounouski, M.K., Vasudevan, V., Xing, L., Yang, Y., 2019. Dose distribution prediction in isodose feature-preserving voxelization domain using deep convolutional neural network. Med. Phys. 46 (7), 2978–2987.
McIntosh, C., Conroy, L., Tjong, M.C., Craig, T., Bayley, A., Catton, C., Gospodarowicz, M., Helou, J., Isfahanian, N., Kong, V., et al., 2021. Clinical integration of machine learning for curative-intent radiation treatment of patients with prostate cancer. Nat. Med. 27 (6), 999–1005.
McIntosh, C., Welch, M., McNiven, A., Jaffray, D.A., Purdie, T.G., 2017. Fully automated treatment planning for head and neck radiotherapy using a voxel-based dose prediction and dose mimicking method. Phys. Med. Biol. 62 (15), 5926.
Milletari, F., Navab, N., Ahmadi, S.-A., 2016. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: 2016 Fourth International Conference on 3D Vision. 3DV, IEEE, pp. 565–571.
Momin, S., Fu, Y., Lei, Y., Roper, J., Bradley, J.D., Curran, W.J., Liu, T., Yang, X., 2021. Knowledge-based radiation treatment planning: A data-driven method survey. J. Appl. Clin. Med. Phys. 22 (8), 16–44.
Nelms, B.E., Robinson, G., Markham, J., Velasco, K., Boyd, S., Narayan, S., Wheeler, J., Sobczak, M.L., 2012. Variation in external beam treatment plan quality: An inter-institutional study of planners and planning systems. Pract. Radiat. Oncol. 2 (4), 296–305.
Nguyen, D., Jia, X., Sher, D., Lin, M.-H., Iqbal, Z., Liu, H., Jiang, S., 2019. 3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture. Phys. Med. Biol. 64 (6), 065020.
Nguyen, D., McBeth, R., Sadeghnejad Barkousaraie, A., Bohara, G., Shen, C., Jia, X., Jiang, S., 2020. 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. Med. Phys. 47 (3), 837–849.
Paddick, I., 2000. A simple scoring ratio to index the conformity of radiosurgical treatment plans. J. Neurosurg. 93 (supplement_3), 219–222.
Pai, S.I., Westra, W.H., 2009. Molecular pathology of head and neck cancer: Implications for diagnosis, prognosis, and treatment. Annu. Rev. Pathol. 4, 49.
Peters, L.J., O'Sullivan, B., Giralt, J., Fitzgerald, T.J., Trotti, A., Bernier, J., Bourhis, J., Yuen, K., Fisher, R., Rischin, D., 2010. Critical impact of radiotherapy protocol compliance and quality in the treatment of advanced head and neck cancer: Results from TROG 02.02. J. Clin. Oncol. 28 (18), 2996–3001.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 234–241.
Scaggion, A., Fusella, M., Cavinato, S., Dusi, F., El Khouzai, B., Germani, A., Pivato, N., Rossato, M.A., Roggio, A., Scott, A., et al., 2023. Updating a clinical knowledge-based planning prediction model for prostate radiotherapy. Physica Medica 107, 102542.
Shan, F., Gao, Y., Wang, J., Shi, W., Shi, N., Han, M., Xue, Z., Shen, D., Shi, Y., 2021. Abnormal lung quantification in chest CT images of COVID-19 patients with deep learning and its application to severity prediction. Med. Phys. 48 (4), 1633–1645.
Shiraishi, S., Moore, K.L., 2016. Knowledge-based prediction of three-dimensional dose distributions for external beam radiotherapy. Med. Phys. 43 (1), 378–387.
Sun, Z., Xia, X., Fan, J., Zhao, J., Zhang, K., Wang, J., Hu, W., 2022. A hybrid optimization strategy for deliverable intensity-modulated radiotherapy plan generation using deep learning-based dose prediction. Med. Phys. 49 (3), 1344–1356.
Tan, S., Tang, P., Peng, X., Xiao, J., Zu, C., Wu, X., Zhou, J., Wang, Y., 2021. Incorporating isodose lines and gradient information via multi-task learning for dose prediction in radiotherapy. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 753–763.
Wang, B., Teng, L., Mei, L., Cui, Z., Xu, X., Feng, Q., Shen, D., 2022. Deep learning-based head and neck radiotherapy planning dose prediction via beam-wise dose decomposition. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 575–584.
Wang, M., Zhang, Q., Lam, S., Cai, J., Yang, R., 2020. A review on application of deep learning algorithms in external beam radiotherapy automated treatment planning. Front. Oncol. 10, 2177.
Webb, S., 2003. The physical basis of IMRT and inverse planning. Brit. J. Radiol. 76 (910), 678–689.
Wu, B., Ricchetti, F., Sanguineti, G., Kazhdan, M., Simari, P., Chuang, M., Taylor, R., Jacques, R., McNutt, T., 2009. Patient geometry-driven information retrieval for IMRT treatment plan quality control. Med. Phys. 36 (12), 5497–5505.
Xu, X., Lian, C., Yap, P., Wang, A., Chera, B., Shen, C., Lian, J., 2021. Prediction of optimal dosimetry for intensity-modulated radiotherapy with a cascaded auto-content deep learning model. Int. J. Radiat. Oncol. Biol. Phys. 111 (3), e113.
Zacharaki, E.I., Shen, D., Lee, S.-K., Davatzikos, C., 2008. ORBIT: A multiresolution framework for deformable registration of brain tumor images. IEEE Trans. Med. Imaging 27 (8), 1003–1017.
Zhang, J., Liu, S., Li, T., Mao, R., Du, C., Liu, J., 2019. Voxel-level radiotherapy dose prediction using densely connected network with dilated convolutions. In: Workshop on Artificial Intelligence in Radiation Therapy. Springer, pp. 70–77.
Zhang, J., Liu, S., Yan, H., Li, T., Mao, R., Liu, J., 2020. Predicting voxel-level dose distributions for esophageal radiotherapy using densely connected network with dilated convolutions. Phys. Med. Biol. 65 (20), 205013.
Zheng, D., Hong, J.C., Wang, C., Zhu, X., 2019. Radiotherapy treatment planning in the age of AI: Are we ready yet? Technol. Cancer Res. Treat. 18, 1533033819894577.
Zimmermann, L., Faustmann, E., Ramsl, C., Georg, D., Heilemann, G., 2021. Dose prediction for radiation therapy using feature-based losses and one cycle learning. Med. Phys. 48 (9), 5562–5566.