Radiation Oncology

RESEARCH

Open Access

Dose prediction of CyberKnife Monte Carlo plan for lung cancer patients based on deep learning: robust learning of variable beam configurations

Yuchao Miao[1,4,6], Jiwei Li[2], Ruigang Ge[3], Chuanbin Xie[3], Yaoying Liu[4], Gaolong Zhang[4], Mingchang Miao[5] and Shouping Xu[1*]

Abstract

Background Accurate calculation of lung cancer dose using the Monte Carlo (MC) algorithm in CyberKnife (CK) is essential for precise planning. We aim to employ deep learning to directly predict the 3D dose distribution calculated by the MC algorithm, enabling rapid and accurate automatic planning. However, most current methods solely focus on conventional intensity-modulated radiation therapy and assume a consistent beam configuration across all patients. This study seeks to develop a more versatile model incorporating variable beam configurations of CK and considering the patient's anatomy.

Methods This study proposed that the AB (anatomy and beam) model be compared with the control Mask (only anatomy) model. These models are based on a 3D U-Net network to investigate the impact of CK beam encoding information on dose prediction. The study collected 86 lung cancer patients who received CK′s built-in MC algorithm plans using different beam configurations for training/validation (66 cases) and testing (20 cases). We compared the gamma passing rate, dose difference maps, and relevant dose-volume metrics to evaluate the model's performance. In addition, the Dice similarity coefficients (DSCs) were calculated to assess the spatial correspondence of isodose volumes.

Results The AB model demonstrated superior performance compared to the Mask model, particularly in the trajectory dose of the beam. The DSCs of the AB model were 20–40% higher than that of the Mask model in some dose regions. We achieved approximately 99% for the PTV and generally more than 95% for the organs at risk (OARs) referred to the clinical planning dose in the gamma passing rates (3 mm/3%). Relative to the Mask model, the AB model exhibited more than 90% improvement in small voxels ($p < 0.001$). The AB model matched well with the clinical plan's dose-volume histograms, and the average dose error for all organs was $1.65 \pm 0.69\%$.

Conclusions Our proposed new model signifies a crucial advancement in predicting CK 3D dose distributions for clinical applications. It enables planners to rapidly and precisely predict MC doses for lung cancer based on patient-specific beam configurations and optimize the CK treatment process.

Keywords Automatic planning, Deep learning, Dose prediction, Monte Carlo, CyberKnife

*Correspondence:
Shouping Xu
shouping_xu@yahoo.com
Full list of author information is available at the end of the article

Background

Current treatment planning systems (TPS) for radiation therapy (RT) typically employ advanced software to address the inverse optimization problem. The primary objective is determining the optimal treatment plan and RT machine parameters by considering predetermined target prescription dose coverage and organ-at-risk (OAR) dose limits [1]. However, due to variations in clinical plans for different patients and differences in computer hardware used to calculate the administered dose, planners still need to manually adjust dose metrics until the patient's dose distribution agrees with the clinical requirements. Treatment planning design is an empirical and time-consuming process, often taking hours to days [2–5], which runs concurrently with vital activities, such as metabolic processes, occurring within the patient's body. Changes in tumor physiology and pathology may also arise, leading to variations in plan quality. Plan quality heavily relies on factors such as time for plan generation, adherence to institutional guidelines, and the planner's expertise. These factors can result in the generation of suboptimal treatment plans, ultimately impacting patient outcomes. Additionally, prolonged treatment planning significantly impedes the implementation of adaptive strategies and may introduce delays in delivering RT to patients. These challenges are commonly encountered in clinical practice and can negatively impact tumor control and patients' life quality.

Research scholars generally desire to expedite the design of RT plans by implementing an automated treatment planning process. This approach aims to minimize the manual plan setting component to address the abovementioned issues. For instance, integrating automatic segmentation subroutines for OARs into TPS has been widely adopted in clinical settings. Although these subroutines have demonstrated favorable clinical performance, they have yet to gain full acceptance from physicians, particularly regarding target area delineation for brain tumors [6, 7]. Currently, several solutions have been proposed in the research field. One such solution is data-driven knowledge-based planning (KBP), which utilizes existing treatment planning datasets to predict dose-volume objectives for new patients [8–10]. RapidPlan (Varian Medical Systems, USA) is a commercially available TPS that exemplifies this approach [11]. These alternative methods for automated planning have significantly expedited the planning process for traditional intensity-modulated radiation therapy (IMRT), reducing the need for human intervention while ensuring the production of high-quality RT plans [12–14].

The KBP-based approach still has certain limitations. The reason is that the KBP-based approach typically utilizes features such as OARs-tumor overlap and distance-to-the PTV to predict a one-dimensional DVH [15]. As a result, these dose-volume objectives are insensitive to the spatial variation of doses within the delineated structures and are even more limited in their assessment of undelineated structures [16]. Consequently, this may lead to suboptimal spatial distribution of patient doses regarding plan design. In such cases, planners may need to introduce additional auxiliary structures and reoptimize the plan manually. In recent years, there has been a surge of research interest in deep convolutional neural network (CNN)-based prediction of patient plan dose, driven by the widespread popularity of open-source deep learning algorithms. The primary methodology involves utilizing anatomical information, such as medical images (e.g., planning CT), and delineated information of regions of interest (ROIs) as inputs to the network model. The network model is trained using the patient's plan dose as the output. Subsequently, only the anatomical information must be inputted into the trained network model for a new patient to obtain the desired clinically planned dose rapidly [17]. This approach eliminates manual feature extraction and significantly expedites the treatment planning process, a crucial step in automated RT planning.

However, implementing artificial intelligence (AI) in clinical practice has been limited to conventional linear accelerators (IMRT) [12–14]. Further research is required to explore the use of AI in high-precision stereotactic radiotherapy (SRT) devices. One unique SRT device is the CyberKnife (CK) system, which enables the precise delivery of high doses to lesions for effective tumor destruction [18–20]. Moreover, the CK system protects surrounding tissues adequately and reduces treatment pain by administering fewer fractions with higher doses. Regarding dose calculation, the CK system employs the ray-tracing and the Monte Carlo (MC) algorithms. However, due to computational limitations and lengthy dose calculation time, clinical practitioners often compromise by favoring ray-tracing algorithms with relatively lower calculation accuracy. This compromise, unfortunately, leads to inaccuracies in dose calculation, particularly in heterogeneous tissues such as the lungs.

In this study, we aimed to address the abovementioned issues by directly predicting the dose for the CK MC plan using a deep learning method. The method allows for the rapid calculation of doses in lung cases. To achieve this,

we gathered a collection of lung cancer cases encompassing CT images, contours, and clinical plan information. The plan design of these cases strictly followed the implementation standards of the TG101 report. The treatment procedures were rigorously based on the clinician's expertise, and all received SRT with the CK MC plan.

Meanwhile, it is worth noting that the performance of deep learning methods for predicting patient voxel doses relies heavily on the training database. In the case of CK patients, the number and orientation of the delivery beams can vary significantly across different central institutions. Consequently, a model that relies solely on anatomical information (i.e., CT images and delineation information) as the input to a CNN may need help to generalize to the more diverse and heterogeneous beam configurations. It poses a significant obstacle to accurate dose prediction for CK patients. We have explored the potential benefits of incorporating anatomical and beam information into a network to address this issue and enable the automation of clinical planning for CK. This approach aims to create a robust model that accommodates variable beam configurations. By doing so, we have developed a generalized model that allows for rapid MC dose prediction using deep CNNs. This approach expedites the planning design process and ensures the accuracy of dose calculations. As a result, the clinical implementation of automatic plans based on this approach becomes more reliable.

Methods

Data preparation

In this retrospective study, we collected a dataset of 92 lung cancer patients who underwent treatment with the CK system at our hospital from 2019 to 2022. All patients' CT images were obtained from the same scanner (SIEMENS Somatom Definition AS64 CT scanner, Germany). The RT doses for these patients were calculated using an MC algorithm, and they received SRT. The uncertainty of the MC calculations was 1%, and the plan optimization was carried out using a fixed cone (typically, two cones are selected for a patient plan). We identified 86 eligible lung cancer cases after screening the case database. This screening was necessary since some patients had a target volume of more than 100 cm^3, which deviated from the typical clinical presentation and could impact the performance of the deep learning model. In addition, the treatment plans were obtained from the CK TPS (Accuray, System Version: CyberKnife 3.3.0, Software Version: Multiplan 4.0.2, USA). According to the guidelines provided by the AAPM TG101 [21], the prescribed dose for the planning target volume (PTV) in lung cancer cases ranged from 42.5 to 60 Gy, with a planned fraction range of 4–5. To mitigate the impact of variations in the prescribed dose on the training results of the model, we normalized the planned dose by dividing it by the corresponding prescribed dose. We randomly divided the lung cancer patients into two datasets: 66 cases for training/validation and 20 cases for testing. In addition, it is essential to note that all patient datasets involved in medical ethical review are listed separately at the end of the article, and the use of this data was strictly for academic scientific research purposes.

Overview and model architecture

Figure 1 illustrates the comprehensive workflow of our study methodology, along with the network model structure of the 3D U-Net. Initially, we employed a binary mask to extract the target area and collected delineation information on the OARs for all cases. These and the patient's CT images formed the Mask baseline control group. Subsequently, we performed dimensional matching between the extracted binary mask of the ROIs and the CT images, aligning them with the output 3D dose matrix. Following this, we employed an encoding algorithm to convert the patient's beam information into a 3D beam matrix. The specifics of this encoding algorithm can be found in the accompanying brief "fake program" shown in Algorithm 1. Subsequently, we incorporated this encoded beam matrix into the training channel, forming the AB experimental group (anatomical and beam information). These inputs were then fed into the 3D U-Net model for training. The model that demonstrated the best performance on the validation set was selected for data testing. During the testing phase, the channel input information from the test set and the selected optimal model were utilized to infer the corresponding 3D dose maps. Finally, we conducted qualitative and quantitative analyses of the predicted dose files to evaluate the model′s performance. It is important to note that our proposed method relies on the fluence volume of CK treatment plan beam information. Consequently, generating executable plans necessitates CK beam optimization and beam delivery.

During our study, we observed significant variations in the contour structure from case to case. These differences arose due to variances in institutional guidelines (across various treatment centers) and in planners′ the individual skills and preferences (within the same treatment center).

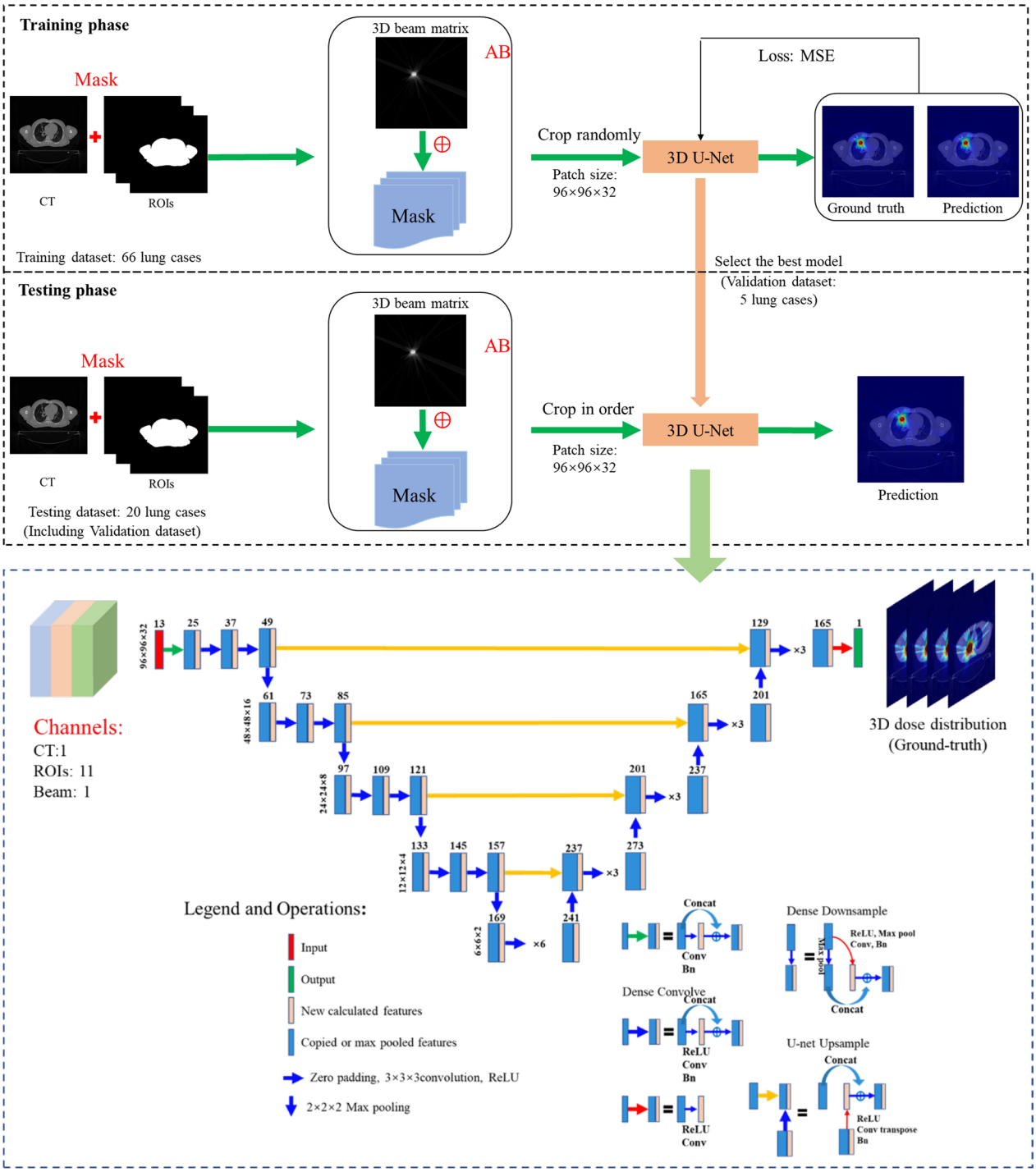

Fig. 1 The overall workflow of our dose prediction method and 3D U-Net model architecture

As a result, achieving a unified input channel for the model became challenging. To address this issue, we replaced the missing portions of the same OARs with an empty matrix (all zeros matrix). Additionally, we explored methods to enhance the original contours under the specific beam configuration. In the next step, we utilized the MIM software (Version: 6.9.5) to modify the contour files to standardize the delineation information for the patient's OARs. Reintroducing the complementary contour structure fostered discussions on enhancing the delineation considering the overall beam configuration.

Moreover, we delineated the halo (Rings with diameters of 0.5 cm, 1 cm, 2 cm, and 5 cm) for each patient's target area, utilizing the unified delineation structure for all patients. It enabled us to discuss the enhancement under the beam configuration by incorporating comprehensive delineation information for the target area halo. To summarize, we designed six comparison tests to evaluate the effectiveness and performance of the proposed approaches:

- The original contouring group, MaskO and ABO.
- The general whole contouring group, Mask and AB.
- The general whole contouring group by adding the target halo, MaskR and ABR.

The U-Net network model has gained widespread popularity since its publication in MICCAI in 2015 [22]. Its well-established structure is familiar, so we will only briefly introduce and focus on the essential details of the 3D U-Net model we constructed. Illustrated in Fig. 1, the model consists of either 12 input channels (Mask baseline control) or 13 channels (AB experimental group), with a 96×96×32 randomly sampled 3D matrix. It employs four-fold upsampling and downsampling and incorporates a densely connected layer at each hierarchical level, creating a 'dense structure' [23–26]. Each dense block encompasses a ReLU activation process, followed by convolution (3×3×3 kernel size), batch normalization, and connection to the previous layer. Zero padding is applied to each convolution, with 12 channels allocated to each layer. The 3D U-Net model undergoes downsampling via max pooling (2×2×2 kernel size) and symmetrical upsampling via deconvolution (2×2×2 kernel size, channel=80). The final hierarchical layer of the convolution generates a single-channel matrix, which serves as the output matrix.

Beam configuration representation

Based on the CK delivery and patient setups, a simplified schematic of the specific beam encoding model process is illustrated (Fig. 2). Simplified linear accelerator (Linac) treatment head modeling was implemented. A dummy 'radioactive source' was considered the initial X-ray excitation point. The distance from this source to the beam exit at the lower end of the secondary collimator (treatment head, point *b*) was set at around 40 cm, as specified in the treatment head parameter specification provided by the manufacturer–the processing steps of transforming CK case beam information into a coding matrix in Algorithm 1. Using the formula, we modeled each point in the beam matrix and saved the result as the "beam encoding matrix."

$$beam\ matrix = \sum_{j=1}^{n}\sum_{i=1}^{m}\left(\mathrm{O}_i + T[0,1]\times \frac{MU_j}{\min\{MU_1, MU_2 \ldots MU_n\}}\prod\nolimits_{voxel_i}\right) \tag{1}$$

where,

$$m = CT\ voxel\ number,\ and$$

$$\angle\left(\overrightarrow{b_j p_i}, \overrightarrow{b_j a_j}\right) \le \arctan\left(\frac{cone/4}{L_{b-source}}\right),$$

$$\prod\nolimits_{voxel_i} = the\ CT\ voxel\ i$$

$$\mathrm{O}_1 = zero\ matrix, \mathrm{O}_{i+1} = \mathrm{O}_i + T[0,1] \times \frac{MU_j}{\min\{MU_1, MU_2 \ldots MU_n\}}\prod\nolimits_{voxel_i},$$

$$MU_j = beam\ or\ MU\ j,\ n = beam\ number,\ T = target\ area,$$

Algorithm 1 Simplified beam coding

Assuming that any point in the CT model is denoted as P (x, y, z), the PTV for the actual beam irradiation is divided into multiple 'small targets,' and the number of beams in the treatment plan determines the quantity. Four critical parameters are extracted from the plan file (plan.xml):

- The 3D coordinates of the accelerator treatment head (point b) and the irradiated target point (point a);
- The size of the secondary collimator cone (light limiting cylinder);
- The accelerator dose in MU (beam's MU count).

Input: 3D coordinates of the three points P, a, b; *Cone size*; *MUs*
Step1: Calculating beam tensor angle： arctan((Cone/4)/$L_{b\text{-source}}$), ($L_{b\text{-source}}$: the distance from the source to the treatment head)
Step2: Loop traversal:
For P in the voxel matrix:
If the angle of the vector *bp* and *beam* (vector *ab*) <= beam tensor angle:
Assign a value to the P point: MUs/MUmin*1, (MUmin: minimum number of MUs in the accelerator's non-zero field)
Step3: Parallel multi-threaded CPUs

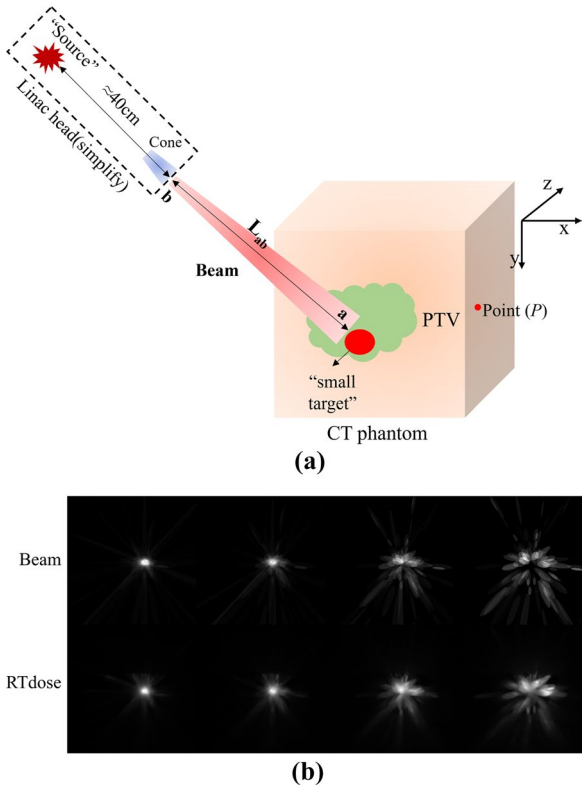

Fig. 2 A simplified schematic of the specific beam encoding model process. **a** Simplified schematic diagram of the treatment process of CK; **b** Transversal visualization of beam encoding and dose matrices

Training and validating the model

For the 86 lung cancer cases, the optimization was performed using the MC algorithm. We randomly selected 66 cases for training and 20 cases for testing. The model input matrix comprised 12 or 13 channels, depending on whether it was the Mask baseline or AB experimental model. The first channel corresponds to CT image information, while channels 2 to 12 encompass the contour information of various ROIs, such as PTV, GTV, body, lung_L, lung_R, heart, esophagus, spinal cord, trachea, and two auxiliary structures. The final channel represents the beam encoding matrix information. The training loss, calculated using the mean square error (MSE), was defined as the difference between the predicted dose (D_{pre}) and the ground truth dose (D_{gt}), as shown in the following Eq.

$$loss(pre, clinical) = \frac{1}{n}\sum_{i=1}^{n}\left(D_{pre}^{i} - D_{gt}^{i}\right)^{2} \quad (2)$$

The variable i represents the voxel index, while n represents the total number of voxels. The 3D U-Net training was performed on a server with a 72-core Intel(R) Xeon(R) Gold 6139 M CPU @ 2.30 GHz and four of 24 GB NVIDIA GeForce RTX 3090s, with TensorFlow-GPU version 2.4.0 platform. The training incorporated the Adam optimizer with an initial learning rate of $1e^{-4}$. The learning rate decayed exponentially from $1e^{-4}$ to $1e^{-6}$ during the CNN network training. One hundred randomly selected slices were collected for each case,

and the batch size was fixed at 2. Adding dense blocks in the network resulted in an extended training process. The model training was halted after 60 epochs, as the loss value for both the training and validation data had stabilized. The training duration for these 60 epochs encompassed approximately ten days. Upon successful model training, the complete 3D dose distribution tensor consisting of 96 × 96 × 32 voxels could be swiftly predicted for clinical deployment in patients requiring CK treatment (within a few minutes). Incorporating the beam encoding information in addition to the binary mask provides a more explicit representation of the dose distribution along the beam trajectory near the target area. Consequently, this reduction in learning difficulty significantly expedited the convergence rate of the network training. The learning rate had typically decreased to 1e-6 within the first 20 epochs.

Model performance

We employed multiple composite metrics to assess the predicted dose distribution of each model group. Initially, we computed the 3D gamma passing rates between the predicted and ground truth doses, considering the overall voxel (Body), target area (PTV), and OARs. Furthermore, the new conformance index (nCI), homogeneity index (HI), and gradient index (GI) were calculated based on the clinical plan reference criteria [27, 28], employing the following Eq.

$$nCI = (TV \times V_{RI})/(TV_{RI} \times TV_{RI}) \tag{3}$$

$$HI = D_5/D_{95} \tag{4}$$

$$GI = \left(volume\ of\ 50\%\ isodose\ line\right)/\left(volume\ of\ prescription\ isodose\ line\right) \tag{5}$$

Additionally, we introduced a simplified measure known as the structural similarity index measure (SSIM) to assess the similarity between two digital images. This measure combines three components, luminance, contrast, and structure, of two digital images (X, Y) to comprehensively assess image quality by comparing the observed image to the reference image. The simplified formula is presented below:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{6}$$

μ_x and μ_y denote the mean values of x and y, while σ_x^2 and σ_y^2 represent the variances of x and y, respectively. Additionally, σ_{xy} denotes the covariance between x and y. Constants are included to ensure numerical stability in the calculations, where $C_1=(k_1L)^2(k_1=0.01)$ and $C_2=(k_2L)^2(k_2=0.03)$. The variable L represents the dynamic range of the pixel values. In this case, it is set to the prescribed dose value for the corresponding case provided by the clinical plan.

We computed the average absolute error in the relative prescription dose for both the mean and maximum doses of the AB and Mask models, compared to the clinical dose (D_c), across all six experimental groups. Additionally, we examined the variations in critical indicators for the high dose region of the target, including D_{98}, D_{95}, D_{90}, D_{50}, D_5, and D_2, as well as isodose volume percentages and isovolumetric dose values (% of prescription dose), for several OARs. We calculated the absolute dose differences between the predicted and ground truth doses

Table 1 The gamma passing rates of the body and target area in lung cancer cases (mean ± SD, *p* value)

Input	Body		PTV	
	3 mm/3%	2 mm/2%	3 mm/3%	2 mm/2%
MaskO	69.35%±13.64%	51.29%±13.92%	80.21%±11.63%	61.19%±15.90%
Mask	72.90%±13.62%	59.43%±14.85%	98.67%±2.01%	91.79%±9.08%
MaskR	82.44%±11.51%	68.47%±12.59%	95.33%±5.44%	88.31%±9.74%
ABO	97.30%±2.87%	90.80%±6.37%	98.83%±1.40%	94.84%±4.93%
AB	97.45%±2.88%	90.93%±6.63%	99.39%±0.90%	95.90%±3.91%
ABR	97.27%±1.85%	88.65%±4.53%	99.47%±0.86%	96.33%±3.36%
Improvement (mean, p value)				
Comparison				
MaskO versus ABO	48.42% (< 0.001)	98.82% (< 0.001)	25.96% (< 0.001)	65.58% (< 0.001)
Mask versus AB	39.63% (< 0.001)	66.56% (< 0.001)	0.77% (0.166)	6.02% (0.078)
MaskR versus ABR	21.03% (< 0.001)	35.42% (< 0.001)	4.71% (< 0.005)	10.51% (< 0.002)

Annotation The gamma passing rate (GPR) for the body is calculated globally, considering the patient as a whole, while the GPR for the PTV (target area) is calculated locally, focusing on a single ROI. Both use a dose threshold of 10% of the reference dose

Table 2 The critical indicators of the target area and SSIM in lung cases (mean ± SD)

Input	nCI	HI	GI	SSIM(Body)	SSIM(PTV)
Clinical plan	1.0173 ± 0.0226	1.3435 ± 0.0210	1.0171 ± 0.0221	–	–
MaskO	1.4943 ± 0.2675	1.4095 ± 0.0414	1.4924 ± 0.2644	0.8836 ± 0.0636	0.9873 ± 0.0069
Mask	1.0201 ± 0.0232	1.3265 ± 0.0266	1.0201 ± 0.0232	0.9088 ± 0.0378	0.9954 ± 0.0027
MaskR	1.0455 ± 0.0623	1.3402 ± 0.0318	1.0453 ± 0.0618	0.9363 ± 0.0240	0.9951 ± 0.0016
ABO	1.0209 ± 0.0232	1.3299 ± 0.0307	1.0208 ± 0.0230	0.9636 ± 0.0107	0.9953 ± 0.0018
AB	1.0141 ± 0.0185	1.3057 ± 0.0276	1.0140 ± 0.0182	0.9554 ± 0.0128	0.9952 ± 0.0019
ABR	1.0085 ± 0.0141	1.3093 ± 0.0312	1.0083 ± 0.0135	0.9794 ± 0.0043	0.9955 ± 0.0020

Annotation The target area's nCI, HI, and GI are critical indicators to evaluate plan quality in clinical practice. The closer the nCI is to 1, the better the result is. The structural similarity index measure (SSIM) indicates the similarity quantifier between the observed and target images, between 0 and 1. The closer to 1, the better the result is

at each voxel. Based on these differences, we generated a dose map and an inter-dose difference map. Finally, we compared the DVHs for the lung cancer patients in the beam configuration group, focusing on each ROI.

Additionally, to account for variations in beam trajectory values and evaluate the accuracy of the Mask and AB model's predicted spatial distribution of doses, we computed the Dice similarity coefficients (DSCs) for isodose volumes ranging from 0 to 110% of the prescribed dose. These coefficients were compared to those obtained for the clinical dose (D_c). To facilitate this analysis, three-dimensional binary masks were generated for each isodose volume, encompassing all voxels with doses equal to or exceeding N% of the prescribed dose in the predicted dose (Y) and the clinical delivery dose (X). Subsequently, the following operations were conducted on these three-dimensional binary masks (X and Y):

$$DSC = \frac{2|X \cap Y|}{|X| + |Y|} = \frac{2 \times V_{iso-p} \times V_{iso-c}}{V_{iso-p} + V_{iso-c}} \quad (7)$$

where V_{iso-p} and V_{iso-c} denote the specific isodose volume of prediction and clinical truth, respectively. Additionally, it is worth noting that all dose deviation values were transformed into relative prescription dose errors.

Results

Gamma analysis of body and target area

Using the predicted and clinical doses, we calculated the 3D gamma passing rates for the overall voxel (Body) and the target area (PTV). The calculation was performed using open-source software 3D Slicer (version 4.11) [29], with evaluation criteria set at 3 mm/3% and 2 mm/2%. The results in Table 1 demonstrate that our proposed methods for incorporating beam configurations (the last three methods) yield accurate dose predictions in both the body and target areas while maintaining robustness in high and low-dose regions. Conversely, the general dose prediction methods for IMRT (MaskO and Mask) employed for dose prediction in CK treatment, particularly for the overall body dose, were found unsuitable. Including beam encoding information enables the model

Table 3 The gamma passing rates of OARs in lung cases (mean ± SD, *p* value)

Input	Lung_L 3 mm/3%	Lung_R 3 mm/3%	Heart 3 mm/3%	Esophagus 3 mm/3%	Spinal cord 3 mm/3%	Trachea 3 mm/3%
MaskO	77.53%±16.75%	75.64%±17.29%	75.05%±10.07%	68.69%±20.33%	69.53%±19.89%	77.78%±21.79%
Mask	78.35%±19.90%	75.85%±20.08%	76.19%±14.59%	64.66%±21.46%	73.66%±21.03%	76.57%±21.38%
MaskR	87.90%±8.49%	84.15%±11.60%	73.45%±16.19%	73.09%±21.01%	83.75%±12.10%	85.59%±12.38%
ABO	96.64%±2.86%	95.73%±6.01%	94.37%±6.90%	96.28%±6.11%	99.69%±0.84%	97.34%±5.81%
AB	98.22%±2.06%	95.09%±6.69%	94.06%±7.42%	97.05%±7.39%	99.81%±0.62%	97.81%±5.47%
ABR	98.46%±1.46%	95.84%±3.05%	97.61%±4.93%	94.70%±10.85%	99.85%±0.55%	96.85%±8.24%
Improvement (mean, p value)						
Comparison						
MaskO versus ABO	32.07% (< 0.001)	35.47% (< 0.001)	28.34% (< 0.001)	59.89% (< 0.001)	62.81% (< 0.001)	44.95% (< 0.001)
Mask versus AB	36.51% (< 0.001)	41.07% (< 0.001)	29.09% (< 0.001)	71.92% (< 0.001)	57.22% (< 0.001)	43.47% (< 0.001)
MaskR versus ABR	13.10% (< 0.001)	16.43% (< 0.001)	42.28% (< 0.001)	43.16% (< 0.001)	22.08% (< 0.001)	16.02% (< 0.005)

Annotation Considering the small size of the OARs and the low-dose region in the lung cancer case, the dose threshold used was 1% of the reference dose for each selected OAR

to learn both the high doses in the target area and the low doses in the trajectory of the treatment beam surrounding the target area.

The comparative results of six methods across three groups reveal that adding beam-encoding information to the original mask significantly improves model performance for the overall voxel and target area (MaskO and ABO). Such improvement is crucial in diverse clinical scenarios. Despite variations in clinical data, our encoded beam information consistently yields accurate prediction results. Furthermore, it is noteworthy that the complete delineation information plays a fundamental role in predicting the dose in the target area (Mask and MaskR). The dose prediction in the target area is highly accurate, regardless of the addition of beam encoding information, and the model's performance improvement diminishes with its inclusion. Incorporating a halo (Rings) around the target area necessitates additional delineation, leading to additional time and effort. However, if the target halos are precisely added during target delineation, this would enhance the overall body dose prediction, especially when no beam encoding information is added (MaskR). Notably, including a target halo input model suggests a potential influence on the prediction of surrounding OAR doses.

Critical metrics and SSIM

In addition to that, we also analyzed three critical metrics in the clinical plan design for the target area. These metrics include the new conformity index (nCI), homogeneity index (HI), and gradient index (GI). The results of the calculations are presented in Table 2. It was observed that the nCI, HI, and GI for all five methods, except the MaskO method, matched well with the clinical plan. On the other hand, models that solely relied on the original mask information exhibited poor performance in all the critical indicators for the target area, as shown in Table 1. However, when the beam encoding information was incorporated (ABO method), there was a significant improvement in the critical indicators for the target area. Meanwhile, the other four methods did not show any significant differences compared to the clinical plan, and the critical indicators for the target area were minimal after the inclusion of the beam encoding information. It emphasizes the importance of having complete delineation information in accurately predicting the dosage for the target area.

In this context, a simplified structural similarity index measure (SSIM) was introduced to assess the similarity between two digital images. This measure was utilized to estimate the image quality of an observed image (predicted dose, X) about a reference image (clinical

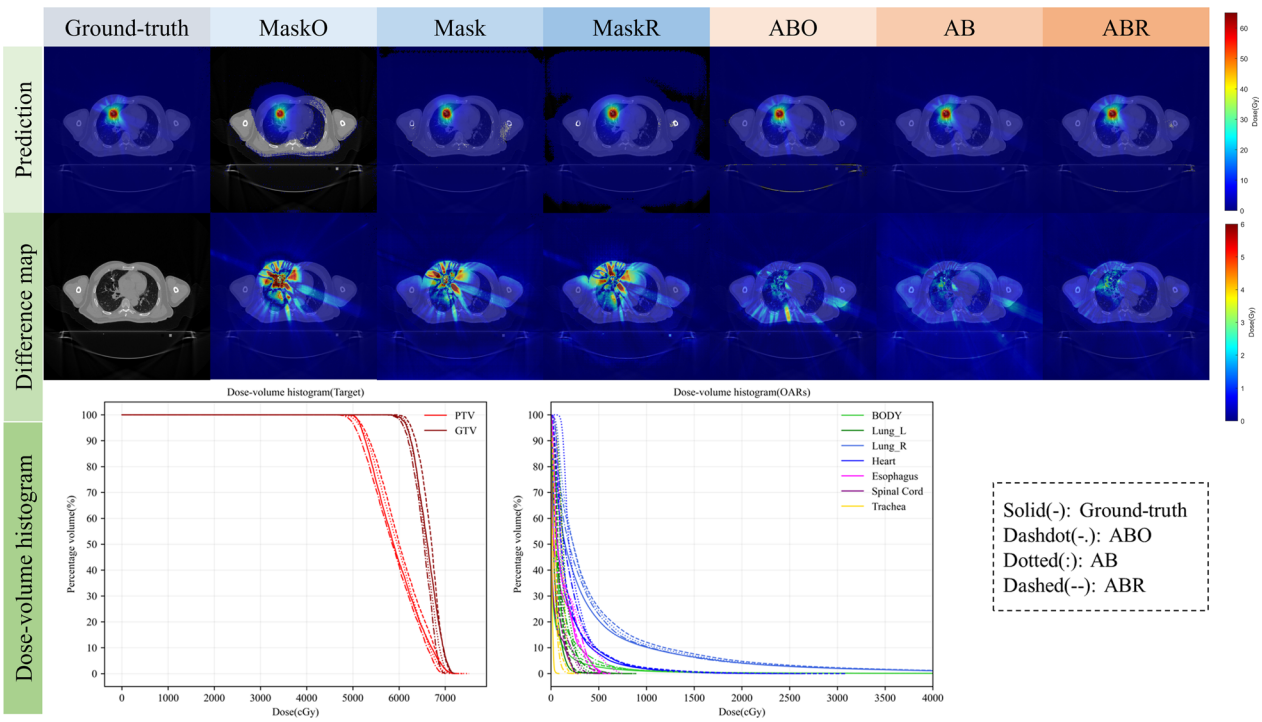

Fig. 3 Quantitative comparison of six methods in three groups of experiments. In the first row, the first image is the dose distribution of a representative clinical plan for lung cancer cases, used as Ground truth. The other images are the predicted model results of the six methods. The second row shows the absolute differences between each method's clinical and predicted plans. The last row shows the DVHs of the clinical and predicted plans in each method, adding the beam encoding information containing the results of target areas and OARs

dose, Y). The overall SSIM was calculated for the three-dimensional dose matrix of the body and PTV regions, as displayed in the last two columns of Table 2. The calculation results indicate that the AB model exhibits an overall enhancement in the body region compared to the Mask model. However, in the PTV region, both models demonstrate similar performance. This finding also aligns with the gamma passing rate calculation, suggesting that the AB model accurately predicts beam traces akin to the clinical plan. Conversely, the Mask model solely predicts the average dose range within the target area.

Gamma analysis of OARs

In addition, we conducted an assessment of the disparities between the predicted and clinical dose distributions for each OAR. Table 3 provides the gamma passing rates calculated at the 3 mm/3% criteria for OARs in lung cancer cases. The longitudinal comparison in Table 3 shows that the Mask and AB models exhibit minimal differences in their respective input structures. However, the MaskR method, which includes the addition of target halos, demonstrates superior performance compared to the first two methods. The model's prediction is satisfactory for the latter three methods, incorporating beam encoding information for immense- and miniature-volume OARs. This observation further verifies the effectiveness of our developed beam encoding algorithm, as it enables accurate prediction of beam trajectory dose.

Furthermore, the transversal comparison reveals that the prediction of small-volume OAR (spinal cord) dose is more accurate than that of the immense-volume OAR (heart). Additionally, among the six methods, the dose distribution of the left lung is better predicted than that of the right lung. This discrepancy could be attributed to the fact that the left lung predominantly serves as the RT target area in the test data, resulting in a superior prediction of its dose distribution compared to that of the OAR.

Moreover, comparing the results of the six methods in the three groups shows that the improvement in model performance when beam encoding information is added to the available full mask is most pronounced for all OARs (Mask and AB). The improvement shows a less pronounced effect for the MaskR and ABR groups compared to the other two groups. It should be noted that the predicted improvement in the irradiated dose for the heart is particularly significant in the MaskR and ABR groups, likely because of the beam trajectory transversing the heart and the constraining effect of target halos delineation.

Quantitative comparison

Based on a quantitative comparison of six methods in three groups of experiments (Fig. 3), it is evident that the predicted dose distributions around the target area (referred to as beam trajectory dose) are well aligned with the clinical plan when the beam encoding information is incorporated into the other three training methods, as opposed to using only the input binary mask. Conversely, the predicted dose around the target area appears blurry for the three methods relying solely on binary mask input. These methods cannot accentuate the dose within the clinical target area, indicating that the network model cannot effectively learn the dose trajectory for multiple beam configurations based solely on the patient's anatomical information. Consequently, the model compromises the data and yields predictions of a fuzzy dose.

The second row displays maps illustrating the absolute difference in dose distribution between the predicted and clinical doses. Analysis of the dose difference maps reveals that while the prediction results for the target area (high dose region) indicate acceptable accuracy, the dose predictions surrounding the target area are extremely poor for the three methods that solely utilize binary mask input. The presence of errors in the mid-dose region within numerous voxel patches significantly increases the likelihood of surpassing the prescribed dose limits for the OARs, rendering it unacceptable within a clinical plan. The AB and ABR methods demonstrate more reasonable predictions, with maximum dose point differences below 6 Gy for lung cancer cases and minor dose differences (order of magnitude in cGy) for most voxels.

DVHs are valuable for optimizing the clinical planning process. The third row visualizes the DVHs of the ROI dose and volume differences for the three methods, including the addition of beam encoding information. The four-line types represent the dose volume histograms of the clinical plan and the three comparison methods. The predicted doses for all three methods, in both the target area and the OARs, closely match the clinical plan with slight volume deviations, primarily in the low-dose region, which falls within clinical acceptability.

In summary, incorporating beam configurations has significantly improved the model's accuracy in predicting dose distribution. It maintains strong robustness in the overall voxel, target area, and OARs, highlighting the model's excellent generality.

Discussion

Prior knowledge-based planning (KBP) techniques have recently been well-developed [15]. The use of deep convolutional neural networks (CNNs) in the RT domain has become increasingly common. However, previous deep-learning-based dose prediction studies have mainly focused on IMRT or VMAT (volumetric modulated arc therapy) plans [12, 30–32], with limited research on SRT

Table 4 Mean absolute error and standard deviation (mean ± SD) for relevant dose-volume metrics on the target and several organs for lung cancer testing for the Mask and AB models. The values are expressed as a percentage of the prescription dose (ranging from 42.5 to 60 Gy) for the metrics reporting the dose received by x % of volume (D_x) and an absolute difference for the metrics reporting the volume (in %) receiving a dose of y Gy (V_y)

	Mask Model			AB Model		
	MaskO	Mask	MaskR	ABO	AB	ABR
PTV						
D_{98} (% of D_{pre})	16.11 ± 3.31	3.68 ± 3.18	2.98 ± 2.82	2.10 ± 1.31	2.05 ± 1.25	2.64 ± 1.92
D_{95} (% of D_{pre})	14.45 ± 2.82	4.36 ± 3.32	2.55 ± 1.80	2.19 ± 1.21	2.39 ± 1.60	3.30 ± 1.86
D_{90} (% of D_{pre})	13.12 ± 2.76	5.15 ± 3.85	3.03 ± 1.75	2.67 ± 1.36	2.92 ± 1.66	3.86 ± 1.86
D_{50} (% of D_{pre})	12.60 ± 3.42	6.08 ± 4.08	4.29 ± 2.97	3.34 ± 2.00	2.93 ± 1.74	4.04 ± 1.55
D_5 (% of D_{pre})	13.67 ± 3.43	4.70 ± 3.04	3.47 ± 2.94	2.27 ± 1.82	1.80 ± 1.55	1.56 ± 1.30
D_2 (% of D_{pre})	12.07 ± 3.79	5.59 ± 3.91	3.52 ± 3.21	2.58 ± 1.98	1.97 ± 1.51	1.70 ± 1.44
Lungs						
D_{mean} (% of D_{pre})	1.47 ± 1.09	1.99 ± 1.43	1.11 ± 0.57	1.13 ± 0.53	1.92 ± 0.45	1.50 ± 0.21
V_5 (% of volume)	5.53 ± 5.52	8.56 ± 6.91	3.30 ± 2.10	3.96 ± 3.35	4.49 ± 3.50	4.02 ± 1.77
V_{20} (% of volume)	0.74 ± 0.74	0.84 ± 0.83	0.47 ± 0.44	0.24 ± 0.17	0.24 ± 0.17	0.46 ± 0.26
V_{30} (% of volume)	0.52 ± 0.57	0.31 ± 0.41	0.22 ± 0.19	0.10 ± 0.07	0.13 ± 0.08	0.23 ± 0.14
Heart						
V_{10} (% of volume)	0.77 ± 1.28	1.02 ± 1.35	0.69 ± 1.30	0.89 ± 1.78	0.83 ± 1.94	0.69 ± 1.53
Esophagus						
D_2 (% of D_{pre})	4.98 ± 3.43	5.27 ± 4.31	3.82 ± 4.71	2.77 ± 2.38	2.67 ± 1.98	2.90 ± 2.67
V_{10} (% of volume)	5.44 ± 8.12	4.29 ± 7.73	5.05 ± 11.80	3.04 ± 5.37	3.36 ± 5.81	5.88 ± 10.19
Spinal cord						
D_2 (% of D_{pre})	5.28 ± 2.90	4.68 ± 2.92	4.75 ± 3.50	1.78 ± 1.22	2.23 ± 1.32	1.24 ± 1.30
Trachea						
D_2 (% of D_{pre})	5.95 ± 3.73	6.63 ± 4.07	5.47 ± 4.47	3.37 ± 2.10	3.26 ± 2.33	3.56 ± 3.28
V_{10} (% of volume)	4.98 ± 8.57	4.54 ± 7.58	4.18 ± 6.84	1.80 ± 3.29	1.88 ± 3.28	3.22 ± 5.19

Annotation Lungs mean left and right lungs as a whole minus target

(e.g., CyberKnife) dose prediction. This study aimed to accurately and efficiently predict the MC dose of heterogeneous tissue tumors, such as lung cancer. To achieve this, we collected lung cancer cases, included beam encoding information from the RT plan in the training phase, and finally accurately predicted the dose distribution for CK lung cancer cases. Through robust learning of beam configuration, the deep learning model enables quick and accurate dose calculation, eliminating the need for time-consuming, redundant TPS calculations. It significantly improves the operational efficiency of planning, providing planners with a reference for setting objective constraints, optimizing the SRT plan design process, and guiding clinical practices. Therefore, this method solves the extended time required for MC dose calculations in clinical practice for patients with lung cancer and other heterogeneous tissue tumors. Additionally, it can serve as a viable and potential platform for plan validation.

The AB model showed improvements in several ROIs, including the target area and OARs, when comparing the isovolumetric dose values (% of prescription dose) and the isodose volume percentages. Table 4 presents the clinically relevant dose-volume metrics commonly used to evaluate CK treatment plans for lung cancer in the target areas and OARs of both models. Furthermore, the AB model demonstrated superior performance to the Mask model in most situations, exhibiting better predictive accuracy across all considered dose-volume metrics. While the mean error differences were generally low for most metrics (approximately 2% for the target area and within 3% for the majority of the OARs), there were specific metrics, such as lungs V_5 and trachea D_2, that showed errors of approximately 4% and within 3.5%, respectively. Heart V_{10} and spinal cord D_2 exhibited significant reductions, with values of less than 1% and 2%, respectively.

To quantitatively analyze the differences in dose between the PTV and OARs in patients using the six methods, we generated histograms of the average absolute error and standard deviation (SD) for the mean and

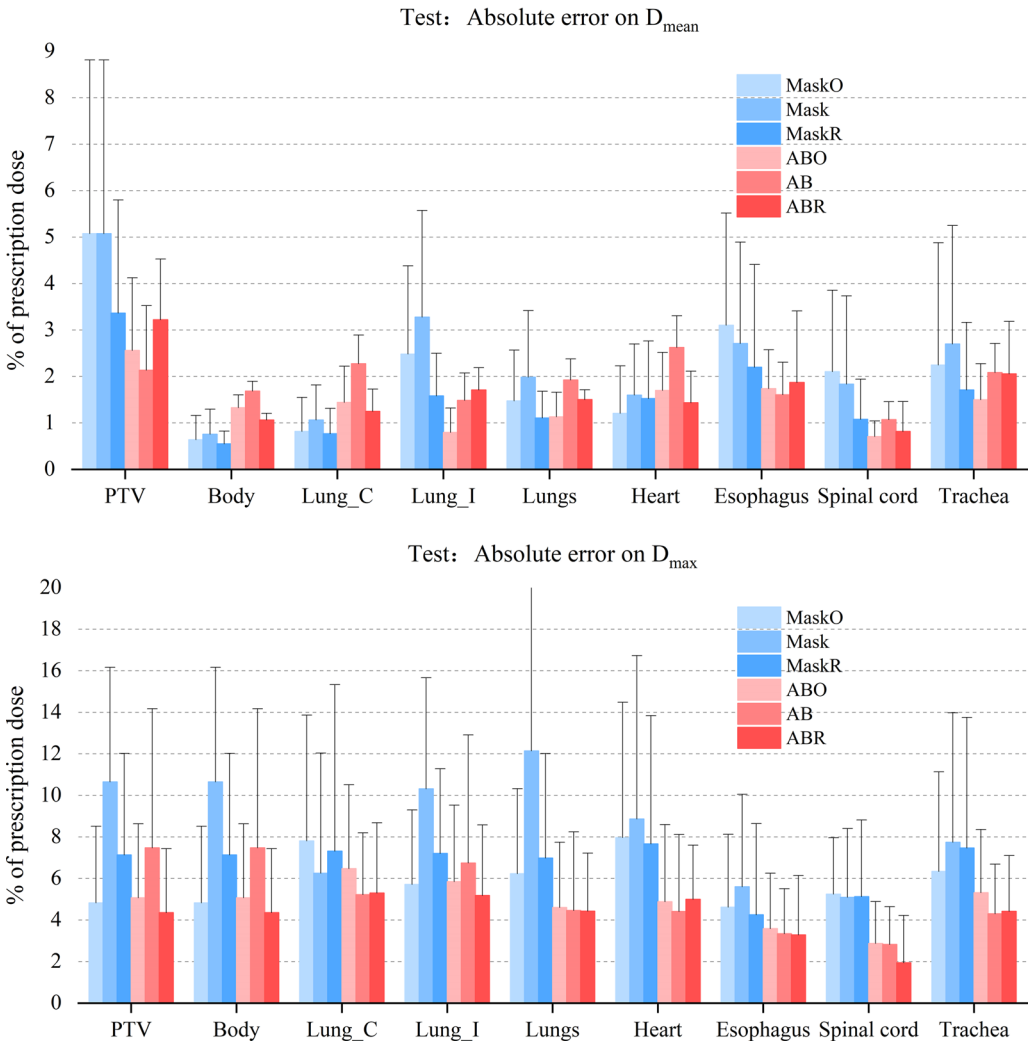

Fig. 4 Average absolute errors on the mean (top) and maximum doses (bottom) for the predictions (AB and Mask models) versus the clinical dose (D_C) of all six experimental groups on the test set for relevant targets and organs. The black lines on top of the bars represent the standard deviation associated with each ROI. Lung_C means contralateral lung, and Lung_I means ipsilateral lung. Considering the abysmal statistical result for PTV on the D_{mean} of the MaskO method, Mask's is used instead

maximum dose errors of the PTV and each OAR. These histograms are presented in Fig. 4. For both scenarios, the mean and maximum dose errors projected by the AB model were, on average, around −0.7–4% and 1.5–4.0% lower compared to the Mask model. The organs with a negative elevation in mean dose error were body, lung_C (contralateral lung), and heart. It can be attributed to the Mask model only learning the overall mean characteristics and the doses of the beam trajectory were not predicted accurately.

The average dose error for all organs in the Mask model was 2.31 ± 1.57%, while 1.65 ± 0.69% in the AB model. The average error of the maximum dose for all organs in the Mask model was 7.08 ± 5.13%, while 4.75 ± 3.40% in the AB model. The lungs and heart showed the most significant differences in the maximum dose, with the AB model exhibiting prediction errors of 4.0% and 3.5% lower than those of the Mask model. The maximum voxel dose difference for the three methods, when beam encoding information was added (i.e. AB model), ranged from 1.0 to 3.0 Gy, corresponding to approximately a 2.0–6.0% difference relative to the prescription dose, while the mean dose difference ranged from 0.3 to 1.0 Gy, accounting for 0.7-2.0% of the prescription dose. The difference in target metrics ranged from 1.0 to 2.0 Gy, corresponding to approximately a 2.0–4.0% difference relative to the prescription dose. These results significantly improve

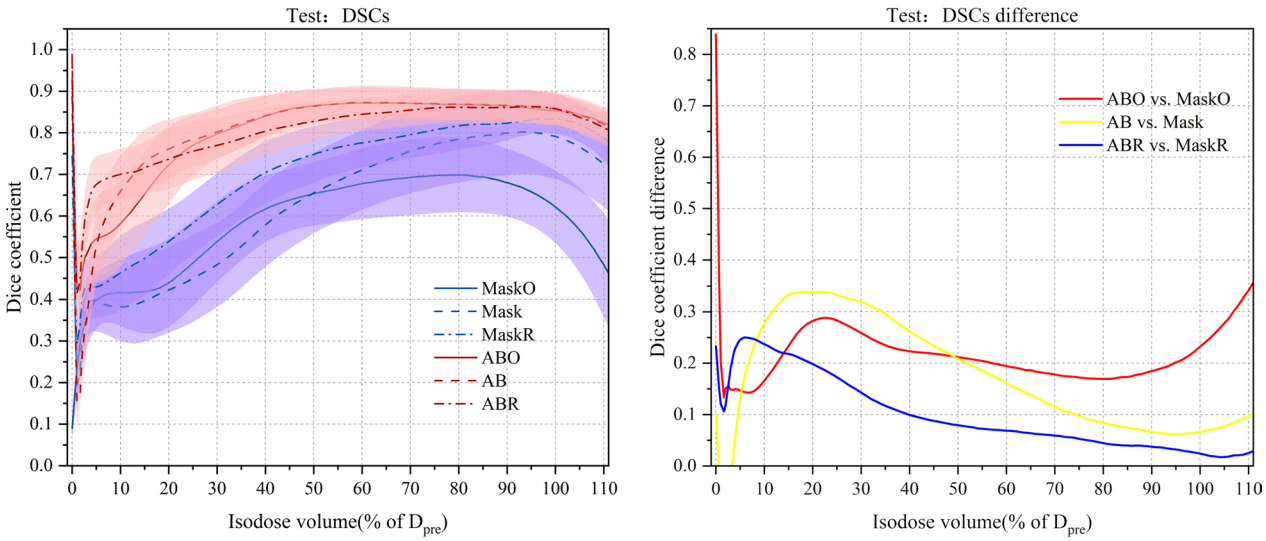

Fig. 5 The left plot contains the DSCs of the isodose volumes from 0–110% of the prescription dose for Mask (three blue lines) and AB (three red lines) model versus clinical dose (Dc), together with their corresponding standard deviation (color wash), for the lung cancer testing set. The right plot contains the difference between the averaged Dice coefficient from the AB model versus the Mask model (three comparison groups are shown by the labels on the figure)

compared to certain IMRT dose prediction studies [1, 33].

To standardize the beam trajectory values and evaluate the accuracy of the predicted dose spatial distribution from both the Mask and AB models, we calculated the DSCs for isodose volumes (% of prescription dose). The results are presented in Fig. 5. The AB model performed superior to the Mask model across all evaluation criteria. Although the difference in prediction errors between the two models is relatively small in the high-dose region (around 90–100% of the prescription dose, on average), significant disparities are observed in the medium to low-dose region (up to 15–40% isodose volume), particularly in the 0–1% isodose volumes where accurate beam encoding information is crucial. For these specific regions, the Dice coefficient of the AB model is 20–40% higher than that of the Mask model. The predictions of the Mask model fail to capture the dose features along the beam path, resulting in a uniform and isotropic dose fall-off. In contrast, the AB model accurately predicts the dose values corresponding to different beam geometries. This finding demonstrates the strong performance of our proposed beam-encoding-based model.

In summary, our proposed method has many advantages. On the one hand, introducing the beam encoding information makes the model′s learning process more efficient. On the other hand, our methods modify the input data for model training without altering the network structure, resulting in a highly compatible model that can be used with various other networks.

Nonetheless, our method still has limitations. Upon comparing the dose maps in Fig. 3, noticeable differences in the dose maps of the ABO method for the beam at the shallow surface of the body phantom (starting point) still exist. These differences necessitate additional optimization and enhancement of our subsequent algorithms for beam encoding. Ideally, the model should be able to effectively learn the dose distribution by incorporating beam configuration information, even when dealing with heterogeneous clinical data. The authors propose an improvement strategy involving pairing the CT values from the patient's CT image with the beam matrix. This convolution is followed by uniform normalization based on the patient's prescribed dose. This approach showcases the randomness of the beam matrix values to some extent as they change with the CT values. The addition of this joint convolutional layer has the potential to optimize the prediction. In the future, we intend to develop a reverse planning system using the current commercial CK TPS to investigate the potential application of our dose prediction approach in expedited clinical planning.

Conclusions

We employed deep CNNs to construct a model that learns from a database comprising CK patients treated with different beam configurations to predict the 3D dose distribution for new patients. Two models were developed and compared: the Mask model solely incorporates the patient's anatomical information, whereas the second model, denoted as AB, incorporates anatomical structure

and beam configuration information. The AB model demonstrates superior accuracy and resilience in the face of varying beam geometries compared to the Mask model. Utilizing the 3D beam matrix, which encompasses essential dose characteristics along the beam path as input, enables more comprehensive automatic planning using deep CNNs. It obviates the need to train distinct models for each beam alignment. Consequently, this optimization enhances the RT process concerning SRT techniques.

Abbreviations

TPS	Treatment planning systems
CK	CyberKnife
RT	Radiation therapy
OAR	Organ-at-risk
KBP	Knowledge-based planning
IMRT	Intensity-modulated radiation therapy
DVH	Dose-volume histogram
CNN	Convolutional neural network
ROIs	Regions of interest
AI	Artificial intelligence
SRT	Stereotactic radiotherapy
MC	Monte Carlo
PTV	Planning target volume
Linac	Linear accelerator
MSE	Mean square error
nCI	new Conformance index
HI	Homogeneity index
GI	Gradient index
SSIM	Structural similarity index measure
DSC	Dice similarity coefficient
VMAT	Volumetric modulated arc therapy
SD	Standard deviation

Acknowledgements
We sincerely thank the editor and all reviewers for their insightful comments and constructive suggestions, which helped us improve the quality of this manuscript. The authors thank Nan Li for reviewing the manuscript in this study.

Author contributions
Yuchao Miao: Conceptualization, Experiment design, Code implementation, Preparation of the Tables and figures, Writing—original draft. Jiwei Li: Resources, Technical support, Software. Ruigang Ge: Data collection, Software. Chuanbin Xie: Data curation. Yaoying Liu: Statistical analysis and interpretation. Gaolong Zhang: Writing—review & editing. Mingchang Miao: Conceptualization, Statistical analysis. Shouping Xu: Supervision, Writing—review & editing.

Funding
This work was supported by the National Natural Science Foundation of China (NSFC Grant No. 81801799, 12375359, 11975040, U2167204).

Availability of data and materials
The datasets generated and analyzed during the current study are not publicly available due to protection of individual patient privacy but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The data used in the study were anonymized. They follow the ethical standards of the institutional research committee and the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. The study was approved with exemption from informed consent by the independent ethics committee of the PLA General Hospital.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China. [2]ACCURAY, China National Nuclear Corporation, Tianjin, China. [3]Department of Radiation Oncology, The First Medical Center of the People's Liberation Army General Hospital, Beijing, China. [4]School of Physics, Beihang University, Beijing, China. [5]Department of Radiation Oncology, the Fourth Hospital of Hebei Medical University, Shijiazhuang, Hebei, China. [6]Department of Radiation Oncology, Fujian Medical University Union Hospital, Fuzhou, Fujian, China.

Received: 13 March 2024 Accepted: 25 September 2024
Published online: 25 November 2024

References

1. Barragán-Montero AM, Nguyen D, Lu W, Lin MH, Norouzi-Kandalan R, Geets X, et al. Three-dimensional dose prediction for lung IMRT patients with deep neural networks: robust learning from heterogeneous beam configurations. Med Phys. 2019;46(8):3679–91. https://doi.org/10.1002/mp.13597.
2. Kandalan RN, Nguyen D, Rezaeian NH, Barragán-Montero AM, Breedveld S, Namuduri K, et al. Dose prediction with deep learning for prostate cancer radiation therapy: model adaptation to different treatment planning practices. Radiotherapy Oncol J Eur Soc Ther Radiol Oncol. 2020;153:228–35. https://doi.org/10.1016/j.radonc.2020.10.027.
3. Guerreiro F, Seravalli E, Janssens GO, Maduro JH, Knopf AC, Langendijk JA, et al. Deep learning prediction of proton and photon dose distributions for paediatric abdominal tumours. Radiotherapy Oncol J Eur Soc Ther Radiol Oncol. 2021;156:36–42. https://doi.org/10.1016/j.radonc.2020.11.026.
4. Momin S, Lei Y, Wang T, Zhang J, Roper J, Bradley JD, et al. Learning-based dose prediction for pancreatic stereotactic body radiation therapy using dual pyramid adversarial network. Phys Med Biol. 2021;66(12):125069.
5. Fan J, Wang J, Chen Z, Hu C, Zhang Z, Hu W. Automatic treatment planning based on three-dimensional dose distribution predicted from deep learning technique. Med Phys. 2019;46(1):370–81. https://doi.org/10.1002/mp.13271.
6. Cao Y, Kunaprayoon D, Xu J, Ren L. AI-assisted clinical decision making (CDM) for dose prescription in radiosurgery of brain metastases using three-path three-dimensional CNN. Clin Transl Radiat Oncol. 2023;39:100565. https://doi.org/10.1016/j.ctro.2022.100565.
7. Tian F, Zhao S, Geng C, Guo C, Wu R, Tang X. Use of a neural network-based prediction method to calculate the therapeutic dose in boron neutron capture therapy of patients with glioblastoma. Med Phys. 2023;50(5):3008–18. https://doi.org/10.1002/mp.16215.
8. Babier A, Mahmood R, McNiven AL, Diamant A, Chan TCY. The importance of evaluating the complete automated knowledge-based planning pipeline. Phys Med. 2020;72:73–9. https://doi.org/10.1016/j.ejmp.2020.03.016.
9. Zimmermann L, Faustmann E, Ramsl C, Georg D, Heilemann G. Technical note: dose prediction for radiation therapy using feature-based losses and one cycle learning. Med Phys. 2021;48(9):5562–6. https://doi.org/10.1002/mp.14774.
10. Liu S, Zhang J, Li T, Yan H, Liu J, Technical Note. A cascade 3D U-Net for dose prediction in radiotherapy. Med Phys. 2021;48(9):5574–82. https://doi.org/10.1002/mp.15034.
11. Kubo K, Monzen H, Ishii K, Tamura M, Kawamorita R, Sumida I, et al. Dosimetric comparison of RapidPlan and manually optimized plans in volumetric modulated arc therapy for prostate cancer. Phys Med PM Int J Devot Appl Phys Med Biol. 2017;44:199–204. https://doi.org/10.1016/j.ejmp.2017.06.026.

12. Feng B, Zhou W, Yang X, Luo H, Zhang X, Yang D, et al. Pseudo-siamese network combined with dosimetric and clinical factors, radiomics features, CT images and 3D dose distribution for the prediction of radiation pneumonitis: a feasibility study. Clin Transl Radiat Oncol. 2023;38:188–94. https://doi.org/10.1016/j.ctro.2022.11.011.
13. Hu J, Song Y, Wang Q, Bai S, Yi Z. Incorporating historical sub-optimal deep neural networks for dose prediction in radiotherapy. Med Image Anal. 2021;67:101886. https://doi.org/10.1016/j.media.2020.101886.
14. Murakami Y, Magome T, Matsumoto K, Sato T, Yoshioka Y, Oguchi M. Fully automated dose prediction using generative adversarial networks in prostate cancer patients. PLoS ONE. 2020;15(5):e0232697. https://doi.org/10.1371/journal.pone.0232697.
15. Ge Y, Wu QJ. Knowledge-based planning for intensity-modulated radiation therapy: a review of data-driven approaches. Med Phys. 2019;46(6):2760–75. https://doi.org/10.1002/mp.13526.
16. Li F, Niu S, Han Y, Zhang Y, Dong Z, Zhu J. Multi-stage framework with difficulty-aware learning for progressive dose prediction. Biomed Signal Process Control. 2023;82:104541. https://doi.org/10.1016/j.bspc.2022.104541.
17. Hedden N, Xu H. Radiation therapy dose prediction for left-sided breast cancers using two-dimensional and three-dimensional deep learning models. Phys Med PM Int J Devoted Appl Phys Med Biology 2021;83:101–7. https://doi.org/10.1016/j.ejmp.2021.02.021.
18. Cheng Y, Lin Y, Long Y, Du L, Chen R, Hu T, et al. Is the CyberKnife(©) radiosurgery system effective and safe for patients? An umbrella review of the evidence. Future Oncol (London Engl). 2022;18(14):1777–91. https://doi.org/10.2217/fon-2021-0844.
19. Hayashi K, Suzuki O, Shiomi H, Ono H, Setoguchi A, Nakai M, et al. Stereotactic ablative body radiotherapy with a central high dose using CyberKnife for metastatic lung tumors. BMC Cancer. 2023;23(1):215. https://doi.org/10.1186/s12885-023-10635-6.
20. Soman C, Alghamdi SRM, Alazemi FNM, Alghamdi AAA. Cyberknife radiosurgery for the treatment of head and neck cancer: a systematic review. Eur J Dent. 2022;16(2):266–73. https://doi.org/10.1055/s-0041-1736330.
21. Benedict SH, Yenice KM, Followill D, Galvin JM, Hinson W, Kavanagh B, et al. Stereotactic body radiation therapy: the report of AAPM task group 101. Med Phys. 2010;37(8):4078–101. https://doi.org/10.1118/1.3438081.
22. Ronneberger O, Fischer P, Brox T, editors. U-net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18; Springer, 2015.
23. Nguyen D, Jia X, Sher D, Lin MH, Iqbal Z, Liu H, et al. 3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture. Phys Med Biol. 2019;64(6):065020. https://doi.org/10.1088/1361-6560/ab039b.
24. Zhang J, Liu S, Yan H, Li T, Mao R, Liu J. Predicting voxel-level dose distributions for esophageal radiotherapy using densely connected network with dilated convolutions. Phys Med Biol. 2020;65(20):205013. https://doi.org/10.1088/1361-6560/aba87b.
25. Gronberg MP, Gay SS, Netherton TJ, Rhee DJ, Court LE, Cardenas CE. Technical note: dose prediction for head and neck radiotherapy using a three-dimensional dense dilated U-net architecture. Med Phys. 2021;48(9):5567–73. https://doi.org/10.1002/mp.14827.
26. Lempart M, Benedek H, Jamtheim Gustafsson C, Nilsson M, Eliasson N, Bäck S, et al. Volumetric modulated arc therapy dose prediction and deliverable treatment plan generation for prostate cancer patients using a densely connected deep learning model. Phys Imag Radiat Oncol. 2021;19:112–9. https://doi.org/10.1016/j.phro.2021.07.008.
27. Patel G, Mandal A, Choudhary S, Mishra R, Shende R. Plan evaluation indices: a journey of evolution. Rep Practical Oncol Radiother J Greatpoland Cancer Cent Poznan Pol Soc Radiat Oncol. 2020;25(3):336–44. https://doi.org/10.1016/j.rpor.2020.03.002.
28. Sun Z, Xia X, Fan J, Zhao J, Zhang K, Wang J, et al. A hybrid optimization strategy for deliverable intensity-modulated radiotherapy plan generation using deep learning-based dose prediction. Med Phys. 2022;49(3):1344–56. https://doi.org/10.1002/mp.15462.
29. Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin JC, Pujol S, et al. 3D slicer as an image computing platform for the quantitative Imaging Network. Magn Reson Imag. 2012;30(9):1323–41. https://doi.org/10.1016/j.mri.2012.05.001.
30. Wang Y, Liu Y, Bai Y, Zhou Q, Xu S, Pang X. A generalization performance study on the boosting radiotherapy dose calculation engine based on super-resolution. Z Med Phys. 2023. https://doi.org/10.1016/j.zemedi.2022.10.006.
31. Sumida I, Magome T, Das IJ, Yamaguchi H, Kizaki H, Aboshi K, et al. A convolution neural network for higher resolution dose prediction in prostate volumetric modulated arc therapy. Phys Med PM Int J Devot Appl Phys Med Biol 2020;72:88–95. https://doi.org/10.1016/j.ejmp.2020.03.023.
32. Ahn SH, Kim E, Kim C, Cheon W, Kim M, Lee SB, et al. Deep learning method for prediction of patient-specific dose distribution in breast cancer. Radiat Oncol (London Engl). 2021;16(1):154. https://doi.org/10.1186/s13014-021-01864-9.
33. Yue M, Xue X, Wang Z, Lambo RL, Zhao W, Xie Y, et al. Dose prediction via distance-guided deep learning: initial development for nasopharyngeal carcinoma radiotherapy. Radiother Oncol J Eur Soc Ther Radiol Oncol. 2022;170:198–204. https://doi.org/10.1016/j.radonc.2022.03.012.

Publisher's note