Physics and Imaging in Radiation Oncology 34 (2025) 100735

Contents lists available at ScienceDirect

Physics and Imaging in Radiation Oncology

journal homepage: www.sciencedirect.com/journal/physics-and-imaging-in-radiation-oncology

Original Research Article

Deep learning-based prediction of Monte Carlo dose distribution for heavy ion therapy

Rui He [a,b,c,d,1], Hui Zhang [a,c,d,1], Jian Wang [a,c,d,e], Guosheng Shen [a,c,d], Ying Luo [a,c,d,e], Xinyang Zhang [a,c,d,e], Yuanyuan Ma [a,c,d], Xinguo Liu [a,c,d], Yazhou Li [a,c,d,e,f], Haibo Peng [b], Pengbo He [a,c,d,*], Qiang Li [a,c,d,e,g,*]

[a] Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou 730000, China
[b] School of Nuclear Science and Technology, Lanzhou University, Lanzhou 73000,China
[c] Key Laboratory of Heavy Ion Radiation Biology and Medicine of Chinese Academy of Sciences, Lanzhou 730000, China
[d] Key Laboratory of Basic Research on Heavy Ion Radiation Application in Medicine, Gansu Province, Lanzhou 730000, China
[e] University of Chinese Academy of Sciences, Beijing 100049, China
[f] Gansu Provincial Hospital, Lanzhou 730000, China
[g] Putian Lanhai Nuclear Medicine Research Center, Putian 351152, China

ARTICLE INFO

Keywords:
Deep learning
Heavy ion therapy
Dose prediction
Monte Carlo simulation
Analytical algorithm

ABSTRACT

Background and purpose: Current methods, like treatment planning system algorithms (TPSDose), lack accuracy, whereas Monte Carlo dose distribution (MCDose) is accurate but computationally intensive. We proposed a deep learning (DL) model for rapid prediction of Monte Carlo simulated dose distribution (MCDose) in heavy ion therapy (HIT).

Materials and methods: We developed a DL model – the Cascade Hierarchically Densely 3D U-Net (CHD U-Net) – to predict MCDose using computed tomography images and TPSDose of 67 head-and-neck patients and 30 thorax-and-abdomen patients. We also compared the results with other proton dose DL models and TPSDose.

Results: Compared to TPSDose, the gamma passing rate (GPR) improved by 16 % (1 %/1 mm). Notably, the model achieved 99 % and 97 % accuracy under clinically relevant criteria (3 %/3 mm) across the whole dose distribution in patients. For head-and-neck patients, the GPRs of the C3D and HD U-Net models in the PTV region were 97 % and 85 %, and in the body were 98 % and 97 %, respectively. For thorax-and-abdomen patients, the GPR of the C3D and HD U-Net models in the PTV region were 71 % and 51 %, and in the body were 95 % and 90 %, respectively.

Conclusions: The proposed CHD U-Net model can predict MCDose in a few seconds and outperforms two alternative DL models. The predicted dose can replace TPSDose in HIT clinical process due to its MC simulation accuracy, thus improving the accuracy of dose calculation and providing a valuable reference for quality assurance.

1. Introduction

The main goal of radiation therapy is to eliminate tumor cells while sparing healthy tissues. Heavy ion radiation, with its Bragg peak and limited side-scattering effects, allows precise tumor cell eradication while minimizing harm to normal tissues. In Heavy Ion Therapy (HIT), ensuring the entire target volume receiving the prescribed dose is crucial for a uniformly effective outcome. However, inter-patient variations and uncertainties in HIT present challenges in accurately calculating dose distribution [1–4].

Two main approaches for dose calculation in HIT exist. The first, the treatment planning system (TPS) analytical dose algorithm (TPSDose), is widely used in clinical practice for its rapid calculations but has compromised accuracy in heterogeneous tissues [5–7]. The second,

* Corresponding authors at: Institute of Modern Physics, Chinese Academy of Sciences, Lanzhou 730000, China.
E-mail addresses: hepengbo@impcas.ac.cn (P. He), liqiang@impcas.ac.cn (Q. Li).
[1] These authors contributed equally to this work and should be considered co-first authors.

https://doi.org/10.1016/j.phro.2025.100735
Received 25 January 2024; Received in revised form 13 February 2025; Accepted 20 February 2025
Available online 5 March 2025

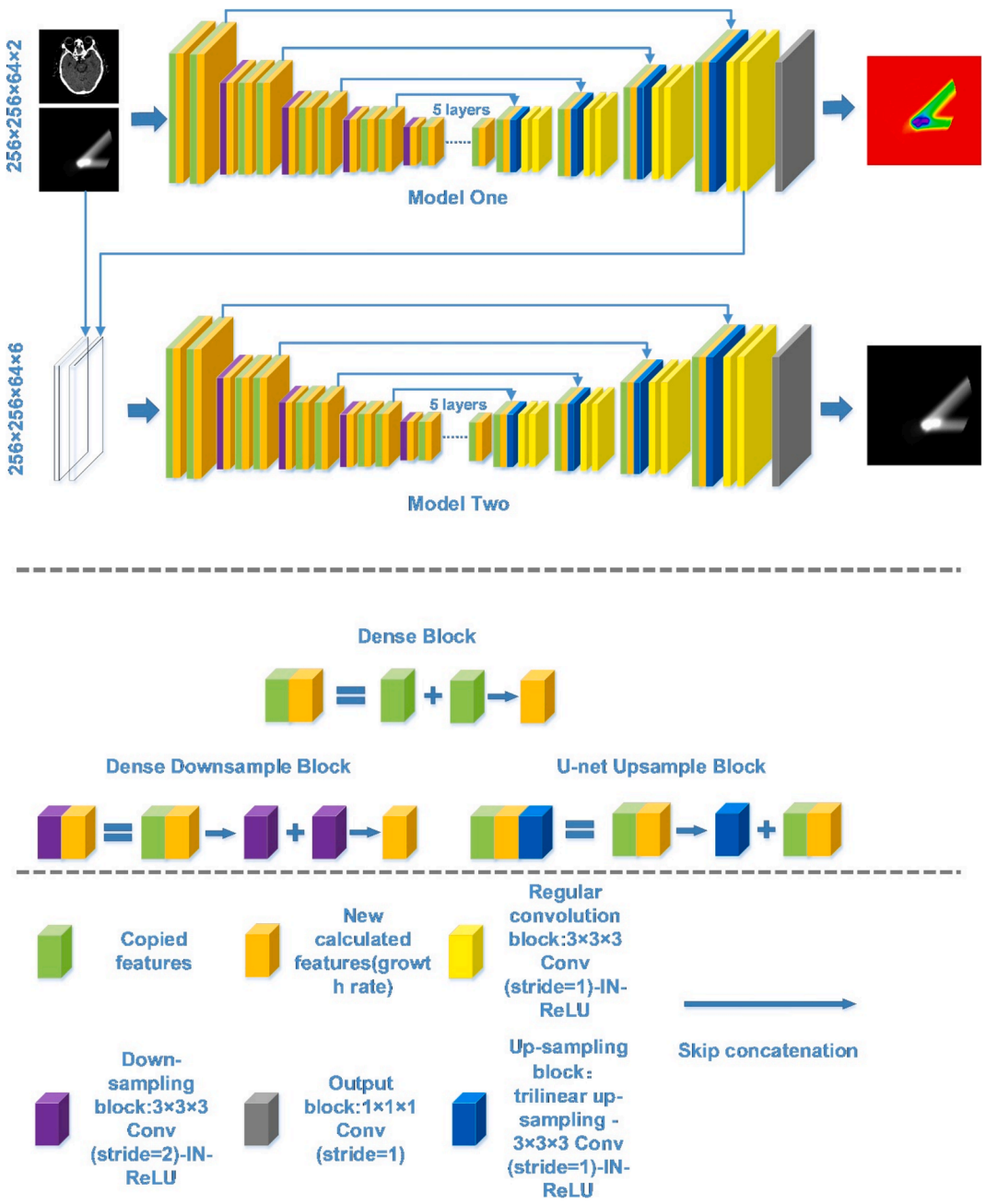

Fig. 1. Diagram of the CHD U-Net Network model.

Monte Carlo simulation dose (MCDose), is the gold standard for dose calculation. However, its extensive computational time limits its widespread clinical use [8–10].

Researchers have focused on speeding up proton and heavy-ion MC dose calculations by exploring the use of central processing units (CPUs) and graphics processing units (GPUs) [11–16]. Despite these efforts, the fastest heavy-ion MC engines still require several minutes for calculations, often limited to specific tasks or scenarios. Some studies limited physical processes in MC simulations to expedite computation time, though at the expense of accuracy [12–16].

In recent years, deep learning (DL), especially convolutional neural Networks (CNNs), has gained traction in medical research. CNNs are used to analyze patient anatomy and predict TPSDose, showing efficiency in individualized predictions [17–20]. DL-based approaches for MCDose prediction offer computational efficiency compared to traditional MC simulations [21–23]. Despite speed advantages, DL's accuracy in predicting dose distribution varies due to data quantity, nature, and inter-patient variations. Efforts have focused on improving DL models and input data to improve accuracy [23–28].

No specific DL MCDose prediction models exist for HIT. Existing proton dose prediction DL models enhance accuracy in body dose distribution but face challenges in predicting specific dose distribution within the PTV crucial for treatment outcomes [17,19,21–23,25–27].

Our study aimed to enhance the accuracy of MC dose prediction for HIT, focusing on the PTV region and overall dose distribution, with the goal of matching the precision of MCDose calculations and potentially replacing the TPSDose currently used in clinical practice. To achieve this, we developed a novel deep learning model to predict MCDose directly from patient CT images and HIT-based TPSDose.

2. Materials and methods

2.1. Patient data and dose calculation

The study involved 67 patients with head-and-neck tumors and 30 patients with thorax-and-abdomen tumors. Data were collected from three centers: Gansu Provincial Cancer Hospital, Gansu Provincial Hospital, and Wuwei Heavy Ion Hospital, following an institutional review board-approved retrospective data collection protocol under the Academic Committee of the Institute of Modern Physics, Chinese Academy of Sciences. (Section S1, Supplementary Material).

The CT images, delineations, and treatment plans were performed at each institute individually. For all patients, TPSDose was calculated utilizing the spot-scanning beamline database of the Heavy Ion Medical Machine situated in Lanzhou, China. All patients received either Single Field Uniform Dose (SFUD) or Intensity-Modulated Proton Therapy (IMPT) plans by an experienced physicist based on the patient's actual condition, and reviewed by an experienced physician, exclusively for research purposes and not implemented on actual patients. All treatment plans were optimized based on the physical dose, adhering to a total prescribed dose of 30 Gy delivered in 30 fractions. (Table S1, Supplementary Material).

MCDose was calculated using the GATE v9.1/Geant4 v10.7.4 simulation platform, incorporating the same spot positions, energies, and weights as calculated in matRad [29]. The GATE simulation employed the same conversion factors for Hounsfield values and stopping power. The production threshold settings for gamma, electron, and positron particles were established at 1 m, 1 mm, and 0.1 mm, respectively. These settings were chosen with consideration of simulation time and calculation accuracy, aiming to strike a balance between the two factors. The physics list, QGSP_BERT_HP_EMY, was used in the simulations, containing the command of hadronic physics processes and EM processes for carbon ions (GateContrib/dosimetry/Radiotherapy/example1) [30]. A total of 10^8 particles were tracked in the simulation, which was executed using 80 parallel calculations on two Intel® Xeon® Gold 6148 CPUs operating at a frequency of 2.40 GHz. Two dose distributions were calculated for each patient to ensure MC noise values around 0.5 % and always below 1 %.

2.2. Data pre-processing

Two sets of Cohorts were conducted: one exclusively with head-and-neck data (Cohort 1) and the other combining head-and-neck and thorax-and-abdomen data (Cohort 2). The first Cohort assessed DL model accuracy with a limited dataset, while the second evaluated its performance across different tumor types for generalization. Each cohort had subsets (training, validation, test), Cohort 1: 41:13:13; Cohort 2 (head-and-neck): 41:13:13, (horax-and-abdomen): 18:6:6. (Table S2, Supplementary Material). Input data included 3D CT images and TPSDose. To optimize GPU resources, CT images and TPSDose resolutions were downsampled to 256x256x64. The CTs, structure sets, and TPSDose DICOM files for each patient were extracted and converted to 3D matrices. To ensure consistency, normalization operations were applied to CT images, cropping them to −1024 to 3500 HU, dividing by 3500 HU, and scaling to [0,1]. Dose values were normalized within [0, 1].

The proposed model used online data augmentation techniques to enhance prediction accuracy. These included random flips (70 % probability) around the x and z axes, random rotations (30 % probability) around the z-axis, and random panning (80 % probability) within the image plane. These augmentations help the model handle diverse orientations, positions, and spatial changes, improving accuracy and generalization. Mirroring input data along different axes increases robustness by generating predictions from multiple perspectives without changing spatial location or orientation.

2.3. Network model structure and model training

The first phase of the CHD U-Net Network used input CT images and TPSDose to predict the output MCDose [31]. The second-stage Network refined predictions using the first stage's input and output to optimize the model parameters. This reduced the overall training burden, allowing the second stage to focus on refining predictions for improved accuracy in the final results [32,33].

The CHD U-Net Network was implemented in PyTorch and trained on an NVIDIA A6000 GPU for efficient image processing. The mean absolute error (MAE) served as the loss function, comparing predicted dose with ground truth dose (MCDose) (Section S2, Supplementary Material).

Fig. 1 illustrates the specific structure of the Network model, which is a two-stage modified HD U-Net. The original architecture was modified to include a dense convolution module, a dense downsampling module, and a standard convolution module within the U-Net upsampling process. Dense Convolution Module: Consists of a $3 \times 3 \times 3$ convolution kernel with a stride of 1, followed by InstanceNorm and ReLU activation. The output is concatenated with the previous feature map. Dense Downsampling Module: Includes a $3 \times 3 \times 3$ convolution kernel with a stride of 1, InstanceNorm, ReLU activation, and a $2 \times 2 \times 2$ max-pooling operation with a stride of 2. The outputs of both operations are concatenated. Standard Convolution Module: Contains a $3 \times 3 \times 3$ convolution kernel with a stride of 1, followed by InstanceNorm and ReLU activation. The Network performs five downsampling and five upsampling operations. During downsampling, each layer includes two dense convolutional modules. In upsampling, two standard convolutional modules replace the dense ones for efficiency. Skip connections link each upsampled feature map with its corresponding downsampled feature map. The final layer applies a $1 \times 1 \times 1$ convolution kernel to produce the final 3D dose prediction.

The model was optimized using the deep supervision technique, which involves optimizing intermediate layers in addition to the final output layer [34]. The MSRA method initialized Network model weights randomly. The Adam optimizer [36] with an initial learning rate of 0.003 and a weight decay of 0.0001 were used. The learning rate was reduced cyclically using the cosine annealing method to improve convergence and avoid local optima. In Cohort 1, the model was trained for 16,400 iterations, and in Cohort 2, training continued for 35,400 iterations. The stopping point was determined based on model convergence and performance. These settings and techniques aimed to ensure effective training and optimization of the CHD U-Net model for accurate dose prediction in HIT.

2.4. Alternative model structures and their parameters

In our Cohorts, we compared our model with two state-of-the-art 3D dose prediction models: (1) C3D [27], a cascaded 3D U-Net architecture that achieved first place in the OPNEKBP competition. This model was specifically designed for dose prediction tasks and demonstrated high performance in previous competitions [35] (Section S3, Fig. S1, Supplementary material). (2) HD U-Net [24], a DL model with excellent performance in proton dose prediction (Section S4, Fig. S2, Supplementary material). Throughout training, no underfitting or overfitting phenomena were observed, and all three DL models eventually converged.

2.5. Network model performance evaluation

In our network performance evaluation we used the root mean square error (RMSE) (Section S5, Supplementary Material) as the distance metric to compare the predicted dose of the DL model, TPSDose, with the "gold standard" MCDose. One-way ANOVA was used to tested if predicted outcomes significantly differed from MCDose, reflecting the stability of DL model predictions. Dose-volume parameters evaluation

Table 1
Mean RMSE values and 3D gamma passing rates (1 %/1 mm and 3 %/3 mm) for all patients Tested in Cohort 1 (mean ± standard deviation).

Cohort 1					
Model	RMSE (p-value)	PTV Gamma Passing Rate (%)		Whole body Gamma Passing Rate (%)	
		1 %/1 mm	3 %/3 mm	1 %/1 mm	3 %/3 mm
CHD U-Net	0.06 ± 0.01 (0.09)	70.0 ± 7.5	99.1 ± 1.3	86.0 ± 3.4	99.3 ± 0.3
C3D	0.13 ± 0.12 (0.26)	60.6 ± 8.7	97.3 ± 3.5	79.3 ± 4.7	97.9 ± 1.5
HD U-Net	0.07 ± 0.016 (0.10)	40.5 ± 14.4	85.4 ± 14.8	76.0 ± 7.3	97.1 ± 2.3
TPSDose	0.28 ± 0.38	54.9 ± 6.9	94.3 ± 1.8	75.7 ± 4.4	96.0 ± 1.8

Abbreviations = PTV: planning target volume, MC: 3D dose distribution from MC simulation, TPSDose: 3D dose distribution from analytical algorithm, RMSE: root mean squared error, the p-value of the RMSE was obtained from a one-way ANOVA comparing the predicted doses of the three DL models with the TPSDose. MCDose was always the reference in the GPR comparisons. The test set included data from 13 patients. We used the Shapiro-Wilk test to assess normality, and the results indicated that the test outcomes followed a normal distribution.

included D_{95} of the PTV area and D_5 of the brainstem, left and right eyeballs, and optic nerve (Section S6, Supplementary material), dose volume histograms (DVH) and using matRad [29] performed 3D gamma analysis. 3D Gamma passing rate (GPR) in the PTV region and overall body region were compared to MCDose using voxel evaluation criteria of 1 %/1 mm and 3 %/3 mm and above the 3 % threshold of maximum MCDose. GPR aimed to comprehensively evaluate the accuracy and reliability of DL model dose prediction. We compared TPSDose and the predicted doses of three DL models with the GPR of MCDose in Cohort 1 and Cohort 2, using MCDose as the reference for all GPR comparisons.

3. Results

3.1. Technical evaluation results

In Cohort 1, the RMSE values for CHD U-Net, C3D, and HD U-Net were 0.06, 0.13, and 0.07, respectively, with p-values of 0.09, 0.26, and 0.10. In Cohort 2, for head-and-neck patients, the RMSE values were 0.07 for both CHD U-Net and C3D, and 0.08 for HD U-Net, with all p-values less than 0.01. For thorax-and-abdomen patients, the RMSE values were 0.06, 0.07, and 0.08, with p-values of 0.22, 0.35, and 0.84, respectively.

3.2. Dosimetric evaluation results

The mean 3D GPR values (1 %/1 mm) for CHD U-Net, C3D, HD U-Net, and TPSDose in the PTV region were 70 %, 61 %, 41 %, and 55 %, respectively, while in the overall body region, they were 86 %, 79 %, 76 %, and 76 %, respectively. Under the stringent criteria of 1 %/1 mm, the CHD U-Net model showed a 15 % better agreement with MCDose compared to TPSDose, while the C3D model had a 6 % improvement. However, the HD U-Net model had a lower mean GPR of 41 % in the PTV region. For overall body 3D GPR, the CHD U-Net showed a 10 % improvement, the C3D model improved by 3 %, and the HD U-Net had a slightly higher GPR compared to TPSDose. Under the 3 %/3 mm standard, the GPR of CHD U-Net reached 99 % in both the PTV region and the whole body (Table 1).

Fig. 2 presents box plots for the PTV region and overall body 3D GPR in Cohort 1. In the PTV region (Fig. 2(a)), the mean GPR for the CHD U-Net was 70 % ($p < 0.001$), for C3D it was 61 % ($p = 0.07$), and for the HD U-Net it was 41 % ($p < 0.01$). In Fig. 2(b), representing the overall body, the CHD U-Net had the highest GPR and the smallest p-value. The GPR of the CHD U-Net is significantly different from that of TPSDose.

Table 2 presents mean 3D GPR for head-and-neck and thorax-and-abdomen patients in Cohort 2. In the PTV region, the CHD U-Net with MCDose had a mean GPR that exceeded TPSDose with MCDose by 14 %. However, the C3D and HD U-Net models showed inconsistent performance, with lower mean GPR values than TPSDose. In the entire head-and-neck region, the CHD U-Net with MCDose had a mean GPR

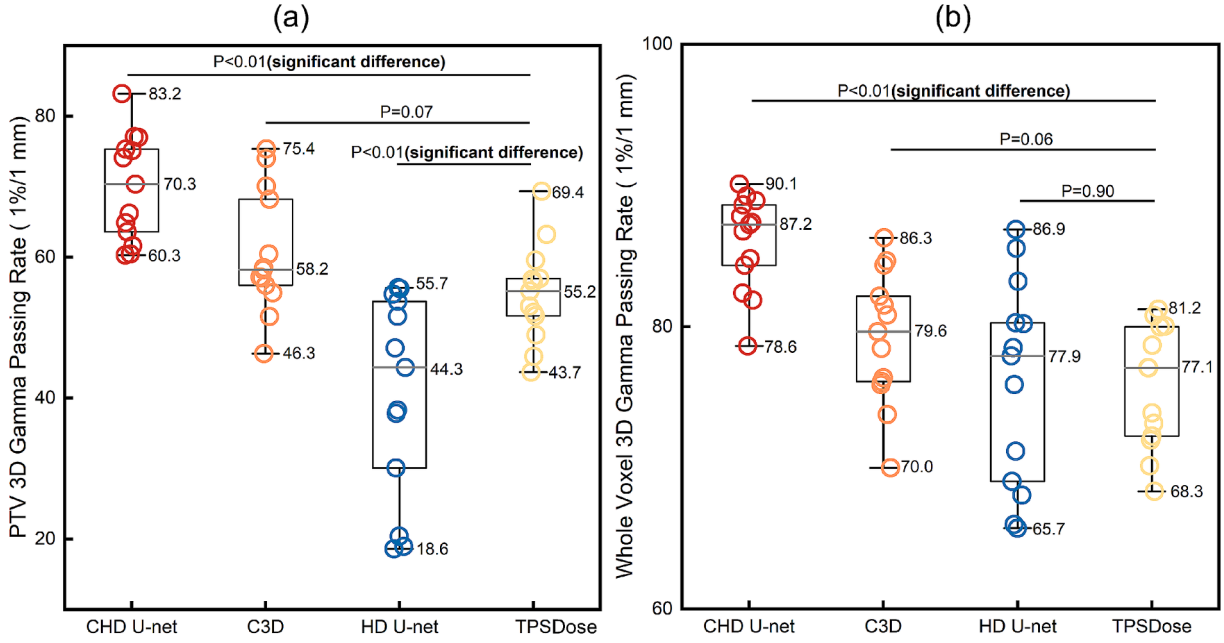

Fig. 2. 3D gamma passing rate BOX plot of all the three DL models and TPS in Cohort 1: (a) 3D gamma passing rate of PTV region (1 %/1 mm), (b) 3D gamma passing rate of whole body region (1 %/1 mm). The center line in the box was the median, the box boundaries are Q1 and Q3, and the whisker line was the range. MCDose was always the reference in GPR comparison.

Table 2
Mean RMSE values and 3D gamma passing rates (1 %/1 mm and 3 %/3 mm) for all patients Tested in Cohort 2 (mean ± standard deviation).

			Cohort 2			
Head & Neck			CHD U-Net	C3D	HD U-Net	TPSDose
	RMSE (P value)		0.07 ± 0.01 ($p < 0.01$)	0.07 ± 0.01 ($p < 0.01$)	0.08 ± 0.02 ($p < 0.01$)	0.10 ± 0.02
	PTV Gamma Passing Rate (%)	1 %/1mm	67.3 ± 5.8	50.2 ± 14.0	42.8 ± 19.0	53.2 ± 7.0
		3 %/3mm	99.2 ± 0.5	92.6 ± 11.0	85.0 ± 20.6	96.9 ± 1.0
	Whole body Gamma Passing Rate (%)	1 %/1mm	82.5 ± 4.0	75.5 ± 6.8	74.3 ± 8.8	74.7 ± 5.0
		3 %/3mm	99.5 ± 0.2	97.6 ± 1.7	97.1 ± 3.2	7.9 ± 1.0
Thorax & Abdomen	RMSE (P value)		0.06 ± 0.01 (0.22)	0.07 ± 0.02 (0.35)	0.08 ± 0.01 (0.84)	0.08 ± 0.01
	PTV Gamma Passing Rate (%)	1 %/1mm	42.4 ± 9.2	23.9 ± 11.2	13.9 ± 14.7	32.2 ± 14.3
		3 %/3mm	95.9 ± 0.9	71.3 ± 20.5	50.8 ± 33.8	89.8 ± 8.3
	Whole body Gamma Passing Rate (%)	1 %/1mm	72.3 ± 3.8	65.4 ± 4.6	56.5 ± 5.5	62.7 ± 3.9
		3 %/3mm	97.7 ± 0.9	95.0 ± 2.5	90.4 ± 5.0	97.0 ± 0.7

Abbreviations = PTV: planning target volume, MC: 3D dose distribution from MC simulation, TPSDose: 3D dose distribution from parsing algorithm, RMSE: root mean squared error, the p-value of the RMSE was obtained from a one-way ANOVA comparing the predicted doses of the three DL models with the TPSDose. MCDose was always the reference in the GPR comparisons. The test set included data from 19 patients, comprising 13 head-and-neck cases and 6 thorax-and-abdomen cases. We used the Shapiro-Wilk test to assess normality, and the results indicated that the test outcomes followed a normal distribution.

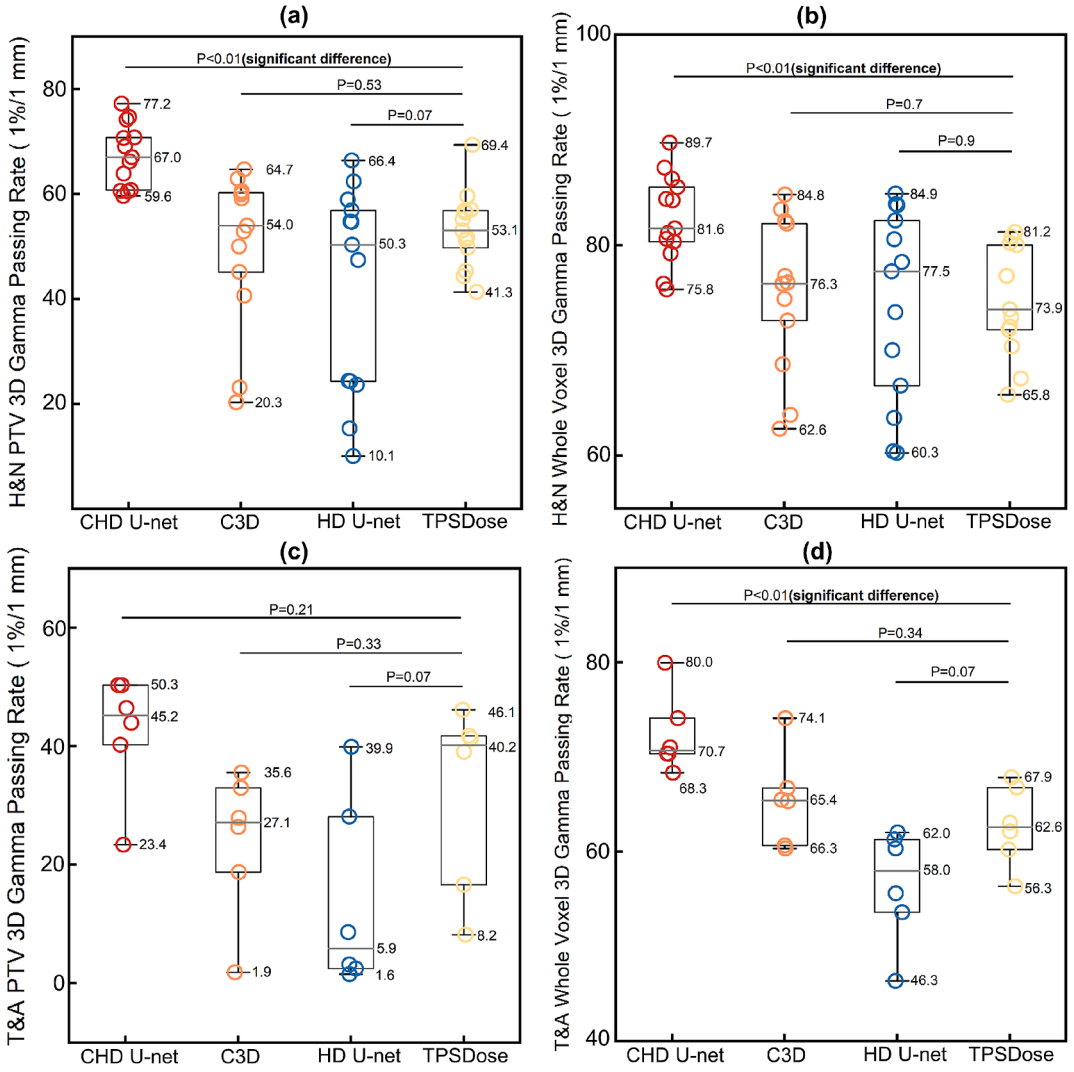

Fig. 3. 3D gamma passing rate BOX plot of the three DL models and TPS of the head-and-neck patients and thorax-and-abdomen patients in Cohort 2: (a) PTV area 3D gamma passing rate for the head-and-neck patients (1 %/1 mm), (b) Whole body area 3D gamma passing rate for the head-and-neck patients (1 %/1 mm), (c) PTV area 3D gamma passing rate for the thorax-and-abdomen patients (1 %/1 mm), (d) Whole body 3D gamma passing rate for the thorax-and-abdomen patients (1 %/1 mm). The center line in the box was the median, the box boundaries are Q1 and Q3, and the whisker line was the range. MCDose was always the reference in GPR comparison.

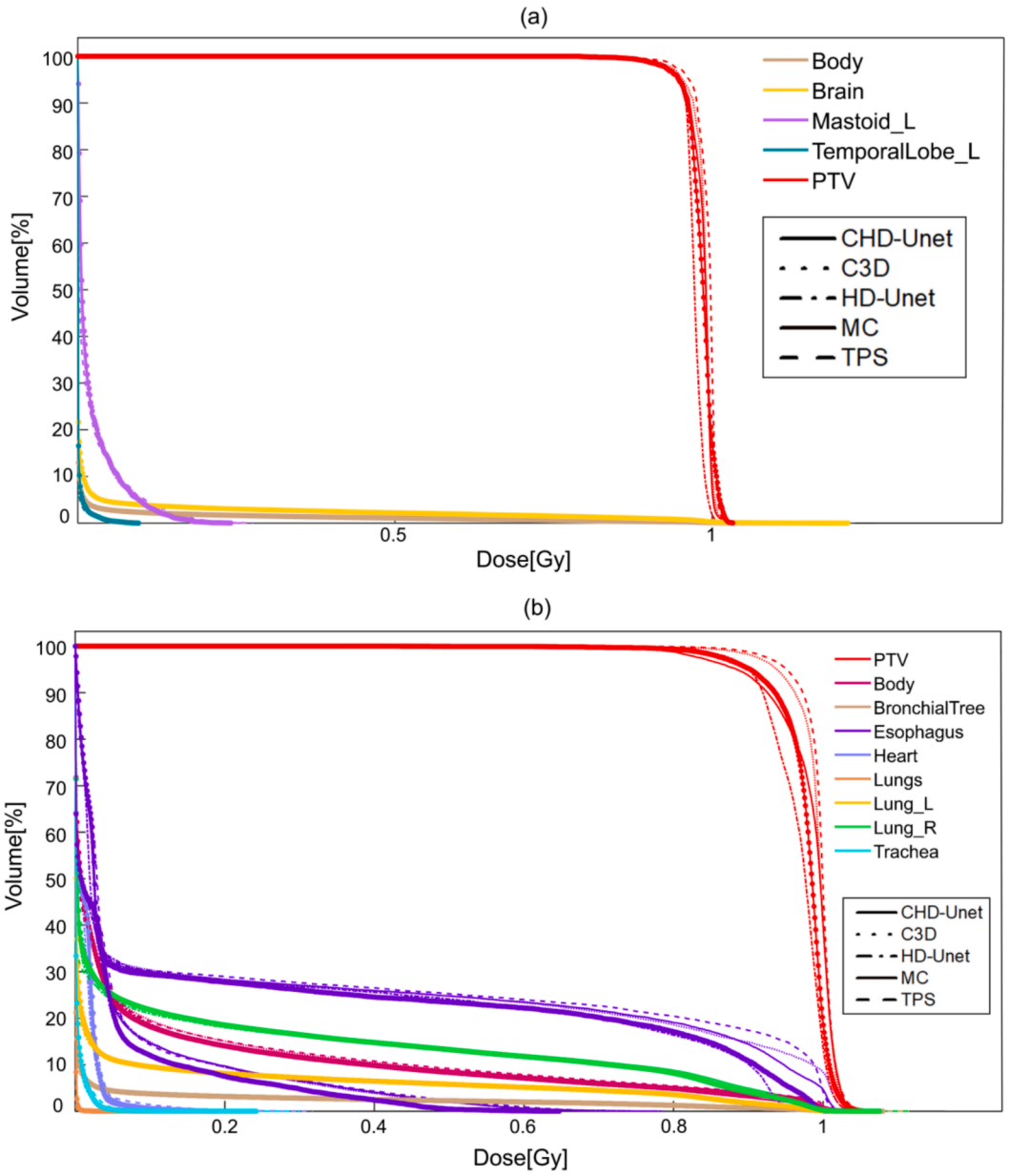

Fig. 4. DVH plots of two Cohort patients: (a) DVH plots of a head-and-neck tumor patient in Cohort 1, (b) DVH plots of a thorax-and-abdomen tumor patient in Cohort 2.

exceeded TPSDose with MCDose by 9 %. The C3D and HD U-Net models had similar mean GPR values to TPSDose, with HD U-Net slightly outperforming C3D. Under the 3 %/3 mm standard, CHD U-Net's GPR reached 99 % in both the PTV area and the entire body.

Fig. 3(a) and 3(b) show box plots for the PTV region and whole body 3D GPR in Cohort 2 for head-and-neck patients. Both plots indicate that the CHD U-Net had the highest GPR and smallest p-value, with 67 % ($p < 0.001$) in the PTV region and 83 % ($p < 0.001$) for the whole body, significantly different from TPSDose. Fig. 3(c) and 3(d) illustrate that for thorax-and-abdomen patients, the CHD U-Net's average GPRs were 42 % and 72 %, respectively. In comparison, C3D had values of 24 % and 65 %, while HD U-Net had 14 % and 57 %.

In Cohort 1, the average relative differences in D_{95} between the CHD U-Net, C3D, HD U-Net, TPSDose and MCDose were 0.2 %, 0.5 %, 1.2 % and 0.6 %. The average relative differences for D_5 were 1.1 %, 1.3 %, 1.9 %, and 3 %. In Cohort 2, for the head-and-neck tumors, the average relative differences in D_{95} were 1.3 %, 1.6 %, 1.9 % and 1.9 %, while for D_5, they were 1.1 %, 2.1 %, 2.8 %, and 2.4 %. For thorax-and-abdomen tumors, the mean relative differences in D_{95} were 1.3 %, 4 %, 4 %, and 2.7 %, and for D_5, they were 3.1 %, 5.2 %, 5.2 %, and 2.9 % (Table S1, Supplementary Material). Fig. 4 shows the DVH plots of the two groups of Cohort patients.

We compared the predicted dose by the three DL models with the MCDose. Fig. S3 provides a plot of the dose differences for a head-and-neck tumor and a thorax-and-abdomen tumor in Cohorts 1 and 2. (Supplementary Material). For head-and-neck tumors, the average MC simulation time was 19.3 ± 12.3 (mean ± standard deviation) hours, while the average prediction times for the CHD U-Net, C3D, and HD U-Net models were 2.1 ± 0.3 s, 2.2 ± 0.5 s, and 2.1 ± 0.2 s, respectively. And for the thorax-and-abdomen tumors, it was 38.4 ± 26.2 h. The prediction times for the three DL models were 2.8 ± 0.5 s, 2.9 ± 0.6 s, and 2.9 ± 0.3 s, respectively. These findings highligh the significantly higher computational cost in terms of MC simulation time for heavy ions compared to photons and protons.

4. Discussion

In this study, we introduced the CHD U-Net, a novel DL model for predicting MCDose distribution in HIT. It accurately predicted dose distribution for the whole body and improved the PTV region, achieving the highest GPR (3 %/3 mm) of 99 % in both areas, with a prediction time of only two seconds.

In clinical HIT, using a 1 %/1 mm GPR is stringent, as the standard is 3 %/3 mm, where a GPR over 95 % indicates a deliverable radiotherapy treatment plan. In Cohorts 1 and 2, the CHD U-Net achieved 99 % GPR for head-and-neck cancer patients in both PTV and whole body areas compared to other models [17,19,21–23,25–27,41].

However, for the thorax-and-abdomen patients in Cohort 2, the GPR did not reach 99 %. This might be due to the smaller data set and more complex incident angles (0°-90°, 180°-360°). These results indicated that both the quantity and consistency of the data (in terms of position and angle) could affect the model's predictive accuracy.

Cohort 1 assessed the DL model's accuracy with a single tumor type, while Cohort 2 evaluated its generalization across various patient data and tumor types. Our findings show potential for adaptability. Future studies will enhance performance by increasing tumor types and integrating data from other locations, expanding clinical applications.

In the DVH index and image evaluations, the CHD U-Net had the smallest relative difference compared to MCDose, excelling in dose prediction for organs at risk (OARs). The D_5 evaluation showed that in both cohorts, the CHD U-Net's relative difference was the smallest among the other DL models and TPSDose, with the lowest at about 1 %. Compared to the previous photon or proton dose prediction studies [17–20], the CHD U-Net significantly improved OAR prediction accuracy. It also had higher accuracy than the widely used clinical analytical algorithm [5–7], suggesting that it could replace TPSDose for precision HIT. Therefore, our method significantly impoved the MC dose prediction accuracy without major modifications to clinical protocols [39,40]. The CHD U-Net's average prediction time for MCDose was about 2 s, much faster than the half-minute calculation time of Zhang et al.'s heavy ion pencil-beam model [10]. Total MCDose prediction time was also about half a minute, shorter than that with MC denoising [37,38], allowing for online adaptation of MC calculation and improving quality with geometric changes. Additionally, the CHD U-Net can predict each patient's MCDose end-to-end, offering clinicians valuable insights to determine the necessity of conventional MC simulation QA, thereby enhancing clinical workflow efficiency. Integrating this model into the HIT process can enhance TPSDose accuracy to match MC simulations and replace the existing system. It can be added as a plug-in to improve heavy-ion dose calculations and assist in clinical QA.

This study had some limitations, such as not integrating relative biological effectiveness (RBE) into dose optimization [41]. In order to keep the data as consistent as possible, the optimization relied only on the physical absorbed dose, the fixed prescription dose (30 Gy/30 fractions) did not consider the dose constraint of OARs. Additionally, we did not verify the model's accuracy when the prescription dose changed. We also had not yet focused on improving the model's efficiency, which could be enhanced through methods like model compression or structural simplification.

In conclusion, our study introduced a novel DL model – CHD U-Net for predicting MCDose in HIT, demonstrating higher accuracy in dose prediction for the PTV region and the overall body compared to existing models. Since our dataset included two types of tumors, the model demonstrated a certain level of generalization ability. Notably, our deep learning model could rapidly predict MCDose in just 2 s.

Funding

This work was jointly supported by the National Key Research and Development Program of China (Grant No. 2022YFC2401503), the Special Project of Science and Technology Cooperation between Hubei Province and Chinese Academy of Sciences (Grant No. 42000021817T300000050), the Key Research and Development Program of Gansu Province (Grant No. 23YFFA0010), the Natural Science Foundation of Gansu Province (Grant No. 22JR5RA125) and the National Natural Science Foundation of China [Grant No. 12005271], the National Natural Science Foundation of China [Grant No. 12375354].

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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.phro.2025.100735.

References

[1] Kraft G. Tumor therapy with heavy charged particles. Prog Part Nucl Phys 2000;45:S473–544. https://doi.org/10.1016/S0146-6410(00)00112-5.
[2] Tsujii H, Kamada T, Shirai T, Noda K, Tsuji H, Karasawa K, et al. Carbon-ion radiotherapy: principles, practices, and treatment. Planning 2014. https://doi.org/10.1007/978-4-431-54561-8.
[3] Jäkel O, Krämer M, Schulz-Ertner D, Heeg P, Karger CP, Didinger B, et al. Treatment planning for carbon ion radiotherapy in Germany: review of clinical trials and treatment planning studies. Radiother Oncol 2004;73:S86–91. https://doi.org/10.1016/S0167-8140(04)80022-7.
[4] Shirai K, Kawashima M, Saitoh J, Wakatsuki M, Ohno T, Shiba S, et al. Clinical outcomes using carbon-ion radiotherapy and dose-volume histogram comparison between carbon-ion radiotherapy and photon therapy for T2b-4N0M0 non-small cell lung cancer—A pilot study. PLoS One 2017;12:e0175589. https://doi.org/10.1371/journal.pone.0175589.
[5] Huang S, Souris K, Li S, Kang M, Barragan Montero AM, Janssens G, et al. Validation and application of a fast Monte Carlo algorithm for assessing the clinical impact of approximations in analytical dose calculations for pencil beam scanning proton therapy. Med Phys 2018;45:5631–42. https://doi.org/10.1002/mp.13231.
[6] Teoh S, Fiorini F, George B, Vallis KA, Van den Heuvel F. Is an analytical dose engine sufficient for intensity modulated proton therapy in lung cancer? Br J Radiol 2020;93:20190583. https://doi.org/10.1259/bjr.20190583.
[7] Parodi K, Mairani A, Sommerer F. Monte Carlo-based parametrization of the lateral dose spread for clinical treatment planning of scanned proton and carbon ion beams. J Radiat Res 2013;54:i91–6. https://doi.org/10.1093/jrr/rrt051.
[8] Kase Y, Kanematsu N, Kanai T, Matsufuji N. Biological dose calculation with Monte Carlo physics simulation for heavy-ion radiotherapy. Phys Med Biol 2006;51:N467. https://doi.org/10.1088/0031-9155/51/24/n03.
[9] Nakaji T, Kanai T, Takashina M, Matsumura A, Osaki K, Yagi M, et al. Clinical dose assessment for scanned carbon-ion radiotherapy using linear energy transfer measurements and Monte Carlo simulations. Phys Med Biol 2022;67:245021. https://doi.org/10.1088/1361-6560/aca003.
[10] Zhang H, Li Q, Liu X, Ma Y, He P, Shen G, et al. Validation and testing of a novel pencil-beam model derived from Monte Carlo simulations in carbon-ion treatment planning for different scenarios. Phys Med 2022;99:1–9. https://doi.org/10.1016/j.ejmp.2022.04.018.
[11] Qin N, Pinto M, Tian Z, Dedes G, Pompos A, Jiang SB, et al. Initial development of goCMC: a GPU-oriented fast cross-platform Monte Carlo engine for carbon ion therapy. Phys Med Biol 2017;62:3682. https://doi.org/10.1088/1361-6560/aa5d43.
[12] De Simoni M, Battistoni G, De Gregorio A, De Maria P, Fischetti M, Franciosini G, et al. A data-driven fragmentation model for carbon therapy GPU-accelerated Monte Carlo dose recalculation. Front Oncol 2022;12:780784. https://doi.org/10.3389/fonc.2022.780784.
[13] Lysakovski P, Kopp B, Tessonnier T, Mein S, Ferrari A, Haberer T, et al. Development and validation of MonteRay, a fast Monte Carlo dose engine for carbon ion beam radiotherapy. Med Phys 2024;51:1433–49. https://doi.org/10.1002/mp.16754.
[14] Choi K, Mein SB, Kopp B, Magro G, Molinelli S, Ciocca M, et al. FRoG—A new calculation engine for clinical investigations with proton and carbon ion beams at CNAO. Cancers (Basel) 2018;10(11):395. https://doi.org/10.3390/cancers10110395.
[15] Souris K, Lee JA, Sterpin E. Fast multipurpose Monte Carlo simulation for proton therapy using multi- and many-core CPU architectures. Med Phys 2016;43:1700–12. https://doi.org/10.1118/1.4943377.
[16] Magro G, Fassi M, Mirandola A, Rossi E, Molinelli S, Russo S, et al. Dosimetric validation of a GPU-based dose engine for a fast in silico patient-specific quality assurance program in light ion beam therapy. Med Phys 2022;49:7802–14. https://doi.org/10.1002/mp.16002.

[17] Li H, Peng X, Zeng J, Jiao J, Nie D, Zu C, et al. Explainable attention guided adversarial deep network for 3D radiotherapy dose distribution prediction. Knowl-Based Syst 2022;241:108324. https://doi.org/10.1016/j.knosys.2022.108324.
[18] 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;46:370–81. https://doi.org/10.1002/mp.13271.
[19] Koike Y, Takegawa H, Anetai Y, Ohira S, Nakamura S, Tanigawa N. Patient-specific three-dimensional dose distribution prediction via deep learning for prostate cancer therapy: Improvement with the structure loss. Phys Med 2023;107:102544. https://doi.org/10.1016/j.ejmp.2023.102544.
[20] Mashayekhi M, Tapia IR, Balagopal A, Zhong X, Barkousaraie AS, McBeth R, et al. Site-agnostic 3D dose distribution prediction with deep learning neural networks. Med Phys 2022;49:1391–406. https://doi.org/10.1002/mp.15461.
[21] Zhang B, Liu X, Chen L, Zhu J. Convolution neural network toward Monte Carlo photon dose calculation in radiation therapy. Med Phys 2022;49:1248–61. https://doi.org/10.1002/mp.15408.
[22] Lee MS, Hwang D, Kim JH, Lee JS. Deep-dose: a voxel dose estimation method using deep convolutional neural network for personalized internal dosimetry. Sci Rep 2019;9:10308. https://doi.org/10.1038/s41598-019-46620-y.
[23] Zhang G, Chen X, Dai J, Men K. A plan verification platform for online adaptive proton therapy using deep learning-based Monte Carlo denoising. Phys Med 2022; 103:18–25. https://doi.org/10.1016/j.ejmp.2022.09.018.
[24] Wu C, Nguyen D, Yang Y, Montero AB, Schuemann J, Shang H, et al. Improving proton dose calculation accuracy by using deep learning. Mach Learn Sci Technol 2021;2:015017. https://doi.org/10.1088/2632-2153/abb6d5.
[25] Mentzel F, Kröninger K, Lerch M, Nackenhorst O, Rosenfeld A, Tsoi AC, et al. Small beams, fast predictions: a comparison of machine learning dose prediction models for proton minibeam therapy. Med Phys 2022;49:7791–801. https://doi.org/10.1002/mp.16066.
[26] Wang W, Chang Y, Liu Y, Liang Z, Liao Y, Qin B, et al. Feasibility study of fast intensity-modulated proton therapy dose prediction method using deep neural networks for prostate cancer. Med Phys 2022;49:5451–63. https://doi.org/10.1002/mp.15702.
[27] Liu S, Zhang J, Li T, Yan H, Liu J. A cascade 3D U-Net for dose prediction in radiotherapy. Med Phys 2021;48:5574–82. https://doi.org/10.1002/mp.15034.
[28] Pastor-Serrano O, Perkó Z. Millisecond speed deep learning based proton dose calculation with Monte Carlo accuracy. Phys Med Biol 2022;67:105006. https://doi.org/10.5281/zenodo.4724125.
[29] Wieser HP, Cisternas E, Wahl N, Ulrich S, Stadler A, Mescher H, et al. Development of the open-source dose calculation and optimization toolkit matRad. Med Phys 2017;44:2556–68. https://doi.org/10.1002/mp.12251.
[30] GateContrib. Dosimetry, Radiotherapy Example. GitHub.https://github.com/0penGATE/GateContrib/tree/master/dosimetry,Radiotherapy/example1. Accessed March 29, 2022.
[31] 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;64:065020. https://doi.org/10.1088/1361-6560/ab039b.
[32] Xue J, He K, Nie D, Adeli E, Shi Z, Lee SW, et al. Cascaded multitask 3-D fully convolutional networks for pancreas segmentation. IEEE Trans Cybern 2019;51(4): 2153–65. https://doi.org/10.1109/tcyb.2019.2955178.
[33] Roth HR, Oda H, Zhou X, Shimizu N, Yang Y, Hayashi Y, et al. An application of cascaded 3D fully convolutional networks for medical image segmentation. Comput Med Imaging Graph 2018;66:90–9. https://doi.org/10.1016/j.compmedimag.2018.03.001.
[34] Babier A, Zhang B, Mahmood R, Moore KL, Purdie TG, McNiven AL, et al. OpenKBP: the open-access knowledge-based planning grand challenge and dataset. Med Phys 2021;48:5549–61. https://doi.org/10.1002/mp.14845.
[35] He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification. In: Proceedings of the IEEE International Conference on Computer Vision; 2015. p. 1026–34.
[36] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980, 2014.
[37] Zhang X, Zhang H, Wang J, Ma Y, Liu X, Dai Z, et al. Deep learning-based fast denoising of Monte Carlo dose calculation in carbon ion radiotherapy. Med Phys 2023;50:7314–23. https://doi.org/10.1002/mp.16719.
[38] Botas P, Kim J, Winey B, Paganetti H. Online adaptation approaches for intensity modulated proton therapy for head and neck patients based on cone beam CTs and Monte Carlo simulations. Phys Med Biol 2018;64(1):015004. https://doi.org/10.1088/1361-6560/aaf30b.
[39] Bazani A, Brunner J, Russo S, Carlino A, Colomar DS, Andersson WI, et al. Effects of nuclear interaction corrections and trichrome fragment spectra modelling on dose and linear energy transfer distributions in carbon ion radiotherapy. Phys Imaging Radiat Oncol 2024;29:100553. https://doi.org/10.1016/j.phro.2024.100553.
[40] Witte M, Sonke J-J. A deep learning based dynamic arc radiotherapy photon dose engine trained on Monte Carlo dose distributions. Phys Imaging Radiat Oncol2 2024;30:100575. https://doi.org/10.1016/j.phro.2024.100575.
[41] Zhang L, Wang W, Li P, Zhang Q, Han R. A deep learning model for predicting the modified micro-dosimetric kinetic model-based dose and the dose-averaged linear energy transfer for prostate cancer in carbon ion therapy. Phys Imaging Radiat Oncol 2024;32:100671. https://doi.org/10.1016/j.phro.2024.100671.