2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

Deep Learning-Based Prediction of Monte Carlo Dose Distributions From Pencil Beam Algorithms for Carbon Ion Radiotherapy

Yongdo Yun[1]; Gahee Son[1,2]; Wonhyeong Lee[1,2]; Dong Min Jung[1]; Hyungjong Yoo[1]; Min Cheol Han[1*]; Jin Sung Kim[1,3*]

[1]Department of Radiation Oncology, Yonsei University College of Medicine, Seoul, Republic of Korea
[2]Department of Nuclear Engineering, Hanyang University, Seoul, Republic of Korea
[3]Oncosoft Inc., Seoul, Republic of Korea

ABSTRACT

Purpose: This study aims to develop a deep learning (DL) model capable of predicting Monte Carlo (MC)-level carbon-ion dose distributions from Pencil Beam Algorithm (PBA) calculations. The objective is to achieve the high accuracy of MC simulations while maintaining the computational efficiency of analytical models for carbon-ion radiotherapy (CIRT).

Methods: A DL model utilizing a 3D Hierarchically Densely Connected U-Net architecture was trained and validated using data from 30 patients, including prostate, lung, and liver cases. The model inputs were 3D dose distributions calculated by a commercial treatment planning system, and the corresponding ground truth distributions were generated using the TOPAS MC code. Evaluation metrics included the gamma passing rate (GPR) with 2 mm/2% criteria and the structural similarity index measure (SSIM)*. The performance of the DL-predicted doses was assessed by comparing them against the MC ground truth, specifically focusing on the accuracy gains over the original PBA calculations.

Results: The DL-predicted doses achieved a mean GPR of 97.21% against the MC ground truth, outperforming the original PBA calculations (GPR: 90.72%). The results show that the developed model effectively reduced dose discrepancies present in the PBA calculations, particularly in the entrance plateau and fragmentation tail regions. The computational time was reduced from 20.4 hours required for full MC simulations to less than 1 second using the DL model.

Conclusions: The developed DL model successfully bridges the gap between the computational speed of PBA and the accuracy of MC simulations. Future studies will focus on extending validation to larger cohorts and additional complex anatomical sites, such as head-and-neck and pancreatic cancers.

*SSIM, listed in the submitted abstract, has been replaced by the range error and dose RMSE — metrics that are dosimetrically interpretable for a dose-correction task.

CONTACT

[name] Min Cheol Han
[organization] Yonsei Cancer Center
[address] 50-1, Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea
[email] mchan@yuhs.ac

[name] Jin Sung Kim
[organization] Yonsei Cancer Center
[address] 50-1, Yonsei-ro, Seodaemun-gu, Seoul, Republic of Korea
[email] jinsung@yuhs.ac

INTRODUCTION

- Carbon-ion radiotherapy (CIRT) offers distinct clinical advantages owing to its favorable physical dose distribution and its elevated relative biological effectiveness (RBE) [1]. To fully utilize these advantages clinically, accurate dose calculation is important.
- Clinical treatment planning systems (TPS) for CIRT rely on analytical pencil beam algorithms (PBA), which are computationally efficient but approximate nuclear fragmentation and lateral scattering. These approximations may introduce dose errors that are pronounced in anatomically heterogeneous regions. Monte Carlo (MC) simulation models these interactions explicitly and is regarded as the gold standard for dose calculation, but its computational cost makes it impractical for routine planning or for iterative processes [2]. Deep learning (DL) has recently emerged as a means of bridging these demands [3].
- This study presents the development and validation of a DL model that predicts MC-level carbon-ion dose distributions from PBA calculations, aiming to combine the accuracy of MC simulation with the computational efficiency of analytical algorithms.

METHODS AND MATERIALS

- The developed DL model was based on a 3D Hierarchically Densely Connected U-Net architecture [4]. The model was trained and validated using data from 30 patients including prostate, lung, and liver cases, partitioned into 24 training, 3 validation, and 3 test patients.
- The model inputs were the RBE-weighted doses from a commercial TPS and the corresponding CT scans. The model was trained to reproduce the ground truth generated with the TOPAS MC code [5]. For RBE calculations, a modified microdosimetric kinetic model was applied to both the PBA input and the MC ground truth.
- Model performance was quantified against the MC ground truth using three metrics: the gamma passing rate (GPR) with global 2 mm/2% criteria and a 10% dose threshold; the range prediction error, defined as the absolute difference in the distal 80% dose fall-off depth; and the voxel-wise root-mean-square error (RMSE) of the RBE-weighted dose. The DL-predicted doses and the original PBA inputs were each compared to the MC reference under identical metrics.

Metric	Method	Prostate	Lung	Liver
Range error (mm)	PBA	1.89	0.91	1.26
	DL	**0.54**	**0.89**	**0.36**
Dose RMSE (Gy)	PBA	0.17	0.10	0.09
	DL	**0.06**	**0.06**	**0.04**
Gamma passing rate (%)	PBA	93.01	86.88	92.27
	DL	**97.28**	**95.13**	**99.21**

Table 1. Quantitative comparison of PBA and DL-predicted dose against the Monte Carlo ground truth for prostate, lung, and liver. PBA, pencil beam algorithm; DL, deep learning.

RESULTS

Quantitative accuracy (Table 1)

- DL correction improved all three quantitative metrics at every site relative to the PBA baseline.
- DL reduced the range error in the both prostate (1.89 → 0.54 mm) and liver (1.26 → 0.36 mm). In lung, however, the range error was essentially unchanged (0.91 → 0.89 mm).
- The voxel-wise RMSE of the RBE-weighted dose fell by 40–65% across the three sites. The reduction was largest in prostate and liver, mirroring the range-error trend.
- The gamma passing rate improved at every site. Liver achieved the highest agreement (92.27 → 99.21%), followed by prostate (93.01 → 97.28%). Lung showed the largest absolute gain (86.88 → 95.13%) yet retained the lowest GPR of the three sites.

Qualitative agreement (Figures 1–2)

- Figures 1 and 2 show representative best- and worst-performing cases.
- Best case – liver (Fig. 1): PBA underestimates dose at the distal edge and fragmentation halo (Fig. 1-(c)). DL removes these differences, including those along the lateral field edge (Fig. 1-(d)).
- Worst case – lung (Fig. 2): DL removes the distal-edge discrepancies, but the deviations remain within the low-density tissue.

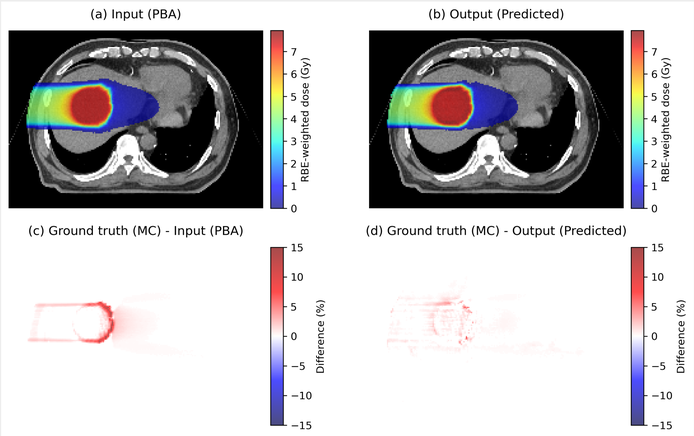

Figure 1. Best case (liver). (a) Input dose distribution calculated by the pencil beam algorithm (PBA). (b) Output dose distribution predicted by the deep learning model. (c) Difference map between the ground truth (Monte Carlo, MC) and the input PBA dose. (d) Difference map between the ground truth and the predicted output.

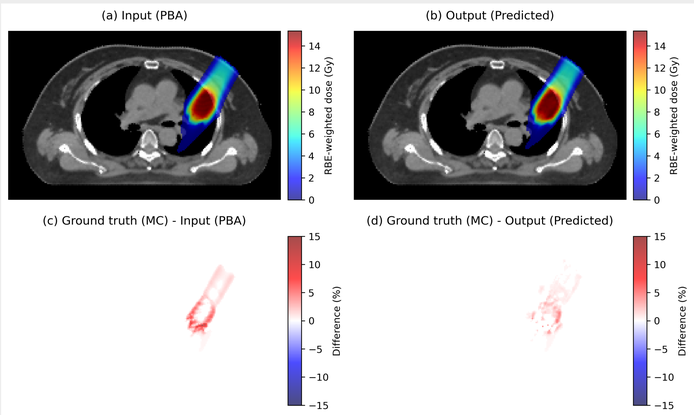

Figure 2. Worst case (lung). Panels (a)–(d) as in Figure 1.

DISCUSSION

- The proposed DL model corrected PBA dose to MC-level accuracy across all three sites, improving GPR and reducing range error and dose RMSE. This supports the feasibility of using a single network to recover the fragmentation and lateral-scatter effects that analytical PBA approximates.
- The results show site dependence. In liver and prostate — relatively homogeneous beam paths — the correction was largest and most complete, eliminating both distal-edge and lateral-edge differences. In lung, the correction improved dose magnitude but left the range error essentially unchanged (0.91 → 0.89 mm), indicating that predicting range across low-density interfaces remains the hard problem. Lung also retained the lowest GPR despite the largest gain, marking it as the limiting case for this approach.
- The sub-second inference time is of practical significance. It positions the model not as a one-off verification step but as a component that can supply MC-quality dose within iterative workflows, including plan optimization and large-scale robustness evaluation, where full MC calculation has been prohibitive.
- Limitations include the cohort size and the scope of the model output: the present network predicts RBE-weighted dose only, and does not provide the dose-averaged linear energy transfer (LET_d).

CONCLUSIONS

- This study developed and validated a DL model that predicts MC-level carbon-ion dose distributions from PBA calculations. By combining MC-level accuracy with analytical-level speed, the model offers a practical route toward integrating high-fidelity dose calculation into routine CIRT planning and verification.
- Future work will extend the framework to the prediction of LET_d and validate it on larger cohorts and to additional, anatomically complex sites such as head-and-neck cancers.

REFERENCES

1. Marx, Vivien. "Sharp shooters." Nature 508.7494 (2014): 133-138.
2. Parodi, Katia, et al. "Monte Carlo simulations to support start-up and treatment planning of scanned proton and carbon ion therapy at a synchrotron-based facility." Physics in Medicine and Biology 57.12 (2012): 3759-3784.
3. He, Rui, et al. "Deep learning-based Monte Carlo dose prediction for heavy-ion online adaptive radiotherapy and fast quality assurance: A feasibility study." Medical Physics 52.4 (2025): 2570-2580.
4. Nguyen, Dan, et al. "3D radiotherapy dose prediction on head and neck cancer patients with a hierarchically densely connected U-net deep learning architecture." Physics in Medicine & Biology 64.6 (2019): 065020.
5. Perl, Joseph, et al. "TOPAS: an innovative proton Monte Carlo platform for research and clinical applications." Medical Physics 39.11 (2012): 6818-6837.

© Poster Template by Genigraphics® 1.800.790.4001 www.genigraphics.com