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

Feasibility of a General-Purpose Deep Learning Dose Engine: A Multi-Site Validation Study

Yao Zhao, Ka Ho Tam, Raphael Douglas, Kyuhak Oh, Xin Wang, Ergys Subashi, Jinzhong Yang, Laurence Court, Dong Joo Rhee
Department of Radiation Physics, The University of Texas MD Anderson Cancer Center, Houston, TX, USA

ABSTRACT

Conventional dose-calculation algorithms trade accuracy against speed and are non-differentiable, limiting online adaptive radiotherapy (ART) and end-to-end automatic planning.

We developed a single, site-independent, differentiable deep-learning dose engine that predicts 3D dose directly from CT and divergent MLC/jaw aperture projections ("beamlets").

Trained on 3,600 IMRT / 3D-CRT plans (6 MV) from 120 patients across seven sites, it was validated by 3D gamma analysis on 70 independent VMAT plans.

The best model (Cascade U-Net, MAE) reached 98.8 ± 1.7% (3%/2 mm), with robust accuracy across all sites and <1.5 min per plan — feasible for autoplanning, online ART, and independent verification.

INTRODUCTION

Accurate dose calculation underpins modern IMRT and VMAT, where algorithms must rigorously model radiation transport in highly modulated fields.

Monte Carlo is the reference standard but is computationally heavy; analytical algorithms (e.g., collapsed-cone) are fast but approximate lateral electron transport, producing errors at density interfaces such as the lung–tumor boundary.

This speed–accuracy trade-off is a key bottleneck for online adaptive radiotherapy (ART), which must re-optimize and recompute dose on daily anatomy while the patient remains on the treatment couch.

Conventional engines are also non-differentiable, which blocks fully end-to-end, gradient-based automatic planning.

Deep-learning dose engines can help[1-3], but most are trained on a single site and rely on TPS-specific or fluence-derived inputs that implicitly encode machine information — limiting its use outside the commercial TPS.

We develop a beamlet/aperture-based, differentiable 3D CNN that predicts dose directly from patient CT and divergent MLC and jaw projections, learning a generalizable map from anatomy and aperture geometry to dose.

Because the engine is differentiable, dosimetric gradients can backpropagate to machine parameters — a foundation for end-to-end autoplanning, independent dose verification, and online ART.

METHODS AND MATERIALS

Data: 120 MD Anderson patients across 7 sites (brain, H&N, breast, thoracic, GI, prostate, cervical); 3,600 step-and-shoot IMRT + 3D-CRT plans (6 MV; RayStation CCC). 201,214 beamlets; patient-level 8:2 train/val.

Inputs (Fig. 1): four 192^3 volumes at 3 mm — CT (electron density), divergent ray-traced MLC and jaw projections, and ground-truth dose per MU; all isocenter-centered.
Networks: 3D U-Net vs. coarse-to-fine 3D Cascade U-Net[4]; MAE and MSE losses; Adam (1e-4), DDP, deep supervision, 100 epochs.

VMAT inference: each arc decomposed into 180 control points (every 2°); per-control-point dose weighted by MU and summed.

Evaluation: 3D global gamma (TG-218)[5], 10% threshold, four criteria. Independent cohort: 70 VMAT plans (40 MDA + 30 RPA[6]), 7 sites.

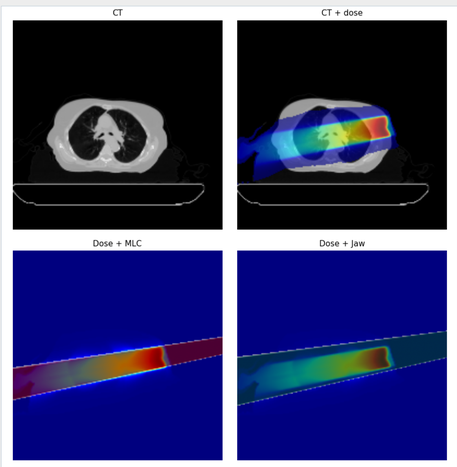

Figure 1. Model inputs for one beamlet — CT, CT+dose, and ray-traced divergent MLC / jaw projections (192^3, 3 mm).

RESULTS

High accuracy across all configurations (Table 1). Best — Cascade U-Net + MAE: 98.8 ± 1.7% at 3%/2 mm; standard U-Net + MAE comparable (98.5 ± 2.1%).

Loss functions (MAE vs MSE) and architectures were statistically comparable — high-resolution divergent inputs, not network depth, drive accuracy.

Site-wise (Fig. 3): pelvic best (cervical & prostate >99%); H&N and thoracic >98.6%; worst-case site minimum >93%.
Open-field water phantom (Fig. 2): FWHM errors <0.4 mm (max 0.74 mm) and central-axis MAE <1% (≤15×15 cm²), meeting AAPM MPPG 5.b[7].

Speed: 84.4 ± 27.4 s per VMAT plan on one NVIDIA A100 (preprocessing 49.9 s, inference 34.5 s).

Table 1. Overall 3D gamma passing rate (mean ± SD, %) on the independent VMAT test set (n = 70).

Model & Loss	3%/2 mm	3%/1 mm	2%/2 mm	2%/1 mm
Cascade U-Net + MAE	**98.8 ± 1.7**	**96.1 ± 3.5**	**96.6 ± 3.2**	**90.8 ± 6.8**
Cascade U-Net + MSE	98.3 ± 2.5	95.1 ± 5.3	96.1 ± 4.3	89.4 ± 9.3
U-Net + MAE	98.5 ± 2.1	95.9 ± 4.2	96.5 ± 3.8	90.4 ± 7.7
U-Net + MSE	97.9 ± 2.6	94.8 ± 4.7	94.9 ± 4.7	87.6 ± 8.8

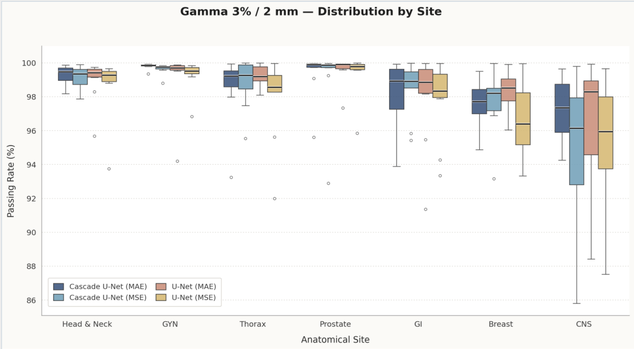

Figure 3. Gamma (3%/2 mm) by anatomical site — robust across all seven sites, best for pelvic anatomy.

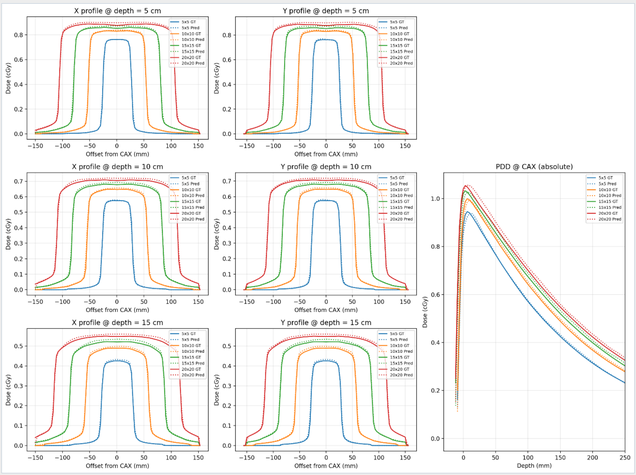

Figure 2. Water-phantom validation: X/Y profiles and central-axis depth dose. Solid = TPS, dotted = network.

DISCUSSION

A single model computes clinically accurate 3D dose (>98% at 3%/2 mm) across seven sites and generalizes across delivery techniques — trained on static IMRT/3D-CRT, tested on VMAT.

Deliverable MLC/jaw inputs make it beam-geometry- and vendor-independent: no TPS pre-dose or fluence features required.

Isocenter-centering lets the network internalize fixed physics (inverse-square fall-off, SAD, beam profile).

Standard U-Net + MAE is preferred: equal accuracy, lower memory — better for time-critical online ART.

Limitations (Fig. 4): localized errors at steep density gradients (skin buildup, surgical clips, lung–tissue interface); 6 MV only; single-institution data.

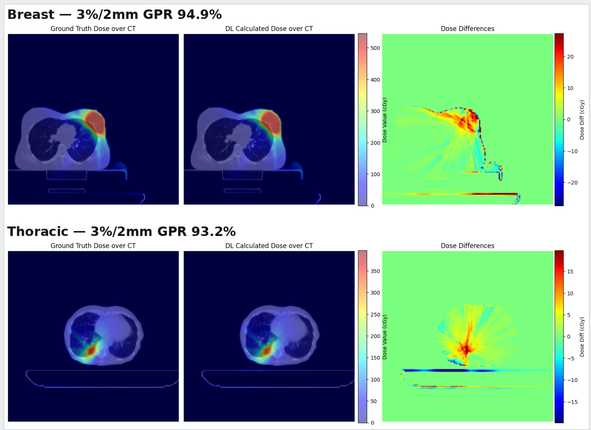

Figure 4. Worst-case plans: ground truth, prediction, and difference for breast (94.9%) and thoracic (93.2%). Errors concentrate at skin and tumor–lung interfaces.

CONCLUSIONS

A beamlet/aperture-based deep-learning dose engine achieves near-clinical accuracy across seven sites with one unified, differentiable model — overcoming the speed–accuracy trade-off and enabling end-to-end autoplanning, online ART, and independent verification.

Future work: extension to additional beam energies (15 MV, FFF) and GPU-accelerated beam projection to further reduce latency.

REFERENCES

1. Fan J, et al. Data-driven dose calculation algorithm based on deep U-Net. Phys Med Biol. 2020;65:245035.
2. Kontaxis C, et al. DeepDose: towards a fast dose calculation engine for radiation therapy using deep learning. Phys Med Biol. 2020;65:075013.
3. Liang B, et al. A deep learning-based dose calculation method for volumetric modulated arc therapy. Radiat Oncol. 2024;19:141.
4. Jiang Z, et al. Two-stage cascaded U-Net: 1st place solution to the BraTS Challenge 2019 segmentation task. In: Brainlesion. Springer; 2020:231–241.
5. Miften M, et al. Tolerance limits and methodologies for IMRT measurement-based verification QA: recommendations of AAPM Task Group No. 218. Med Phys. 2018;45:e53–e83.
6. Court LE, et al. Radiation Planning Assistant — a streamlined, fully automated radiotherapy treatment planning system. J Vis Exp. 2018;(134):57411.
7. Geurts MW, et al. AAPM Medical Physics Practice Guideline 5.b: commissioning and QA of treatment-planning dose calculations — megavoltage photon and electron beams. J Appl Clin Med Phys. 2022;23:e13641.

CONTACT

Dong Joo Rhee, PhD
Department of Radiation Physics
MD Anderson Cancer Center
Houston, TX, USA
drhee1@mdanderson.org

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