NEURIPS 2026 · EVALUATIONS & DATASETS TRACK

PortPy.AI

A Benchmark for AI and Optimization in Cancer Radiotherapy Planning

Gourav Jhanwar[1] · Mojtaba Tefagh[2] · Qijie Huang[1] · Alireza Golkarieh[3] · Linda Hong[1] · Hai Pham[1] · Vicki Trier Taasti[4] · Seppo Tuomaala[5] · Anthony Magliari[5] · Michael Folkerts[5] · Ying Zhou[1] · Jie Yang[1] · Saad Nadeem[1] · Masoud Zarepisheh[1]

[1] Memorial Sloan Kettering Cancer Center [2] University of Edinburgh [3] Oakland University [4] Aarhus University Hospital [5] Varian Medical Systems, Siemens Healthineers

github.com/PortPy-Project/PortPy · huggingface.co/datasets/PortPy-Project/PortPy_Dataset

Radiotherapy is half of cancer care

20 M

new cancer cases every year worldwide

1 in 2

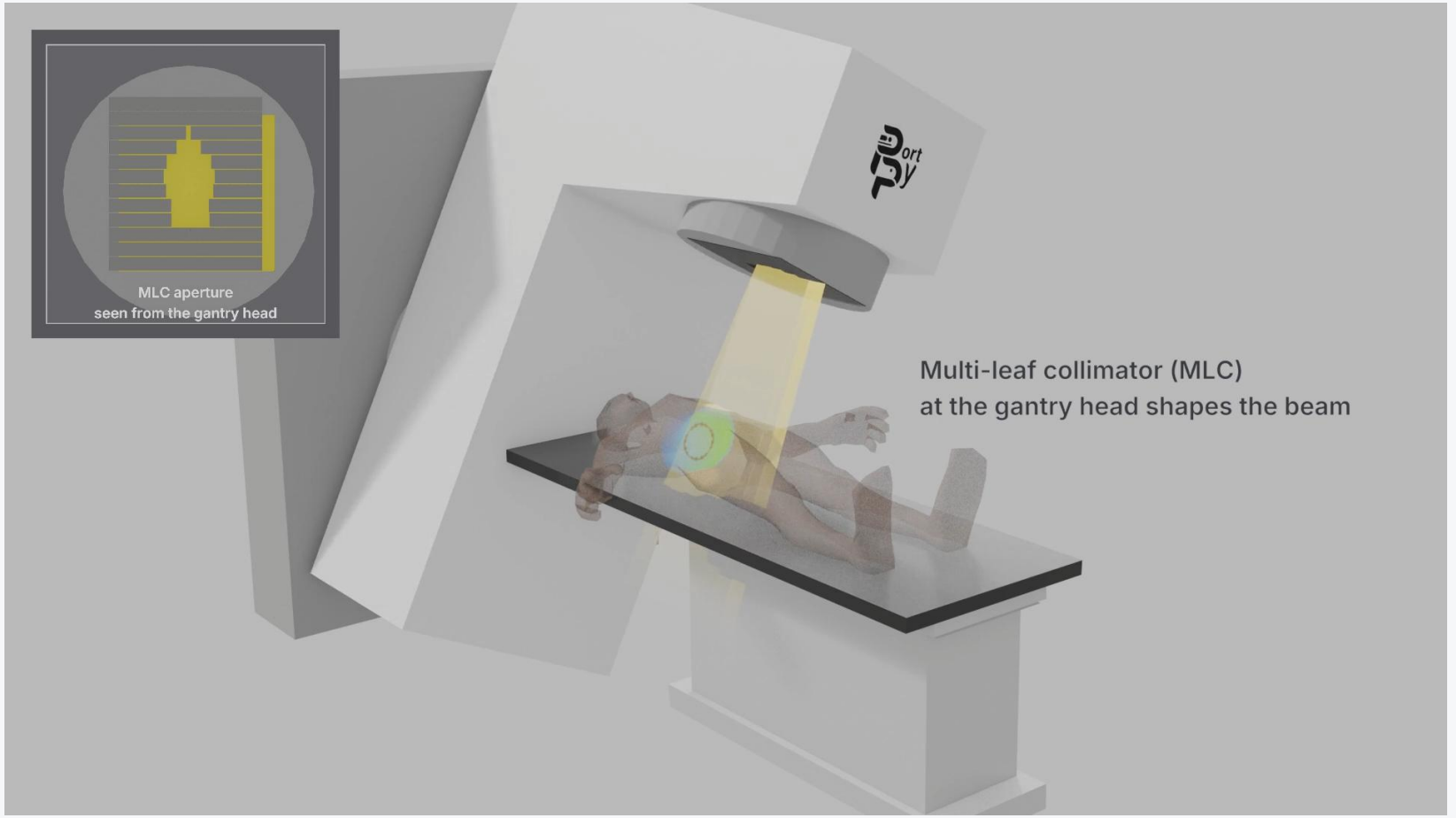
cancer patients are treated with radiation

1 plan

individually optimized for every patient

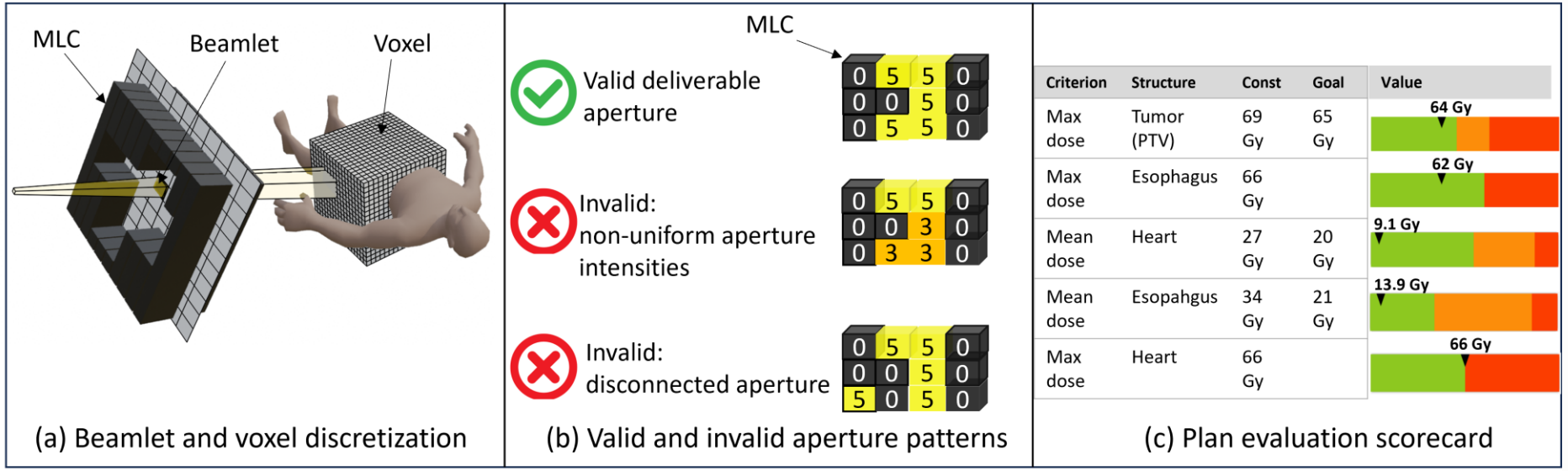
Roughly 10 million treatment plans a year, each a hard optimization problem solved under clinical time pressure.

Radiotherapy: many shaped beams, one dose map



Planning = choosing the shape and intensity of every beam so the tumor gets a high dose and healthy organs do not.

Why planning is hard



Expensive physics

Beamlet-to-voxel dose can take hours per patient

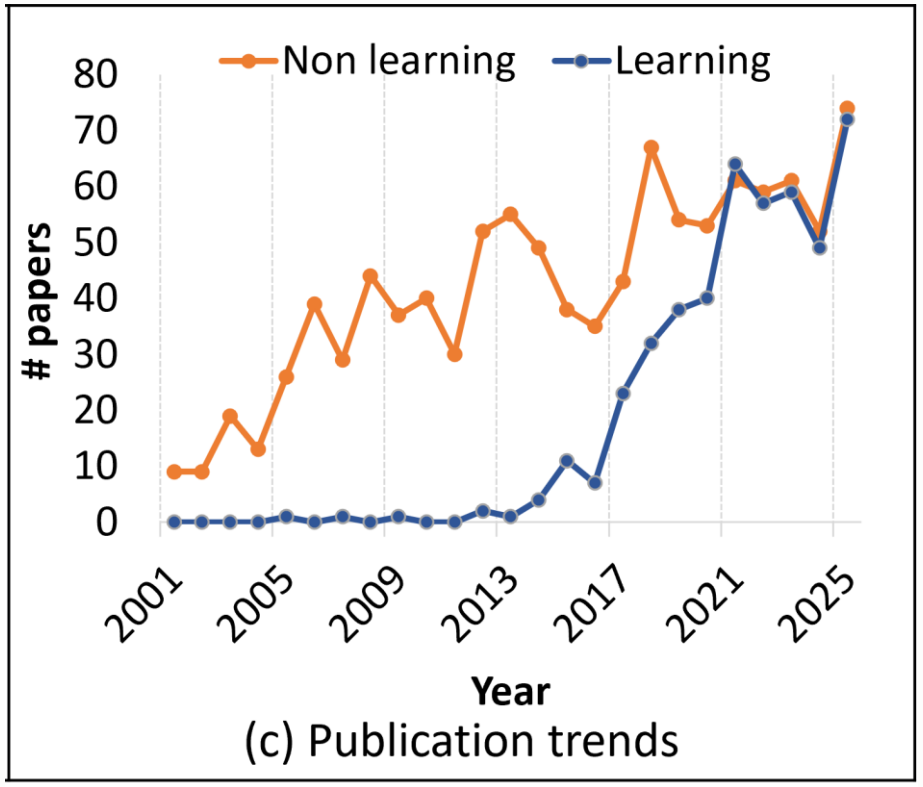
Deliverability

Valid apertures only: combinatorial, non-convex

Competing objectives

Weights tuned by hand, patient by patient

AI tools have shown promise for these challenges



Publications per year on radiotherapy planning optimization (Scopus, 2001–2025): learning-based work has grown from almost none in 2010 to match classical optimization today.

The gap: no common ground for research

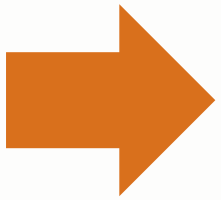
TODAY

DATA
Locked in vendor systems

TASKS & BASELINES
None shared

TOOLS
Commercial software only

Not reproducible, not comparable



WITH PORTPY

DATA
Open, from 329 real patients

TASKS & BASELINES
Three standard tasks + baselines

TOOLS
Open Python toolkit

Reproducible and comparable

What PortPy provides

Dataset

329 patients · 8 TB · clinical criteria · expert plans

Toolkit

Python · Integration with commercial/open planning system

Three benchmark tasks

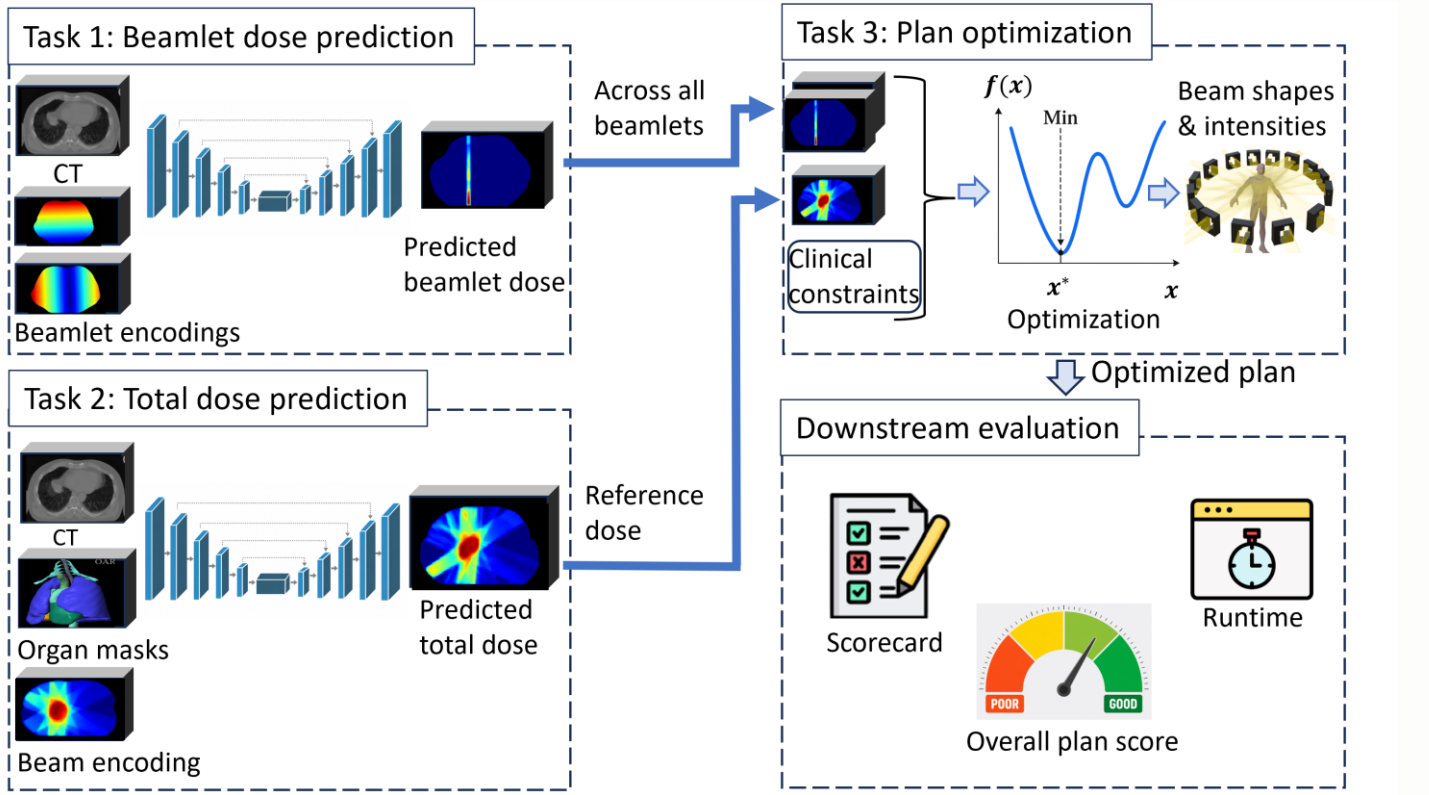
Deep learning: to learn the computationally expensive physics-based beamlet dose calculation

Deep learning: to learn the expert multi-criteria clinical-trade-offs between tumor irradiation and healthy tissue sparing

Optimization: to identify the beam shapes and intensities

First open platform grounded in clinical practice · 40,000+ downloads

Three tasks, one end-to-end evaluation



Task 1 · AI dose engine

Plans from AI dose engine are close to physics-based plans

Task 2 · Total Dose prediction

Close to expert plans, with clear room to improve

Task 3 · Optimization

Clinical solver matches the global optimum, 7,000× faster than MIP

Every prediction is pushed through the optimizer and judged by the plan it produces, not by its error alone.

Thank you

A clinically grounded benchmark, open to the AI, optimization and medical physics communities.

TOOLKIT

github.com/PortPy-Project/PortPy

pip install portpy

DATASET

huggingface.co/datasets/PortPy-Project/PortPy_Dataset

JSON + HDF5, optimization-ready