

## INTRODUCTION

- Deep learning models achieve strong performance in radiotherapy fluence prediction
- Most existing evaluations assume stable and idealized data distributions
- Clinical deployment introduces uncertainty from imaging noise, anatomical variation, and geometric misalignment
- Small perturbations in fluence prediction may produce clinically significant non-linear deviations
- Robustness analysis for direct fluence map prediction remains limited

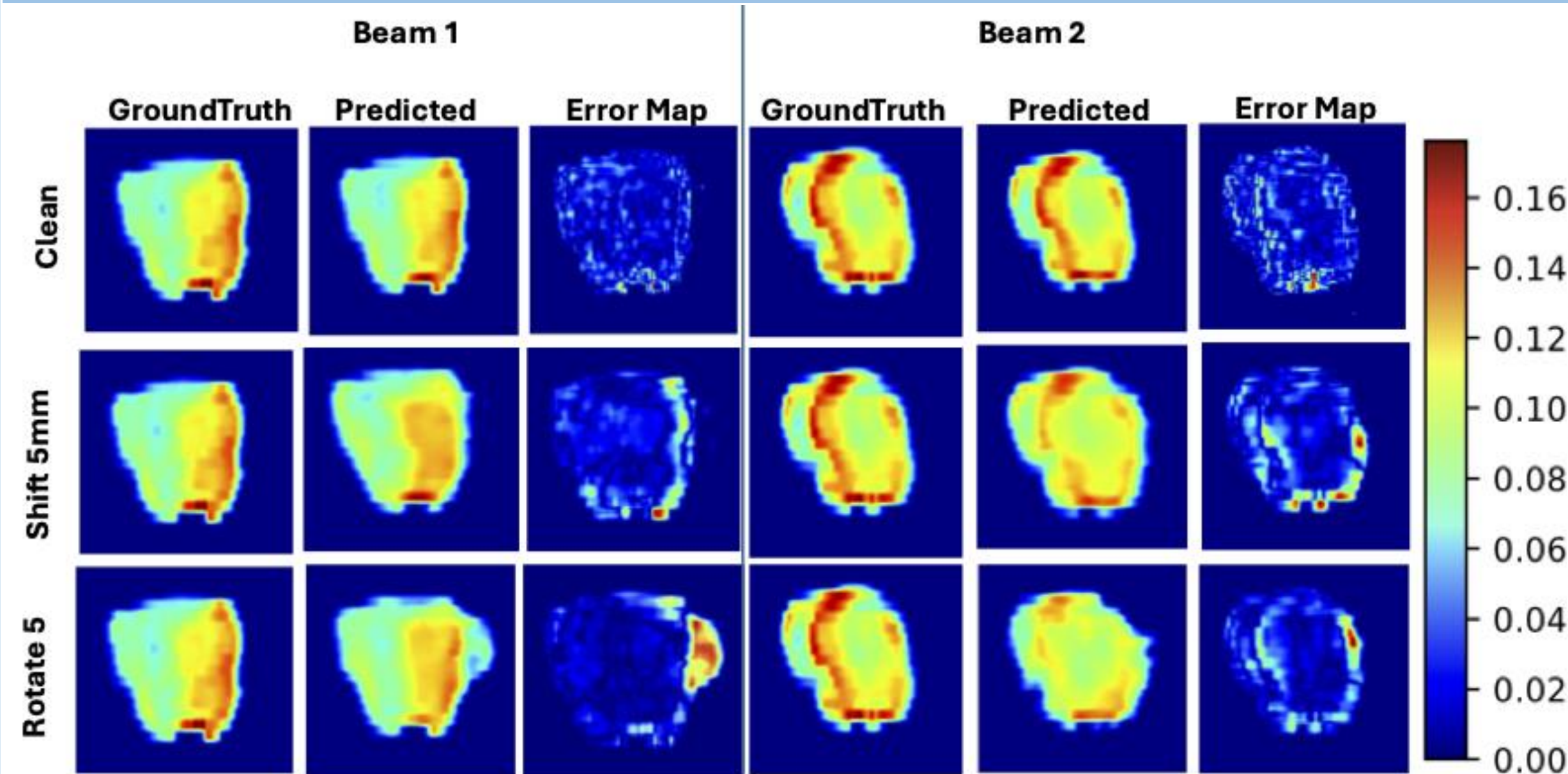
## OBJECTIVES

- To evaluate FluenceFormer robustness under deployment-relevant perturbations
- To analyze degradation under geometric, radiometric, and domain shifts
- To assess robustness under limited training data conditions
- To compare architectural sensitivity to perturbations
- To investigate physics-based metrics for detecting clinically relevant failure modes

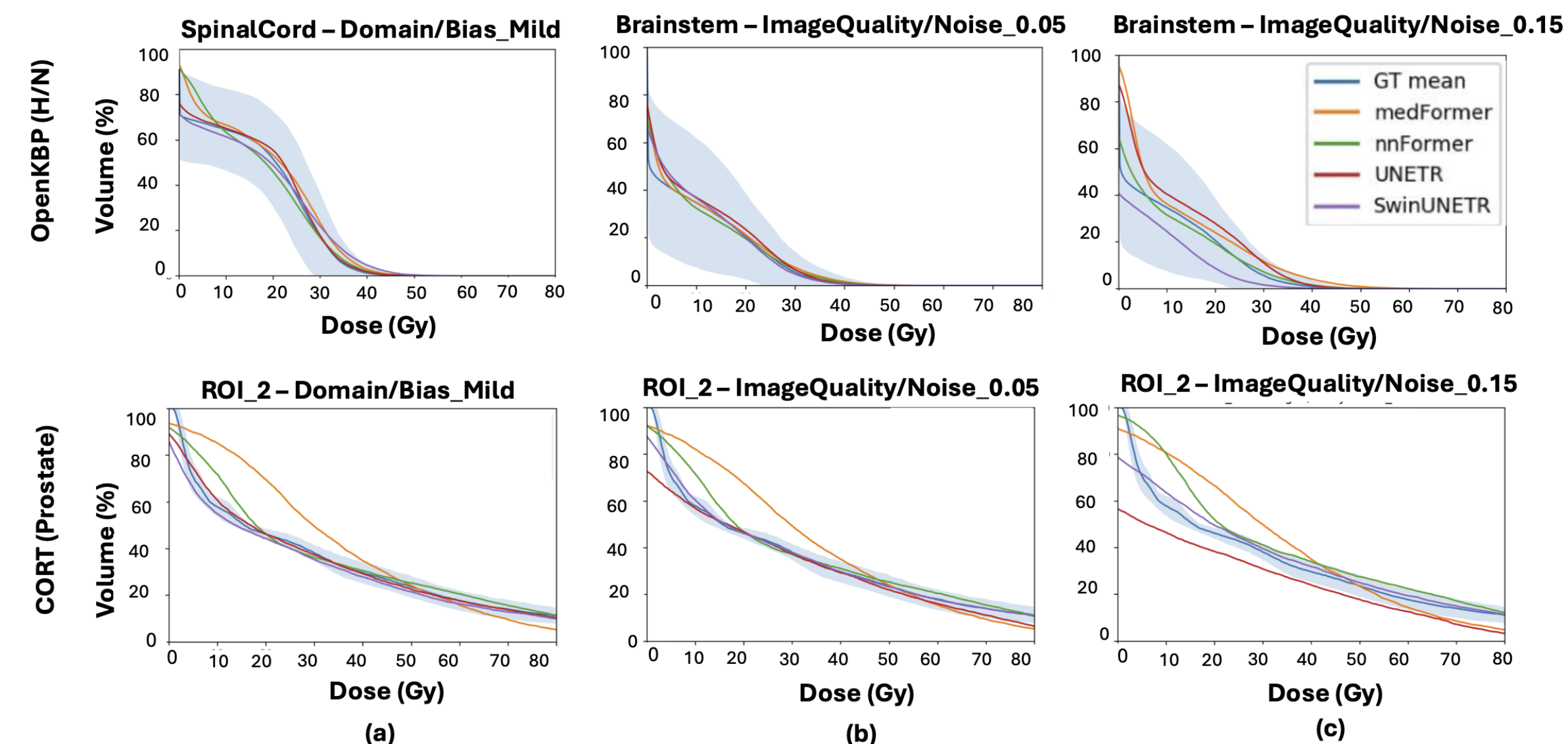
## MATERIAL & METHODS

- Two-stage IMRT pipeline: dose prediction followed by fluence prediction
- Compared transformer backbones: SwinUNETR, UNETR, nnFormer, and MedFormer
- Training performed using Fluence-Aware Regression (FAR) loss
- Evaluation on 99-case prostate IMRT dataset
- Cross-dataset testing on OpenKBP and CORT

## RESULTS



**Figure 1:** Qualitative geometric robustness. Ground-truth fluence maps (repeated for visual reference) remain invariant, as geometric perturbations are applied only at inference.



**Figure 2:** DVH robustness on public datasets. DVH curves on OpenKBP (top) and CORT (bottom) under bias and noise. Shaded: ground truth mean±std; lines: model predictions.

## RESULTS

**Table 1:** Data efficiency and scaling. SSIM and upper-quartile energy error EQ75 error as a function of training data fraction.

Backbone	Metric	25%	50%	75%	100%
SwinUNETR	SSIM↑	0.65±0.09	0.71±0.08	0.74±0.07	<b>0.76±0.08</b>
	$E_{err}^{Q75}$ ↓	10.56	7.32	5.43	<b>4.67</b>
UNETR	SSIM↑	0.58±0.09	0.62±0.08	0.65±0.07	<b>0.67±0.07</b>
	$E_{err}^{Q75}$ ↓	12.56	10.89	10.56	<b>8.67</b>
nnFormer	SSIM↑	0.40±0.09	0.52±0.08	0.55±0.07	<b>0.58±0.06</b>
	$E_{err}^{Q75}$ ↓	12.78	11.89	11.56	<b>10.34</b>
MedFormer	SSIM↑	0.52±0.08	0.58±0.08	0.62±0.07	<b>0.65±0.06</b>
	$E_{err}^{Q75}$ ↓	12.26	11.92	11.45	<b>10.67</b>

## ROUBUSTNESS EVALUATION

### Geometric

- Setup shifts (3–5 mm) and rotations (2°–5°)

$$X_{geom}(\mathbf{r}) = X(\mathbf{Rr} + \mathbf{t})$$

### Radiometric

- Gaussian noise ( $\sigma = 0.05-0.20$ )

$$X_{noise}(\mathbf{r}) = X(\mathbf{r}) + \epsilon(\mathbf{r}), \quad \epsilon(\mathbf{r}) \sim \mathcal{N}(0, \sigma^2)$$

### Data Scarcity

- Evaluated robustness under limited data availability

$$\mathcal{D}_\alpha \subset \mathcal{D}$$

### Domain Shift

- Bias-field intensity perturbations

$$X_{domain}(\mathbf{r}) = b(\mathbf{r}) X(\mathbf{r}) + c,$$

## CONCLUSION/SUMMARY

- Fluence prediction degrades under realistic perturbations
- Physics-based metrics reveal clinically relevant errors missed by SSIM
- SwinUNETR shows greater robustness to perturbations

## CONTACT/RESOURCES



Research Lab



Paper Link