Development and Preliminary Evaluation of a Deep Learning Pipeline for Target Segmentation in MRI-Guided HDR Brachytherapy

Soroush Ghomashchi[1,2], Ruiyan Ni[1,2], Alexandra Rink[1-3]
[1]Department of Medical Biophysics, University of Toronto, [2]Princess Margaret Cancer Centre, UHN, [3] Department of Radiation Oncology, University of Toronto

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

ABSTRACT

Purpose: High-Dose-Rate Brachytherapy (HDR-BT) is a critical part of cervical cancer treatments, with MRI-guidance serving as the gold standard for defining the Gross Tumor Volume (GTV) and Clinical Target Volumes (CTV_{HR}, CTV_{IR}). While MRI offers superior soft-tissue contrast, manual segmentation remains time-intensive. This study evaluates the efficacy of a 3D Deep Learning (DL) framework for MRI-based segmentation, specifically comparing the performance of single-view models versus multi-view ensemble architectures.

Methods: The study cohort includes T2-weighted MRI planning scans from 164 patients (369 total fractions) with cervical cancer. A self-configuring 3D U-Net (nnU-Net) was utilized to establish a single-view axial baseline. To assess the value of contextual information from orthogonal planes, separate models were trained on coronal and sagittal views and integrated into a multi-view ensemble.

Results: The single-view axial model demonstrated equivalent performance to multi-view fusion strategies across all targets. For the challenging GTV, the axial model achieved a DSC of 0.46 ± 0.27. For larger targets, the axial model achieved high performance (CTV_{HR} DSC: 0.83 ± 0.06; CTV_{IR} DSC: 0.84 ± 0.06).

Conclusion: This study demonstrates that a well-optimized single-view 3D axial architecture provides robust segmentation for MRI-guided HDR brachytherapy, outperforming complex multi-view ensembles. Future work will focus on addressing GTV class imbalance through generalized Dice loss and focal loss, increasing oversampling rate, cascade training, and integration of additional metrics. We will also compare this baseline model against other state-of-the-art DL architectures to further benchmark segmentation performance.

INTRODUCTION

- High-Dose-Rate Brachytherapy (HDR-BT; **Figure 1**) is a critical component of cervical cancer treatment.

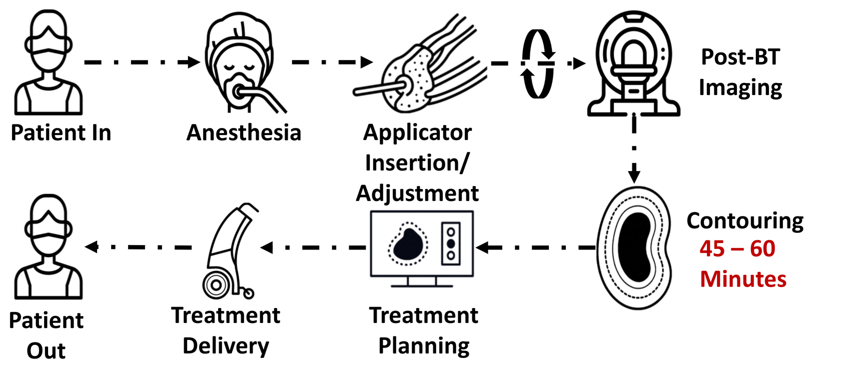


Figure 1. Brachytherapy Workflow Procedure

- MRI serves as the gold standard for defining the Gross Tumor Volume (GTV) and Clinical Target Volumes (CTV_{HR}, CTV_{IR}) due to its superior soft-tissue contrast compared to CT.
- Manual segmentation (**Figure 2**) is highly time-intensive.
- Real gap in automatic contouring of harder to segment targets such as GTV[1-2].

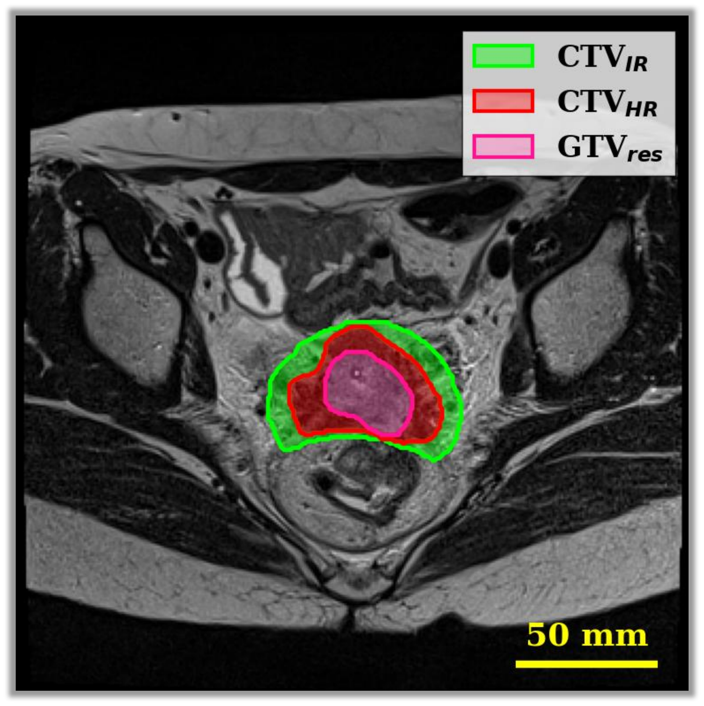


Figure 2. Clinical targets in HDR-BT

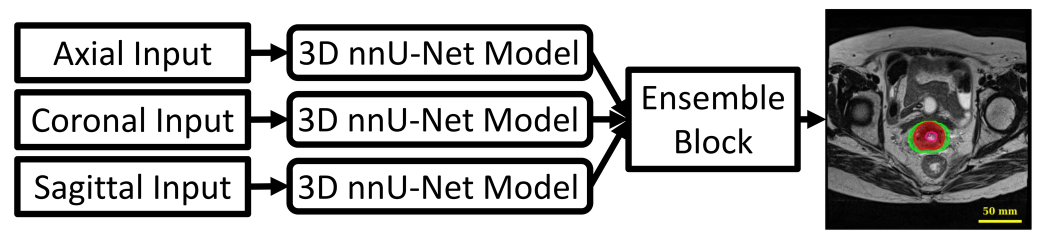
- T2W MRI scans are anisotropic. Image resolution and contextual information offered by each view is not identical.
- Multi-view 2D-based studies demonstrate superior contouring performance.

Objectives

- To develop and evaluate a 3D deep learning framework for MRI-based target segmentation in HDR Brachytherapy.
- To compare the geometric accuracy and clinical utility of a robust single-view axial baseline model against multi-view ensemble architectures.

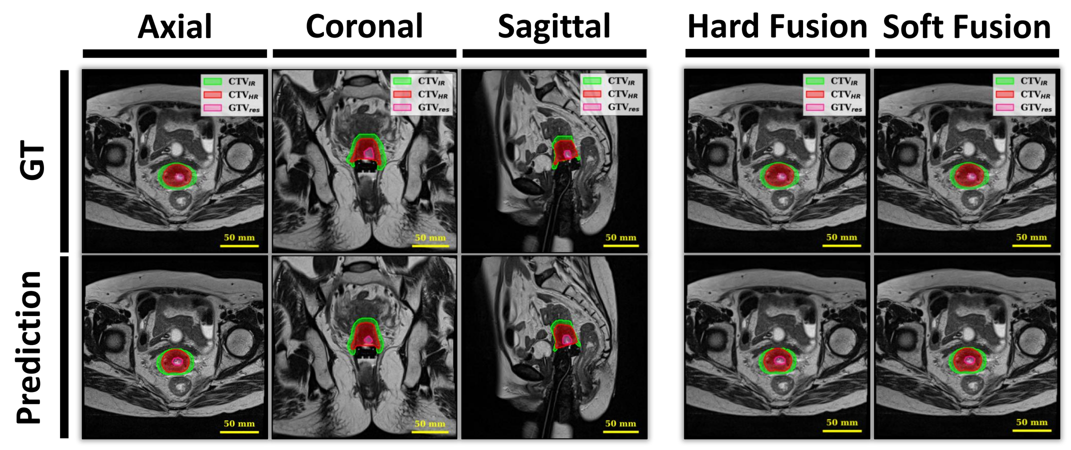
Methods

- A total of 369 fractions from T2-weighted MRI scans of 164 patients with cervical cancer (FIGO stages IB–IIIC2).
- Patient-level Ground Truth (GT) target contours were manually delineated and reviewed by radiation oncologists in the axial view, then propagated and manually corrected to generate coronal and sagittal GT contours.
- **Multi-View Architecture Design:**
  - Segmentation: 3D nnU-Net v2
  - Train/Test Split: 80/20 – 130/34 Patients (295/74 Fractions)
- Three independent models trained on each view



- Combine the outputs of the three single-view models
- **Metrics:** Volumetric and Surface Dice Similarity Coefficients (**vDSC** & **sDSC**), Hausdorff Distance (**HD95**), and Added Path Length (**APL**)[3].
- **Statistical Analysis:** Wilcoxon signed-rank test

RESULTS



- Single-view axial model demonstrates equivalent performance to multi-view fusion strategies across all targets.
- **CTV_{HR} & CTV_{IR}:** Axial model achieved high performance (CTV_{HR} vDSC: 0.83 ± 0.06; CTV_{IR} vDSC: 0.84 ± 0.06), with no statistically significant improvement from multi-view fusion ($p > 0.05$).
- **GTV:** Axial model matching (vDSC = 0.46 ± 0.27) all fusion ensembles (Figure 3).

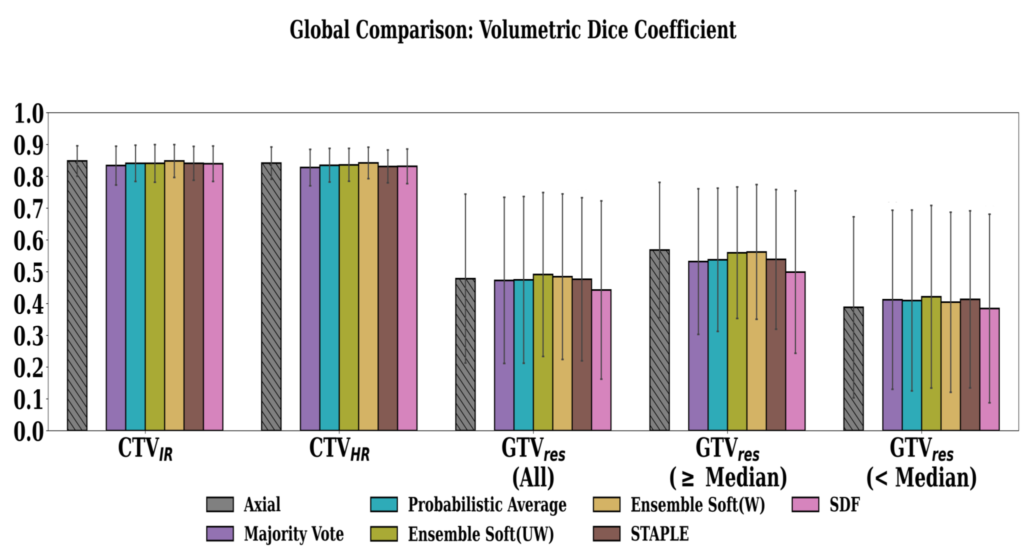


Figure 3. Volumetric Dice Coefficients for Axial vs. Fusion ensembles

- For sDSC and HD-95, axial model also matched all fusion ensembles.
- Axial model yielded a significantly lower ($p < 0.05$) correction burden (lower APL) across all targets compared to the fusion ensembles (**Figure 4**).

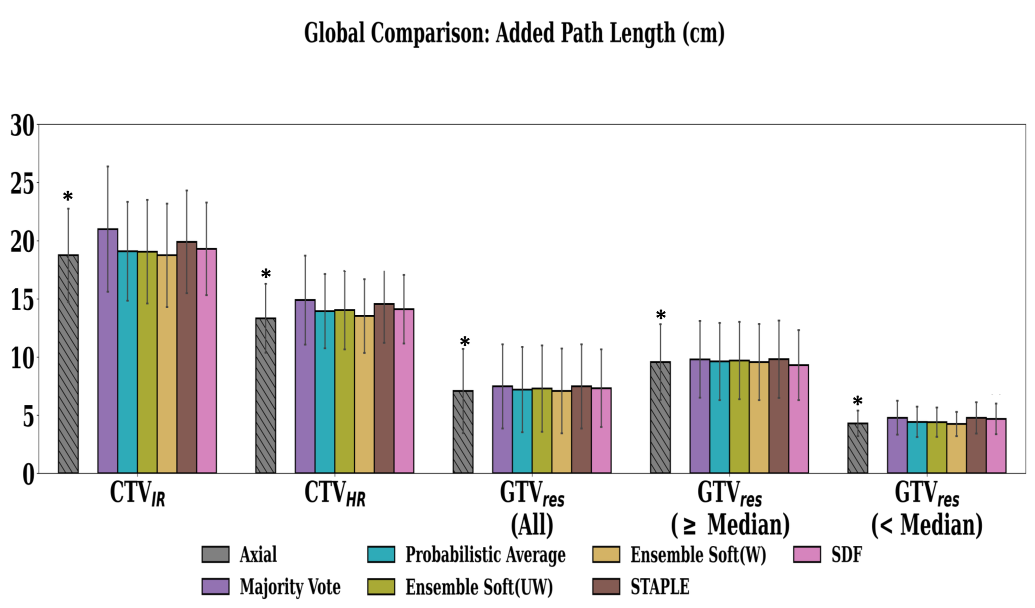


Figure 4. APL values for Axial vs. Fusion ensembles

DISCUSSION

- Single-view axial model demonstrates equivalent performance to multi-view fusion strategies across all targets.
- Fusion ensembles introduce anisotropic noise that outweigh the benefits of multi-view augmentation observed with 2D-based studies
- The lower quantitative performance for GTV compared to other targets is attributed to the high surface-area-to-volume ratio of small targets, where minor boundary deviations at the voxel level disproportionately penalize volumetric metrics.
- Beyond volumetric overlap, the fusion approaches, specifically Majority Vote, resulted in predicted contours with "choppy" or jagged edges.
- Fusion, multi-view, ensembles offer no statistically significant geometric improvement over a robust, well-optimized single-view axial model.

CONCLUSIONS

- A single-view 3D axial architecture provides robust segmentation for MRI-guided HDR brachytherapy; outperforming/matching complex multi-view ensembles for target delineation.
- To address the GTV Imbalance within nnU-Net:
  - Increasing Oversampling Rate
  - Generalized dice loss and focal loss
  - Cascade Training
- Compare this baseline model against other DL architectures:
  - nnU-Net v2: Convolution vs ResNet Blocks
  - Prompt-Based
- Implement additional metrics:
  - Entropy/Uncertainty Map
  - Absolute Volume Error (AVE)
- Deployment of a full segmentation architecture – include OARs.
- Multi-Modal Approach – Integration of the clinical notes and diagnostic reports.

ACKNOWLEDGEMENTS

This work was funded through the Canadian Institute of Health Research (Project Grant Number: PJT-186028).

REFERENCES

1. Shi J, et al. Artificial intelligence in high-dose-rate brachytherapy treatment planning for cervical cancer: a review.
2. Yoganathan S, et al. Automatic segmentation of magnetic resonance images for high-dose-rate cervical cancer brachytherapy using DL
3. Vaassen F, et al. Evaluation of measures for assessing time-saving of automatic organ-at-risk segmentation in radiotherapy