A Generalizable Deep Learning Framework for Real-Time CT Reconstruction from a Single Projection for Lung Cancer Radiotherapy

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

Xiangyu Zhang; Qiuyi Ye; Weige Wei; Yongchang Wu; Lian Duan; Jing Li; Sen Bai; Guangjun Li
West China Hospital, Sichuan University, Chengdu, China | The University of Texas MD Anderson Cancer Center, Houston, USA

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

Abstract

Purpose: To enable real-time volumetric guidance during lung radiotherapy using only one CBCT projection.
Methods: We developed SP2V, a cross-patient CNN-Transformer framework that combines a reference CT with paired reference/target projections to infer a 3D deformation vector field (DVF). SP2V was compared with 3DUnet, Elastix, and TransMorph using image quality, anatomical accuracy, and deformation regularity metrics.
Results: SP2V reconstructed thoracic CT in <50 ms without patient- or fraction-specific training. It achieved 9.88 HU MAE, 0.813 GTV DSC, 0.975 lung DSC, and 24.9 negative-Jacobian voxels, outperforming 3DUnet across all metrics.
Conclusion: Single-projection, motion-resolved CT reconstruction is feasible for real-time lung radiotherapy guidance.

Introduction

Clinical challenge: Respiratory motion changes lung tumor and organ-at-risk positions during irradiation, increasing dose-delivery uncertainty.
Unmet need: CBCT/4DCBCT require multi-angle acquisition, while surface and MR guidance have surrogate or spatiotemporal limitations. Existing projection-based methods often require patient- or fraction-specific optimization.
Objective: Develop a cross-patient framework for real-time thoracic CT reconstruction from one CBCT projection without individualized retraining.

Methods and Materials

Cohort: 150 lung cancer patients with ten-phase 4DCT; 105/15/30 patients for training/validation/testing.
Pre-processing: CTs were resampled to 256 x 256 x 128; paired 0-degree projections were generated with DeepDRR.
Inputs: End-expiration reference CT plus reference and target CBCT projections.
Architecture: A 2D encoder extracts motion features, expands them to 3D, and fuses them with the reference CT. Swin-UNETR predicts the DVF used to reconstruct the target CT.
Training: TransMorph DVFs supplied motion labels. Six combinations of HU MSE, DVF MSE, and DVF-gradient L1 losses were evaluated.
Evaluation: SP2V was compared with 3DUnet, Elastix, and TransMorph using MAE/PSNR/SSIM, GTV HD95/DSC, lung DSC, centroid error, and negative-Jacobian voxels.

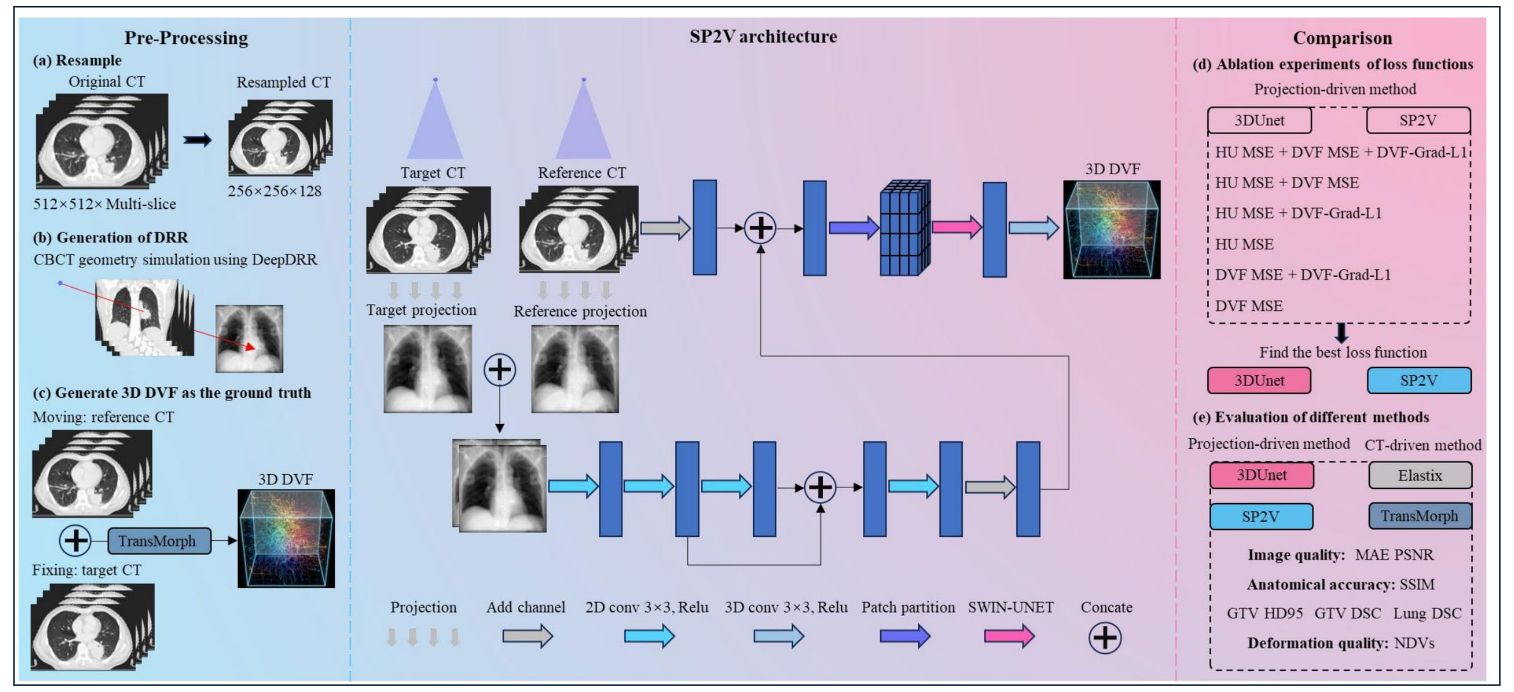


Figure 1. SP2V workflow and evaluation design.

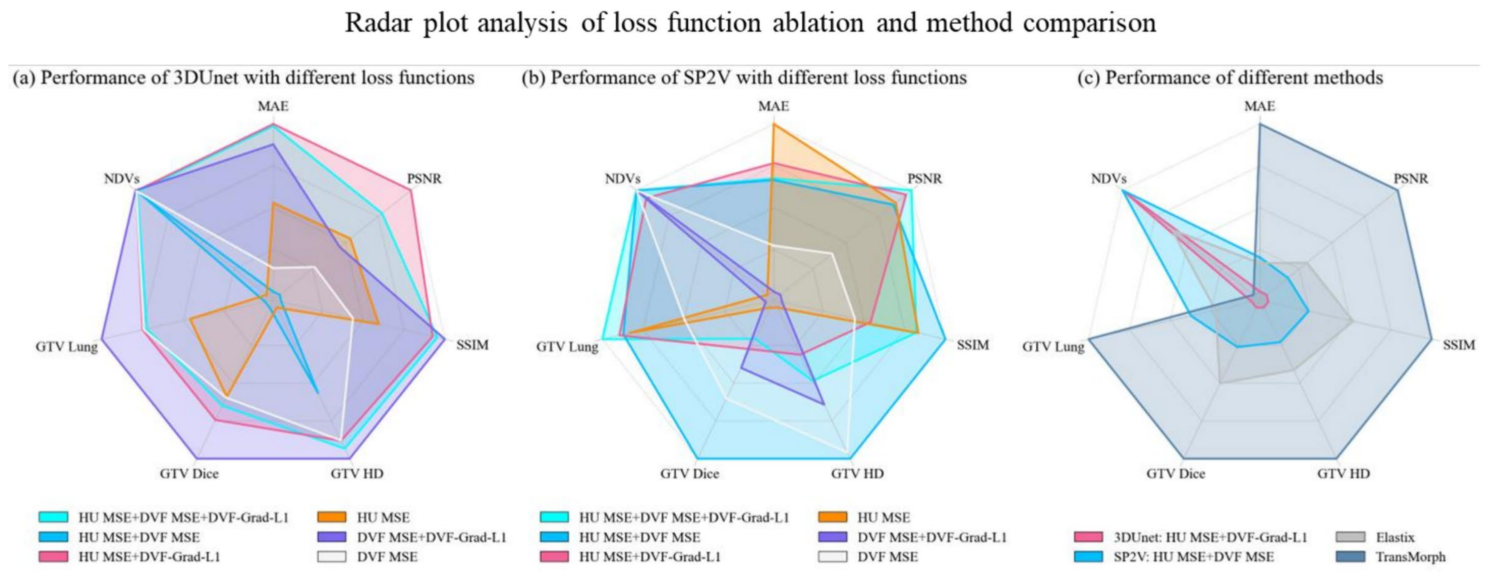


Figure 2. Loss-function ablation and method comparison.

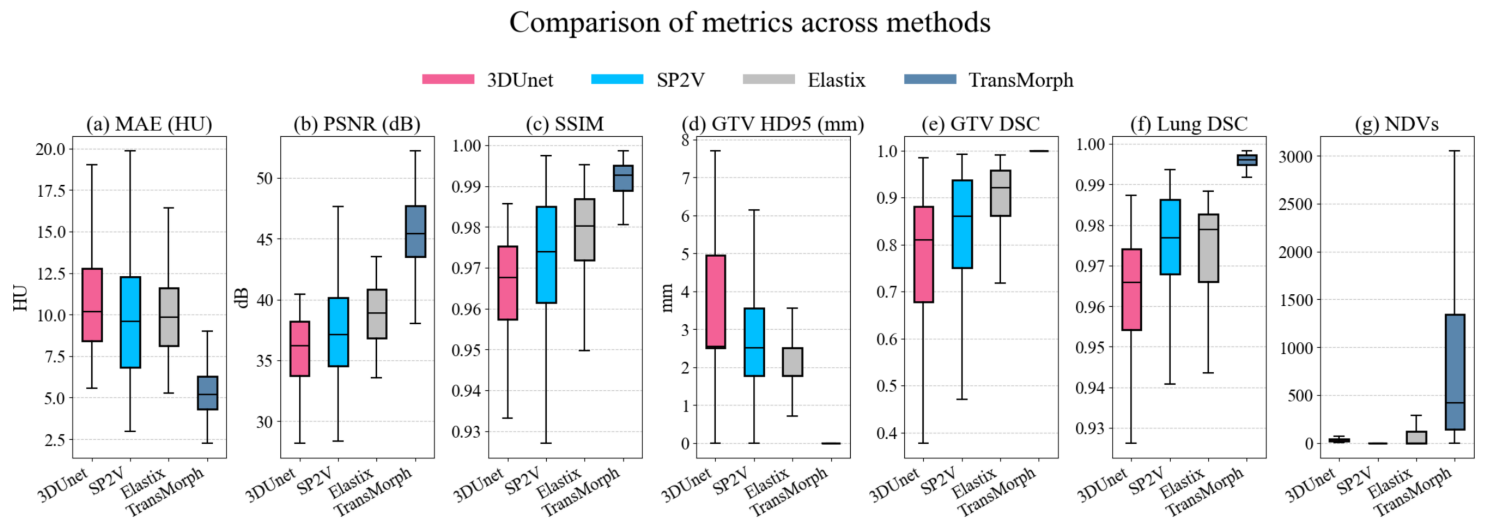


Figure 3. Quantitative metric comparison across methods.

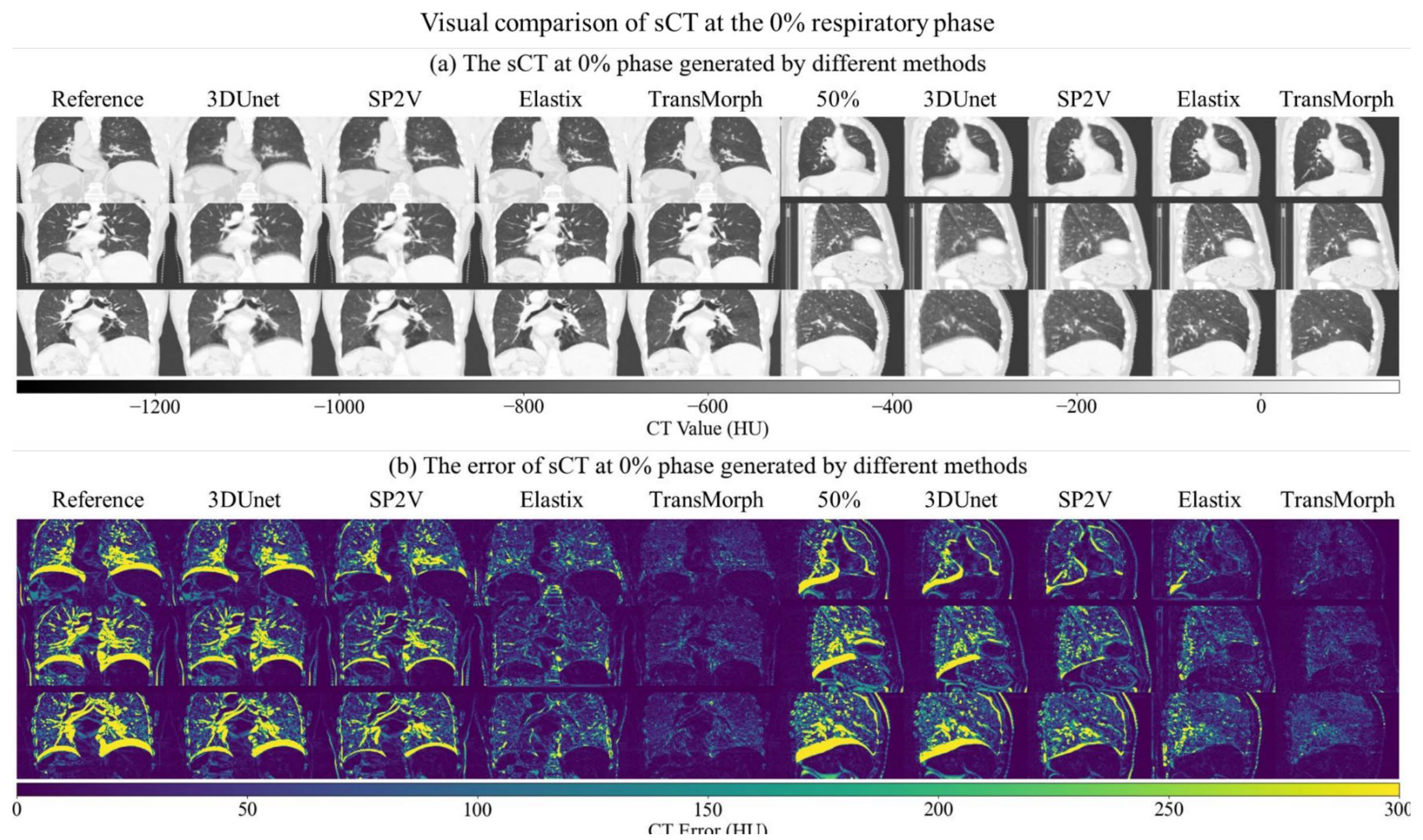
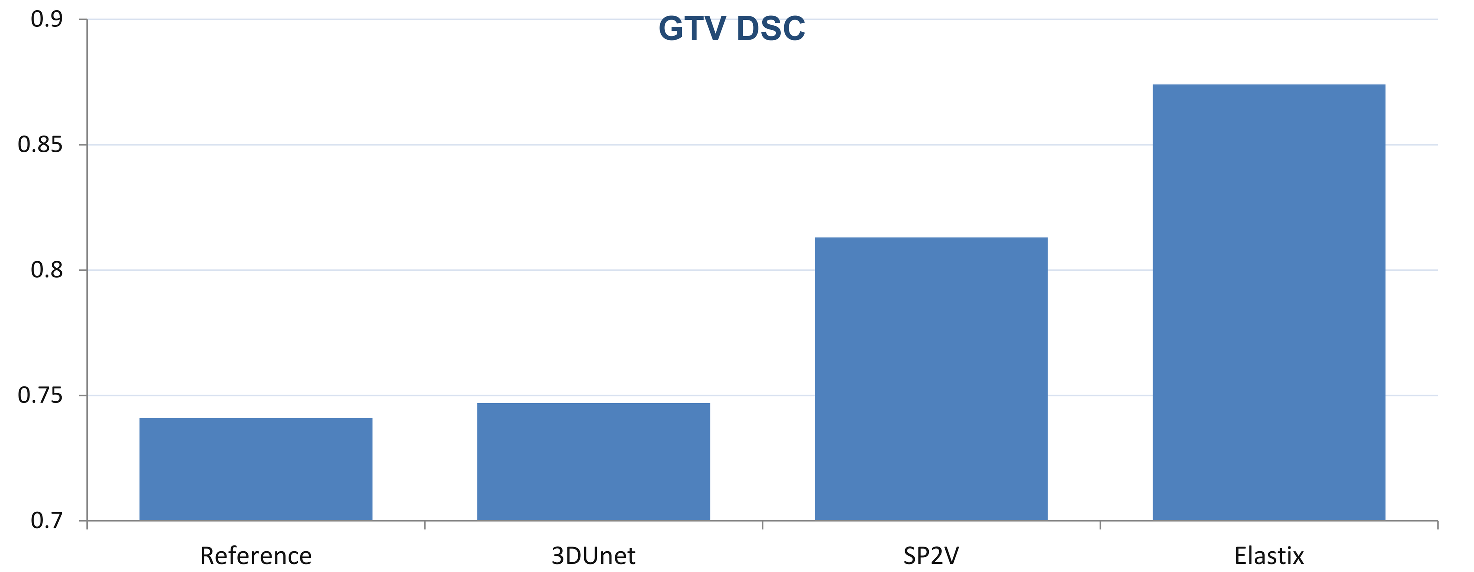


Figure 4. sCT and HU error comparison at the 0% phase.



Results

Best loss: HU MSE + DVF MSE gave the strongest overall SP2V performance.
Real-time inference: <50 ms per reconstruction.
Image quality: MAE 9.88 HU, PSNR 37.4 dB, and SSIM 0.971.
Anatomical accuracy: GTV HD95 3.14 mm, GTV DSC 0.813, and lung DSC 0.975.
Deformation regularity: Only 24.9 negative-Jacobian voxels on average.
Baseline comparison: SP2V outperformed 3DUnet across every reported metric and produced fewer folding voxels than Elastix and TransMorph.

Discussion

Why it works: Paired projections encode respiratory motion; the reference CT supplies patient-specific anatomy. 2D motion extraction and Swin-UNETR preserve both local and long-range deformation patterns.
Clinical relevance: SP2V uses standard on-board CBCT, reconstructs in <50 ms, and requires no patient- or fraction-specific retraining.
Limitations: Validation used simulated 0-degree projections from one center. Multi-platform clinical validation and arbitrary-angle extension are needed.

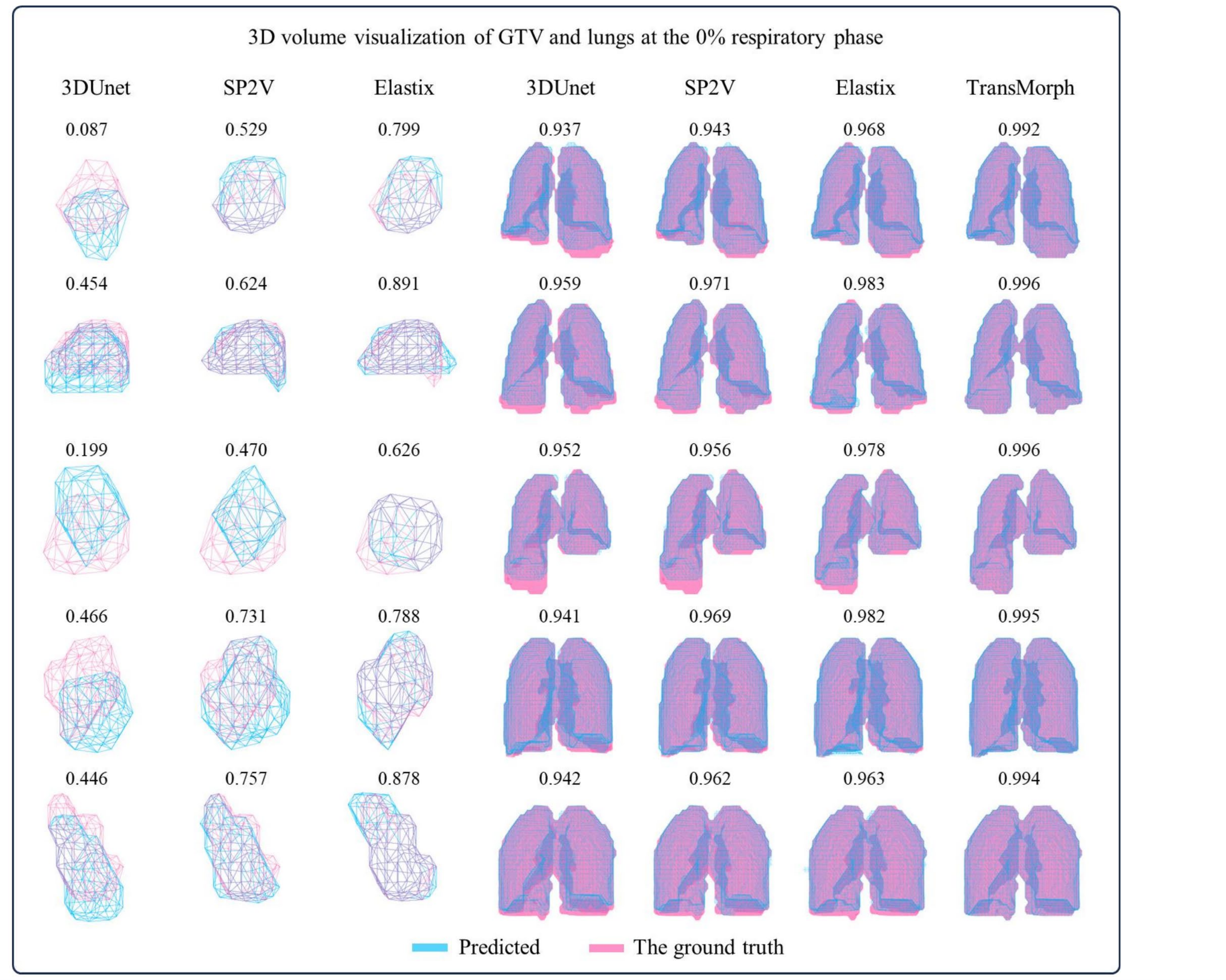


Figure 5. GTV and lung visualization at the 0% phase.

Conclusions

SP2V reconstructs motion-resolved thoracic CT from a single CBCT projection by learning a cross-patient projection-to-deformation mapping constrained by reference anatomy.
High image quality, accurate GTV/lung representation, and physically plausible DVFs support single-projection volumetric guidance for real-time lung radiotherapy.

Contact

Guangjun Li, PhD
West China Hospital, Sichuan University
Email: liguangjun@scu.edu.cn
Chengdu 610041, China

References

1. 1. Keall PJ et al. Med Phys. 2006;33:3874-3900.
2. 2. Dai J et al. Med Image Anal. 2024;91:102998.
3. 3. Miura R et al. IEEE TRPMS. 2025;9:69-82.
4. 4. Chen J et al. Med Image Anal. 2022;82:102615.
5. 5. Unberath M et al. MICCAI. 2018:98-106.