A Patient-Specific Deep Learning Model for Angle-Agnostic Tumor Tracking in Projections for Lung Cancer Radiotherapy

Qiuyi Ye, BEng[1], Xiangyu Zhang, MM[1], Weige Wei, MM[1], Yongchang Wu[1], Sen Bai, MSc[2] and Guangjun Li, PhD[2]
[1] West China Hospital of Sichuan University, Chengdu, China, [2]Department of Radiation Oncology, West China Hospital of Sichuan University, Chengdu, China

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

ABSTRACT

Purpose: Radiotherapy for lung cancer is challenged by respiratory motion and inter-fractional anatomical variations. The respiratory motion management based on tumor tracking can reduce the discrepancies between planned and absorbed doses. However, current clinical tumor tracking methods in clinical practice rely on patients' tolerance and specialized hardware. We developed a patient-specific framework based on deep learning for angle-agnostic tumor tracking in projections of CBCT.

Methods: Multi-fraction 4DCT data from eight lung cancer patients were used. First-fraction 4DCT generated patient-specific training data through DVF-based respiratory motion augmentation and localized Gaussian tumor deformation. Projection images were first generated from CT volumes using DeepDRR and then corrected using CUT. SiamMCF-UNet incorporated adjacent-frame Siamese encoding, multi-modal change fusion, and a CBAM-enhanced decoder.

Results: SiamMCF-UNet achieved the best performance on both subsequent two fractions, with DSC of 0.937 ± 0.013 and HD95 of 3.23 ± 0.96 mm on F2, and DSC of 0.914 ± 0.040 and HD95 of 5.45 ± 4.83 mm on F3.

Conclusion: SiamMCF-UNet showed promising stability and precision for markerless lung tumor tracking from CBCT projections.

CONTACT

Qiuyi Ye
West China Hospital, Sichuan University
yeqiuyi32@gmail.com

INTRODUCTION

Background:
Respiratory motion and inter-fraction anatomical changes remain key challenges in lung cancer radiotherapy. Tumor regression, progression, and surrounding tissue variation can cause planned-delivered dose discrepancies, affecting target coverage and organs-at-risk sparing.

Challenge:
CBCT projections markerless tracking is attractive for conventional linacs, but is limited by weak tumor contrast, anatomical overlap, angle-dependent appearance, and scarce patient-specific training data.

Objective.
We develop a patient-specific, angle-agnostic, projection-domain tracking framework for accurate and real-time markerless lung tumor contour prediction.

Contributions:
(1) **Deformable data generation.**
DVF interpolation and localized Gaussian tumor deformation augment limited 4DCT data with respiratory motion, inter-fraction variation, and tumor morphology changes.
(2) **Adjacent-frame tracking.**
SiamMCF-UNet uses neighboring projections during gantry rotation to exploit temporal tumor-motion cues beyond single-projection tracking.
(3) **Multi-modal change fusion.**
The MCF module jointly models feature difference, interaction, and channel correlation to enhance tumor-related changes and suppress overlapping anatomy.
(4) **Multi-fraction validation.**
Testing on the second and third fractions (F2/F3) evaluates tracking accuracy, inter-fraction robustness, and real-time processing speed.

METHODS AND MATERIALS

Patient-specific data generation.
Multi-fraction 4DCT data from eight lung cancer patients were used[1]. For each patient, F1 generated training data, while F2 and F3 were used for inter-fraction testing.
The 50% respiratory phase was registered to the remaining phases using Elastix, and the resulting deformation fields were used to synthesize intermediate respiratory states.

$$D_{[t,t+1]}(\alpha) = (1-\alpha)D_t + \alpha D_{t+1}$$

A localized Gaussian deformation around the tumor centroid simulated tumor morphology changes. Augmented CT and GTV volumes were projected by DeepDRR [2].DRRs were corrected using CUT [4], and thresholded GTV DRRs served as masks.

SiamMCF-UNet.
Two adjacent projections were input to a shared-weight Siamese encoder to extract paired multi-scale features. The MCF module modeled three inter-frame cues:

$$D_t = F_t - F_{t-1}$$
$$I_t = F_t \odot F_{t-1}$$
$$C_t = Corr(F_t, F_{t-1})$$

The three cues are adaptively fused.

$$\hat{F}_t = W_D D_t + W_I I_t + W_C C_t,$$
$$[W_D, W_I, W_C] = Softmax([w_D, w_I, w_C])$$

This fusion enhances tumor-related changes and suppresses overlapping anatomy. A CBAM-enhanced decoder predicts the projection-domain tumor mask.

Evaluation.
Performance was tested on F2 and F3 using **DSC** and **HD95** across gantry angles and respiratory phases, compared with U-Net and improved U-Net.

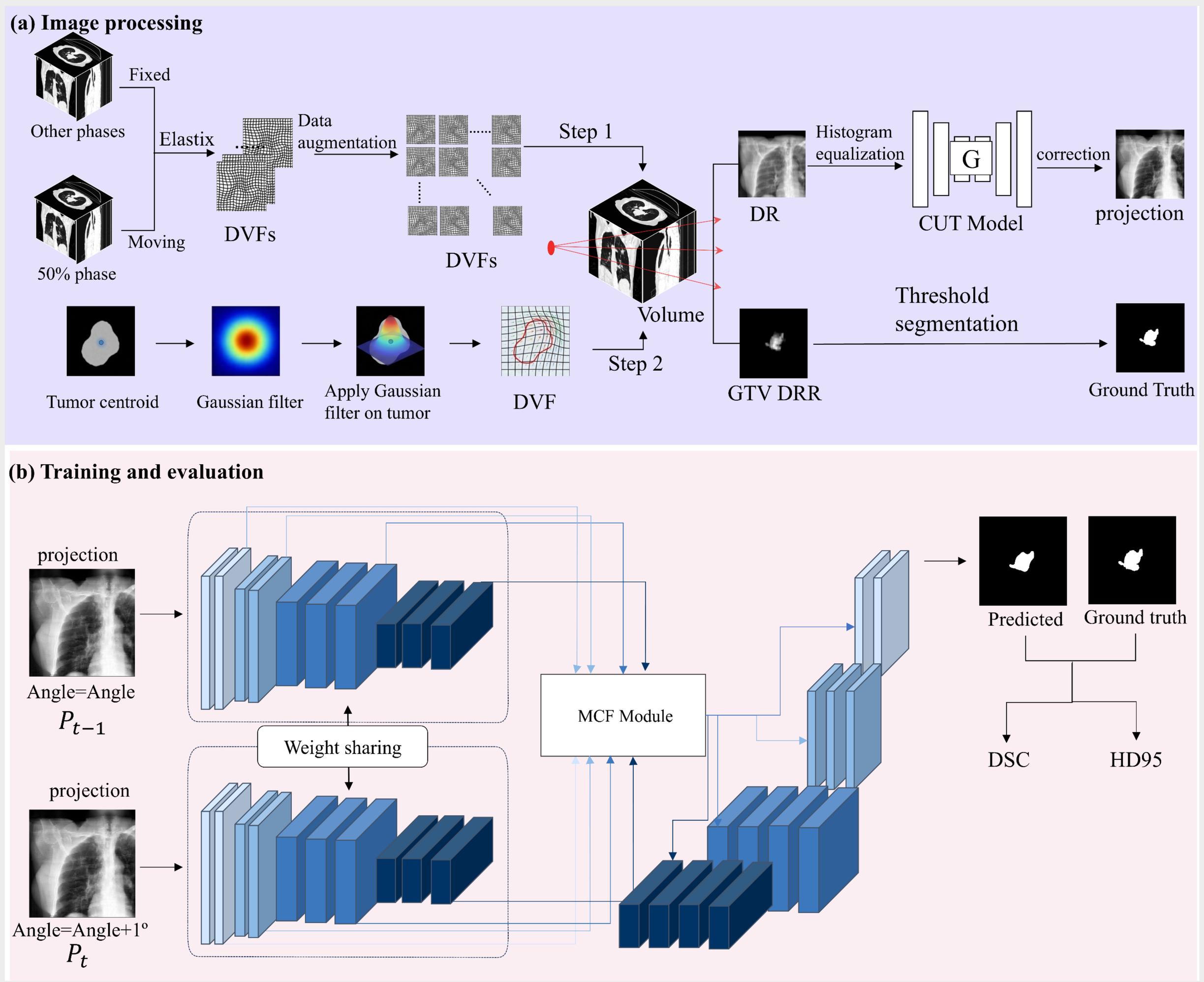

Figure 1. Overall workflow of patient-specific data generation and SiamMCF-UNet tracking.

RESULTS

SiamMCF-UNet achieved the best performance on both F2 and F3, with higher DSC and lower HD95 than U-Net[3] and improved U-Net[5].
The model remained robust across gantry angles and respiratory phases, while showing angle and phase dependent tracking patterns. Accuracy decreased when high-density anatomy overlapped the tumor region.

Fraction	Model	DSC	HD95 (mm)
F2	U-Net	0.761 ± 0.161	16.94 ± 13.96
	Improved U-Net	0.894 ± 0.079	6.10 ± 6.19
	SiamMCF-UNet	**0.937 ± 0.013**	**3.23 ± 0.96**
F3	U-Net	0.727 ± 0.175	19.99 ± 14.37
	Improved U-Net	0.864 ± 0.080	7.92 ± 6.88
	SiamMCF-UNet	**0.914 ± 0.040**	**5.45 ± 4.83**

Table 1. Average tracking performance across eight patients on F2 and F3.

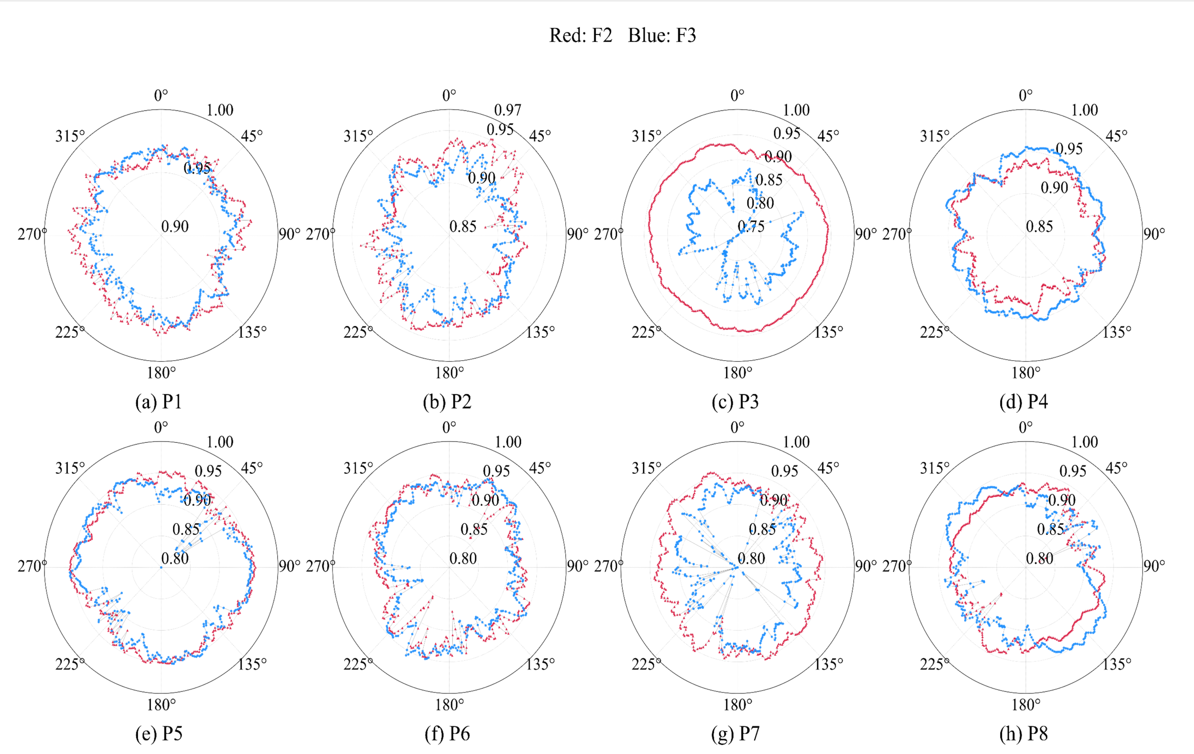

Figure 2. Patient-specific tracking performance across gantry angles.

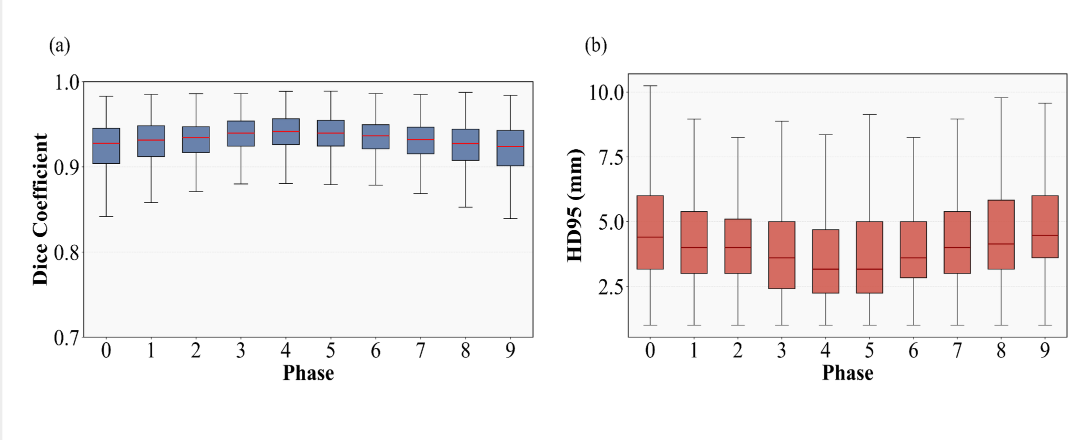

Figure 3 . Tracking performance across respiratory phases.

DISCUSSION

Projection-domain tumor tracking is intrinsically challenging because a single projection provides limited and non-unique tumor information under anatomical overlap. SiamMCF-UNet addresses this task by combining **patient-specific motion priors** from deformable data generation with **adjacent-frame change information** from neighboring projections.
The deformable data generation strategy injects patient-specific respiratory motion and tumor morphology variation into training, rather than relying only on the limited original 4DCT phases. This provides the model with motion and shape patterns that are more consistent with treatment-course variation.
The MCF module explicitly models inter-frame changes using difference, interaction, and channel-correlation features. These complementary cues allow the network to capture tumor-related motion patterns while reducing interference from static or overlapping projection structures.
The maintained performance on both F2 and F3 suggests that SiamMCF-UNet does not merely fit the training fraction, but preserves tracking ability under inter-fraction anatomical variation. This supports its potential for markerless, angle-agnostic, CBCT projection-based tumor tracking on conventional linear accelerators.

CONCLUSIONS

SiamMCF-UNet showed promising performance for markerless lung tumor tracking from arbitrary-angle projections.
With real-time processing speed, this framework shows potential for future CBCT-guided motion-adaptive radiotherapy on conventional linear accelerators.

REFERENCES

[1] Hugo GD, Weiss E, Sleeman WC, Balik S, Keall PJ, Lu J, Williamson JF. A longitudinal four-dimensional computed tomography and cone beam computed tomography dataset for image-guided radiation therapy research in lung cancer. Medical Physics. 2017;44(2):762-771.

[2] Unberath M, Zaech JN, Lee SC, Bier B, Fotouhi J, Armand M, Navab N. DeepDRR: A Catalyst for Machine Learning in Fluoroscopy-Guided Procedures. MICCAI. 2018:98-106.

[3] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI. 2015:234-241.

[4] Park T, Efros AA, Zhang R, Zhu JY. Contrastive Learning for Unpaired Image-to-Image Translation. ECCV. 2020:319-345

[5] Grama D, Dahele M, van Rooij W, Slotman B, Gupta DK, Verbakel WFAR. Deep learning-based markerless lung tumor tracking in stereotactic radiotherapy using Siamese .