Surface-Guided 3D Liver Tumor Tracking for Respiratory Motion Management Using SLOT-Net

Xinyi Fu; Hengjie Liu; Mary Feng; William C. Chen; Alex Hotca-Cho; Mekhail Anwar; Jessica Scholey; Huiming Dong; Ahmad Sakaamini; Wensha Yang; Ke Sheng

Department of Radiation Oncology, University of California, San Francisco

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

University of California San Francisco

TAKE-HOME MESSAGE

Whole anterior-surface motion can localize 3D liver tumors with sub-2-mm accuracy - without ionizing imaging

STUDY AT A GLANCE

157 patients with liver tumors

8 respiratory phases from 4D-MRI

3D tumor mask + centroid prediction

BEST TEST RESULTS

1.54 ± 1.29 mm centroid error

0.86 ± 0.11 Dice score

2.90 ± 1.05 mm HD95

WHAT PRIORS FROM REFERENCE PHASE HELP MOST?

1. Jacobian determinant
2. End-exhale diaphragm mask

WHICH SURFACE REGIONS HELP MOST?

1. Abdominal center
2. Lateral body

CONTACT

Xinyi.Fu@ucsf.edu
Ke.Sheng@ucsf.edu

Keywords: liver tumor · SGRT · markerless tracking · 4D-MRI · uncertainty

INTRODUCTION

CLINICAL NEED Sparse markers and small surface patches fail to reliably reflect internal liver motion across patients or breathing states

RESEARCH GAPS Limited high-quality 4D MRI data for 3D liver motion research

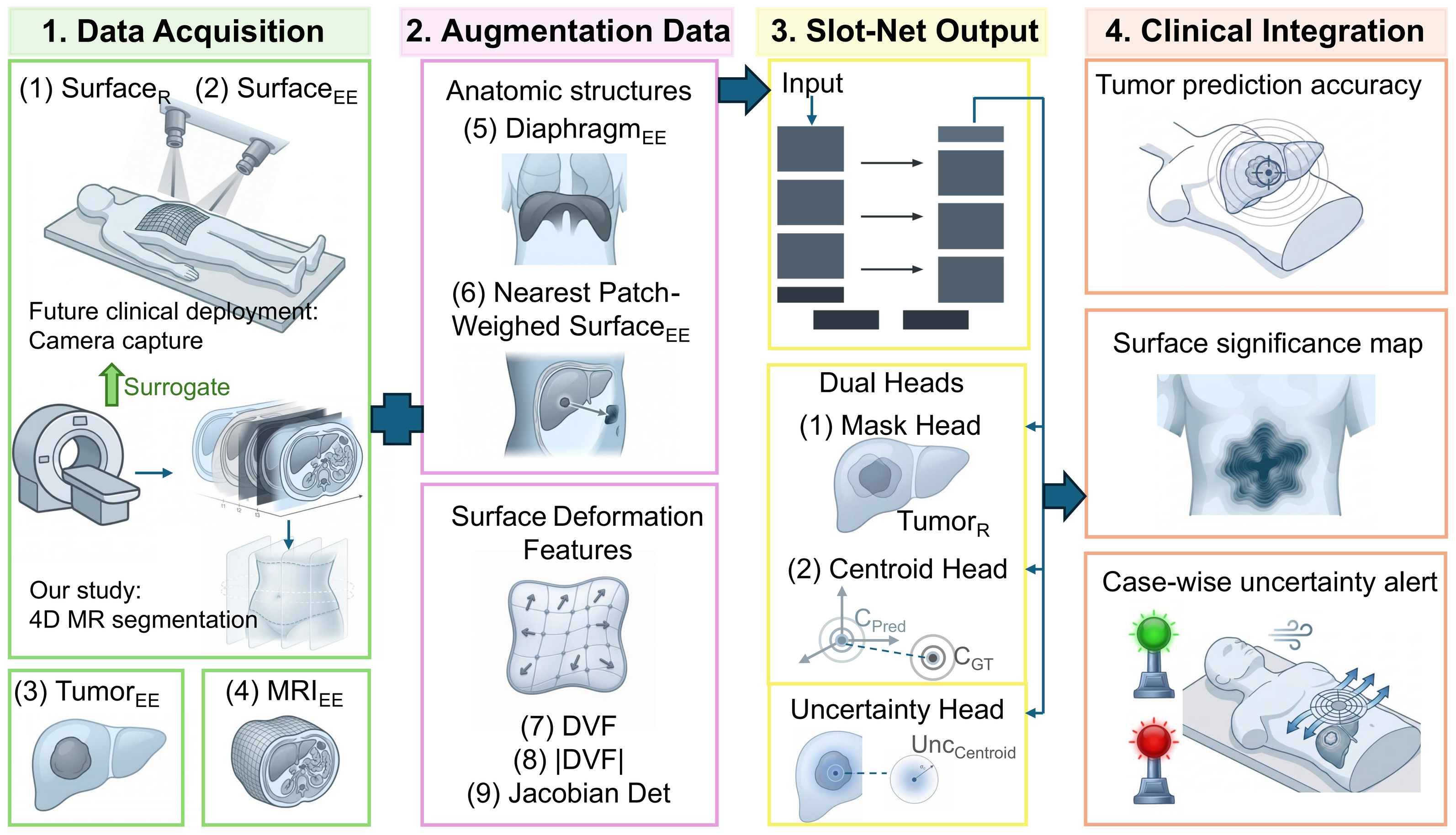


Fig 1. Overview of the proposed framework for 3D internal liver tumor localization using skin surface. The workflow includes multimodal data acquisition and processing, Slot-Net training with muti-head outputs, and downstream analyses of localization accuracy, surface contribution and prediction uncertainty. Abbreviations: R=real-time; EE=end-exhalation (reference phase).

METHODS

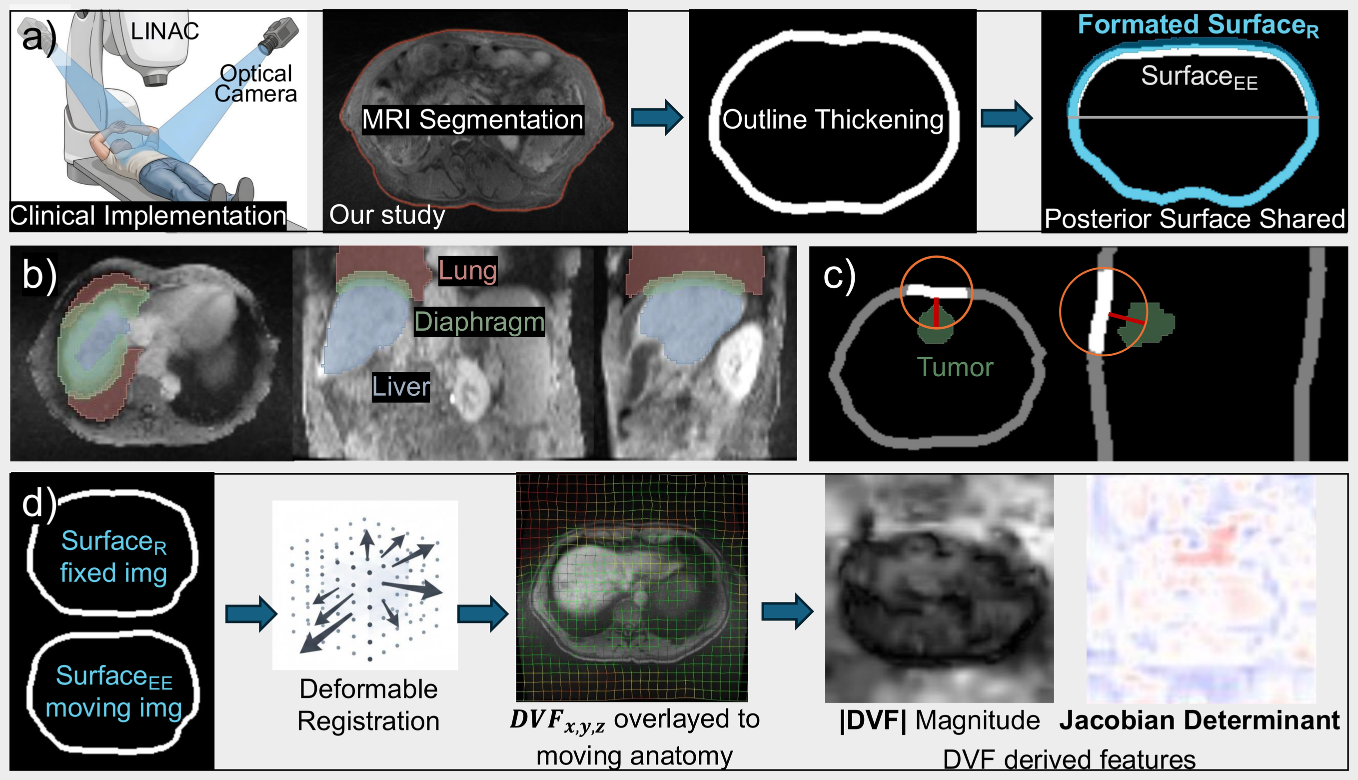


Fig 2. Multimodal data preparation: a) Skin surface. b) Diaphragm mask. c) Nearest patch weighted surface. d) Surface deformable registration derived features.

LOSS FUNCTION

$$L_{total} = L_{Mask} + L_{Centroid} + L_{Uncertainty} \quad (1)$$

$$L_{Mask} = w_{BCE} \times L_{BCE} + w_{Dice} \times L_{Dice} + w_{CenDistM} \times L_{CenDistM} \quad (2)$$

$$L_{Centroid} = w_{KL} \times L_{KL} + w_{CenDistC} \times L_{CenDistC} \quad (3)$$

$$L_{Uncertainty} = w_{Uncertainty} \times \left(\frac{(y-\mu)^2}{\exp(s)} + s\right) \quad (4)$$

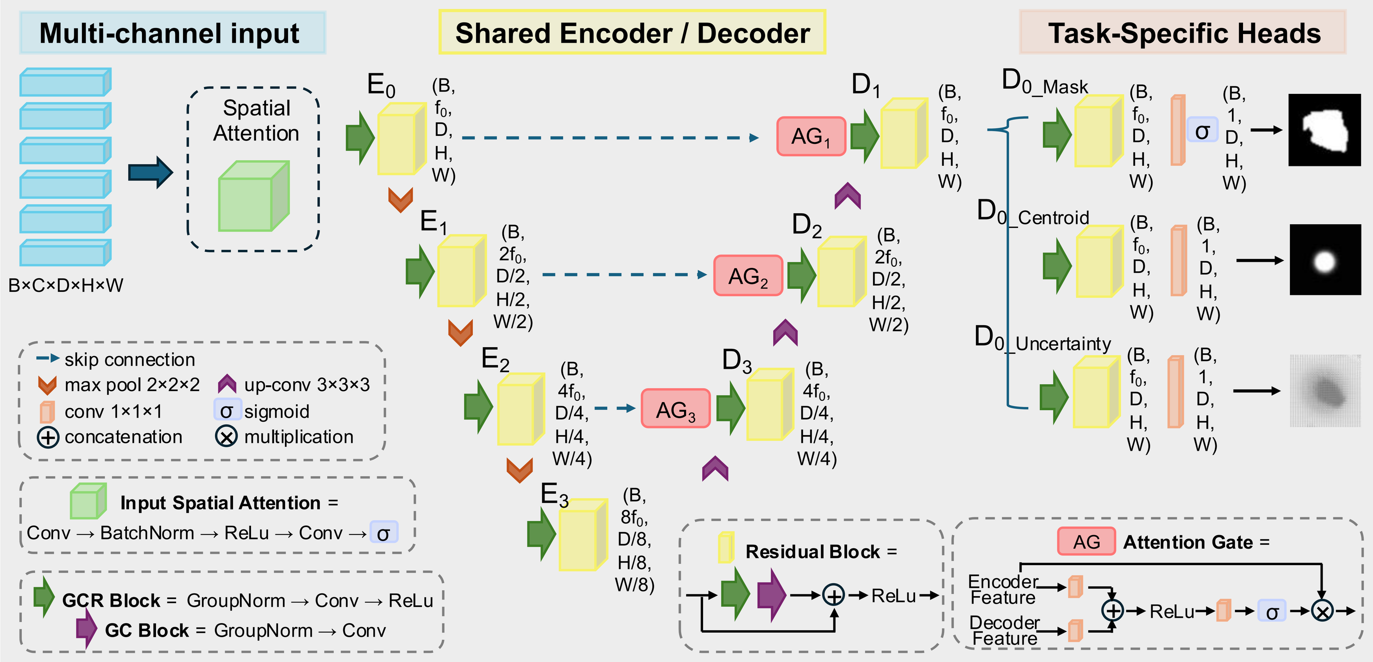


Fig 3. Illustration of the SLOT-Net. The model is based on a 3D U-Net structure with integrated attention gates, inspired by the design described in (Zhang et al., 2025). Optional network variants include the input-attention module and the uncertainty head.

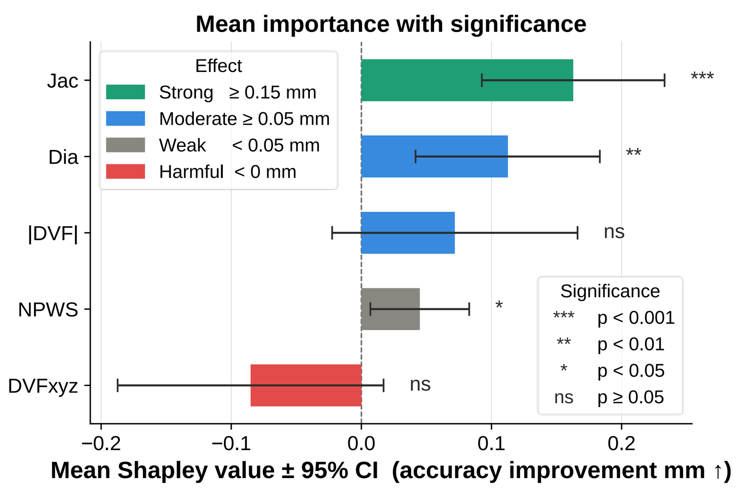
REFERENCES

1. Mao, W., et al, 2022. Association Between Internal Organ/Liver Tumor and External Surface Motion From Cine MR Images on an MRI-Linac. *Front. Oncol.*
2. Zhang, G., et al, 2025. Optical surface imaging-driven tumor tracking with deformable image registration-enhanced deep learning model for surface-guided radiotherapy. *Biomedical Signal Processing and Control*
3. Akinci D'Antonoli, et al, 2025. TotalSegmentator MRI: Robust Sequence-independent Segmentation of Multiple Anatomic Structures in MRI. *Radiology*

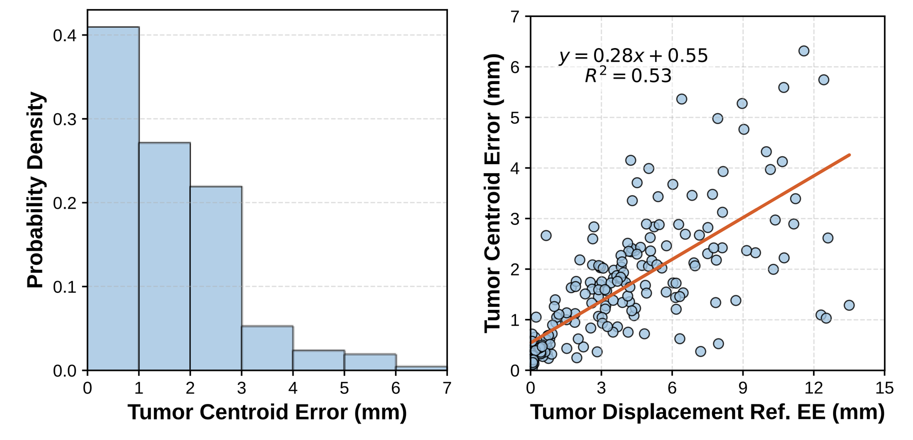
RESULTS

TUMOR PREDICTION ACCURACY (30×7=210 test cases)

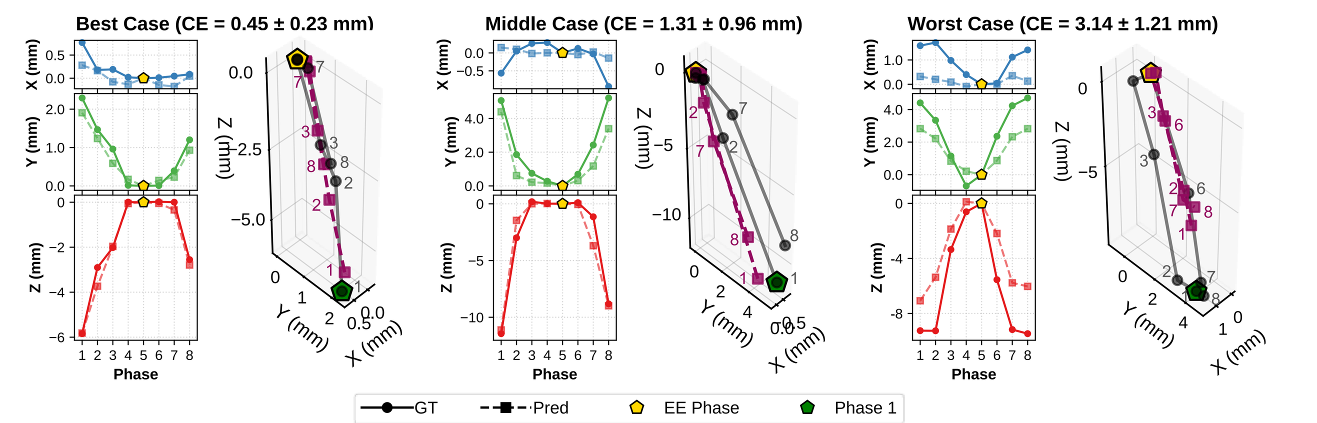
Feature importance analysis



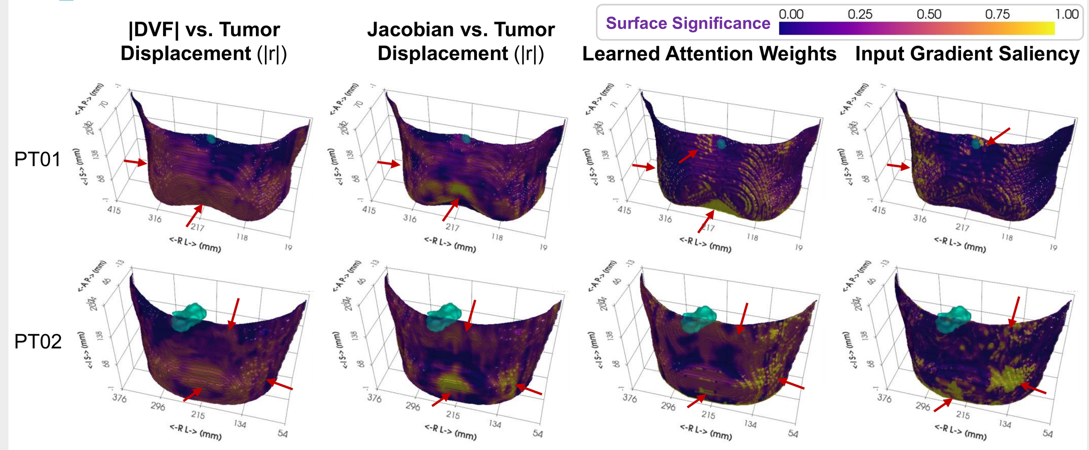
Tumor centroid error analysis from the best model



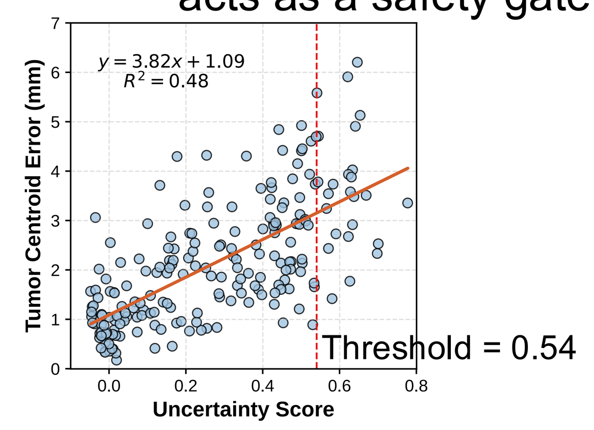
Trajectories of 3 test tumor cases



SURFACE SIGNIFICANCE is patient-specific



UNCERTAINTY acts as a safety gate



DISCUSSION

Unlike lung tumors, which typically exhibit a consistent correlation with chest wall motion, liver tumor motion demonstrates a complex, nonlinear, and patient-specific relationship with the abdominal wall. Using 4D MRI, this study demonstrates that the entire anterior abdominal surface can be used to predict liver tumor motion with high spatial accuracy (tumor centroid error of only 1.54 ± 1.29 mm) across all respiratory phases.

This high accuracy was largely driven by the proposed dual-head output architecture. The mask head focuses on high-frequency local information (edges and textures), while the Gaussian centroid distribution head—optimized via KL divergence—imposes a global topological constraint. Our inclusion of an attention mechanism also helps.

Evaluating the input configurations revealed that the Jacobian determinant derived from surface-mask registration (quantifies local surface expansion and contraction) provided the greatest performance boost, followed by the end-exhale diaphragm mask (separates the chest and abdomen and establishes the tumor's relative location).

Anatomically, the correlation between internal organ motion and the *center abdominal* and *lateral skin surfaces* is primarily driven by the mechanical influence of the diaphragm and internal abdominal pressure, but their exact distribution varies based on individual patient anatomy and respiratory mechanic patterns. Liver tumors are subjected to complex internal pressures and tissue inertia, which often results in a severe phase shift between the external surface motion and the internal tumor trajectory.

Limited-marker or small-patch approaches may perform well only in selected patients in whom the chosen surface region maintains a strong and stable respiratory correlation but can fail substantially in others. In contrast, using the entire anterior surface provides a more robust, generalized representation of respiratory motion.

Limitation: 4D MRI fails to capture instantaneous irregularities (extreme breathing, baseline drift, intestinal gas)

Future workflow: optical cameras → continuous tracking → confidence-aware beam adaption

CONCLUSIONS

- We developed a deep learning framework to robustly infer 3D positions of liver tumors based on the entire anterior abdominal skin surface extracted from 4D MRI.
- By statistically investigating the spatially variant surface contributions, this work sheds light on patient-specific surface selection or weighting for future clinical applications.
- Case-wise uncertainty is quantified by assigning a confidence score to each case, providing additional information for improved safety in motion management.