Quantitative MR Fingerprinting for Synthetic CT via a Deformably-Registered Fusion Transformer Network

Yimin Ni[1], Chenyang Liu[1], Wen Li[1], Ho-Fun Victor Lee[2], Aya El Helali[2], Yat-Lam Wong[3], Xiang Wang[1], Peilin Wang[1], Ge Ren[1], Jing Cai[1], Tian Li[1]

[1]Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong SAR, China
[2]Department of Clinical Oncology, The University of Hong Kong, Hong Kong SAR, China
[3]Department of Clinical Oncology, Queen Mary Hospital, Hong Kong

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

ABSTRACT

Purpose: To develop and validate a deep learning framework for high-accuracy pseudo-CT synthesis from rapid, multi-parametric Magnetic Resonance Fingerprinting (MRF) for liver radiotherapy. This work addresses the technical challenge of translating quantitative tissue property maps into electron density information while ensuring spatial fidelity through integrated motion compensation.

Methods: We present a Deformably-Registered Fusion Transformer (DRFT) network for MRF-to-CT synthesis. The DRFT architecture uniquely combines two key components: 1) a cascaded transformer-based fusion encoder that jointly processes T1, T2, and proton density (PD) maps to extract complementary features, and 2) a trainable, diffeomorphic registration sub-network that compensates for respiratory motion between MRF and planning CT scans within the training pipeline. The model was developed using a cohort of 24 hepatocellular carcinoma patients. Performance was benchmarked against U-Net and pix2pix cGAN baselines using structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), and mean absolute error in Hounsfield Units (MAE).

Results: The DRFT framework achieved state-of-the-art synthetic CT quality, yielding an SSIM of 0.914 ± 0.011, a PSNR of 28.55 ± 1.31, and an MAE of 63.2 ± 7.8 HU, outperforming all baseline models. Multi-parametric MRF input was essential; the full T1+T2+PD model significantly surpassed the best single-parameter (PD-only SSIM: 0.885) and dual-parameter models (best dual-input SSIM: 0.894). Integrating the trainable registration module was critical, improving SSIM by 0.028 compared to an unregistered variant. Visual analysis confirmed enhanced preservation of bone anatomy and soft-tissue interfaces.

Conclusion: This study demonstrates the efficacy of a transformer-based fusion network with embedded deformable registration for generating synthetic CT from quantitative MRF data.

CONTACT

Yimin Ni
The Hong Kong Polytechnic University
Yimin01.ni@connect.polyu.hk

INTRODUCTION

This study presents a significant technical innovation for MR-guided radiotherapy by introducing the first deep learning framework for generating synthetic CT (sCT) from quantitative Magnetic Resonance Fingerprinting (MRF). The model, termed the Deformably-Registered Fusion Transformer (DRFT), uniquely addresses two critical technical barriers. First, it leverages the intrinsic, reproducible multi-parametric tissue properties (T1, T2, PD) from a single rapid MRF scan, moving beyond the scanner-dependent qualitative contrasts of conventional MRI. Second, it integrates a trainable, diffeomorphic registration sub-network (Fig. S1) directly into the training pipeline to proactively correct for respiratory motion between the MRF and CT scans, ensuring the synthesis network learns from anatomically aligned data. **The clinical impact is the establishment of a robust, unified pathway toward a truly quantitative MR-only workflow for abdominal sites.** By generating high-fidelity sCTs, this framework has the potential to eliminate systematic errors from MR-CT registration, spare patients an extra CT simulation dose, and utilize the functional tissue characterization of MRF for potentially adaptive treatment strategies, beginning with hepatocellular carcinoma.

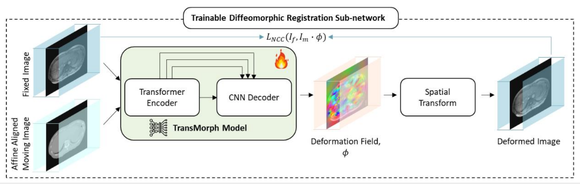

Figure 1. Schematic of the trainable, diffeomorphic registration sub-network based on the TransMorph architecture.

METHODS AND MATERIALS

The model was developed using a cohort of 24 hepatocellular carcinoma patients. Each patient underwent a planning CT and a 2D MRF acquisition using a fast imaging with steady-state precession (FISP) sequence with a spiral readout [1]. The MRF data was matched to a simulated dictionary to yield coregistered quantitative T1, T2, and Proton Density (PD) parameter maps. Data was split into training (n=17) and testing (n=7) sets.

The proposed DRFT network consists of two integrated components: (1) Trainable Diffeomorphic Registration Sub-network (Fig. 1): Prior to synthesis, this module takes the moving MRF-derived features and the fixed CT as input. Based on TransMorph [2], it computes a smooth, invertible diffeomorphic deformation field, which warps the MRF features into spatial alignment with the CT domain. This sub-network was trained end-to-end with the synthesis network using a combination of local normalized cross-correlation and bending energy loss. (2) Cascaded Transformer-based Fusion Encoder (Fig. 2): The T1, T2, and PD maps are processed through separate initial convolutional pathways. Their features are then fused and refined through a cascaded series of transformer blocks [3]. This architecture is designed for modality-aware feature extraction and synergistic fusion of complementary quantitative information from the three parametric maps.

RESULTS

The DRFT framework demonstrated superior performance in generating synthetic CTs from MRF data. Quantitatively, it achieved state-of-the-art results with an SSIM of 0.914 ± 0.011, a PSNR of 28.55 ± 1.31, and an MAE of 63.2 ± 7.8 HU, outperforming both the U-Net (SSIM: 0.863) and pix2pix cGAN (SSIM: 0.884) baselines.

The ablation study on multi-parametric input confirmed the necessity of utilizing the full suite of quantitative MRF maps. The model using all three parameters (T1+T2+PD) significantly surpassed configurations using only PD maps (SSIM: 0.885) or the best dual-parameter combination of PD+T1 (SSIM: 0.894).

The ablation on registration module demonstrated that the integrated registration sub-network provided a critical improvement in anatomical fidelity. A variant of the DRFT trained without this module exhibited notably lower performance (SSIM: 0.886). The inclusion of the module improved the SSIM by 0.028, underscoring that explicit, learned compensation for respiratory motion is essential for accurate synthesis in the abdomen, where misalignment between scans is substantial.

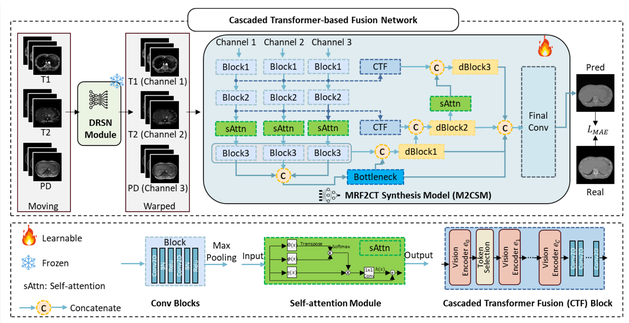

Figure 2. Diagram of the proposed cascaded transformer-based fusion encoder for synergistic multi-parametric MRF feature integration.

The visual quality of the sCTs generated by the full DRFT model showed strong correspondence with the ground truth. As shown in a representative case (Fig. 3), the synthesized image preserved the overall anatomical structure.

The general morphology of bony anatomy, such as the spine and ribs, was captured, and the major soft-tissue interfaces, including the liver boundary, were identifiable. However, some localized inaccuracies were observed, particularly in regions of fine anatomical detail or heterogeneous tissue texture, where the synthetic image exhibited minor blurring or intensity variations compared to the planning CT.

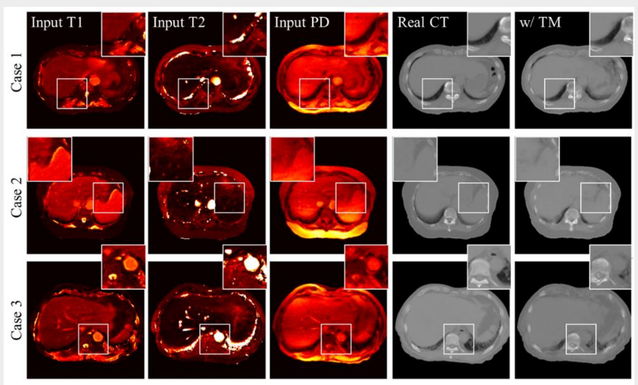

Figure 3. Representative synthetic CT generated by the DRFT model compared to the ground-truth planning CT and the input MRF parameter maps.

DISCUSSION

We also evaluated the dosimetric accuracy of the synthetic CT by performing dose calculation in clinical treatment planning system (TPS).

Figure 4 presents the dosimetric comparison between treatment plans calculated using the clinical planning CT (ground truth) versus those calculated using our synthetic CT. Across axial, sagittal, and coronal views, the dose distributions are nearly identical, exhibiting excellent spatial agreement in terms of isodose line alignment, high-dose region localization, and steep dose gradient positions—particularly at the tumor-normal tissue interface.

Quantitatively, the gamma passing rate between the two plans reaches 98.99% under the standard clinical criterion of 3%/3 mm, and 98.18% under the more stringent 3%/2 mm criterion—both substantially exceeding the typical clinical tolerance of 95%. Even at the aggressive 2%/2 mm criterion, the passing rate remains high at 94.61%, approaching clinical acceptability for institutions employing stricter quality assurance protocols.

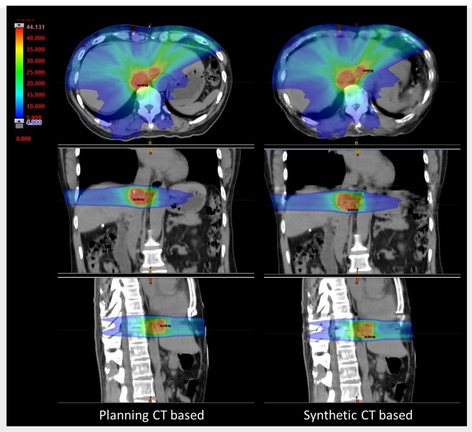

Figure 4. Comparison between planning CT-based and synthetic CT-based dose distribution of a test patient.

CONCLUSIONS

This study demonstrates the efficacy of a transformer-based fusion network with embedded deformable registration for generating synthetic CT from quantitative MRF data.

The results confirm the necessity of synergistic multi-parametric guidance and explicit motion modeling for anatomical accuracy in the abdomen.

This framework provides a robust technical foundation for quantitative MR-only treatment planning for liver cancers.

REFERENCES

1. Jiang et al., MR Fingerprinting Using FISP with Spiral Readout, Magn Reson Med 2015.
2. Chen et al., TransMorph: Transformer for Unsupervised Medical Image Registration, Med Image Anal 2022.
3. Zhang et al., ASCFormer: Adaptive Structure-Aware Cascaded Transformer, IEEE TCSVT 2025.