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

Towards Registration-Free PET-CT Enhancement via Resolution-Aware Latent Diffusion

Yike Guo[1], Yi Luo[1], Junqi Liu[1], Zongwei Zhou, PhD[1] and Kai Ding, PhD[1]
[1]Johns Hopkins University

ABSTRACT

Purpose: While PET-CT imaging holds promise for simulation-free radiotherapy workflows, its inherent image resolution limits its use for accurate tumor and organ-at-risk (OAR) contouring. This study aims to enhance the spatial resolution of PET-CT by leveraging a resolution-aware latent diffusion model guided by high-resolution but limited field of view (FOV) chest CT images, without requiring spatial alignment between modalities.

Methods: We propose a three-stage latent diffusion framework. Firstly, a medical variational autoencoder (VAE) was trained to encode PET-CT and chest CT slices into a compact latent space. Secondly, a resolution-aware latent diffusion model was trained using textual prompts highlighting image resolution learn semantic representations of image resolution. Chest CT and synthetically degraded chest CT slices were used to learn high- and low-resolution priors, while PET-CT slices served as additional low-resolution examples. Thirdly, a conditional diffusion model was trained to enhance PET-CT latent embeddings. To preserve anatomical consistency during enhancement, we incorporated structural guidance losses, including a segmentation loss, the sum of squared vesselness measure difference (SSVMD) and a cycle consistency loss. Enhanced PET-CT images were then decoded using the pretrained autoencoder. Although chest CT and PET-CT were not spatially registered, the semantic priors learned from the second stage enabled registration-free enhancement. The model was trained on 150 patients and tested on 20 patients.

Results: Enhanced PET-CT images demonstrated improved visual quality compared with raw PET-CT, with clearer anatomical boundaries, reduced noise, and better structural continuity in lung regions. These improvements suggest the proposed framework can enhance spatial resolution while preserving anatomical fidelity in a registration-free setting. Conclusions: We present a registration-free PET-CT enhancement framework based on resolution-aware latent diffusion. By leveraging semantic resolution priors and structural guidance losses, the proposed method improves image resolution and anatomical consistency, offering a promising direction toward simulation-free radiotherapy planning.

CONTACT

Kai Ding and Yike Guo
Johns Hopkins University
733 N. Broadway, Baltimore, MD 21205
kai@jhu.edu and yguo122@jh.edu
4109557391

INTRODUCTION

Simulation-free radiotherapy workflows increasingly rely on PET-CT for treatment planning, since it can eliminate a separate CT simulation step and streamline the imaging pipeline.[1] However, PET-CT's limited spatial resolution restricts its effectiveness for accurate delineation of tumors and organs-at-risk (OARs). Chest CT, in contrast, offers high spatial resolution but typically covers only a limited field of view (FOV) and is not always co-registered with PET-CT.

This creates a practical gap: the modality most convenient for simulation-free workflows (PET-CT) is not resolution-adequate for contouring, while the modality with adequate resolution (chest CT) is not always available or spatially registered. Our goal is to bridge this gap by enhancing PET-CT resolution using complementary information from chest CT, without requiring the two modalities to be registered, which removes a major practical barrier to combining information across scans with different fields of view.

METHODS AND MATERIALS

We developed a three-stage latent diffusion framework for registration-free PET-CT enhancement (Figure 1).

Stage I: Latent Encoding: A medical variational autoencoder (MedVAE) is trained to project PET-CT and chest CT slices into a shared, compact latent space.

Stage II: Resolution-Aware Training: High-resolution chest CT slices are paired with textual prompts such as "high-resolution CT slice," while degraded chest CT and PET-CT slices are paired with "low-resolution" prompts. A latent diffusion model[2] is trained on these prompt-image pairs, learning a latent embedding space that encodes resolution semantics without spatial alignment between chest CT and PET-CT.

Stage III: PET-CT Enhancement: PET-CT latents are enhanced using a diffusion model conditioned on the resolution-aware prompt embeddings from Stage II. To preserve anatomical fidelity, three structural losses are applied: (1) a segmentation loss using nnU-Net-generated organ masks, (2) the Sum of Squared Vesselness Measure Difference (SSVMD) to preserve fine vascular structures in the lungs,[3] and (3) a cycle consistency loss to stabilize the latent space.

The framework was trained on 150 paired PET-CT/chest CT cases from Johns Hopkins Hospital and evaluated on 20 held-out patients. PET-CT and chest CT were not spatially registered at any stage, so the model learns cross-modality resolution priors in a fully registration-free setting.

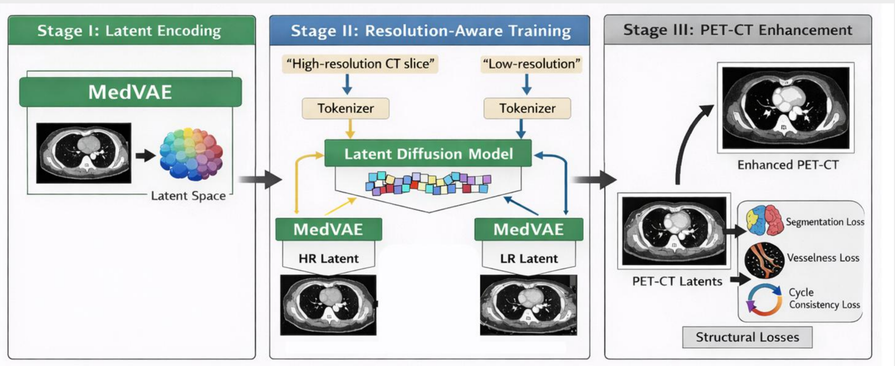

Figure 1. Three-stage resolution-aware PET-CT enhancement framework. Stage I encodes PET-CT and CT into a shared latent space; Stage II learns resolution semantics via prompt-guided training; Stage III enhances PET-CT latents under structural guidance, without registration to chest CT.

RESULTS

Enhanced PET-CT images showed consistently improved visual quality compared with raw PET-CT (Figure 2). Across the 20 held-out test cases, enhanced images exhibited clearer anatomical boundaries, reduced background noise, and improved soft-tissue contrast, particularly in the lung and mediastinal regions. Fine vascular structures appeared more continuous, and organ boundaries were more clearly distinguishable. These qualitative improvements were achieved despite PET-CT and chest CT never being spatially registered, supporting the framework's ability to transfer resolution information across modalities using semantic, prompt-guided priors alone.

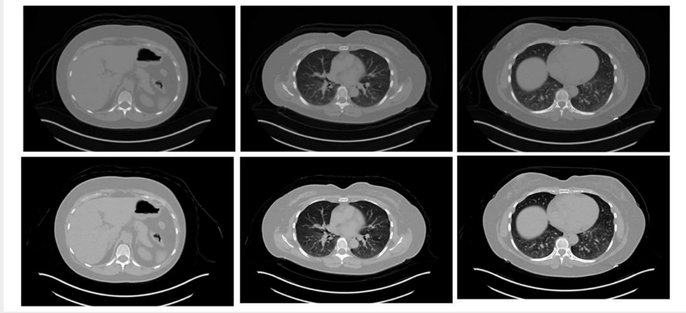

Figure 2. Visuaal comparison between raw and enhanced PET-CT slices. Top: Raw PET-CT slices showing limited contrast and fuzzy anatomical boundaries. Bottom: Enhanced PET-CT slices generated by our resolution-aware latent diffusion model, exhibiting improved anatomical clarity, enhanced vascular visibility, and noise reduction across lung and mediastinal regions.

DISCUSSION

These results suggest that resolution priors can be transferred between imaging modalities purely through semantic, prompt-guided learning, without deformable or rigid registration, a step that is often unreliable or unavailable when two modalities differ substantially in field of view, as with PET-CT and chest CT. By combining prompt-conditioned diffusion with explicit structural guidance (segmentation, vesselness, and cycle-consistency losses), the framework aims to enhance apparent spatial resolution while constraining the enhancement to remain anatomically faithful to the original PET-CT, an important safeguard against hallucinated structures in generative enhancement methods.

Future work will focus on quantitative validation using image quality metrics and downstream task performance (e.g., segmentation accuracy on enhanced versus raw PET-CT), as well as evaluation across a larger, more diverse patient cohort. If validated, this approach could support more accurate contouring directly on PET-CT, reducing reliance on an additional high-resolution CT simulation scan.

CONCLUSIONS

We present a registration-free PET-CT enhancement framework based on resolution-aware latent diffusion. By leveraging semantic resolution priors learned through prompt-guided training and structural guidance losses, the proposed method improves PET-CT spatial resolution while preserving anatomical consistency — offering a promising direction toward simulation-free radiotherapy planning and more accurate PET-CT-based contouring.

REFERENCES

1. Hooshangnejad H, Chen Q, Feng X, Zhang R, Farjam R, Voong KR, Hales RK, Du Y, Jia X, Ding K. DAART: A deep learning platform for deeply accelerated adaptive radiation therapy for lung cancer. Front Oncol. 2023;13. doi:10.3389/fonc.2023.1201679
2. Rombach R, Blattmann A, Lorenz D, Esser P, Ommer B. High-resolution image synthesis with latent diffusion models. arXiv. 2021. https://arxiv.org/abs/2112.10752
3. Cao K, Ding K, Christensen GE, Reinhardt JM. Tissue volume and vesselness measure preserving nonrigid registration of lung CT images. In: Dawant BM, Haynor DR, eds. Medical Imaging 2010: Image Processing. Vol 7623. SPIE; 2010:762309. doi:10.1117/12.844541