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

SU-315-GPD-T-657

Generative Priors for CBCT Reconstruction: A 3D Conditional Diffusion Model for Prostate Cancer CT

LiCheng Kuo, Yabo Fu, Jason Fan, Jiening Zhu, Jean M Moran, Laura A Cervino, Xiang Li, Tianfang Li

Department of Medical Physics, Memorial Sloan Kettering Cancer Center, New York, NY

ABSTRACT

Purpose: Inferior CBCT quality can compromise anatomy visualization during Image-Guided Radiotherapy (IGRT), increasing uncertainty in target localization and organ-at-risk positioning. Improving CBCT reconstruction can enable more reliable daily guidance and adaptive decisions, potentially improving target coverage and normal tissue sparing. This study develops a 3D conditional patch diffusion model trained on planning CT (PCT) to generate robust priors for CBCT reconstruction.

Methods: A total of 500 PCT scans were resampled to 256 × 256 pixels in the axial plane with 2-mm isotropic voxel spacing and used to train a 3D conditional diffusion model that outputs 64 × 64 × 64 patches. The model starts from a noisy image and iteratively denoises the input, conditioned on an input patch that may correspond to a clean image, a noisy image, or an image with artifacts, such as those from a limited-angle CBCT scan. In this study, the model was trained to learn realistic prostate anatomical priors. Evaluation was performed on an independent set of 20 clinical cases, including 10 PCT and 10 CBCT cases, under clean and 90° limited-angle conditions. Image quality was quantified using the structural similarity index measure (SSIM), peak signal-to-noise ratio (PSNR), mean squared error (MSE), mean absolute error (MAE), and rigid registration error relative to the corresponding clean ground truth.

Results: The model demonstrated high fidelity and noise robustness across modalities. For clean images, the PCT cohort achieved SSIM 0.84 ± 0.08 and PSNR 14.63 ± 6.47 dB, while CBCT cohort achieved SSIM 0.78 ± 0.12 and PSNR 14.55 ± 8.33 dB. Under 90^0 limited angle conditions, PCT cohort yielded SSIM 0.86 ± 0.11 and PSNR 19.64 ± 12.58 dB; CBCT cohort achieved SSIM 0.83 ± 0.114 and PSNR 19.8 ± 12.74 dB. MSE and MAE remained consistently low for the outputs.

Conclusion: We developed a 3D conditional patch diffusion model that generates high-quality prostate anatomical priors with consistent performance on PCT and CBCT, including under noise. This approach is well-suited for regularizing iterative CBCT reconstruction to improve image guidance in prostate radiotherapy, supporting more reliable localization and adaptation.

Memorial Sloan Kettering Cancer Center

CONTACT

LiCheng Kuo, M.Sc.
Memorial Sloan Kettering Cancer Center
Email: kuol@mskcc.org
Phone: (516) 559-1557

This research was funded in part through the NIH/NCI Cancer Center support Grant/Core Grant (P30 CA008748)

INTRODUCTION

Background:

Cone-beam CT (CBCT) enables 3D visualization of patient anatomy and is crucial in modern image-guided radiotherapy (IGRT). However, it often suffers from inferior image quality due to:

- Severe noise and scatter artifacts
- Limited angle / sparse-view acquisitions
- Over-smoothing from conventional regularization (e.g., total variation)

These limitations increase uncertainty in target localization and OAR positioning during IGRT, compromising patient care.

Our Approach:

We investigate a 3D conditional diffusion model (CDM) [1] to generate realistic volumetric anatomy priors conditioned on degraded inputs. By training on high-quality planning CT (PCT) data, our model captures complex anatomical dependencies in the prostate region — enabling robust prior generation even from noisy or incomplete CBCT data.

Goal:

Improve CBCT reconstruction quality to support reliable daily image guidance and adaptive radiotherapy decisions.

METHODS AND MATERIALS

Data & Preprocessing:

Training: 500 prostate PCT scans
- Resampled to 256×256 pixels axially at 2 mm isotropic resolution
- HU: range :-1024 to 3072, normalize between 0.0 to 1.0
- 3D patches of 64×64×64 voxels randomly sampled
- Training Conditions:
  - Clean: (Original PCT image)
  - Noisy: add Random Gaussian Noise
  - Limited Angle : Simulated 90^0 limit angle reconstructed PCT

Evaluation: 20 independent clinical cases
- 10 PCT + 10 CBCT scans
- Tested under clean, and limited angle conditions

Model Architecture and Sampling:

- Patch provided as second input channel alongside initial random noise (Figure 1)
- Iterative denoising conditioned on PCT patch data
- Trained on both clean, noisy (Gaussian noise added), and 90^0 limited angle inputs to learn robust anatomy priors
- DDIM sampling with 100 steps

Model Evaluation: Image Quality Metrics

- SSIM — Structural Similarity Index Measure
- PSNR — Peak Signal-to-Noise Ratio (dB)
- MSE — Mean Squared Error
- MAE — Mean Absolute Error
- 6 degrees Rigid Image Registration (RIR) metrics

RESULTS

The model demonstrated exceptional stability and noise robustness across modalities.

For clean images, the PCT cohort achieved SSIM 0.84 ± 0.08 and PSNR 14.63 ± 6.47 dB, while CBCT cohort achieved SSIM 0.86 ± 0.11 and PSNR 19.64 ± 12.58 dB.

Under 90^0 limited angle conditions to test robustness as a prior, PCT priors yielded SSIM 0.78 ± 0.12 and PSNR 14.55 ± 8.33 dB; CBCT priors achieved SSIM 0.83 ± 0.11 and PSNR 19.83 ± 12.74 dB.

Mean Squared Error (MSE) and Mean Absolute Error (MAE), translation and rotation error remained consistently low, showing effective noise suppression while preserving critical anatomical structures.

Quantitative performance is summarized in Table 1, and visual performance on 90^0 limited angle PCT and CBCT input is shown in Figure 2.

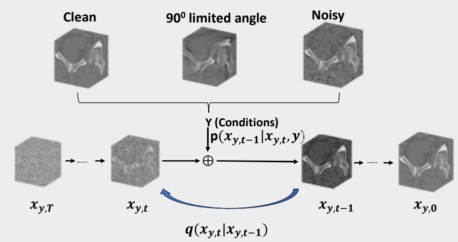
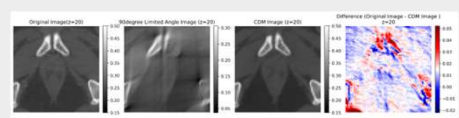
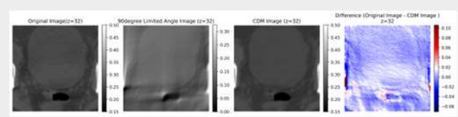


Figure 1. Schematic architecture of the proposed conditional diffusion model for prostate CT/CBCT reconstruction. The framework utilizes conditional inputs Y—comprising clean, 90^0 limited-angle, and noisy image states—to guide the reverse transition $p(x_{y,t-1}|x_{y,t}, y)$.
Through this conditioned reverse process, the model iteratively mitigates noise and severe angular artifacts from the pure noise state $x_{y,T}$ to reconstruct the final high-fidelity prostate image $x_{y,0}$



(a)



(b)

Figure 2. Example of a 90^0 limited angle inputs, (a) PCT and (b) CBCT, used in the evaluation. The conditional diffusion model (CDM) used these inputs to generate clean, high-fidelity 3D priors for reconstruction.

DISCUSSION

Key Findings:

The 3D conditional diffusion model demonstrated exceptional performance across modalities and conditions:

- High structural fidelity (SSIM > 0.83) maintained under both clean and 90^0 limited angle inputs
- Consistently low MSE and MAE, translation and rotation errors across modalities confirm generalizability
- PCT and CBCT priors perform near-identically under clean conditions (SSIM 0.86 v.s. 0.83)

Noise Robustness

A key innovation of this work is evaluating the model as a prior for iterative reconstruction — where input data is inherently degraded. The diffusion process effectively filters noise while retaining critical boundary information required for accurate registration. This robustness is particularly important for low-dose CBCT acquisitions where signal quality is limited.

Future Directions

The current study establishes a clean/noisy baseline. Future work will incorporate additional degradation conditions during training:

- ❑ More training data to further improve stability
- ❑ Motion artifacts from respiratory/cardiac motion
- ❑ More Sparse-view and limited-angle acquisition patterns
- ❑ Metal artifact reduction scenarios
- ❑ Integration with iterative reconstruction pipeline for full end-to-end validation

CONCLUSIONS

- ❖ A 3D conditional patch diffusion model was successfully developed and trained on 500 prostate PCT scans for anatomy prior generation.
- ❖ The model achieved high structural fidelity (SSIM > 0.83 and low MSE/MAE) across clean and 90^0 limited angle PCT and CBCT inputs.
- ❖ Performance was consistent across imaging modalities (PCT vs. CBCT), supporting cross-modality applicability.
- ❖ The approach is well-suited for regularizing iterative CBCT reconstruction to improve image guidance in prostate radiotherapy
- ❖ Future work will extend conditioning to motion, sparse-view, and metal artifact scenarios for comprehensive clinical validation.

REFERENCES

1. [1] Wolleb, J., Sandkühler, R., Bieder, F., Valmaggia, P., & Cattin, P. C. Diffusion models for implicit image segmentation ensembles. In MIDL (pp. 1336-1348). PMLR (2022).

Cohort / Condition	SSIM	PSNR (dB)	MSE	MAE	RIR: Translation Errors (mm)	RIR: Rotation Errors (degree)
Clean Condition - PCT	0.84 ± 0.08	14.63 ± 6.47	0.02 ± 0.02	0.14 ± 0.06	(0.04 ± 0.04,0.00±0.05,-0.07±0.05)	(-0.02 ± 0.06,-0.02±0.02,0.05±0.04)
Clean Condition - CBCT	0.78 ± 0.12	14.55 ± 8.33	0.02 ± 0.02	0.11 ± 0.08	(0.03 ± 0.05,0.09±0.09,-0.06±0.07)	(-0.03 ± 0.1,-0.03±0.04,0.05±0.06)
90^0 limit angle Condition - PCT	0.86 ± 0.11	19.64 ± 12.58	0.02 ± 0.02	0.11 ± 0.08	(0.03 ± 0.02,0.04±0.02,0.00±0.04)	(-0.01 ± 0.03,0.00±0.02,-0.01±0.03)
90^0 limit angle Condition - CBCT	0.83 ± 0.11	19.83 ± 12.74	0.02 ± 0.02	0.11 ± 0.08	(0.04 ± 0.04,0.05±0.04,0.02±0.06)	(-0.02 ± 0.05,-0.02±0.02,-0.01±0.03)

Table 1. Quantitative performance of the conditional diffusion model under clean and 90^0 limit angle conditions