Synthetic CT Generation from Cone-Beam CT for Online Adaptive Radiotherapy Using Flow Matching and Diffusion Schrödinger Models

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

Ronja Stern[1,2]; Muheng Li[1,2]; Antony John Lomax, PhD[1,2] ; Ye Zhang, PhD[1]
[1]Paul Scherrer Institute (PSI), [2]ETH Zurich Switzerland

Motivation

Generating synthetic CT from cone-beam CT can support the accurate delivery of adaptive radiotherapy online by accounting for daily anatomical and positional variations. Generative transport models, including flow-matching and diffusion-based approaches, demonstrate strong performance in the generation of natural images and offer promising solutions for CBCT-to-sCT translation.

Key Findings

- Both models significantly improve target accuracy across all metrics.
- Error analysis revealed systematic discrepancies linked to residual anatomical misalignment between the CBCT and reference CT.
- Single-step inference significantly reduced the usual limitations of diffusion-based methods in terms of inference time, but maintaining high prediction accuracy.

Open questions

How can model-generated errors be distinguished from anatomical misalignment between the input CBCT and target CT?

How can reliance on accurately aligned CBCT–CT pairs during training be reduced?

Project funded by

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
State Secretariat for Education, Research and Innovation SERI

This work has received funding from the Swiss State Secretariat for Education, Research and Innovation (SERI).

Funded by the European Union

The UPLIFT project is co-funded by the European Union under Grant Agreement No. 101168955.

www.uplift-project.eu

Contact

Ronja Stern
Switzerland
Paul Scherrer Institute (PSI)
ronja.stern@psi.ch

METHODS AND MATERIALS

The study included 67 paired head-and-neck CT and CBCT scans. The dataset was divided at the patient level into 52 training, 3 validation, and 12 test samples. CT and CBCT pairs were deformably registered [1], with a **mean NCC of 0.98±0.01**. Furthermore, the CBCT–CT pairs in the test set had a **median MAE of 115.86±24.88 HU**, a **PSNR of 30.61±1.87 dB** and an **SSIM of 0.92±0.03**. The flow-matching model learns a continuous velocity field that transports the CBCT distribution directly towards the CT distribution. During inference, the ODE is numerically integrated using the explicit Euler method with two integration steps. The Diffusion Schrödinger Bridge model (DSBM) learns a stochastic transport process between the CBCT and CT distributions over 250 discrete timesteps. Both models used a 3D U-Net architecture incorporating transformer blocks as the network backbone. Because full-resolution medical-image volumes exceed the available GPU memory, patches of size (4,128,128) voxels were used. Training patches were sampled randomly from the image volumes with the constraint to contain a portion of the patient body to avoid sampling patches composed exclusively of air. Data augmentation included rotations of up to 10°, isotropic scaling by factors between 0.9 and 1.1, and Gaussian blurring.

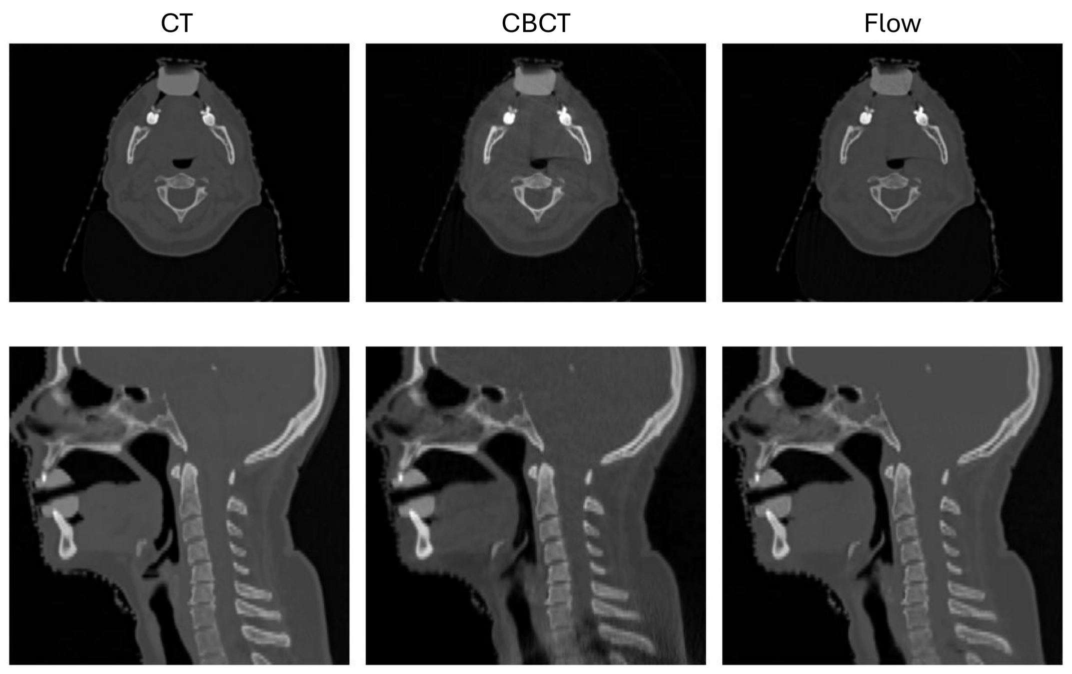

Figure 1: Example target CT, input CBCT and predicted sCT using the Flow model

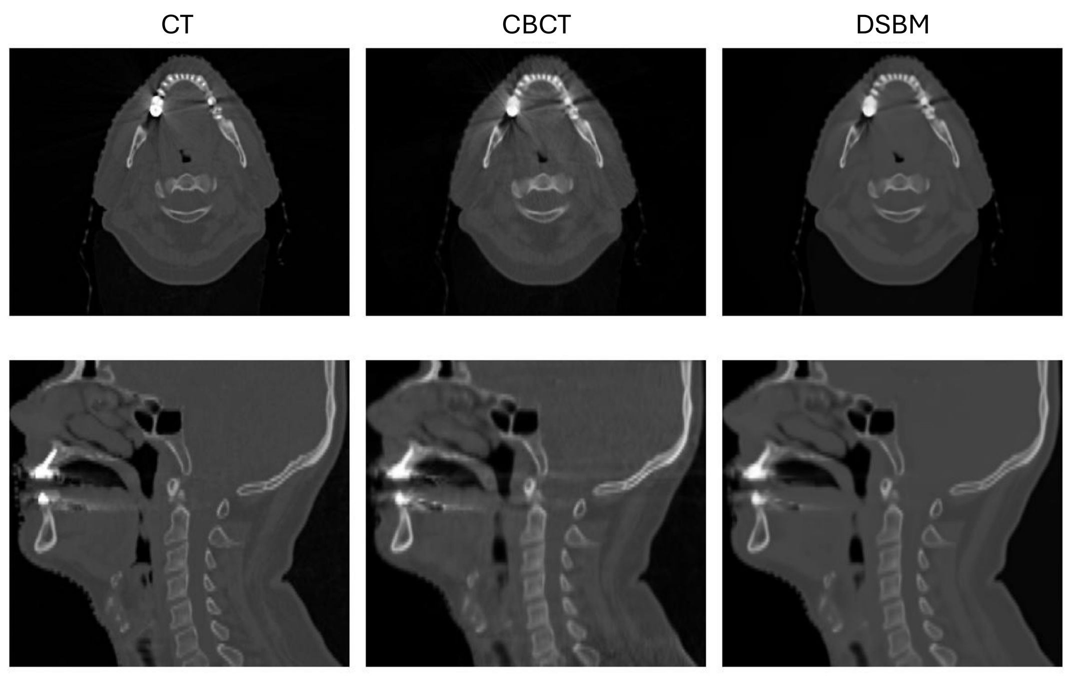

Figure 2: Example target CT, input CBCT and predicted sCT using the DSBM model

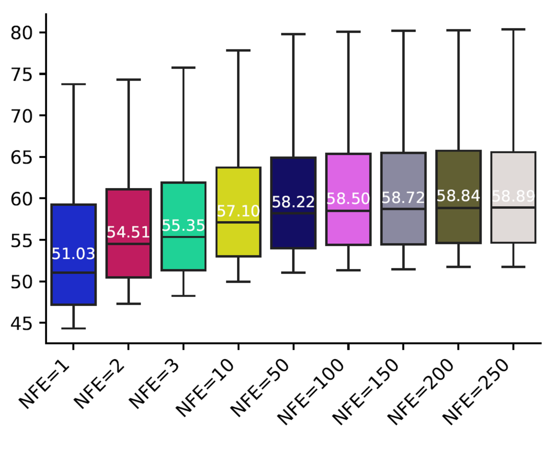

Figure 3: MAE results for different numbers of sampling steps (NFE) in the DSBM model

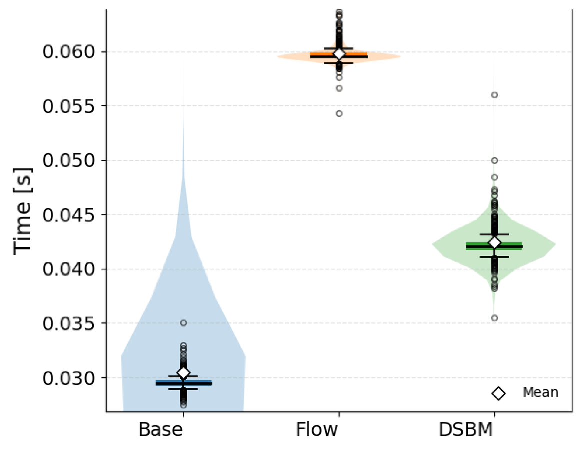

Figure 4: Inference time comparison between the Flow and DSBM model compared to their common backbone 3D Unet.

RESULTS

The Flow model achieves a **median MAE of 51.70 HU**, a **median PSNR of 34.91** and an **inference time of 0.06 seconds per patch**. The single-step DSBM achieves a **median MAE 51.31 HU, a median PSNR of 35.12**, and an **inference time reduced to 0.04 seconds per patch.** (Figure 6)

A decrease in sampling steps for the DSBM model did lead to a speed up in inference time and at the same time improved the prediction accuracy across all voxel-based metrics. (Figure 3)

DISCUSSION

Visual inspection show no systematic under or over estimation of HU values but region specific differences often related to the dental artifacts or anatomical differences in the input CBCT and the target CT image (despite registration). This highlights that deformable registration errors (to the ground truth CT) limit voxel-wise evaluation, suggesting that standard metrics may underestimate true model performance as we can not distinguish between anatomical misalignments and model errors.

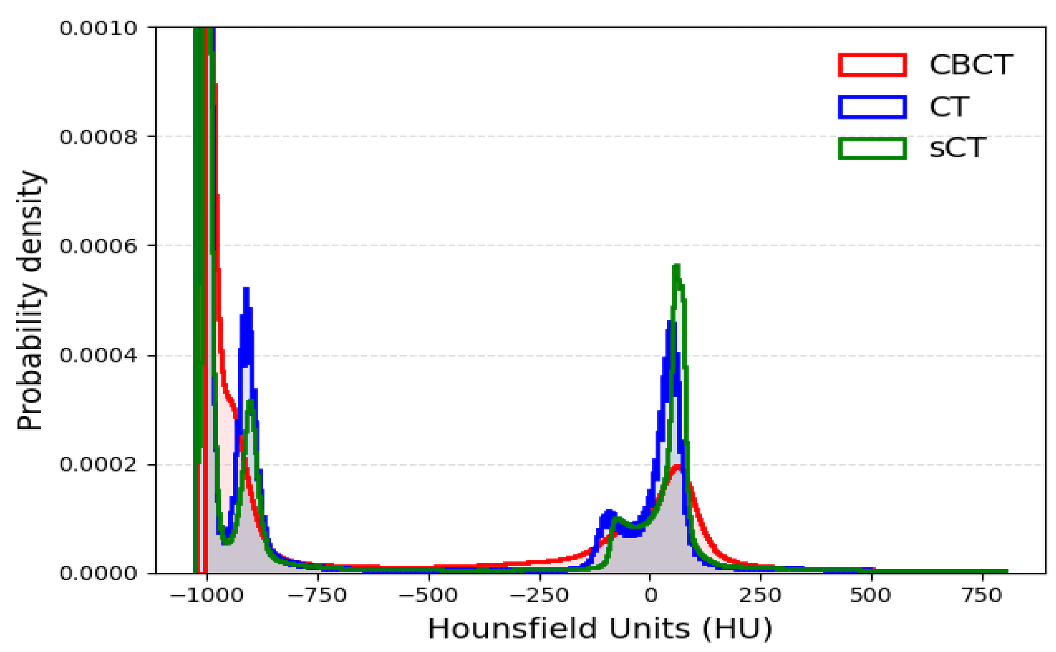

Figure 5: Voxel intensity distributions of CBCT, CT and sCT (Flow), zoomed in to the relevant HU values in the body

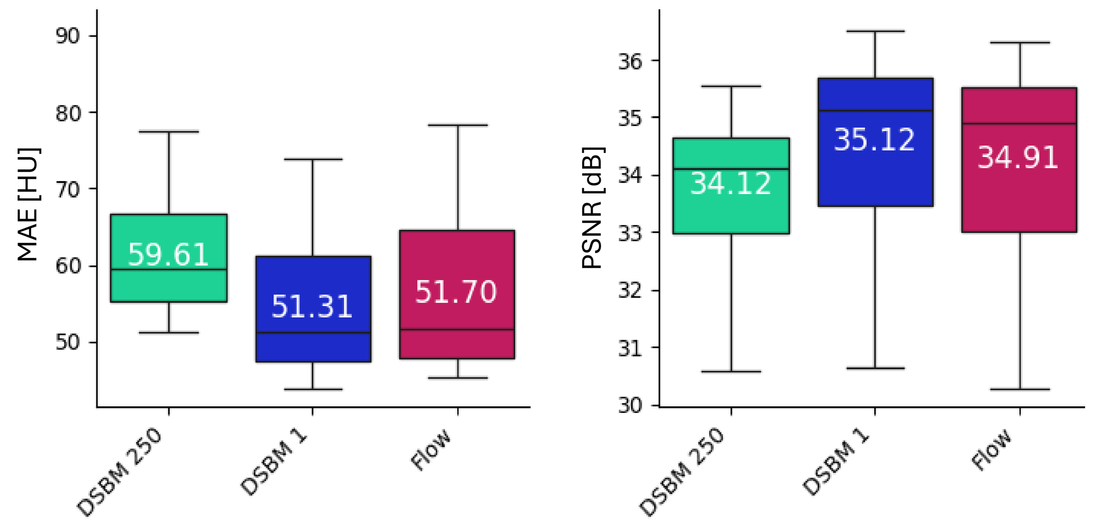

Figure 6: Voxel-based evaluation metrics (MAE, PSNR) comparing the DSBM model using 250 / 1 sampling step and the flow matching method solving the ODE using two integration steps

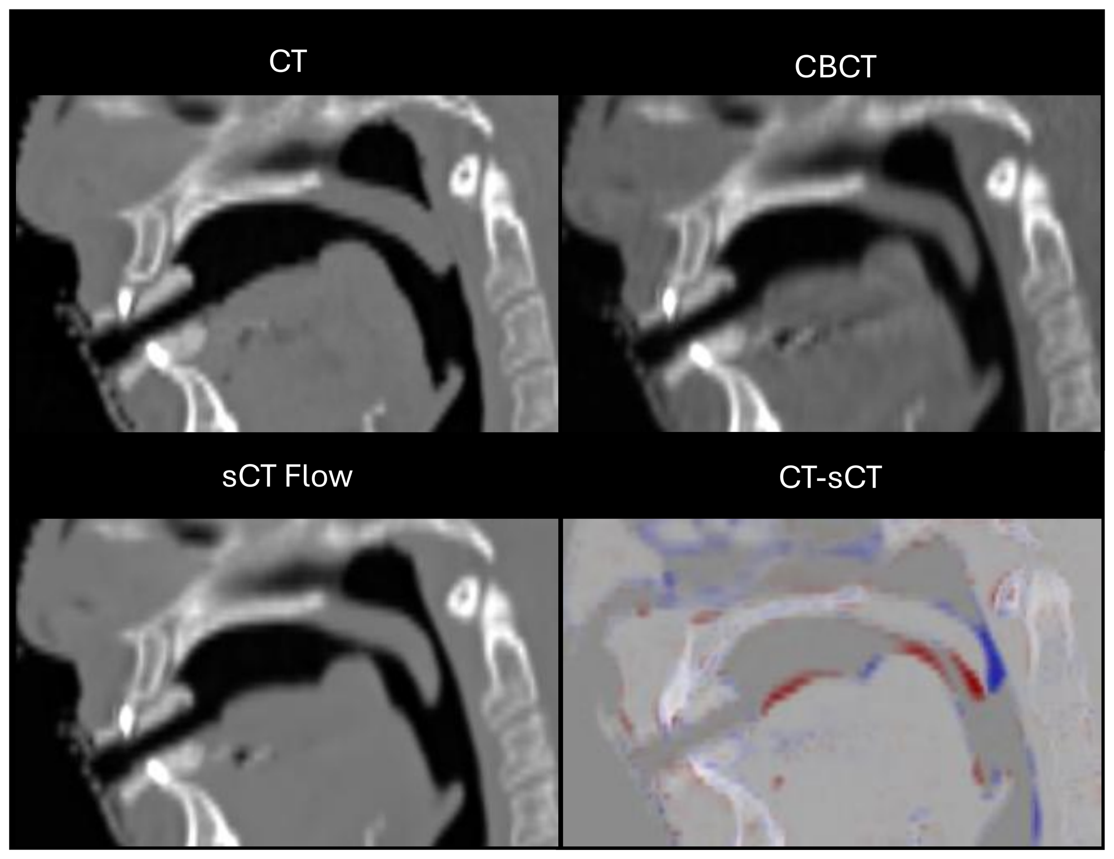

Figure 7: Anatomical misalignments between sCT and ground truth CT

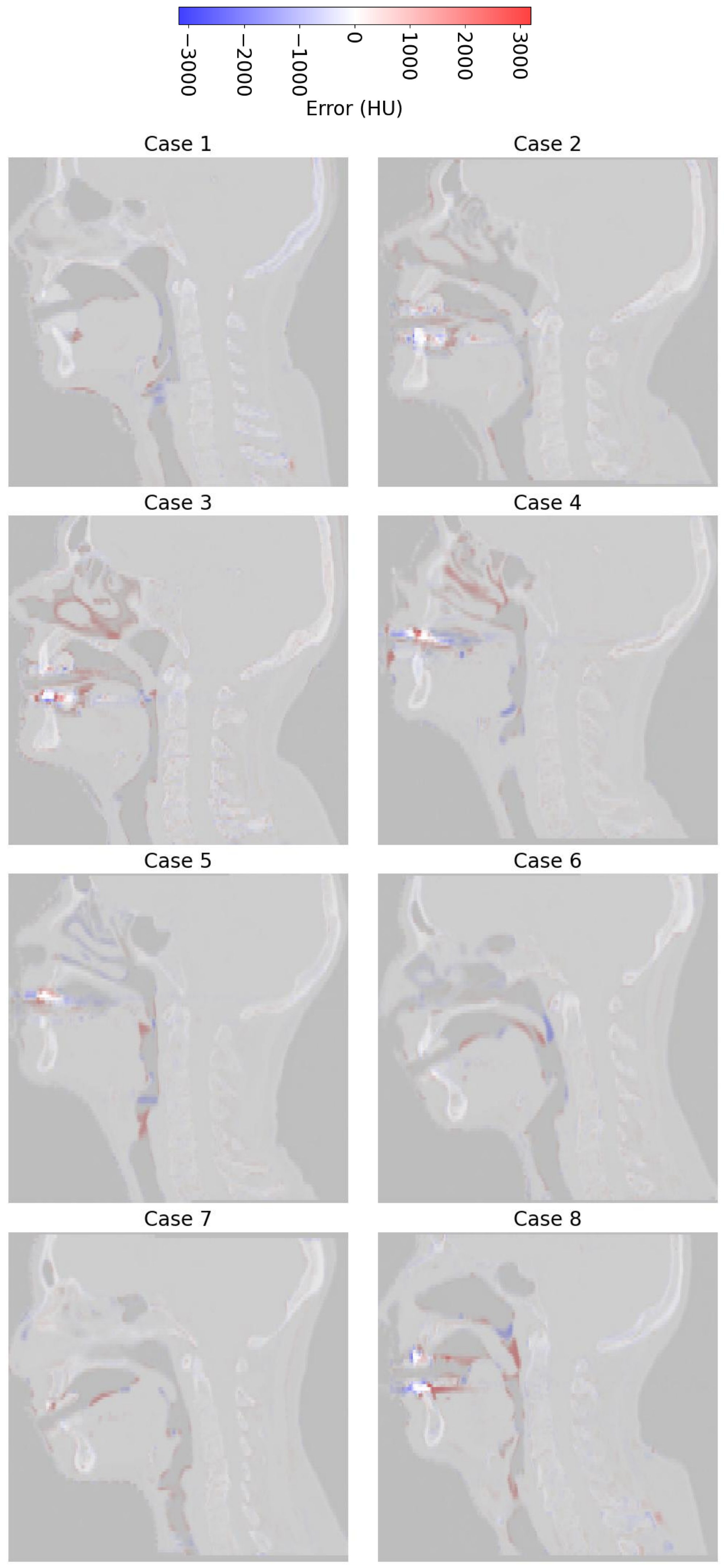

Figure 8: Signed errors between sCT Flow and CT for 8 different test cases

CONCLUSIONS

Both flow-matching and DSBM models demonstrate strong potential for CBCT-to-sCT generation, as they can directly transform the input modality while preserving anatomical structure. Notably, the DSBM model achieves optimal performance with a single sampling step, substantially reducing the time limitations typically associated with diffusion-based inference. However, residual errors in deformable registration affect voxel-wise evaluation, highlighting the need for improved reference alignment, more reliable quantification of anatomical misalignment and training strategies that are less dependent on paired images.

REFERENCES

1. Li X, Yang R, Li M, et al. Continuous sPatial-temporal deformable image registration and 4D frame interpolation. Med Phys. 2026; 53:e70248. https://doi.org/10.1002/mp.70248
2. Li M, Li X, Safai S, Lomax AJ, Zhang Y. Diffusion Schrödinger bridge models for high-quality MR-to-CT synthesis for proton treatment planning. Med Phys. 2025; 52:e17898. https://doi.org/10.1002/mp.17898