Velocity-Corrected Rectified Flow for Volumetric Cross-Modality Medical Image Synthesis

Mobina Rafieipour[1], Pouya Shiri[1], Xin Yi, Neel P. Mistry, Seok-Bum Ko, Amirali Baniasadi , Scott J. Adams

1. M.R. and P.S, contributed equally to this work.

BACKGROUND

- Contrast-enhanced CT improves visualization of vascular and soft-tissue structures, helping clinicians better identify abnormalities [1]. However, obtaining contrast-enhanced scans requires injecting contrast agents, which may not always be preferred or available.
- Diffusion models iteratively denoise images over 50–1000 steps [2], making inference slow. GANs [3] are faster but difficult to train.

- SCBBDM [1] shows promising performance but fails to adequately enhance contrast, producing oversmoothed images with loss of structural detail.
- VCRF [4] uses rectified flow to learn a deterministic neural ODE mapping in only 3 steps.
- A lightweight 2D velocity correction aggregates neighboring slices for 3D volumetric consistency, avoiding the high computational cost of full 3D models , $x_{t-1} \leftarrow x_{t-1} - \gamma_k \bar{v}_\theta$

Where $\bar{v}_\theta$ averages velocity for 3D consistency.

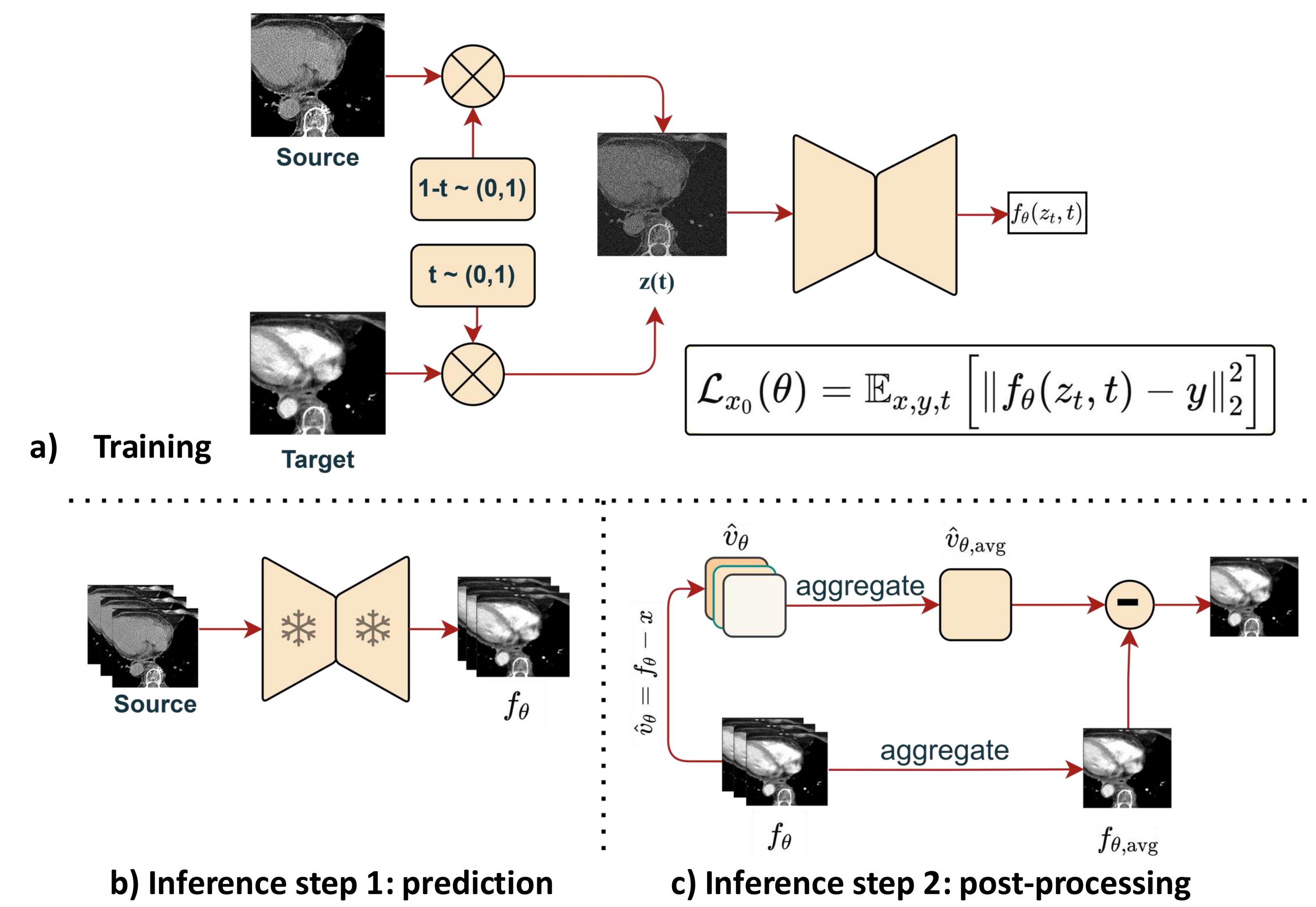
OBJECTIVE

To generate contrast-enhanced CT images from non-contrast images, applying velocity-corrected rectified flow for volumetric cross-modality medical Image synthesis.

METHOD

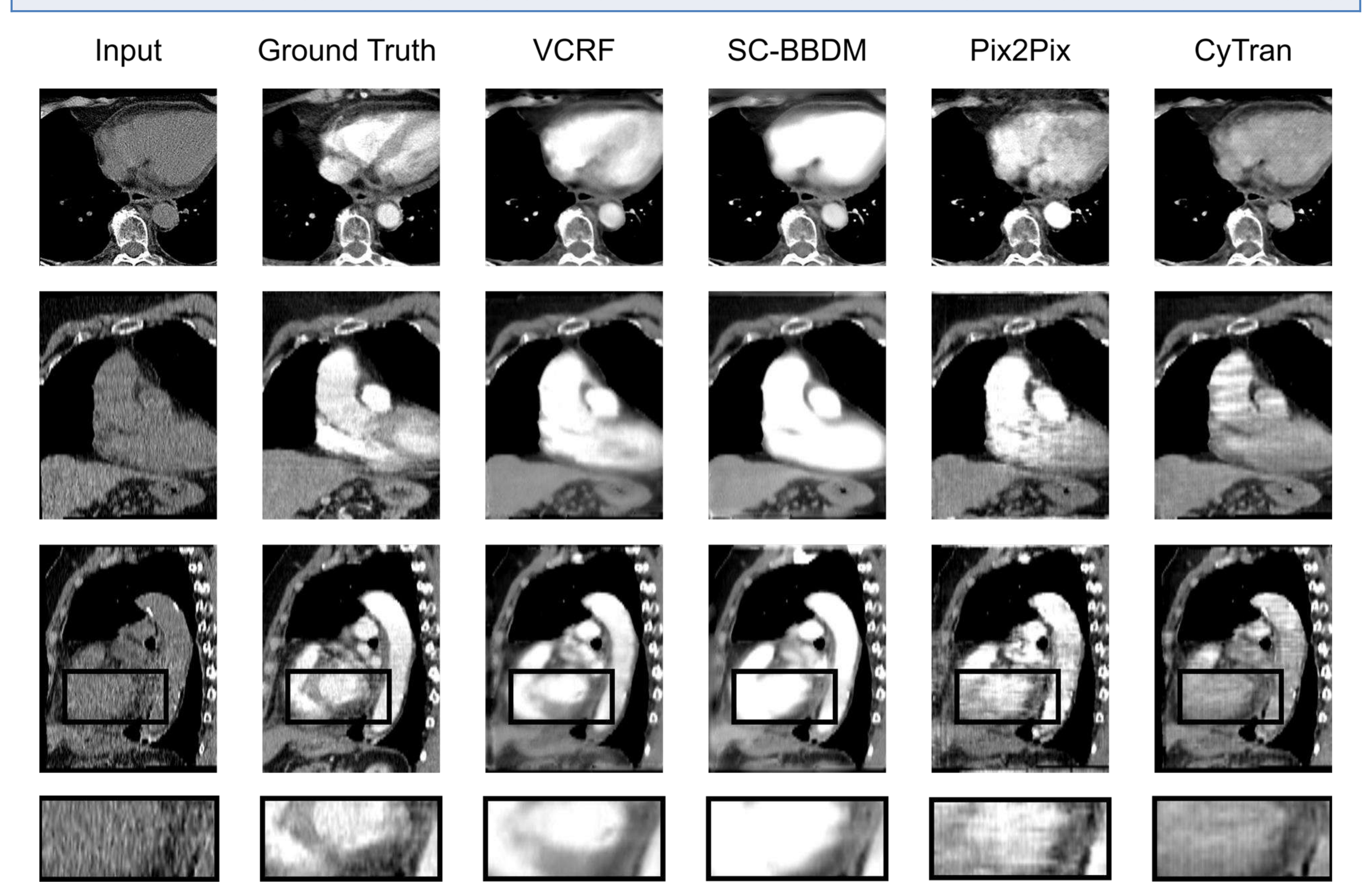
- Linear interpolation paths: $x_t = (1-t)x_0 + ty,\ t \in [0,1]$
- Train $f_\theta$to predict target volume: $\mathcal{L}_{x_0} = E_{xy,t}[\|f_\theta(x_t,t) - y\|_2^2]$
- Solve ODE backward in 3 discrete steps.
- Velocity Correction

1) Recompute velocity: $\tilde{v}_\theta = \hat{v}_\theta(x_{t-1}, t-1)$
2) Aggregate neighboring slices.
3) Apply corrective update: $\gamma_k = \alpha \Delta t_k \frac{\sqrt{HW}}{\|\bar{v}_\theta\|_2^2}$



Results

Model	FID↓			LPIPS$_{Rad}$ ↓			SSIM$_{2D}$ ↑			NMRSE↓	PSNR↑
	AXI	COR	SAG	AXI	COR	SAG	AXI	COR	SAG		
CyTran	37.70	51.84	35.03	0.50	0.44	0.39	0.77	0.73	0.58	0.085	22.36
Pix2Pix	18.07	117.05	96.48	0.39	0.44	0.40	0.81	0.76	0.71	0.074	22.93
SC-BBDM	40.63	41.63	41.49	0.32	0.35	0.22	**0.90**	0.87	0.88	0.042	27.83
VCRF	**32.11**	**31.85**	**32.93**	**0.29**	**0.24**	**0.21**	**0.90**	**0.88**	**0.89**	**0.037**	**29.65**



All visualizations are displayed using the soft-tissue window setting (window level: 50 HU, window width: 350 HU).

DATASET

Name	Coltea-Lung-CT
#samples	100 paired 3D CT volumes
Input Modality	Native (Non-Contrast)
Target	Arterial (Contrast-Enhanced)
Train/Test/Val	75/15/15 CT volumes
Resolution	256 × 256 × 96 voxels
Preprocessing	clip [−1000 HU, 1000 HU], normalize [0,1]

INFERENCE STRATEGY

- 2D slice-wise inference with velocity aggregation for 3D consistency.
- Neighboring-slice averaging for drift correction.

KEY ADVANTAGES

✓100 training steps
✓3 interference steps
✓17.04s inference time per volume inference on H100 GPU.
✓ deterministic and reproducible output.
✓ improved structural consistency.

REFERENCES

[1] Shiri et al., Slice-consistent BBDM for chest CT, 2025.
[2] Ho et al., DDPM, NeurIPS 2020.
[3] Isola et al., Pix2Pix, CVPR 2017.
[4] Liu et al., Rectified Flow, ICLR 2023.

CONTACTS

Pouya Shiri
pouya.shiri@usask.ca
+1 (587) 969-7180

Mobina Rafieipour
mobinar@uvic.ca
+1 (778) 700-3160