Team KoalAI: Locally-enhanced 3D Pix2Pix GAN for Synthetic CT Generation

Bowen Xin, Aaron Nicolson, Hilda Chourak, Gregg Belous, and Jason Dowling
The Australian e-Health Research Centre, CSIRO Health and Biosecurity, Brisbane, Australia

HEALTH & BIOSECURITY

Introduction

Background. In radiation therapy planning, synthesis of Computed tomography (CT) from Magnetic Resonance Imaging (MRI) and Cone-Beam CT (CBCT) is an important task as it reduces the additional clinical workload as well as additional radiation to the patient.

Challenges. However, issues such as spatial misalignment, imaging protocol varieties, presented noise, etc. leads to blur in local details in generated synthetic CT.

Related work. Generative adversarial networks: supervised GANs such as Pix2pix [1], RegGAN, and unsupervised GANs such as CycleGAN, contrastive unsupervised translation (CUT).
Diffusion networks: 2D DDPM [2], thick-slice DDPM.

Contributions. In this paper, we proposed a locally-enhanced 3D pix2pix GAN to improve fine details in generated synthetic CT. The experimental results show the proposed model achieved the best results to compared to other implemented methods including 2D DDPMs and unsupervised GANs.

Methodology

Preprocessing. In task1, MR images were preprocessed with 1) histogram matching, 2) N4 bias field correction, 3) smoothing with a gradient anisotropic filter, 4) arm removing and masking with body mask, and 5) min-max normalization. In task2, CBCT images were preprocessed with 1) noise removal using connected component algorithm, 2) masking with body mask, 3) intensity scaling to [-1024, 3000] and then normalize to [-1, 1].

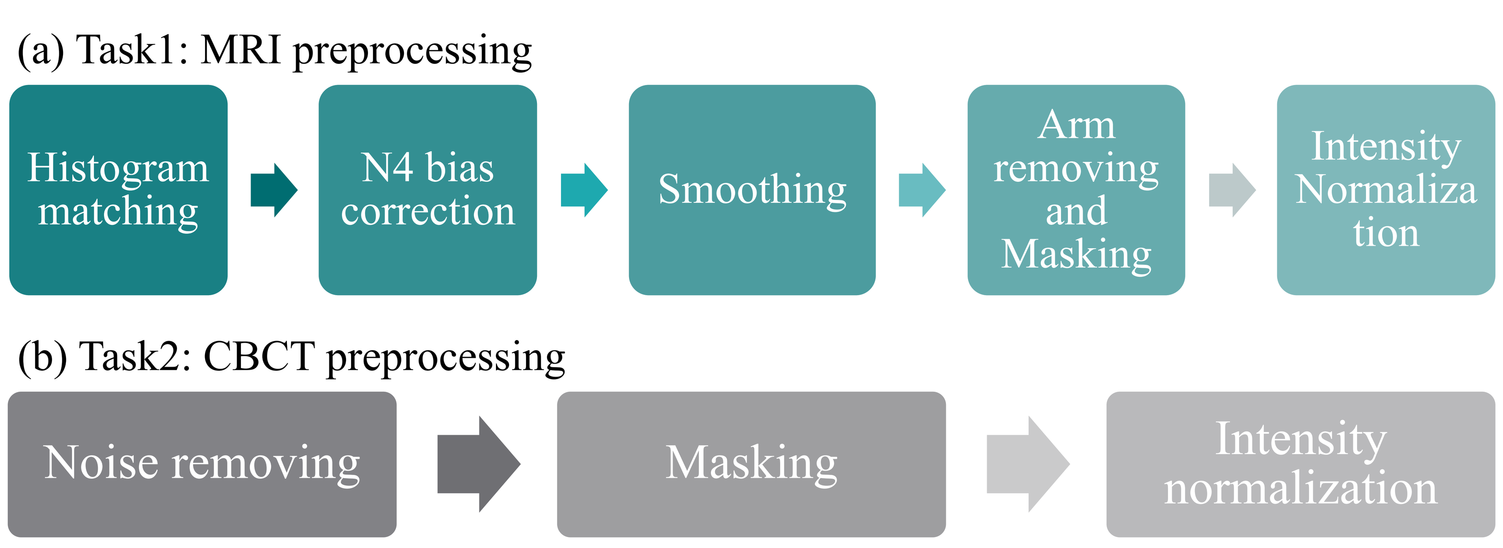

Fig. 1 Preprocessing for MR and CBCT images in task 1and 2.

Locally-enhanced 3D Pix2Pix GAN. Fig. 2 illustrates the proposed locally-enhanced 3D Pix2Pix GAN. In terms of model architecture, it consists of a 3D Generator and a locally-enhanced 3D discriminator. The 3D generator takes 3D patches as input and produces translated 3D images. In locally-enhanced 3D discriminator, a 3D discriminator is employed to capture global fidelity while an additional 2D discriminator is incorporated to improve local fine details. Image iterator is used to generate 3D and 2D patches from translated images and target images. In terms of the loss function, MSE loss was used for adversarial training, while L1 loss was used for pixel-wise supervision.

Ensemble. An ensemble of models consisting of different 3D generators or generators taking different patch sizes is used to enhance generative performance.

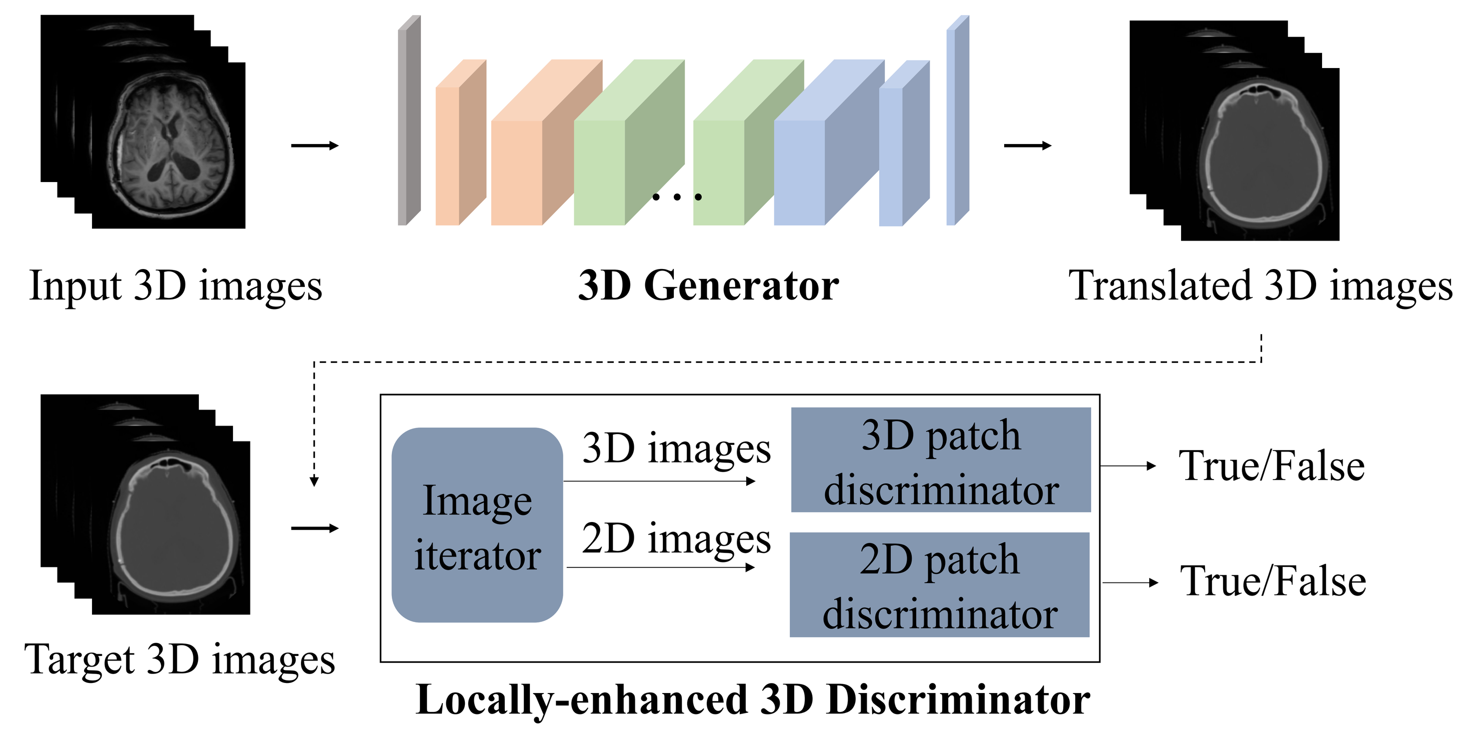

Fig. 2 Flowchart of Locally-enhanced 3D Pix2pix GAN.

Experimental Results

Results of the final model. Table 1 shows (1) the implementation of the final models for task 1 and 2 and (2) the results in validation leaderboard.

Table 1. Final model implementation with testing leaderboard results. Abbr: LE – locally-enhanced.

Data	Generator	Discriminator	Patch size	Val-MAE
Task1_pelvis	Resnet	LE Discriminator	(256, 256, 56)	69.41
Task1_head	Resnet	LE Discriminator	(256, 256, 56)	74.28
	Resnet	Patch Discriminator	(256, 56, 256)	
	Resnet	LE Discriminator	(56, 256, 256)	
Data	**Generator**	**Discriminator**	**Patch size**	**Val-MAE**
Task2_pelvis	DynUnet	Patch Discriminator	(128, 128, 128)	62.48
	Resnet	LE Discriminator	(256, 256, 56)	
	Unet	Patch Discriminator	(448, 448, 64)	
Task2_head	Resnet	LE Discriminator	(256, 256, 56)	52.89
	Resnet	LE Discriminator	(256, 56, 256)	
	Resnet	LE Discriminator	(56, 256, 256)	

Comparison of synthetic CT models. Fig. 3 shows that our proposed method outperformed the other implemented comparison models in local validation (MAE).

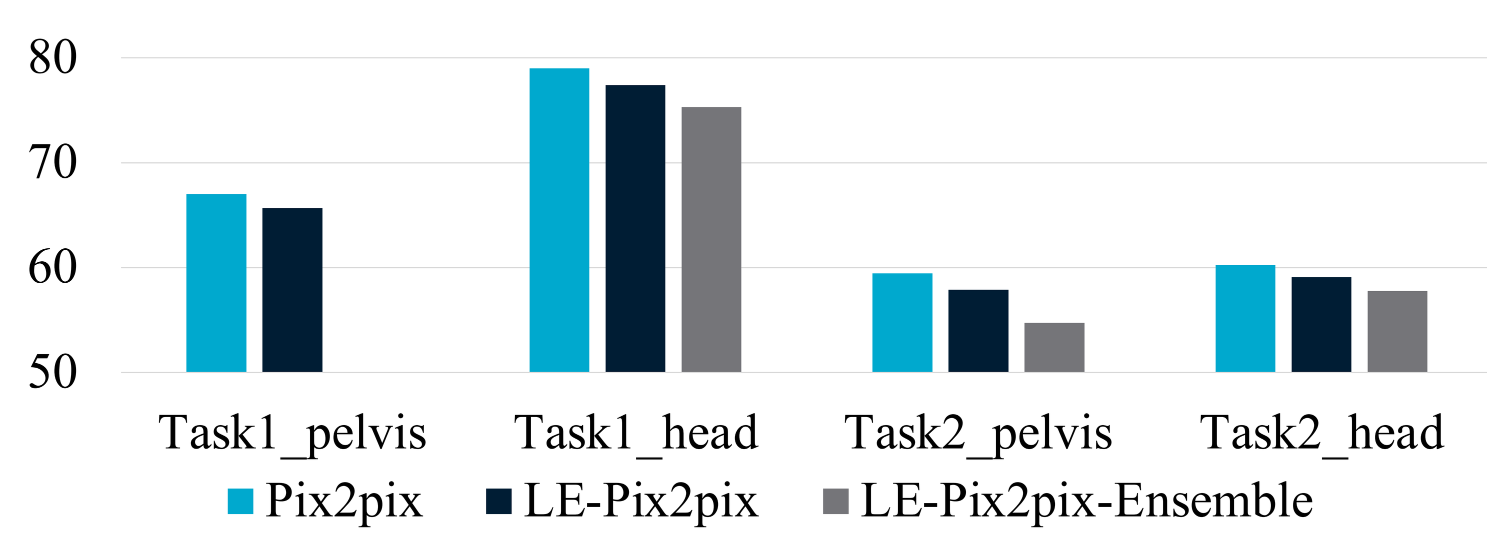

Fig. 3 Comparison of different models in validation leaderboard (MAE)

Conclusions

We present a novel Locally-enhanced 3D Pix2pix GAN for synthesizing CT from both MRI and CBCT. It outperformed other supervised and unsupervised GANs and 2D DDPM, showing promise to contribute to radiation therapy planning.

Reference
[1] Isola, Phillip et al. "Image-to-Image Translation with Conditional Adversarial Networks." 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2016): 5967-5976.
[2] Ho, Jonathan et al. "Denoising Diffusion Probabilistic Models." ArXiv abs/2006.11239 (2020): n. pag.

FOR FURTHER INFORMATION
Dr. Bowen Xin
Email: bowen.xin@csiro.au
Website： https://aehrc.com/

THE AUSTRALIAN E-HEALTH RESEARCH CENTRE