Synthetic CT imaging using limited patch-based CycleGAN for adaptive radiotherapy

Seungwan Lee[1,2]*, Youngeun Choi[3], Gwangsoo Han[2], Seonghee Kang[4]

[1]Department of Radiological Science, Konyang University, Daejeon, South Korea; [2]Department of Medical Science, Konyang University, Daejeon, South Korea ; [3]National Institute for Mathematical Science, Daejeon, South Korea; [4]Department of Radiation Oncology, Seoul National Univ. Hospital, Seoul, South Korea

*Corresponding author: Seungwan Lee (slee1@konyang.ac.kr)

INTRODUCTION

- **Cycle generative adversarial network (GAN)**
  - Having a difficulty in maintaining the structural integrity in a synthesized image
  - With a whole image-based training, causing the degradation of the spatial resolution in an output image
  - With a patch-based training, inducing the inconsistency of global structures and the fluctuation of CT numbers
- **Location-limited patch-based loss calculation**
  - Compensating for the limitations of both the a whole image-based and patch-based approaches
  - Focusing on crucial image features during network training
- **Multi-scale block (MSB) & attention gate (AG)**
  - MSB: Clearly conveying the trainable features of an image during network training
  - AG: Emphasizes essential image features while reducing redundant image features
- **Objective**
  - Proposing a CycleGAN with the MSBs and AGs for improving the accuracy of synthetic CT (sCT) images in adaptive radiotherapy
  - Applying the location-limited patch-based approach for the precise loss calculation

DATA PREPARATION

- **Database**
  - Planning CT (pCT) and cone-beam CT (CBCT) images for 58 patients from The Cancer Imaging Archive (TCIA)
  - pCT: 120 kVp, 92-503 mA, 512×512, 1 mm voxel, 3mm slice thickness
  - CBCT: 120 kVp, 40-80 mA, 512×512, 1 mm voxel, 3mm slice thickness
  - Training: 4,802 / Test: 120

NETWORK ARCHITECTURE

- **Generator** (Fig. 1)
  - U-Net structure including 8 MSBs and 4 AGs
  - MSB: Five convolution layers with the kernel sizes of 1×1, 3×3 and 7×7, concatenation layer
  - AG: Three convolution layers with a kernel size of 1×1, ReLU & sigmoid activation functions, addition & multiplication procedures, resampler

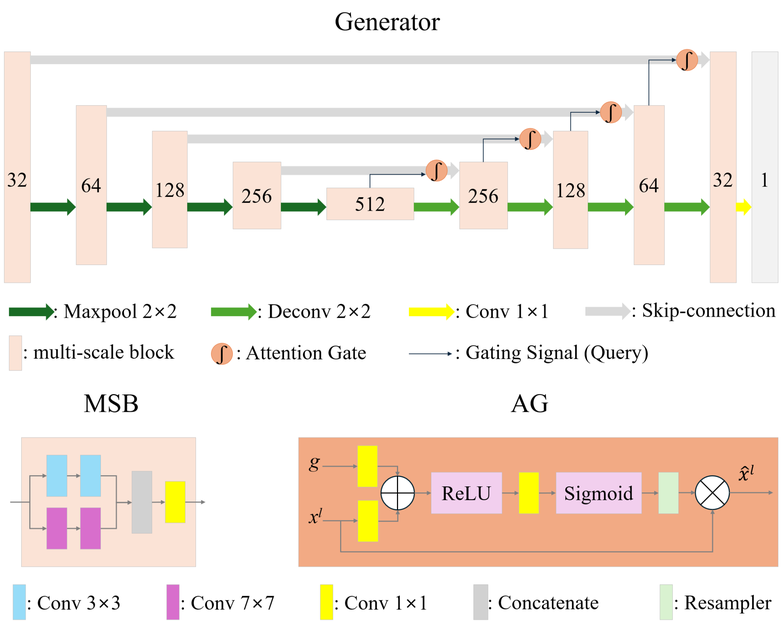

Fig. 1. Generator architecture of the proposed CycleGAN including the MSBs and AGs

- **Discriminator**
  - Five convolution layers for a binary classification with 16×16 patches

LOSS CALCULATION & TRAINING

- **Loss calculation** (Fig. 2)
  - 8 Patches: Location-limited extraction for rejecting a background
  - Network: Encoder of the generator, two convolution layers with a kernel size of 1×1, up-sampling layer
  - Mean squared error (MSE)

$$L(G_{CBCT}, D_{CBCT}, P) = (P(G_{CT}(I_{CBCT})^{(i,j)}) - I_{CBCT}^{(i,j)})^2 + (P(G_{CBCT}(I_{CT})^{(i,j)}) - I_{CT}^{(i,j)})^2$$

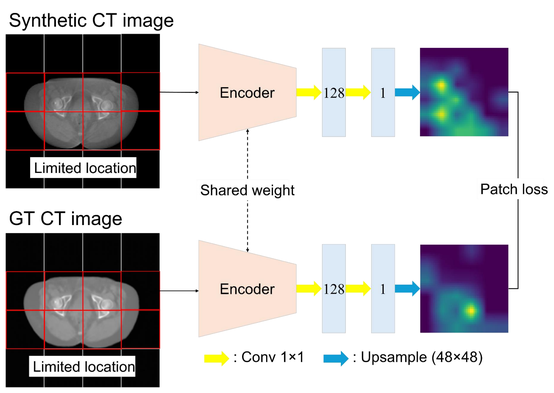

Fig. 2. Strategy for calculating the loss between the GT and output CT image using the location-limited patches.

- **Training**
  - Adam optimizer with a learning rate of 2×10^{-4}
  - 200 epochs
  - Environment: Intel Xeon Scalable Gold 6526Y CPUs, NVIDIA RTX 4090 GPU, Python 3.10, PyTorch 2.3, and TensorFlow 2.17

RESULTS

- **Output images** (Fig. 3)
  - Compared to the other models, the visual similarity between the GT and output CT images was highest for the proposed model.
  - The structural distortion was decreased by the proposed model, while the other models distorted the details.

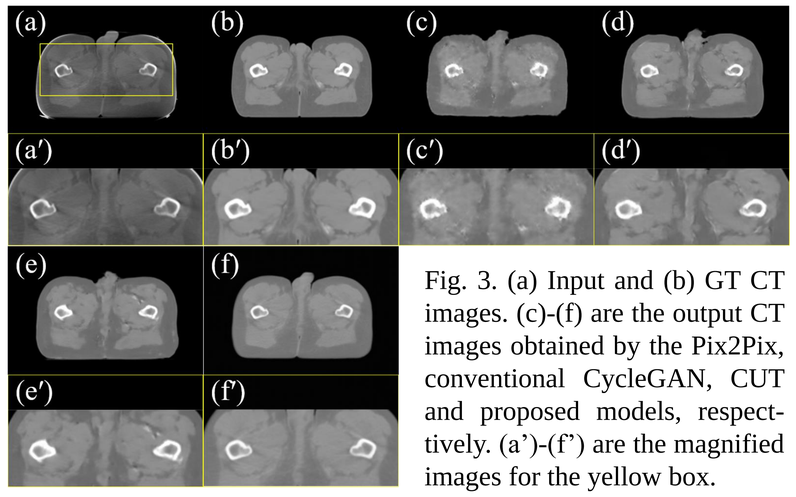

Fig. 3. (a) Input and (b) GT CT images. (c)-(f) are the output CT images obtained by the Pix2Pix, conventional CycleGAN, CUT and proposed models, respectively. (a')-(f') are the magnified images for the yellow box.

- **Quantitative analysis** (Fig. 4)
  - Vs. the input image: PSNR 10.20% ↑, SSIM 9.12% ↑, NRMSE 63.12% ↓
  - Vs. the other models: PSNR 6.02% ↑, SSIM 13.20% ↑, NRMSE 42.13% ↓

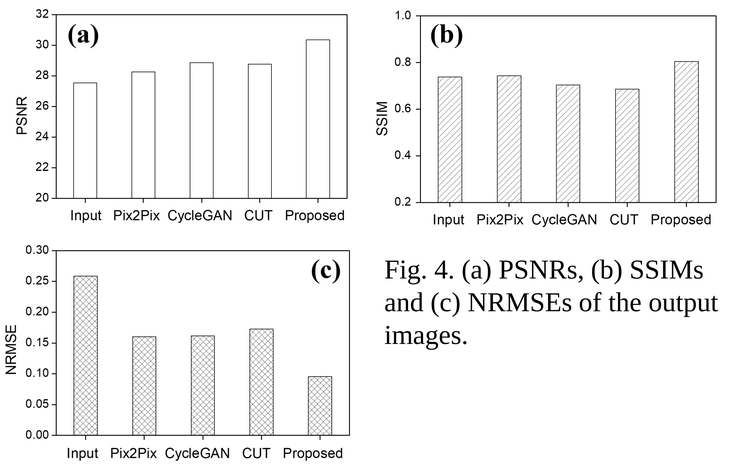

Fig. 4. (a) PSNRs, (b) SSIMs and (c) NRMSEs of the output images.

CONCLUSION

- The proposed CycleGAN model can be potentially utilized for improving the quantitative accuracy of synthetic CT images.
- The performance of the proposed model was superior to the other synthetic CT imaging models.
- The proposed model is expected to precisely provide the anatomical information of a patient for adaptive radiotherapy.