



## OPEN ACCESS

## EDITED BY

Xiaodong Wu,  
The University of Iowa, United States

## REVIEWED BY

Sara Lyons Hackett,  
University Medical Center Utrecht,  
Netherlands  
Renjie He,  
University of Texas MD Anderson Cancer  
Center, United States

## \*CORRESPONDENCE

Jaehee Chun  
✉ cjhsmile@gmail.com  
Jin Sung Kim  
✉ JINSUNG@yuhs.ac

†These authors have contributed equally to  
this work

RECEIVED 09 August 2024

ACCEPTED 09 December 2024

PUBLISHED 03 January 2025

## CITATION

Lee C, Yoon YH, Sung J, Kim JW, Cho Y,  
Kim J, Chun J and Kim JS (2025) Abdominal  
synthetic CT generation for MR-only  
radiotherapy using structure-conserving  
loss and transformer-based cycle-GAN.  
*Front. Oncol.* 14:1478148.  
doi: 10.3389/fonc.2024.1478148

## COPYRIGHT

© 2025 Lee, Yoon, Sung, Kim, Cho, Kim, Chun  
and Kim. This is an open-access article  
distributed under the terms of the [Creative  
Commons Attribution License \(CC BY\)](#). The  
use, distribution or reproduction in other  
forums is permitted, provided the original  
author(s) and the copyright owner(s) are  
credited and that the original publication in  
this journal is cited, in accordance with  
accepted academic practice. No use,  
distribution or reproduction is permitted  
which does not comply with these terms.

# Abdominal synthetic CT generation for MR-only radiotherapy using structure-conserving loss and transformer-based cycle-GAN

Chanwoong Lee<sup>1,2</sup>, Young Hun Yoon<sup>1,2,3</sup>, Jiwon Sung<sup>4</sup>,  
Jun Won Kim<sup>4</sup>, Yeona Cho<sup>4</sup>, Jihun Kim<sup>4</sup>, Jaehee Chun<sup>5\*†</sup>  
and Jin Sung Kim<sup>1,2,5\*†</sup>

<sup>1</sup>Department of Radiation Oncology, Yonsei Cancer Center, Heavy Ion Therapy Research Institute, Yonsei University College of Medicine, Seoul, Republic of Korea, <sup>2</sup>Medical Physics and Biomedical Engineering Lab (MPBEL), Yonsei University College of Medicine, Seoul, Republic of Korea, <sup>3</sup>Department of Radiation Oncology, Washington University in St. Louis, St Louis, MO, United States, <sup>4</sup>Department of Radiation Oncology, Gangnam Severance Hospital, Yonsei University College of Medicine, Seoul, Republic of Korea, <sup>5</sup>Oncosoft Inc., Seoul, Republic of Korea

**Purpose:** Recent deep-learning based synthetic computed tomography (sCT) generation using magnetic resonance (MR) images have shown promising results. However, generating sCT for the abdominal region poses challenges due to the patient motion, including respiration and peristalsis. To address these challenges, this study investigated an unsupervised learning approach using a transformer-based cycle-GAN with structure-preserving loss for abdominal cancer patients.

**Method:** A total of 120 T2 MR images scanned by 1.5 T Unity MR-Linac and their corresponding CT images for abdominal cancer patient were collected. Patient data were aligned using rigid registration. The study employed a cycle-GAN architecture, incorporating the modified Swin-UNETR as a generator. Modality-independent neighborhood descriptor (MIND) loss was used for geometric consistency. Image quality was compared between sCT and planning CT, using metrics including mean absolute error (MAE), peak signal-to-noise ratio (PSNR), structure similarity index measure (SSIM) and Kullback-Leibler (KL) divergence. Dosimetric evaluation was evaluated between sCT and planning CT, using gamma analysis and relative dose volume histogram differences for each organ-at-risks, utilizing treatment plan. A comparison study was conducted between original, Swin-UNETR-only, MIND-only, and proposed cycle-GAN.

**Results:** The MAE, PSNR, SSIM and KL divergence of original cycle-GAN and proposed method were 86.1 HU, 26.48 dB, 0.828, 0.448 and 79.52 HU, 27.05 dB, 0.845, 0.230, respectively. The MAE and PSNR were statistically significant. The global gamma passing rates of the proposed method at 1%/1 mm, 2%/2 mm, and 3%/3 mm were  $86.1 \pm 5.9\%$ ,  $97.1 \pm 2.7\%$ , and  $98.9 \pm 1.0\%$ , respectively.

**Conclusion:** The proposed method significantly improves image metric of sCT for the abdomen patients than original cycle-GAN. Local gamma analysis was slightly higher for proposed method. This study showed the improvement of sCT using transformer and structure preserving loss even with the complex anatomy of the abdomen.

#### KEYWORDS

MR-linac, abdominal synthetic CT, structure consistency loss, transformer, unsupervised learning

## 1 Introduction

Magnetic resonance (MR) images are widely used in radiotherapy, which could identify target and organ-at-risks with excellent soft tissue contrast (1). Recent studies reported the potential benefits of MR-guided radiation therapy (MRgRT) than computed tomography (CT) based image-guided radio therapy (2–4). However, lack of electron density information of MR precludes dose calculation, requiring additional CT scans for treatment planning. The use of multimodal imaging could result in geometric error of 2–5mm during the registration process between CT and MR (5–11). This error could result in systematic geometric deviations, leading to potential underdosage or overdosage of the tumor area and thus compromising the effectiveness of tumor control (12). Especially for the abdomen, significant organ motion and frequent changes in intestinal gases, such discrepancies are further amplified, increasing uncertainty throughout the treatment (13). Additionally, the acquisition of MR images from MR-Linac is more susceptible to degradation due to the B0 field inhomogeneity induced by Linac components (14). However, for clinically streamlined MR-only radiotherapy, the use of MR-Linac is necessary for sCT.

MR-only radiotherapy has been proposed in several studies to mitigate geometric discrepancies (15). By eliminating planning CT imaging and relying solely on MR, MR-only radiotherapy reduces uncertainties from the registration process, decreases the workload of medical professionals, and protects patients from additional radiation exposure from additional CT scans (16). However, reconstructing synthetic CT (sCT) images is necessary to obtain the electron density required for treatment dose calculation. Conventional approaches of sCT are bulk density override and atlas-based methods (17). The bulk density override approach divides the MR into several classes, such as air, bone, soft tissue, and fat, assigning a homogeneous electron density to each segment (18). However, this method is time-consuming when performed manually and does not consider tissue heterogeneity (19). The atlas-based method uses co-registered MR-CT in an atlas to obtain the sCT for the desired MRI, but it can lose robustness when the

anatomical structure significantly differs from the existing atlas, and due to the numerous registrations required, it can be extremely time-consuming (19, 20).

Recently, since its introduction by Han in 2017 (21), the deep learning method has proven to be much faster and more accurate than the previously mentioned methods. Building on this study, many studies explored sCT generation for head and neck or pelvis (21–26). However, few studies investigated abdominal sCT due to challenges such as organ motion and the presence of air bubbles, which degrade MR images (27). Furthermore, existing studies for synthetic CT generation for MR-Linac systems have predominantly utilized convolutional neural networks (CNNs), which, despite their effectiveness, often fall short in capturing the complex dynamics of abdominal anatomy (28–30). Transformer, initially applied in natural language processing, was introduced to computer vision as the Vision Transformer (ViT) (31, 32). ViT successfully overcame the limitations of CNNs, which were widely used in the medical image field, by capturing strong correlations among global features of an image through the multi-headed self-attention mechanism (31). Transformer-based networks for synthetic CT generation across various modalities were reported superiority than CNN (33–35).

This study introduces a novel approach for generating abdominal sCT for MR-only radiotherapy. First, this method integrates the Shifted Window U-net Transformer (SwinUNETR) with an unsupervised cycle generative adversarial network (cycle-GAN). Unlike conventional methods that rely solely on CNNs structures, our approach leverages the strengths of both transformers and u-net in capturing detailed anatomical features and global context. Second, we applied a structure-conserving loss to maintain geometric consistency between the MR and sCT images. We employed the modality independent neighborhood descriptor (MIND) loss to extract geometric features that are consistent across different modalities (36). The aim of this study is to assess the feasibility and performance of this hybrid model with structure conserving loss in improving sCT quality and accuracy, particularly in the challenging abdominal region (36–39).

## 2 Materials and Methods

### 2.1 Patient data characteristics

Table 1 describes the characteristics of the MR and CT images used in this study. We collected 120 abdominal cancer patients who underwent radiotherapy using Elekta Unity between September 2, 2021, and June 1, 2023, including their T2 MR and CT images. The CT images were scanned with the SOMATOM Definition AS (Siemens Healthcare, Erlangen, Germany). The MR images were scanned using the Ingenia 1.5T MR (Philips, Amsterdam, Netherlands) integrated in Unity (Elekta, Stockholm, Sweden). The range of patients age was 31 to 91. The range of volume size for MR images was  $(480 - 800) \times (480 - 800) \times (250 - 300)$ , and for CT images was  $512 \times 512 \times (117 - 543)$ .

### 2.2 Data preprocessing for CT and MR

Deformable registration was performed using Python 3.9 and Simple-ITK. All images were normalized to have a size of  $256 \times 256$  and a resolution of  $0.83 \times 0.83 \text{ mm}^2$ . The intensity of MR images was normalized using histogram matching. For CT images, the Hounsfield unit (HU) values were clipped to the range from -1024 to 3071 HU. The largest connected component within the CT image were identified, and an algorithm creating a body mask through binary processing was used to remove external structures. After removing external structures from the images, all patients were aligned based on the coordinates of the spine. The datasets were split into 80, 20, 20 for training, validation, and testing, respectively.

### 2.3 Training details of proposed sCT

#### 2.3.1 Baseline architecture

The overall architecture is illustrated in Figure 1. This study utilized the cycle-GAN (39) comprised of two generators, which produce CT and MR images, respectively. Additionally, there are two Discriminators, which discriminate between generated and planning CT and MR images. Both sets, are trained in a competitive manner. The primary goal of CT generator is to generate sCT images that are indistinguishable from real ones. Conversely, CT discriminator aims to discern whether a given image is genuine or artificially created. The

TABLE 1 Data characteristics of magnetic resonance (MR) and computed tomography (CT) images.

	MR	CT
Number of patients	120	120
System	Ingenia 1.5T	SOMATOM Definition AS
Mean volume size	$800 \times 800 \times 250$	$512 \times 512 \times 203$
Mean resolution ( $\text{mm}^3$ )	$1.02 \times 1.02 \times 2.60$	$0.57 \times 0.57 \times 1.19$
Sequence	T2 3D Tra 5min	-

generator and discriminator of MR operate under similar principles. The network hyperparameters for training the generator and discriminator are as follows. The input size for the model was set to  $256 \times 256$  pixels to standardize the resolution of the data. Training was conducted over a total of 100 epochs, using the Adam optimizer with a learning rate of 0.0001. The learning rate was gradually reduced in a linear manner from 30th epochs until it reached zero at the end of final epoch.

#### 2.3.2 Modified generator network of cycle-GAN

Figure 2 depicted architecture of the modified generator used in this study. Recently, Swin-UNETR has gained prominence in medical image segmentation tasks, achieving state-of-the-art results on datasets such as the Medical Decathlon and the Multi-Atlas Labeling Beyond the Cranial Vault segmentation challenge (37). For this study, a modified Swin-UNETR was employed as the primary generator network, tailored for 2D operations. To enhance feature extraction, a  $7 \times 7$  convolutional layer was added both before and after the main Swin-UNETR network.

#### 2.3.3 Loss functions

Cycle-GAN, being unsupervised learning, lacks a direct ground-truth, rendering one-to-one mapping unfeasible. To overcome this limitation, the following loss functions are incorporated as proposed by Zhu et al. (39).

1) The adversarial loss employed is the least square loss. The goal of discriminator is to classify real images as 1 and fake images as 0, whereas the objective of generator is to make discriminator classify the generated images as Equation 1. The loss functions to be minimized for generator and discriminator are as follows:

$$L_{G_{MR-CT}} = \lambda_G((D_{CT}(G_{MR-CT}(MR)) - 1)^2)$$

$$L_{D_{CT}} = \lambda_D((D_{CT}(G_{CT}(MR)))^2 + (D_{CT}(CT) - 1)^2) + \tag{1}$$

$$L_{G_{CT-MR}} = \lambda_G(D_{MR}(G_{CT-MR}(CT)) - 1)^2$$

$$L_{D_{MR}} = \lambda_D((D_{MR}(G_{MR}(CT)))^2 + (D_{MR}(MR) - 1)^2) \tag{2}$$

$G_{MR-CT}, G_{CT-MR}, D_{CT}, D_{MR}$  refer to CT generator, MR generator, CT discriminator, and MR discriminator, respectively.

2) The definition of cycle consistency loss is to compare the image reconstructed back to the original domain from the synthetic image with the real input image using the L1 loss. The equation is as follows:

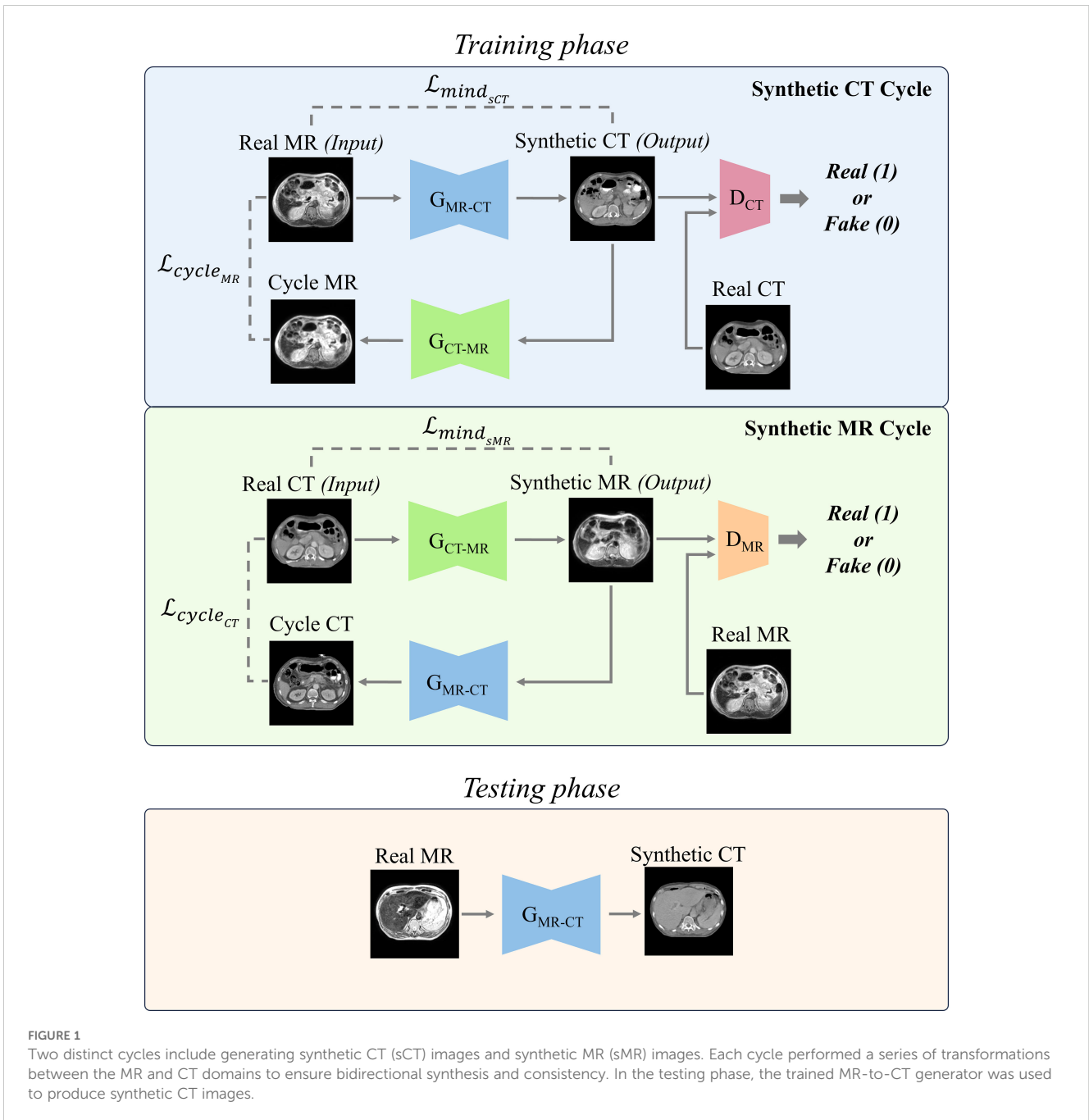
$$L_{cycle_{MR}} = \lambda_{cycle} \|MR - G_{CT-MR}(G_{MR-CT}(MR))\|_1$$

$$L_{cycle_{CT}} = \lambda_{cycle} \|CT - G_{MR-CT}(G_{CT-MR}(CT))\|_1 \tag{3}$$

3) Identity loss in cycle-GAN measures how well the generator preserves the original image's features when transforming images within the same domain. It ensures that when an image is processed by its corresponding generator, the resulting output image remains similar to the input. This similarity is quantified using the L1 loss:

$$L_{idt_{CT}} = \lambda_{idt} \|CT - G_{MR-CT}(CT)\|_1$$

$$L_{idt_{MR}} = \lambda_{idt} \|MR - G_{CT-MR}(MR)\|_1 \tag{4}$$



**FIGURE 1** Two distinct cycles include generating synthetic CT (sCT) images and synthetic MR (sMR) images. Each cycle performed a series of transformations between the MR and CT domains to ensure bidirectional synthesis and consistency. In the testing phase, the trained MR-to-CT generator was used to produce synthetic CT images.

values for  $\lambda_D$ ,  $\lambda_G$ ,  $\lambda_{cycle}$ , and  $\lambda_{idt}$  were set to 0.5, 1, 10 and 5, respectively.

### 2.3.4 Structure Conserving Loss function

This task involves performing style transformation while preserving the geometric structure between MR and sCT. In the abdominal region, obtaining paired data is often challenging, and since this study is conducted using unsupervised learning, there is no explicit constraint between MR (input) and sCT (output). To increase the geometric consistency between the input image and the target image, Modality Independent Neighborhood (MIND) loss was applied (36). This method has been shown to improve performance in generating Head and Neck sCT (40). MIND

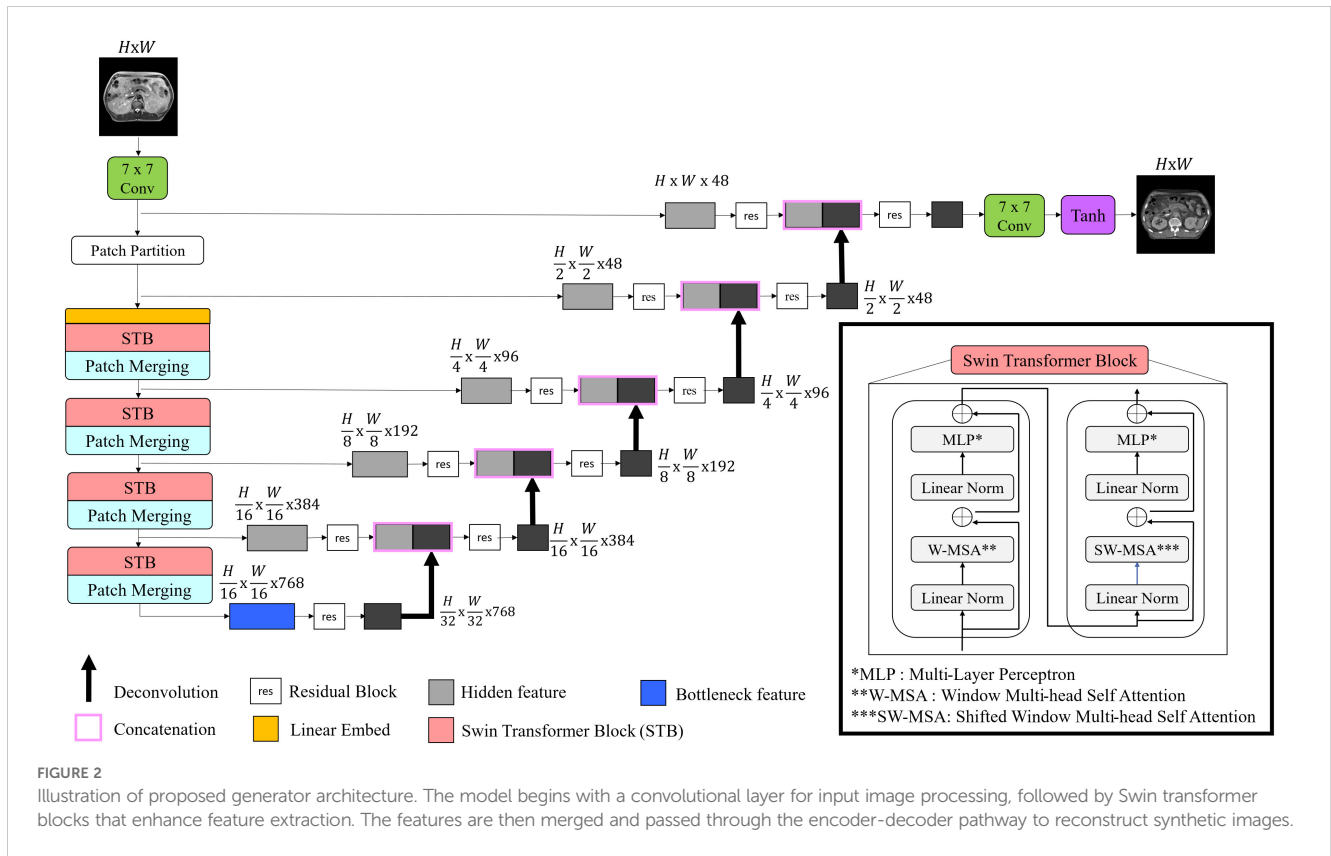
compares local image structures instead of intensity-based comparisons. MIND for an image  $I$  is defined as follows (36).

$$MIND(I, x, r) = \frac{1}{n} \exp\left(-\frac{D_p(I, x, x+r)}{V(I, x)}\right) \quad (5)$$

$$D_p(I, x, x+r) = \sum_{p \in P} (I(x+p) - I(x+r+p))^2 \quad (6)$$

$$V(I, x) = \frac{1}{N} \sum_{n \in N} D_p(I, x, x+n) \quad (7)$$

Here,  $N$  is the number of pixels surrounding pixel  $x$ , which was set to 8.  $D_p$  represents the distance between patches, and  $V$  is the



mean of the distances of the  $N$  neighboring patches. Direct computation of  $D_p$  is computationally expensive; thus, it was implemented using convolution operations as follows.

$$D_p(I, x, x + a) = C * (I - I'(a)) \tag{8}$$

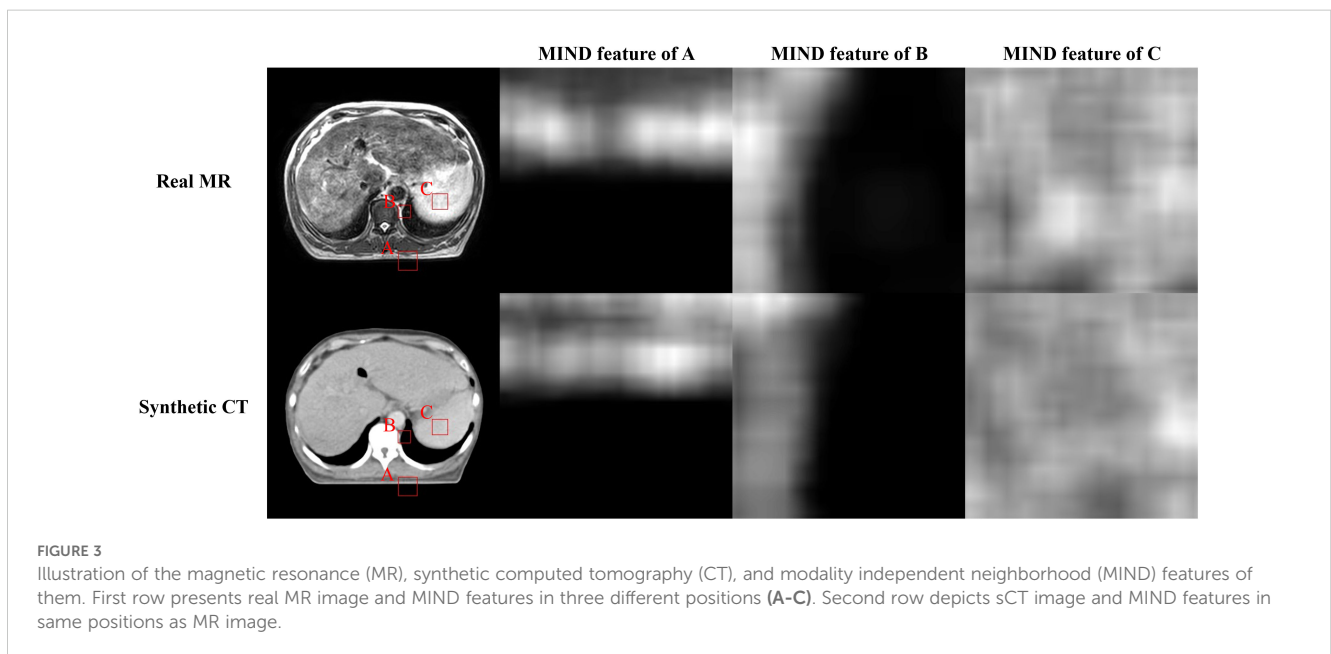
$C$  is a kernel with all weights set to 1 and the same size as  $P$ , and  $I'(a)$  is the image  $I$  translated by  $a$ . This operation simplifies the

calculation of the derivative. A visual example of MIND features is depicted in Figure 3.

Structure conserving loss that employing MIND is defined as follows:

$$L_{mind_{CT}} = \|MIND(MR) - MIND(G_{MR-CT}(MR))\|_1 \tag{9}$$

$$L_{mind_{MR}} = \|MIND(CT) - MIND(G_{CT-MR}(CT))\|_1$$



### 2.3.5 Total Loss

$G_{MR-CT}$  and  $G_{CT-MR}$  were ultimately trained to minimize the follows:

$$\begin{aligned} L_{sCT} &= L_{G_{MR-CT}} + L_{cycle_{MR}} + L_{idt_{CT}} + L_{mind_{sCT}} \\ L_{sMR} &= L_{G_{CT-MR}} + L_{cycle_{CT}} + L_{idt_{MR}} + L_{mind_{sMR}} \end{aligned} \quad (10)$$

The CT discriminator and MR discriminator were trained to minimize  $L_{D_{CT}}$  and  $L_{D_{MR}}$  as presented in (1).

## 2.4 Evaluation of proposed method

### 2.4.1 Image quality

The similarity between the sCT and planning CT images was quantitatively evaluated using commonly used metrics: mean absolute error (MAE), peak signal-to-noise ratio (PSNR), and the structural similarity measure index (SSIM). MAE provides an overall difference by comparing the voxel-wise value between the sCT and planning CT, and its value decreases as the image gets closer to the real one. PSNR is an indicator that measures the amount of noise in the sCT relative to the planning CT signal, with a higher value indicating better image quality. SSIM compares luminance, contrast, and structure between planning CT and sCT images. It produces a value between 0 and 1, with values closer to 1 indicating better similarity and image quality. Additionally, the histogram of intensity for CT and sCT images were compared using the Kullback-Leibler (KL) divergence (41). KL divergence measures the difference between two distributions, quantifying how much one distribution diverges from the other. A lower KL divergence implies that the those have similar distribution. Image metrics were calculated only within the external mask of the planning CT mask. Wilcoxon rank-sum test was used for statistical analysis (42). To qualitatively evaluate the sCT images, two certified radiation oncologists from authors' institution rated the images using a 5-grade scale.

### 2.4.2 Dosimetric evaluation of synthetic CTs

For each of the 20 test patients, dose calculations on sCT were performed using the same clinical plan applied to the planning CT, utilizing the Monte Carlo algorithm in the treatment planning system MONACO 5.51.11 (Elekta, Stockholm, Sweden) for Unity. The dose grid resolution was  $2.0 \times 2.0 \times 2.0 \text{ mm}^3$ , and the statistical uncertainty per calculation was 1%. For dosimetric evaluation, gamma analysis was conducted between the dose distributions based on planning CT and sCT (43). The delivery quality assurance (DQA) criteria of authors' institution was local gamma analysis with a 3%/3 mm, 10% dose threshold. Additionally, we conducted 1%/1 mm, 2%/2 mm, and 3%/3 mm gamma analysis for further evaluation, with the same dose threshold. To investigate the impact of anatomical differences between CT and MR on dose distribution, the 20 cases were divided into two groups: 10 for Group 1, where MR and CT showed good anatomical alignment, and 10 for Group 2, with less anatomical alignment. Subsequently, gamma analysis was conducted for each group, followed by a comparison of intensity distributions utilizing KL divergence.

Additionally, for 10 patients with same organ-at-risks (OARs), planning target volume (PTV) and gross tumor volume (GTV) the average dose volume histogram (DVH) differences for each structure were calculated and evaluated. The OARs, PTV and GTV delineated on the planning CT by a certified radiation oncologist were rigidly copied to the sCT for assessment.

### 2.4.3 Ablation study

Ablative study was performed for structure-conserving loss and generator. The comparisons were made between the baseline (original cycle-GAN), Swin only (modify cycle-GAN generator as Swin-UNETR), MIND only (cycle-GAN with MIND loss), and the proposed method (modify generator and use MIND loss).

## 3 Results

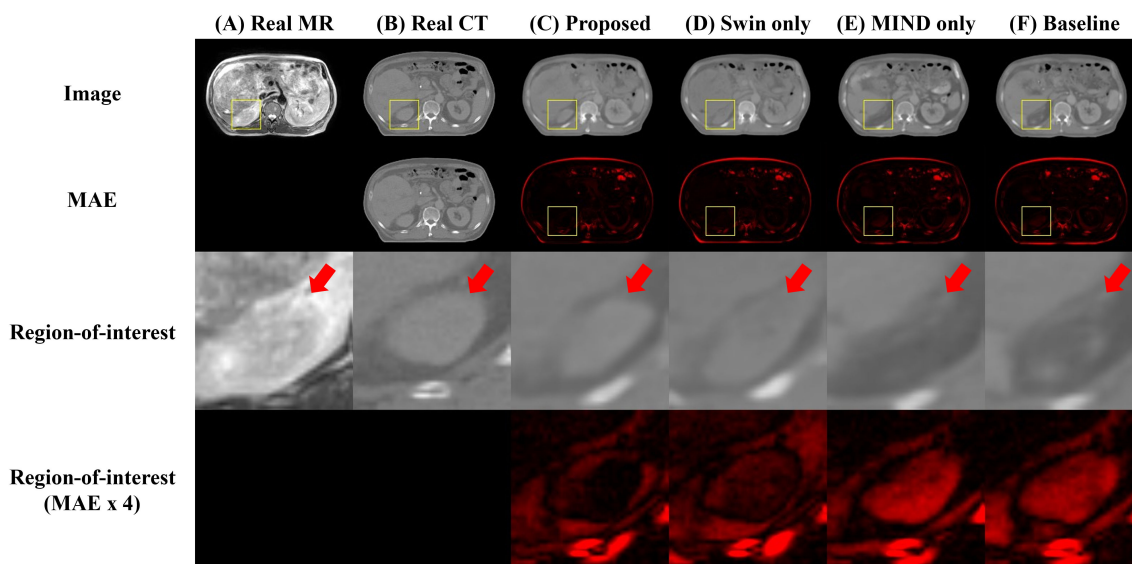
### 3.1 Image quality

Figure 4 compared scanned MR, scanned CT, and generated sCT images of various methods. The proposed model produced sCT images with greater geometric consistency relative to the MR image and improved texture homogeneity with the planning CT image, especially in regions such as the kidney. The MAE, PSNR, and SSIM of proposed method were highest than other methods. Table 2 indicates that applying Swin-UNETR and MIND loss individually did not result in statistically significant differences compared to the baseline. However, the combination of both methods led to statistically significant improvements in MAE and PSNR. The SSIM values were better with the proposed method, although the differences were not statistically significant. The KL divergence between the intensity histograms of CT and sCT demonstrated statistically significant differences from the baseline sCT across all cases: when using the proposed method, applying only Swin-UNETR, and applying only MIND loss. Figure 5 depicted the histograms of MR, CT, and generated sCT images from various methods. Table 3 presents the results of the qualitative evaluation of each sCT, conducted by two certified radiation oncologists.

### 3.2 Dose evaluation

Dosimetric evaluation was performed by comparing dose distribution of the planning CT and sCT generation methods. Table 4 describes the results of local and global gamma analysis for 20 patients based on a 10% threshold, conducted at 1%/1 mm, 2%/2 mm, and 3%/3 mm criteria.

Figures 6 and 7 presents the relative DVH differences for the same OARs and GTV structures in 10 patients. The proposed method demonstrated a relative DVH difference within 5% compared to the planning CT, except for the spinal cord and stomach. This indicates that the dose distributions based on sCT from the proposed method closely matched those from the planning CT-based clinical plans, showing high consistency in dosimetric accuracy across various structures.



**FIGURE 4**  
 Comparison of computed tomography (CT) images as follows: (A) real MR, (B) real planning CT, sCT images generated by (C) proposed method, (D) Swin only, (E) MIND only, and (F) baseline models. The first and second row depict images and mean absolute error (MAE) with planning CT. The third and fourth row illustrate magnified right kidney (indicated by yellow square box) and MAE with planning CT. Red arrows highlighted the anatomical differences.

Figure 8 presents a comparison of dose distributions for (A) Real planning CT, (B) Proposed sCT, (C) Swin only sCT, (D) MIND only sCT, and (E) Baseline sCT methods. The figure shows the axial slices of the dose distribution with overlaid contour lines for critical structures and the target region.

Table 5 illustrates the impact of anatomical differences on dose distribution. In local gamma analysis, no statistically significant differences were found between the two groups at the 3%/3 mm and 2%/2 mm criteria. However, in global gamma analysis, statistically significant differences were observed at the 3%/3 mm and 2%/2 mm levels, except for the baseline sCT. Figure 9 provides a qualitative comparison of dose distribution differences related to anatomical variation. In contrast, as illustrated in Table 6, there was no statistically significant difference between the two groups in the comparison of KL divergence results.

## 4 Discussion

In this study, we developed sCT generation from MR images of abdominal cancer patients using the unpaired data from 1.5 T MR-

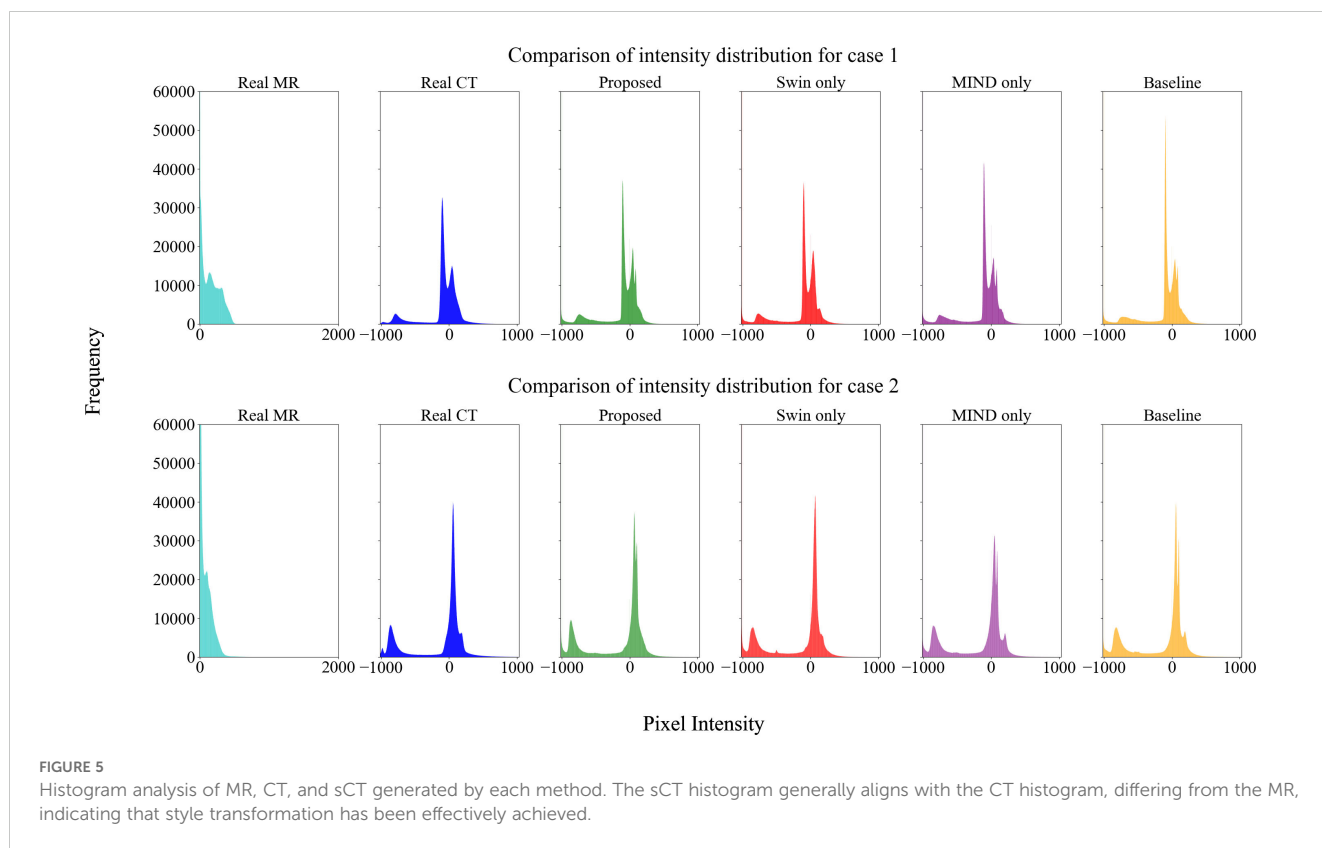
Linac. The primary task is image translation process that transforms style while preserving the content of MR images. However, image synthesis in abdominal region is often challenging than other body parts due to anatomical changes such as peristalsis and intestinal gas. To reduce anatomical difference, minimizing the time interval between the CT and MR scans are crucial. However, this study utilized retrospective data, and due to the predefined clinical protocol, it is procedurally challenging to acquire additional data beyond this framework. As shown in Figure 10, we conducted deformable registration as a preliminary step to mitigate anatomical differences between CT and MR.

This study focuses on stabilizing GAN training, addressing challenges of training instability and optimization difficulties in both the generator and discriminator (44). To achieve this, we employed the Adam optimizer with a learning rate of 0.0001, maintaining for the first 30 epochs. Subsequently, the learning rate was linearly decayed over the remaining 70 epochs. These configurations were designed to ensure smoother convergence during training. Additionally, patient data varied considerably in shape, resolution, and setup which could adversely affect the stability of cycle-GAN. To address this, we aligned the back

**TABLE 2** Mean  $\pm$  standard deviation of MAE, PSNR, SSIM and KL divergence for synthetic CT images compared to planning CT images.

	MAE (HU) ↓	PSNR (dB) ↑	SSIM↑	KL divergence ↓
Proposed	<b>79.5 <math>\pm</math> 11.7*</b>	<b>27.1 <math>\pm</math> 0.1*</b>	<b>0.845 <math>\pm</math> 0.034</b>	0.230 $\pm$ 0.070*
Swin only	83.6 $\pm$ 9.7	26.4 $\pm$ 0.8	0.832 $\pm$ 0.031	<b>0.225 <math>\pm</math> 0.053*</b>
MIND only	80.5 $\pm$ 11.3	26.9 $\pm$ 1.0	0.841 $\pm$ 0.033	0.339 $\pm$ 0.090*
Baseline	86.1 $\pm$ 11.5	26.4 $\pm$ 0.9	0.828 $\pm$ 0.034	0.448 $\pm$ 0.070

Asterisks denote the statistically significant difference of metrics between proposed and the baseline method.(p-value < 0.05; Wilcoxon rank sum test). Bold values indicate the best values.



positions of all patients based on the head-first supine position. Additionally, deformable registration was performed to enhance geometric correspondence between MR and planning CT images, as demonstrated in Figure 10. These preprocessing steps were crucial in improving the stability and reliability of our GAN-based approach, allowing network to focus on specific features while maintaining consistency in other aspects of the data.

Effective utilization of both global and local features was made using Swin-UNETR as a generator. Swin-UNETR combines the transformer structure of UNETR, which integrates a transformer into the traditional U-Net, with the Swin transformer (37, 38, 45). The Swin transformer extract both global and local features by leveraging both global and local attention (32, 45). Especially, as shown in Figure 4, applying Swin-UNETR as a generator maintained MR content in sCT images. Additionally, in the histogram analysis of MR, planning CT, and sCTs generated by each method in Figure 5, the proposed method qualitatively

demonstrates successful style transformation between MR and planning CT.

Additionally, we employed MIND loss to further preserve structure since Kieselmann et al. (46) reported that cycle-GAN alone can result in subtle differences between MR and sCT. This is particularly important for this study since abdominal MR and CT often exhibit anatomical discrepancies due to factors such as imaging modality and peristalsis. While deformable registration can partially alleviate, it was challenging to mitigate them entirely (47). Therefore, it is necessary to utilize constraints on the geometric information between the input and output. MIND utilizes features extracted independently from MR and sCT intensities as in Figure 3, allowing direct structural comparison between MR and sCT (36). This method was previously applied in generating head and neck synthetic CTs using cycle-GAN and demonstrated better performance compared to when MIND loss was not applied (40). The proposed method showed a slight improvement in gamma analysis and DVH difference. Specifically, Figure 7 indicates that the use of a Transformer structure helped in mitigating outliers, enhancing the overall robustness of the generated sCT images, leading to statistically significant improvements in image quality.

There are several related studies of sCT for MR-only radiotherapy. Cusumano et al. (28) utilized a conditional GAN (cGAN) on 20 test patients. They evaluated the image metrics using MAE and mean error (ME), achieving an MAE of 78.71 HU in the abdominal body region and an ME of 10.83 HU. In the dose evaluation, they achieved a gamma passing rate of 99% under the

**TABLE 3** Results of the qualitative evaluation of sCTs generated by each method.

	Proposed	Swin-only	MIND-only	Baseline
Expert 1	2.85 ± 0.85	2.15 ± 0.85	2.50 ± 0.74	2.65 ± 0.79
Expert 2	3.30 ± 1.14	2.95 ± 1.07	3.25 ± 1.13	3.15 ± 1.11
Total	3.08 ± 1.03	2.55 ± 1.05	2.88 ± 1.03	2.90 ± 0.99

Certified radiation oncologists assessed the images using a grade scale ranging from 1 (very bad) to 5 (very good). The results are presented as the mean ± standard deviation of the scores.



**TABLE 4** Results of local and global gamma analysis for dose distribution obtained from proposed, Swin only, MIND only, and baseline sCT.

	Gamma passing rate	p-value
<b>Local gamma analysis (3%/3 mm)</b>		
<i>Proposed</i>	0.974 ± 0.012	-
<i>Swin only</i>	0.970 ± 0.013	0.43
<i>MIND only</i>	0.976 ± 0.011	0.32
<i>Baseline</i>	0.975 ± 0.013	0.53
<b>Global gamma analysis (3%/3 mm)</b>		
<i>Proposed</i>	0.989 ± 0.008	-
<i>Swin only</i>	0.988 ± 0.008	0.70
<i>MIND only</i>	0.991 ± 0.007	0.59
<i>Baseline</i>	0.990 ± 0.009	0.63
<b>Local gamma analysis (2%/2 mm)</b>		
<i>Proposed</i>	0.923 ± 0.027	-
<i>Swin only</i>	0.915 ± 0.029	0.37
<i>MIND only</i>	0.928 ± 0.027	0.47
<i>Baseline</i>	0.925 ± 0.029	0.73
<b>Global gamma analysis (2%/2 mm)</b>		
<i>Proposed</i>	0.971 ± 0.027	-
<i>Swin only</i>	0.967 ± 0.018	0.48
<i>MIND only</i>	0.974 ± 0.015	0.34
<i>Baseline</i>	0.972 ± 0.018	0.50
<b>Local gamma analysis (1%/1 mm)</b>		
<i>Proposed</i>	0.727 ± 0.064	-
<i>Swin only</i>	0.716 ± 0.065	0.53
<i>MIND only</i>	0.739 ± 0.067	0.47
<i>Baseline</i>	0.731 ± 0.066	0.79
<b>Global gamma analysis (1%/1 mm)</b>		
<i>Proposed</i>	0.861 ± 0.059	-
<i>Swin only</i>	0.850 ± 0.062	0.50
<i>MIND only</i>	0.870 ± 0.060	0.47
<i>Baseline</i>	0.864 ± 0.061	0.68

Best value highlighted in bold. The p-values between the proposed method and the other methods were listed in the table.

2%/2 mm criteria and the mean dose difference in the PTV was -0.28%. Fu et al. (48) employed both cGAN and cycle-GAN networks, validating their results using leave-one-out cross-validation on 12 abdominal tumor patients. They evaluated the image metrics using MAE and PSNR, with the cGAN achieving an MAE of 89.8 HU and a PSNR of 27.4 dB, while the cycle-GAN

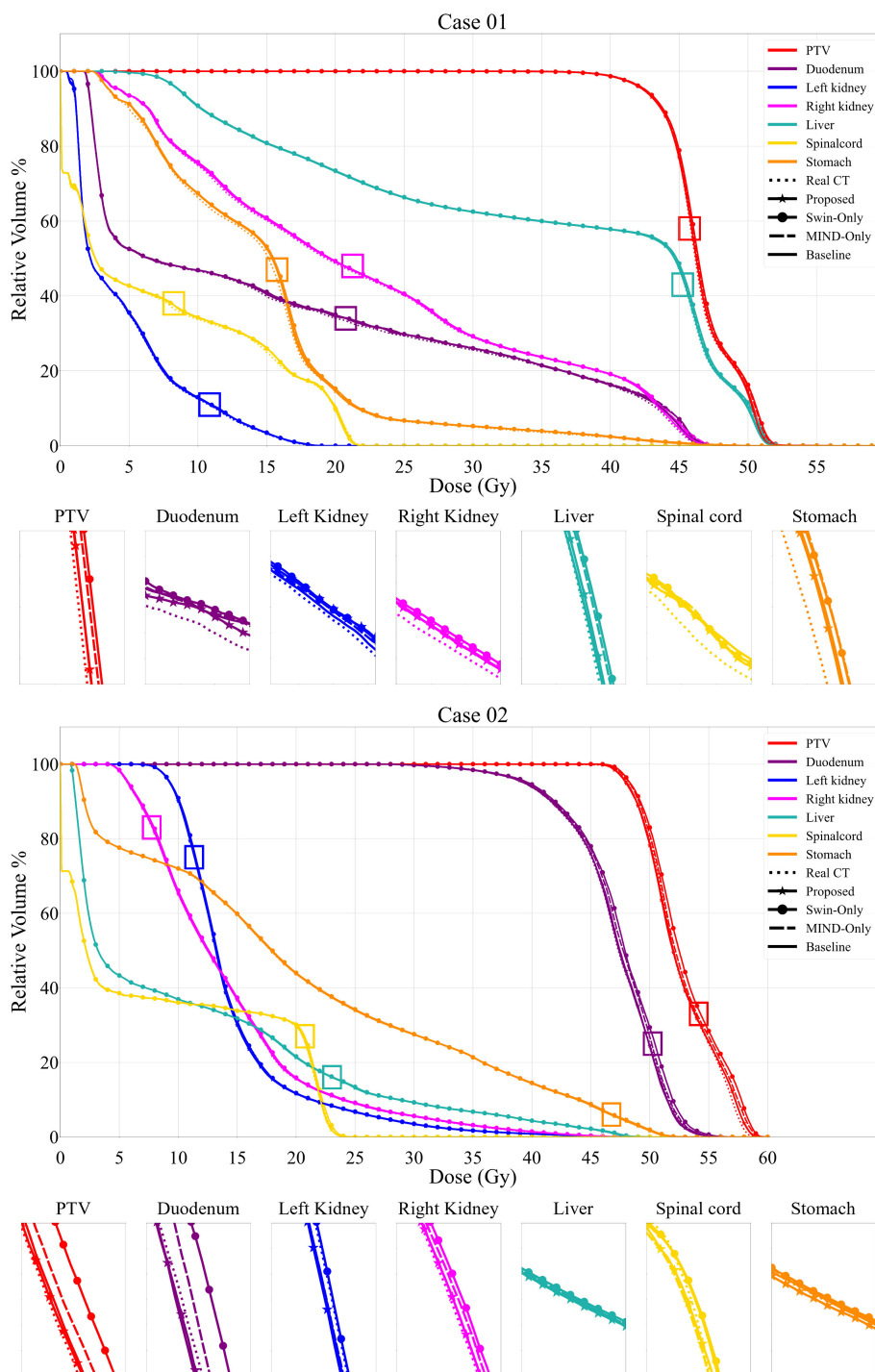
achieved an MAE of 94.1 HU and a PSNR of 27.2 dB. In the dose evaluation, both cGAN and cycle-GAN achieved a gamma passing rate of 99% under the 3%/3 mm criteria and the mean dose difference in the PTV was -0.17%. The MAE values of proposed method (79.5 ± 11.7 HU) were comparable to those in Cusumano et al. (28) and slightly higher than the 89.8 ± 18.7 HU reported by Fu et al. (48). However, the gamma passing rates were not as high as the 99.8 ± 0.2% reported by Cusumano et al. or the 99.5 ± 0.7% Fu et al. (28) (48). There are several possible reasons for the differences. First, the regional difference between MAE and dose calculation. While the MAE was calculated within the patient body, dose calculation was conducted in relatively smaller regions than MAE due to dose thresholding. Second, MR imaging sequence. The previous studies predominantly utilized breath-hold MR images (28, 30, 48). In contrast, this study employed free-breathing MR images without breath-hold which includes more artifacts. Despite of this, our results demonstrate a comparable quantitative results. Lastly, local gamma pass ratio. This study utilized local gamma analysis, providing a stricter assessment of geometric alignment and intensity consistency compared to the conventional global gamma analysis (49).

In the case of DVH difference, 5% is generally considered as the action level according to the TG-119 report and TG-218 report (50, 51). However, to the best of the authors' knowledge, no specific guidelines for acceptable differences between calculation methods are clearly defined. Therefore, in this study, the dose criteria were referenced for different treatment plans for each patient, and Figure 6 shows that the synthetic CT satisfies the dose criteria used in the planning CT. This judgement is informed by radiation oncologists' and physicists' expertise, clinical experience, and the specific anatomical and functional considerations relevant to each case. Thus, the observed differences in DVH values are interpreted within the context of these patient-specific clinical priorities, allowing for variability in assessment depending on the unique circumstances of each treatment plan. Furthermore, as illustrated in Figure 7, the proposed method demonstrated robust performance, with differences within 5% for PTV and OAR overall except stomach and GTV. The mean dose differences were -0.09 ± 1.16%, -2.10 ± 2.96%, -0.98 ± 1.40%, and -0.02 ± 1.10 for Proposed, Swin-only, MIND-only, and baseline, respectively, in the PTV; -0.38 ± 1.20%, -2.46 ± 3.37%, -1.31 ± 1.68%, and -0.30 ± 1.28% in the GTV; and 0.58 ± 2.99%, -1.28 ± 3.59%, -0.87 ± 2.95%, and -0.04 ± 3.44% in the duodenum. One case in the stomach showed an outlier for all methods. This was attributed to the limitations of the unpaired dataset, where a high signal in the CT intestine resulted in a dose discrepancy in the stomach.

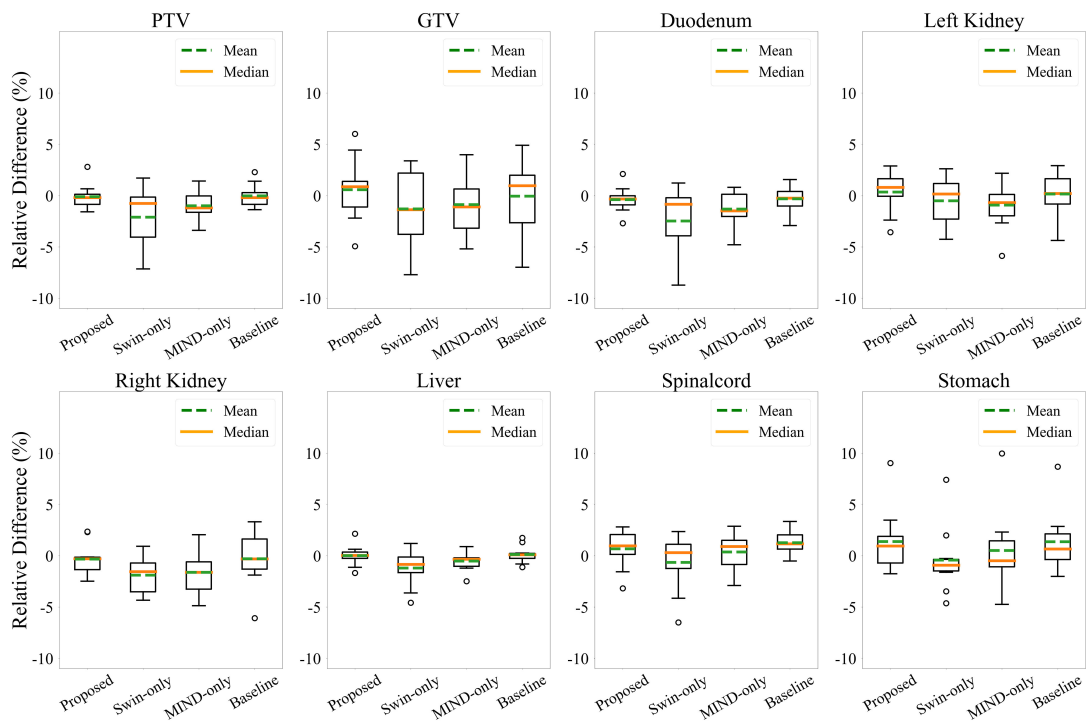
In this study, we referred to the manufacturer's recommendations for acceptable gamma passing rates due to the lack of specific criteria for gamma passing rates among calculations. They recommend a gamma criterion of 3%/3mm with a 10% threshold and global gamma analysis for delivery quality assurance, noting that a passing rate above 95% is considered acceptable under these criteria (52). Under the 3%/3mm criterion with global gamma analysis, we achieved a gamma passing rate of

99%, exceeding the 95% threshold typically used for comparing calculations and measurements. In our study, the gamma passing rates for global gamma analysis were as follows:  $86.1 \pm 5.9\%$  for 1%/1mm,  $97.1 \pm 2.7\%$  for 2%/2mm, and  $98.9 \pm 0.8\%$  for 3%/3mm. These results showed consistency with previous studies, even

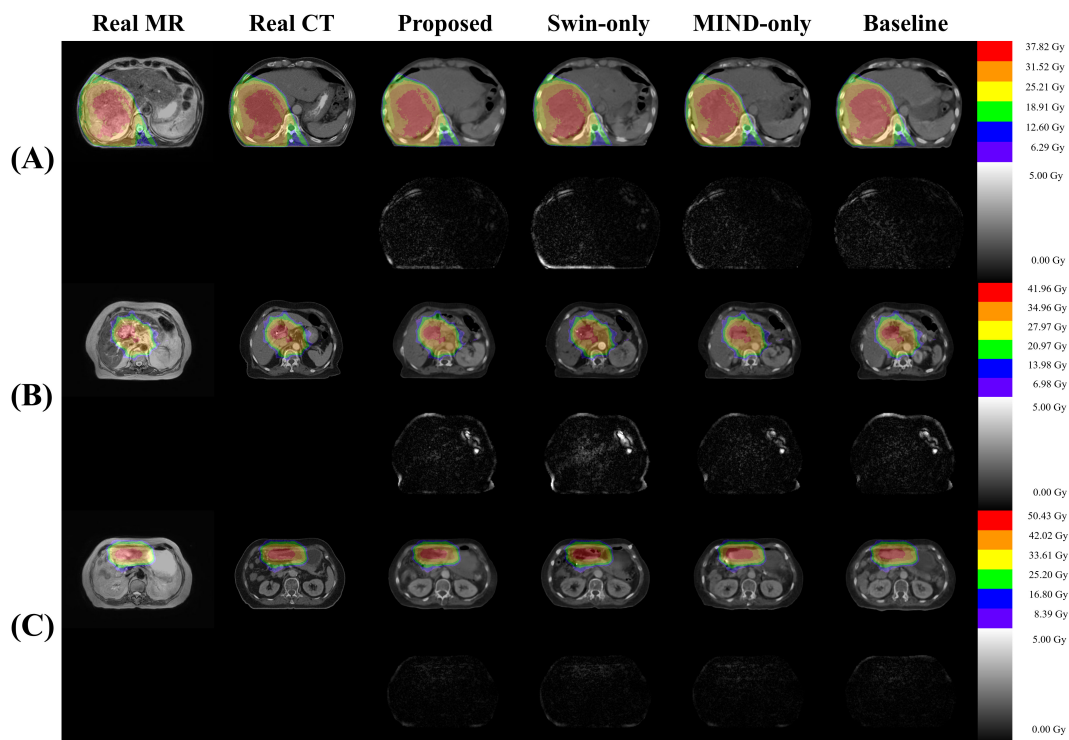
though our studies employed free-breathing abdominal MR which have more artifacts. For example, Olberg et al. (30) reported a gamma passing rate of  $98.3\% \pm 1.3\%$  for the 3%/3mm criterion. Similarly, Fu et al. (48) demonstrated passing rates of  $98.5\% \pm 2.8\%$  for 2%/2mm and  $99.5\% \pm 0.7\%$  for 3%/3mm.



**FIGURE 6** Comparison of dose volume histograms (DVH) for 2 cases, showing the comparison of the proposed method, Swin only, MIND only and baseline sCT and planning CT (solid lines) against the planning CT for the same OARs and each PTV. The PTV dosimetric criteria for Case 1 were  $V_{4.75Gy} > 95\%$ . The  $V_{4.75Gy}$  in real CT was 99.20%, while the values in Proposed, Swin-only, MIND only, and Baseline were 99.25%, 99.43%, 99.38%, and 99.11%, respectively. In Case 2, with the same criteria, the  $V_{4.75Gy}$  in real CT was 98.91%, while the values in Proposed, Swin-only, MIND only, and Baseline were 98.09%, 99.11%, 98.82%, and 98.31%, respectively.



**FIGURE 7** Relative dose volume histogram (DVH) differences for the PTV, GTV and organ-at-risks across the 10 patients. The DVH from proposed method had less than 5% average differences for all structures except the spinal cord and stomach.



**FIGURE 8** Comparison of dose distributions for magnetic resonance (MR), computed tomography (CT), and synthetic CT (sCT) images for three subjects (A–C). Each subpanel includes an overlay of dose distribution on the images in first row and the dose differences in the second row.

TABLE 5 Gamma passing rates for Group 1 (CT and MR with good anatomical alignment) and Group 2 (CT and MR with poor anatomical alignment), with p-values indicating differences between the groups.

	Group 1 Gamma passing rate	Group 2 Gamma passing rate	p-value
Local gamma analysis (3%/3 mm)			
<i>Proposed</i>	0.976 ± 0.011	0.971 ± 0.012	0.29
<i>Swin only</i>	0.973 ± 0.012	0.967 ± 0.013	0.29
<i>MIND only</i>	0.978 ± 0.011	0.974 ± 0.010	0.26
<i>Baseline</i>	0.978 ± 0.012	0.972 ± 0.013	0.26
Global gamma analysis (3%/3 mm)			
<i>Proposed</i>	<b>0.993 ± 0.006</b>	0.986 ± 0.008	<b>0.04</b>
<i>Swin only</i>	<b>0.992 ± 0.006</b>	0.984 ± 0.008	<b>0.03</b>
<i>MIND only</i>	<b>0.994 ± 0.006</b>	0.988 ± 0.007	<b>0.04</b>
<i>Baseline</i>	<b>0.994 ± 0.006</b>	0.987 ± 0.008	< <b>0.05</b>
Local gamma analysis (2%/2 mm)			
<i>Proposed</i>	0.924 ± 0.028	0.921 ± 0.026	0.55
<i>Swin only</i>	0.918 ± 0.029	0.913 ± 0.029	0.50
<i>MIND only</i>	0.929 ± 0.027	0.927 ± 0.026	0.65
<i>Baseline</i>	0.927 ± 0.029	0.922 ± 0.029	0.45
Global gamma analysis (2%/2 mm)			
<i>Proposed</i>	<b>0.978 ± 0.013</b>	0.963 ± 0.017	<b>0.04</b>
<i>Swin only</i>	<b>0.975 ± 0.014</b>	0.959 ± 0.018	< <b>0.05</b>
<i>MIND only</i>	<b>0.981 ± 0.012</b>	0.967 ± 0.015	< <b>0.05</b>
<i>Baseline</i>	0.980 ± 0.013	0.964 ± 0.018	0.07

Statistically significant results are highlighted in bold.

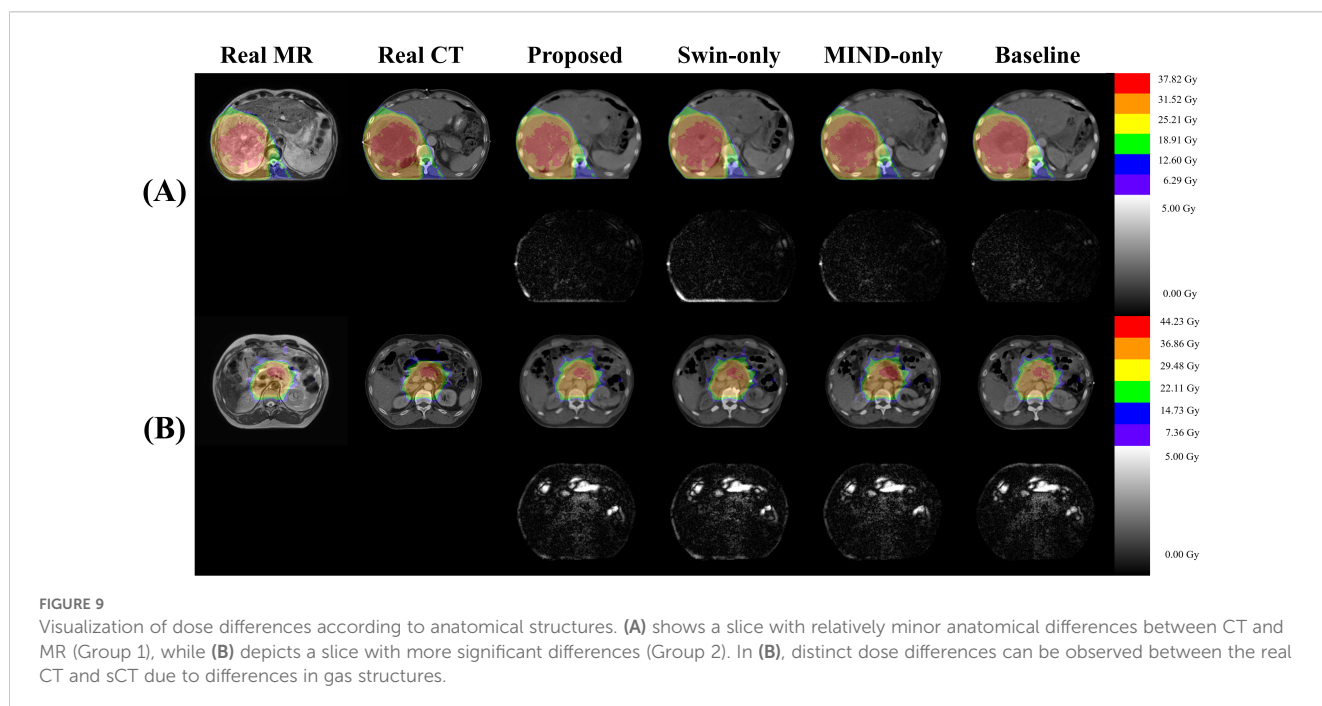


FIGURE 9

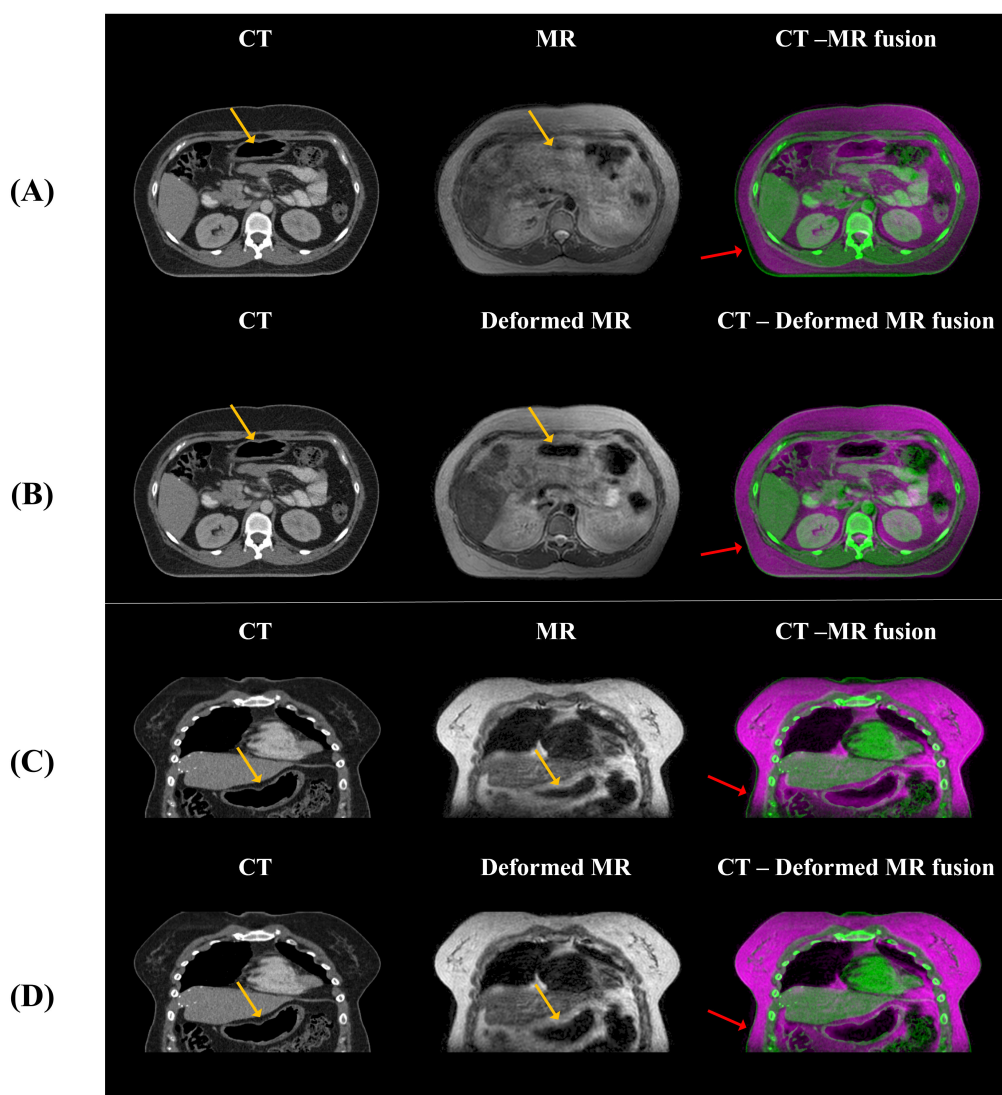
Visualization of dose differences according to anatomical structures. (A) shows a slice with relatively minor anatomical differences between CT and MR (Group 1), while (B) depicts a slice with more significant differences (Group 2). In (B), distinct dose differences can be observed between the real CT and sCT due to differences in gas structures.

**TABLE 6** KL divergence for Group 1 (CT and MR with good anatomical alignment) and Group 2 (CT and MR with poor anatomical alignment), with p-values indicating differences between the groups.

	Group 1 KL divergence	Group 2 KL divergence	p-value
Proposed	0.231 ± 0.089	0.229 ± 0.044	0.33
Swin only	0.225 ± 0.067	0.225 ± 0.034	0.13
MIND only	0.343 ± 0.093	0.335 ± 0.086	0.82
Baseline	0.430 ± 0.058	0.466 ± 0.075	0.23

Cusumano et al. (28) reported gamma passing rates of 90.8% ± 4.5%, 98.7% ± 1.1%, and 99.8% ± 0.2% for 1%/1mm, 2%/2mm, and 3%/3mm, respectively. However, they did not specify their gamma analysis global or local.

Although proposed method successfully generated sCT, there are few limitations. First, uncertainties of evaluation between sCT and CT. Since MR images are not scanned immediately after the planning CT, there are inherent geometric discrepancies between the sCT generated from online MR and the planning CT (22). Those anatomical discrepancies including intestinal gas and weight change result in inaccurate evaluation of image and dosimetric evaluations. Figure 9 and Table 5 demonstrates the dose differences attributable to anatomical discrepancies between MR and CT, but Table 6, which compares the intensity distributions of the images, shows no statistically significant differences between groups 1 and 2. Second, artifacts in the MR images limited the quality of the sCT. The MR data used in this study were online MR images captured during abdominal treatment with Unity. However, artifacts caused by respiratory and organ movements were still present, and the kidneys were not well distinguished from surrounding organs in the



**FIGURE 10** The anatomical differences between CT and MR, and between CT and deformed MR, are shown. (A, B) display axial views, while (C, D) show coronal views, illustrating that anatomical differences with CT are reduced when deformation is applied to the MR compared to the original MR. The yellow arrows indicate alignment of internal gas structures between CT and MR after deformable registration, while the red arrows show approximate alignment of the body external contour between CT and MR. However, the blue arrows highlight regions where intestinal gas structures still do not align between CT and MR.

MR images. These limitations in the input MR conditions resulted in a decrease in the overall quality of the sCT. These artifacts could be mitigated by breathing considered sequences (e.g., shallow breathing, breath holding) or using software-based corrections (53–56). Lastly, while the proposed method's sCT showed the best results in terms of image metrics and qualitative comparison, there were no statistically significant differences in gamma passing rates compared to the other methods. Although correlation between the MAE and the gamma passing rate was reported, unpaired dataset and different calculation region between MAE and gamma pass ratio due to the dose thresholding, the correlation between gamma pass rate and MAE could be weaker (57). Similarly, in the comparison of cGAN and cycle-GAN conducted by Fu et al. (48), the MAE values were  $89.8 \pm 18.7$  HU and  $94.1 \pm 30.0$  HU, respectively, while the 3%/3 mm gamma analysis with a 30% dose threshold showed passing rates of  $99.5 \pm 0.8\%$  and  $99.5 \pm 0.7\%$ , respectively.

## 5 Conclusion

This study proposed the generation of sCT images from MR images obtained from 1.5T MR-Linacs using cycle-GAN. By modifying the generator to Swin-UNETR and incorporating a structure-conserving loss, proposed method was able to enhance both image quality and dosimetric accuracy for abdominal cancer patients. An ablation study demonstrated that the proposed method improves geometric consistency and texture homogeneity of sCT images compared to other models. This study underscores the importance of considering both local and global features, and structure preservation for sCT.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Ethics statement

The studies involving humans were approved by Yonsei University Institutional Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

## Author contributions

CL: Data curation, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. YY: Investigation, Methodology, Software, Validation, Writing – original draft, Writing – review & editing. JS: Resources, Software, Validation, Writing – original draft, Writing – review & editing. JWK: Resources, Validation, Writing – review & editing. YC: Resources, Validation, Writing – review & editing. JK: Resources, Software, Validation, Writing – review & editing. JC: Conceptualization, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. JSK: Conceptualization, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

## Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2022R1A2C2008623). This research was supported by the Korea Health Industry Development Institute (KHIDI) grant funded by the Ministry of Health & Welfare, Republic of Korea (No. HI23C0730000023). This research was financially supported by the Ministry of Trade, Industry and Energy (MOTIE) and Korea Institute for Advancement of Technology (KIAT) through the International Cooperative R&D program (Project No.P0019304).

## Conflict of interest

Author JC was employed by the company Oncosoft Inc. Author JSK is a co-founder of Oncosoft Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Schmidt MA, Payne GS. Radiotherapy planning using mri. *Phys Med Biol.* (2015) 60:R323. doi: 10.1088/0031-9155/60/22/R323
- Corradini S, Alongi F, Andratschke N, Belka C, Boldrini L, Cellini F, et al. Mr-guidance in clinical reality: current treatment challenges and future perspectives. *Radiat Oncol.* (2019) 14:1–12. doi: 10.1186/s13014-019-1308-y
- Boldrini L, Cusumano D, Chiloiro G, Casà C, Masciocchi C, Lenkowitz J, et al. Delta radiomics for rectal cancer response prediction with hybrid 0.35 T magnetic resonance-guided radiotherapy (Mrgt): A hypothesis-generating study for an innovative personalized medicine approach. *La radiol Med.* (2019) 124:145–53. doi: 10.1007/s11547-018-0951-y
- Henke LE, Contreras J, Green O, Cai B, Kim H, Roach M, et al. Magnetic resonance image-guided radiotherapy (Mrgt): A 4.5-year clinical experience. *Clin Oncol.* (2018) 30:720–7. doi: 10.1016/j.clon.2018.08.010
- Daisne J-F, Sibomana M, Bol A, Cosnard G, Lonnewux M, Grégoire V. Evaluation of a multimodality image (Ct, mri and pet) coregistration procedure on phantom and head and neck cancer patients: accuracy, reproducibility and consistency. *Radiation Oncol.* (2003) 69:237–45. doi: 10.1016/j.radonc.2003.10.009
- Dean C, Sykes J, Cooper R, Hatfield P, Carey B, Swift S, et al. An evaluation of four ct-mri co-registration techniques for radiotherapy treatment planning of prone rectal cancer patients. *Br J Radiol.* (2012) 85:61–8. doi: 10.1259/bjr/11855927
- Edmund JM, Nyholm T. A review of substitute ct generation for mri-only radiation therapy. *Radiat Oncol.* (2017) 12:1–15. doi: 10.1186/s13014-016-0747-y
- Liu Y, Lei Y, Wang Y, Wang T, Ren L, Lin L, et al. Mri-based treatment planning for proton radiotherapy: dosimetric validation of a deep learning-based liver synthetic ct generation method. *Phys Med Biol.* (2019) 64:145015. doi: 10.1088/1361-6560/ab25bc
- Nyholm T, Nyberg M, Karlsson MG, Karlsson M. Systematisation of spatial uncertainties for comparison between a mr and a ct-based radiotherapy workflow for prostate treatments. *Radiat Oncol.* (2009) 4:1–9. doi: 10.1186/1748-717X-4-54
- Roberson PL, McLaughlin PW, Narayana V, Troyer S, Hixson GV, Kessler ML. Use and uncertainties of mutual information for computed tomography/magnetic resonance (Ct/mr) registration post permanent implant of the prostate. *Med Phys.* (2005) 32:473–82. doi: 10.1118/1.1851920
- Ulin K, Urie MM, Cherlow JM. Results of a multi-institutional benchmark test for cranial ct/mr image registration. *Int J Radiat Oncol Biol Phys.* (2010) 77:1584–9. doi: 10.1016/j.ijrobp.2009.10.017
- Van Herk M ed. Errors and Margins in Radiotherapy. In: *Seminars in radiation oncology.* Amsterdam, Netherlands: Elsevier.
- Brandner ED, Wu A, Chen H, Heron D, Kalnicki S, Komanduri K, et al. Abdominal organ motion measured using 4d ct. *Int J Radiat Oncol Biol Phys.* (2006) 65:554–60. doi: 10.1016/j.ijrobp.2005.12.042
- Gach HM, Curcuru AN, Mutic S, Kim T. B(0) field homogeneity recommendations, specifications, and measurement units for mri in radiation therapy. *Med Phys.* (2020) 47:4101–14. doi: 10.1002/mp.14306
- Owringi AM, Greer PB, Glide-Hurst CK. Mri-only treatment planning: benefits and challenges. *Phys Med Biol.* (2018) 63:05TR1. doi: 10.1088/1361-6560/aaaca4
- Florkow MC, Guerreiro F, Zijlstra F, Seravalli E, Janssens GO, Maduro JH, et al. Deep learning-enabled mri-only photon and proton therapy treatment planning for paediatric abdominal tumours. *Radiation Oncol.* (2020) 153:220–7. doi: 10.1016/j.radonc.2020.09.056
- Spadea MF, Maspero M, Zaffino P, Seco J. Deep learning based synthetic-ct generation in radiotherapy and pet: A review. *Med Phys.* (2021) 48:6537–66. doi: 10.1002/mp.v48.11
- Johnstone E, Wyatt JJ, Henry AM, Short SC, Sebag-Montefiore D, Murray L, et al. Systematic review of synthetic computed tomography generation methodologies for use in magnetic resonance imaging-only radiation therapy. *Int J Radiat Oncol Biol Phys.* (2018) 100:199–217. doi: 10.1016/j.ijrobp.2017.08.043
- Largent A, Barateau A, Nunes J-C, Lafond C, Greer PB, Dowling JA, et al. Pseudo-ct generation for mri-only radiation therapy treatment planning: comparison among patch-based, atlas-based, and bulk density methods. *Int J Radiat Oncol Biol Phys.* (2019) 103:479–90. doi: 10.1016/j.ijrobp.2018.10.002
- Huynh T, Gao Y, Kang J, Wang L, Zhang P, Lian J, et al. Estimating ct image from mri data using structured random forest and auto-context model. *IEEE Trans Med Imaging.* (2015) 35:174–83. doi: 10.1109/TMI.2015.2461533
- Han X. Mr-based synthetic ct generation using a deep convolutional neural network method. *Med Phys.* (2017) 44:1408–19. doi: 10.1002/mp.12155
- Lei Y, Harms J, Wang T, Liu Y, Shu HK, Jani AB, et al. Mri-only based synthetic ct generation using dense cycle consistent generative adversarial networks. *Med Phys.* (2019) 46:3565–81. doi: 10.1002/mp.13617
- Maspero M, Savenije MH, Dinkla AM, Seevinck PR, Intven MP, Jurgenliemk-Schulz IM, et al. Dose evaluation of fast synthetic-ct generation using a generative adversarial network for general pelvis mr-only radiotherapy. *Phys Med Biol.* (2018) 63:185001. doi: 10.1088/1361-6560/aaada6
- Emami H, Dong M, Nejad-Davaran SP, Glide-Hurst CK. Generating synthetic cts from magnetic resonance images using generative adversarial networks. *Med Phys.* (2018) 45:3627–36. doi: 10.1002/mp.2018.45.issue-8
- Chen S, Qin A, Zhou D, Yan D. U-net-generated synthetic ct images for magnetic resonance imaging-only prostate intensity-modulated radiation therapy treatment planning. *Med Phys.* (2018) 45:5659–65. doi: 10.1002/mp.2018.45.issue-12
- Arabi H, Dowling JA, Burgos N, Han X, Greer PB, Koutsouvelis N, et al. Comparative study of algorithms for synthetic ct generation from mri: consequences for mri-guided radiation planning in the pelvic region. *Med Phys.* (2018) 45:5218–33. doi: 10.1002/mp.2018.45.issue-11
- Rehman A, Khan FG. A deep learning based review on abdominal images. *Multimed Tools Appl.* (2021) 80:30321–52. doi: 10.1007/s11042-020-09592-0
- Cusumano D, Lenkowitz J, Votta C, Boldrini L, Placidi L, Catucci F, et al. A deep learning approach to generate synthetic ct in low field mr-guided adaptive radiotherapy for abdominal and pelvic cases. *Radiation Oncol.* (2020) 153:205–12. doi: 10.1016/j.radonc.2020.10.018
- Nousiainen K, Santurio GV, Lundahl N, Cronholm R, Siverson C, Edmund JM. Evaluation of mri-only based online adaptive radiotherapy of abdominal region on mrlinac. *J Appl Clin Med Phys.* (2023) 24:e13838. doi: 10.1002/acm2.13838
- Olberg S, Chun J, Choi BS, Park I, Kim H, Kim T, et al. Abdominal synthetic ct reconstruction with intensity projection prior for mri-only adaptive radiotherapy. *Phys Med Biol.* (2021) 66:204001. doi: 10.1088/1361-6560/ac279e
- Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. *Adv Neural Inf Process Syst.* (2017) 30. doi: 10.48550/arXiv.1706.03762
- Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, et al. An image is worth 16x16 words: transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929.* (2020). doi: 10.48550/arXiv.2010.11929
- Dalmaz O, Yurt M, Çukur T. Resvit: residual vision transformers for multimodal medical image synthesis. *IEEE Trans Med Imaging.* (2022) 41:2598–614. doi: 10.1109/TMI.2022.3167808
- Li J, Qu Z, Yang Y, Zhang F, Li M, Hu S. Tcgan: A transformer-enhanced gan for pet synthetic ct. *Biomed Optics Express.* (2022) 13:6003–18. doi: 10.1364/BOE.467683
- Zhao B, Cheng T, Zhang X, Wang J, Zhu H, Zhao R, et al. Ct synthesis from mr in the pelvic area using residual transformer conditional gan. *Computer Med Imaging Graphics.* (2023) 103:102150. doi: 10.1016/j.compmedimag.2022.102150
- Heinrich MP, Jenkinson M, Bhushan M, Matin T, Gleeson FV, Brady M, et al. Mind: modality independent neighbourhood descriptor for multi-modal deformable registration. *Med Image Anal.* (2012) 16:1423–35. doi: 10.1016/j.media.2012.05.008
- Hatamizadeh A, Nath V, Tang Y, Yang D, Roth HR, Xu D eds. Swin Unetr: Swin Transformers for Semantic Segmentation of Brain Tumors in Mri Images. In: *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries: 7th International Workshop, BrainLes 2021, Held in Conjunction with MICCAI 2021, Virtual Event, September 27, 2021, Revised Selected Papers, Part I.* New York, USA: Springer.
- Hatamizadeh A, Tang Y, Nath V, Yang D, Myronenko A, Landman B, et al. (2022). Unetr: transformers for 3d medical image segmentation, in: *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, Piscataway, New Jersey, USA: IEEE.
- Zhu J-Y, Park T, Isola P, Efros AA. (2017). Unpaired image-to-image translation using cycle-consistent adversarial networks, in: *Proceedings of the IEEE international conference on computer vision*, Piscataway, New Jersey, USA: IEEE.
- Yang H, Sun J, Carass A, Zhao C, Lee J, Prince JL, et al. Unsupervised mr-to-ct synthesis using structure-constrained cyclegan. *IEEE Trans Med Imaging.* (2020) 39:4249–61. doi: 10.1109/TMI.42
- Kullback S, Leibler RA. On information and sufficiency. *Ann Math Stat.* (1951) 22:79–86. doi: 10.1214/aoms/1177729694
- Wilcoxon F. Individual Comparisons by Ranking Methods. In: *Breakthroughs in Statistics: Methodology and Distribution.* New York, USA: Springer (1992). p. 196–202.
- Low DA, Harms WB, Mutic S, Purdy JA. A technique for the quantitative evaluation of dose distributions. *Med Phys.* (1998) 25:656–61. doi: 10.1118/1.598248
- Kodali N, Abernethy J, Hays J, Kira Z. On convergence and stability of gans. *arXiv preprint arXiv:170507215.* (2017). doi: 10.48550/arXiv.1705.07215
- Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, et al. Swin transformer: hierarchical vision transformer using shifted windows. *Proc IEEE/CVF Int Conf Comput Vision.* (2021) 10012–22. doi: 10.1109/ICCV48922.2021.00986
- Kieselmann JP, Fuller CD, Gurney-Champion OJ, Oelfke U. Cross-modality deep learning: contouring of mri data from annotated ct data only. *Med Phys.* (2021) 48:1673–84. doi: 10.1002/mp.14619
- Padgett KR, Stoyanova R, Pirozzi S, Johnson P, Piper J, Dogan N, et al. Validation of a deformable mri to ct registration algorithm employing same day planning mri for surrogate analysis. *J Appl Clin Med Phys.* (2018) 19:258–64. doi: 10.1002/acm2.2018.19.issue-2
- Fu J, Singhrao K, Cao M, Yu V, Santhanam AP, Yang Y, et al. Generation of abdominal synthetic cts from 0.35 T mr images using generative adversarial networks for mr-only liver radiotherapy. *Biomed Phys Eng Express.* (2020) 6:015033. doi: 10.1088/2057-1976/ab61ef

49. Rajasekaran D, Jeevanandam P, Sukumar P, Ranganathan A, Johnjothi S, Nagarajan V. A study on correlation between 2d and 3d gamma evaluation metrics in patient-specific quality assurance for vmat. *Med Dosimetry*. (2014) 39:300–8. doi: 10.1016/j.meddos.2014.05.002
50. Ezzell GA, Burmeister JW, Dogan N, LoSasso TJ, Mechalakos JG, Mihailidis D, et al. Imrt commissioning: multiple institution planning and dosimetry comparisons, a report from aapm task group 119. *Med Phys*. (2009) 36:5359–73. doi: 10.1118/1.3238104
51. Miften M, Olch A, Mihailidis D, Moran J, Pawlicki T, Molineu A, et al. Tolerance limits and methodologies for imrt measurement-based verification qa: recommendations of aapm task group no. 218. *Med Phys*. (2018) 45:e53–83. doi: 10.1002/mp.12810
52. Powers M, Baines J, Crane R, Fisher C, Gibson S, Marsh L, et al. Commissioning measurements on an Elekta Unity MR-Linac. *Phys Eng Sci Med*. (2022) 45:457–73. doi: 10.1007/s13246-022-01113-7
53. Zaitsev M, Dold C, Sakas G, Hennig J, Speck O. Magnetic resonance imaging of freely moving objects: prospective real-time motion correction using an external optical motion tracking system. *Neuroimage*. (2006) 31:1038–50. doi: 10.1016/j.neuroimage.2006.01.039
54. von Siebenthal M, Szekely G, Gamper U, Boesiger P, Lomax A, Cattin P. 4d mr imaging of respiratory organ motion and its variability. *Phys Med Biol*. (2007) 52:1547. doi: 10.1088/0031-9155/52/6/001
55. Hirokawa Y, Isoda H, Maetani YS, Arizono S, Shimada K, Togashi K. MRI artifact reduction and quality improvement in the upper abdomen with propeller and prospective acquisition correction (Pace) technique. *Am J Roentgenol*. (2008) 191:1154–8. doi: 10.2214/AJR.07.3657
56. Baumgartner CF, Kolbitsch C, McClelland JR, Rueckert D, King AP. Autoadaptive motion modelling for mr-based respiratory motion estimation. *Med Image Anal*. (2017) 35:83–100. doi: 10.1016/j.media.2016.06.005
57. Huijben EM, Terpstra ML, Pai S, Thummerer A, Koopmans P, Afonso M, et al. Generating synthetic computed tomography for radiotherapy: synthrad2023 challenge report. *Med Image Anal*. (2024) 97:103276. doi: 10.1016/j.media.2024.103276