



OPEN MAN-GAN: a mask-adaptive normalization based generative adversarial networks for liver multi-phase CT image generation

Wei Zhao^{1,2,3,8}, Wenting Chen^{4,8}, Li Fan¹, Youlan Shang¹, Yisong Wang¹, Weijun Situ¹, Wenzheng Li⁵, Tianming Liu⁶, Yixuan Yuan⁷✉ & Jun Liu^{1,3}✉

Liver multiphase enhanced computed tomography (MPECT) is vital in clinical practice, but its utility is limited by various factors. We aimed to develop a deep learning network capable of automatically generating MPECT images from standard non-contrast CT scans. Dataset 1 included 374 patients and was divided into three parts: a training set, a validation set and a test set. Dataset 2 included 144 patients with one specific liver disease and was used as an internal test dataset. We further collected another dataset comprising 83 patients for external validation. Then, we propose a Mask-Adaptive Normalization-based Generative Adversarial Network with Cycle-Consistency Loss (MAN-GAN) to achieve non-contrast CT to MPECT translation. To assess the efficiency of MAN-GAN, we conducted a comparative analysis with state-of-the-art methods commonly employed in diverse medical image synthesis tasks. Moreover, two subjective radiologist evaluation studies were performed to verify the clinical usefulness of the generated images. MAN-GAN outperformed the baseline network and other state-of-the-art methods in all generations of the three phases. These results were verified in internal and external datasets. According to radiological evaluation, the image quality of generated three phase images are all above average. Moreover, the similarities between real images and generated images in all three phases are satisfactory. MAN-GAN demonstrates the feasibility of liver MPECT image translation based on non-contrast images and achieves state-of-the-art performance via the subtraction strategy. It has great potential for solving the dilemma of liver CT contrast canning and aiding further liver interaction clinical scenarios.

Keywords Multiphase enhanced computed tomography, Generative adversarial networks, Image synthesis

Abbreviations

HCC	Hepatocellular carcinoma
MPECT	Multiphase-enhanced computed tomography
AP	Arterialphase
PVP	Portal venous phase
DP	Delayed phase
ESRD	End-stage renal disease
DL	Deep learning
MAN-GAN	Mask-adaptive normalization-based generative adversarial network with cycle-consistency loss
MANorm	Mask-adaptive normalization
SR	Superresolution

¹Department of Radiology, The Second Xiangya Hospital, Central South University, Changsha 410011, China. ²Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China. ³Clinical Research Center for Medical Imaging in Hunan Province, Changsha 410011, China. ⁴Department of Electrical Engineering, City University of Hong Kong, Kowloon Tong 999077, Hong Kong SAR, China. ⁵Department of Radiology, and National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Changsha 410008, China. ⁶School of Computing, The University of Georgia, Athens 30602, USA. ⁷Department of Electronic Engineering, The Chinese University of Hong Kong, Ma Liu Shui 999077, Hong Kong SAR, China. ⁸Wei Zhao and Wenting Chen are co-first authors. ✉email: yxyuan@ee.cuhk.edu.hk; junliu123@csu.edu.cn

PSNR	Peak signal-to-noise ratio
SSIM	The structural similarity index

Liver lesions are commonly encountered in clinical practice¹, which accounts for approximately 2 million deaths annually with cirrhosis and hepatocellular carcinoma (HCC) being primary causes². Accurately determining the biological nature of lesions can help clinicians select the most appropriate treatment regimen and facilitate better management of liver lesions. Multiphase-enhanced computed tomography (MPECT) provides arterial (AP), portal venous (PVP), and delayed phase (DP) enhancement images, adequately characterizing the blood supply to the liver. Compared to non-contrast CT, MPECT improves the qualitative diagnosis of lesions and the accuracy of tumor staging, as most liver lesions have a specific enhancement pattern³. For example, primary hepatocellular carcinoma is characterized by a “fast-in, fast-out” enhancement pattern. In contrast, most metastatic hepatocellular carcinomas have little blood supply, and the lesions have no enhancement in the AP, whereas changes such as the “bull’s-eye sign” can be observed in the portal venous phase⁴. Benign lesions, like hepatic hemangiomas, typically exhibit an enhancement pattern summarized as ‘fast in, slow out.’ Therefore, MPECT is a necessary imaging modality for avoiding misdiagnoses that could lead to unnecessary interventions.

However, MPECT’s applicability is sometimes limited by contraindications^{5,6}, such as patients who are allergic to contrast agents, patients with end-stage renal disease (ESRD), or patients with unstable vital signs are not recommended for MPECT. MRI offers an alternative but comes with drawbacks like high expense, long scanning time and also not suitable for certain patient groups. Moreover, MPECT or MRI is sometimes unavailable in primary-level hospitals due to the lack of professional nurses and radiological technologists. In this context, generating synthetic MPECT images from non-contrast CT images may constitute a feasible way to overcome this dilemma.

Generating other medical images based on one specific medical image presents a significant advancement in addressing clinical requirements, e.g., reducing scanning time, improving image quality, and solving specific clinical issues⁷. Moreover, synthetic medical images are also considered a data augmentation technique for addressing the issue of medical data scarcity. Recently, as deep learning (DL) has been widely used in image analysis, numerous works have applied deep neural networks for image synthesis^{8–13}. Researchers have extended the application of deep neural networks to CT synthesis, particularly in the context of synthesizing CT images from MRI or positron emission tomography, i.e., intermodal CT translation^{14–20}. Furthermore, other methods were proposed to achieve CT image translation, i.e., intramodal CT translation^{21–23}. However, as a generic intramodal CT translation task, the generation of MPECT images from non-contrast CT images has not been investigated thus far. The challenge lies in the inherent differences between these modalities, known as “domain distance.” CT and MPECT are two medical image modalities with considerable differences in terms of data representation, feature extraction, and information content. Because of these modality differences, their respective data distributions have a considerable domain distance, and a direct transformation may face the challenge of domain adaptation, as the model needs to learn effective feature mappings between two distinct distributions²⁴.

Although several studies have synthesized contrasted images from non-contrast CT images²⁵, none of these works has fully utilized the information of the subtraction image in their approach. Several works have demonstrated that subtraction images play an important role in the detection and diagnosis of lesions^{26,27}. Thus, inspired by the subtraction strategy, our initial step involved subtracting MPECT images from non-contrast CT images as fake subtraction images to guide the generation of MPECT images in our work.

In this paper, we propose a Mask-Adaptive Normalization-based Generative Adversarial Network with Cycle-Consistency Loss (MAN-GAN) to achieve non-contrast CT-to-MPECT translation. To use subtraction images to guide MPECT synthesis, we introduce MaskNet to predict subtraction image-like mask images and integrate them into our network. To the best of our knowledge, this is the first work to apply subtraction imaging for non-contrast CT to MPECT translation. Aiming to fuse mask images with non-contrast CT image features, we propose a generator with mask-adaptive normalization (MANorm) as a normalization layer to highlight contrast-enhanced regions. For image fidelity, a discriminator is adopted to distinguish the generated and real MPECT images to preserve discriminative details. To reduce the space of possible mapping functions, we utilize cycle consistency loss to ensure that the learned mapping functions are cycle consistent. To the best of our knowledge, this is the first study to generate liver multiphase CT images using only non-contrast CT images.

Materials and methods

The research was performed in accordance with the principles expressed in the Declaration of Helsinki. This retrospective study was approved by the hospital Institutional Review Board (IRB), which waived the requirement for informed consent. All methods were performed in accordance with the relevant guidelines and regulations.

Datasets

A total of 374 patients (Dataset 1) who successfully underwent MPECT at one hospital were retrospectively and successively included. We randomly divided this dataset into three subsets at a 6:2:2 ratio: training set (224 patients, 14,756 slices), validation set (75 patients, slices) and test set (75 patients, 5,183 slices). The second independent dataset (Dataset 2), which included 144 patients who underwent MPECT and had one of the four most common diseases in the liver, was utilized as the internal test dataset. Dataset 2 included 144 patients who underwent MPECT and had one of the four most common diseases in the liver, i.e., HCC (45 patients, 2,753 slices), cavernous hemangioma (37 patients, 1,969 slices), hepatic adenoma (16 patients, 834 slices), or cysts (46 patients, 3,083 slices). To verify the generalizability of the proposed model, we further collected an external test dataset comprising 83 MPECT patients (23,594 slices) from another hospital (Dataset 3).

MPECT scanning protocols

For the internal dataset, Somatom Definition Flash, Siemens (Kernel, B30f. medium smooth), German, UCT780, United Imaging, Shanghai (Kernel, B_SOFT_E), Somatom Perspective 128, Siemens, German (Kernel, B31s medium smooth), Somatom Definition Force, Siemens, German (Kernel, Br40), with the following scanning parameters: CT section thickness 1–5 mm, pitch 0.8–1.1, 90–120 kVp, and dynamic mA based on scout images. For the external dataset, all CT scans were performed with one of four scanners: Aquilion ONE, Canon, Japan (Kernel, Std); Revolution, GE, USA (Kernel, Std); Somatom Drive, Siemens, German (Kernel, Bf37); uCT 960+, United Imaging, China (Kernel, B_SOFT_B) with the following scanning parameters: CT section thickness 1–5 mm, pitch 0.8–1.1, 120 kVp, and dynamic mA based on the scout image. The dynamic scanning protocols were as follows: MPECT consisted of non-contrast and contrast-enhanced phases. The iodine contrast agent was administered via a power injector at a flow rate of 4 mL/s, and the AP, portal venous phase (PVP) and delayed phase (DP) were acquired after the injection of the iodine contrast agent at 25 s, 55 s and 3 min, respectively.

Data preprocessing

To align different phase CT scans with standard non-contrast CT scans, a two-step process is employed. First, affine and deformable registration is carried out using advanced normalization tools²⁸ with mutual information as the optimization metric and elastic regularization. This process is conducted separately on two datasets. Subsequently, the aligned CT images are transformed from their initial grayscale range of [0, 4095] to a display range of [0, 255] using a window transformation technique. This transformation involved the use of a window center of 60 HU and a window width of 500 HU to enhance visualization. Then, we resized all the CT images to 512×512 as standard input for training.

Algorithm description

Figure 1 demonstrates the proposed MAN-GAN with cycle-consistency loss. Given a non-contrast CT image x , the forward generator attempts to translate it to a real CT image y by leveraging MANorm to fuse non-contrast CT images with the mask image generated by MaskNet M . To reduce the searching space of the mapping functions, as illustrated in Fig. 1a, we make the forward G and backward F generators cycle consistent by bringing the translated CT image y' to the original image x , i.e., $F(G(x), M(x)) \approx x$, where M denotes the MaskNet. Similarly, the opposite direction also obeys cycle consistency, i.e., $G(F(y), M(F(y))) \approx y$. Since only the forward generator is adopted for MPECT generation and the backward generator is useless during inference, the backward generator does not apply the mask image for CT translation. To achieve this, we apply the cycle-consistency loss:

$$L_{cyc} = \|F(G(x), M(x)) - x\|_1 + \|G(F(y), M(F(y))) - y\|_1 \quad (1)$$

During forward translation, as shown in Fig. 1b, MAN-GAN consists of a MaskNet, a generator and a discriminator. MaskNet was applied to map non-contrast CT images to mask images, subsequently directing the synthesis of contrast-enhanced CT images. MaskNet is an encoder-decoder-based network with three residual blocks in the bottleneck to predict masks. The generator is a variation of the pix2pix generator network, which replaces the residual blocks with a mask-adaptive block (MABlk). Compared to the residual block, MABlk replaces the batch normalization layer with the MANorm layer, as shown in Fig. 1b. MANorm was used to better visualize the fusion mask images and non-contrast CT images. In each MANorm layer, the image features are first normalized and processed with a convolutional layer. The predicted mask is concatenated with the processed image feature to predict the affine parameters γ and β for normalization. Then, we perform elementwise multiplication of

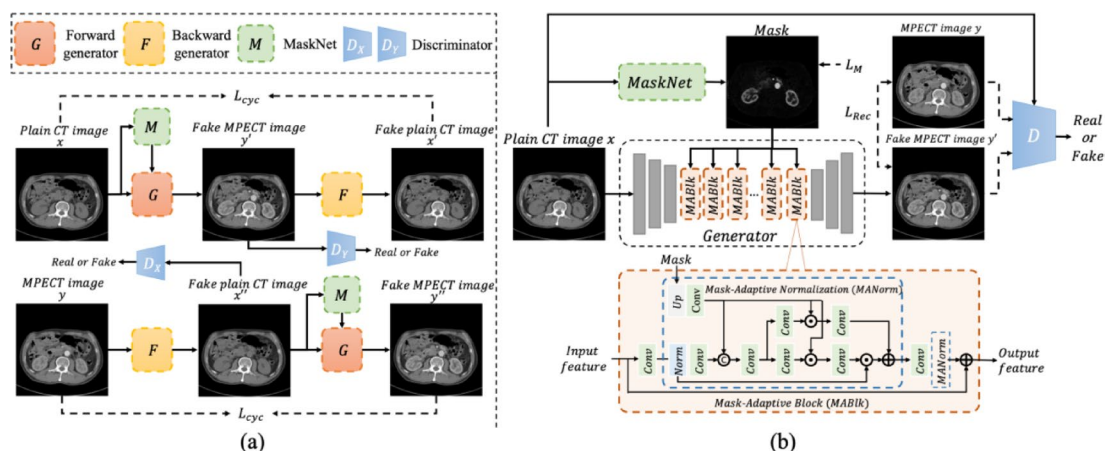


Fig. 1. The overview of Mask-Adaptive Normalization based Generative Adversarial Networks (MAN-GAN) with Cycle-Consistency Loss for CT translation. (a) The forward and backward cycle-consistency loss are utilized to reduce the space of possible mapping functions. (b) The forward generation includes a MaskNet to synthesize subtraction-image-like mask, a generator to fuse mask with non-contrast CT image and a discriminator to distinguish the real from fake MPECT image.

the affine parameters and the predicted mask to further strengthen the feature. Afterward, we calculate the elementwise multiplication of the normalized image feature and γ . Finally, the elementwise addition of the calculated results and β is computed as the final output of the MANorm layer. We apply the discriminator of Pix2Pix in our framework, which concatenates the input CT image and the output/ground-truth CT image as input to distinguish whether the image is a real or fake pair.

The non-contrast CT images were fed as input data, and the synthesis of three phases, i.e., non-contrast CT images to AP images (P to AP), non-contrast CT images to PVP images (P to V), and non-contrast CT images to DP images (P to D), was successively conducted using MAN-GAN. We trained and tested the main dataset using the generator of pix2pix as the baseline; this generator is composed of three convolutions, nine residual blocks, two fully convolutional layers, and a convolution that maps features to RGB images.

To optimize the proposed networks, the L1 loss is utilized as a reconstruction loss L_{Rec} to compute the distance between the output and ground-truth CT image. In addition, we compute the L1 loss between the generated and real masks, where the real mask is predefined by subtracting the ground-truth CT image from the input CT image. For the training strategy, an Adam optimizer with a learning rate of 1×10^{-4} and a batch size of 8 was used. We implemented our method with PyTorch²⁹ and trained it for approximately 3 days. All the experiments were conducted on an NVIDIA Tesla V100 GPU with 32 GB of memory. The code and data will be released once the paper is accepted.

Quantitatively evaluation

To further verify the efficiency of MAN-GAN, we compared it with state-of-the-art methods used for different medical image synthesis tasks, including pix2pix, CycleGAN, CUT, and F-LSeSim. All of these previous methods were trained and tested on our main dataset. To quantitatively evaluate the superresolution (SR) performance, we employ the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM). The PSNR computes the pixelwise difference between the generated CT image and the real CT image, while the SSIM evaluates the structural and perceptual similarity. For these metrics, a higher value indicates greater similarity. The PSNR is defined as:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2, \quad (2)$$

$$PSNR = 10 \cdot \log_{10} \left(\frac{255^2}{MSE} \right), \quad (3)$$

where I and K denote the real and fake CT images, respectively, and m and n represent the height and width, respectively, of the CT images. The SSIM is formulated as follows:

$$SSIM(I, K) = \frac{(2\mu_I\mu_K + c_1)(2\sigma_{IK} + c_2)}{(\mu_I^2 + \mu_K^2 + c_1)(\sigma_I^2 + \sigma_K^2 + c_2)}, \quad (4)$$

where μ_I and μ_K denote the average of I and K , σ_I^2 and σ_K^2 represent the variance of I and K , σ_{IK} denotes the covariance of I and K , and both c_1 and c_2 are constant values. To implement the CT translation for each phase, we train different independent MAN-GAN models for each phase. Moreover, we evaluate the mask images generated by MAN-GAN and compare the performances of the various methods for different phrases.

Radiological evaluation

To verify the efficiency of our proposed model, we performed two subjective evaluation studies on the test set. *Evaluation study 1 (quality analysis)*: Two radiologists (with 6 and 8 years of experience in liver imaging) who were unaware of the circumstances in which the tested images were generated independently evaluated the image quality using a 5-point scale. The image sets were scored as 1, unacceptable; 2, substandard; 3, acceptable; 4, above average; or 5, superior. The qualitative image quality score was based on image noise, soft-tissue contrast, and the sharpness of organ boundaries for the liver, adrenal glands, kidneys, pancreas, and abdominal wall³⁰. *Evaluation study 2 (similarity analysis)*: Two radiologists (with 2 and 4 years of experience in liver imaging) independently evaluated the image similarity between the ground truth and generated images based on the following five parts: structural deformation, image artifacts, degree of enhancement, homogeneity of enhancement, and imaging distortion. There were 20 total points in each part, and the total number of points in the evaluation was 100. The evaluators were informed of which group of images was ground truth, and the images in the ground truth group were selected as the reference for the evaluation. In evaluation studies 1 and 2, a re-evaluation to achieve a consensus was performed to resolve any disagreements between the two radiologists.

Statistical analysis

The statistical analysis was performed using SPSS software. The mean values and standard deviations are expressed for continuous variables. ANOVA and the LSD test were used to compare the differences between and among the groups. $P < 0.05$ indicated a significant difference.

Methods	Mode conversion					
	P to AP		P to PVP		P to DP	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Non-contrast CT Images	26.02 ± 2.72	0.74 ± 0.07	24.73 ± 2.81	0.73 ± 0.07	24.84 ± 3.28	0.73 ± 0.08
pix2pix ¹⁰	26.41 ± 2.63	0.78 ± 0.06	25.57 ± 2.77	0.77 ± 0.06	25.01 ± 3.11	0.76 ± 0.07
CycleGAN ¹³	26.81 ± 2.73	0.79 ± 0.06	25.88 ± 2.98	0.78 ± 0.07	25.57 ± 3.44	0.77 ± 0.07
CUT ¹²	26.58 ± 2.68	0.79 ± 0.06	24.51 ± 2.31	0.75 ± 0.06	24.90 ± 3.03	0.75 ± 0.07
F-LSeSim ¹¹	25.54 ± 2.29	0.78 ± 0.06	25.59 ± 2.72	0.77 ± 0.06	25.44 ± 3.25	0.77 ± 0.07
DINO ⁹	26.88 ± 2.79	0.79 ± 0.06	25.21 ± 3.24	0.76 ± 0.06	24.96 ± 3.27	0.75 ± 0.07
CHAN ⁸	26.67 ± 2.75	0.79 ± 0.06	25.67 ± 2.80	0.78 ± 0.06	25.14 ± 3.25	0.77 ± 0.07
Sun et al. ²¹	26.48 ± 2.50	0.78 ± 0.06	25.52 ± 2.72	0.77 ± 0.06	25.05 ± 3.13	0.76 ± 0.07
Chen et al. ²²	27.45 ± 2.43	0.81 ± 0.03	26.49 ± 3.20	0.80 ± 0.06	25.86 ± 3.50	0.79 ± 0.07
MAN-GAN	27.71 ± 2.96	0.82 ± 0.06	26.76 ± 3.10	0.80 ± 0.06	26.17 ± 3.62	0.79 ± 0.07

Table 1. Comparison with existing methods on dataset 1. Significant values are in bold. Values were presented as mean ± SD.

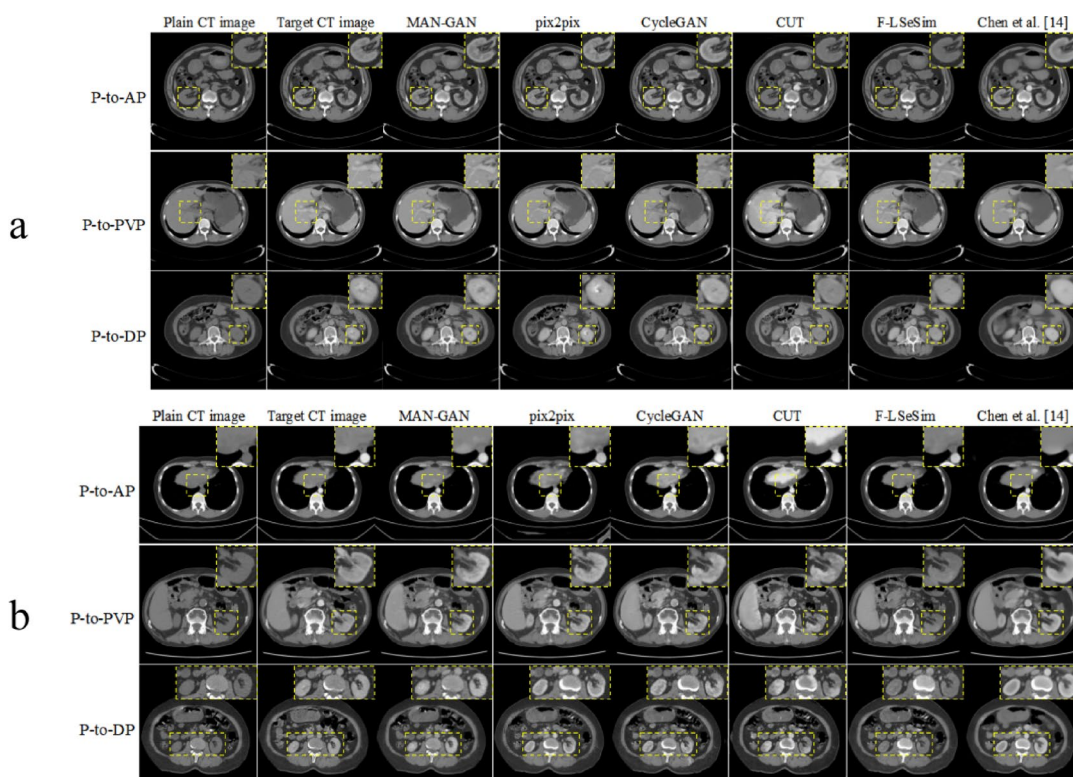


Fig. 2. Visual comparison of existing methods and MAN-GAN on Dataset 1 (a) and Dataset 3 (b).

Results

Comparison with state-of-the-art methods

To prove the effectiveness of our proposed method, we compare the proposed MAN-GAN with the existing methods on the main test dataset. As shown in Table 1, our method obtained the best PSNR and SSIM in all generations of the three phases in comparison to the current methods. Figure 2a shows the enhanced CTs generated from the main dataset for different phases of CT translation. In P-to-AP translation, as shown in the magnified patches, MAN-GAN can synthesize CT images with more discriminative features than can other methods. When translating from plain to PVP CT images, the synthetic CT image generated by MAN-GAN is contrast-enhanced more accurately than that generated by other methods, suggesting the superiority of the proposed MAN-GAN to existing methods. As shown in Table 2, the mask images for P-to-AP translation preserve the best performance, with a PSNR of 25.2738, in comparison with the mask images for other phrases. In Fig. 3, we qualitatively evaluate the generated mask images for CT translation by computing the PSNR and

	PSNR	SSIM
P-to-AP	25.27 ± 2.80	0.57 ± 0.07
P-to-PVP	24.37 ± 2.76	0.59 ± 0.07
P-to-DP	24.13 ± 2.96	0.59 ± 0.08

Table 2. Performance of the generated mask images. Values were presented as mean ± SD.

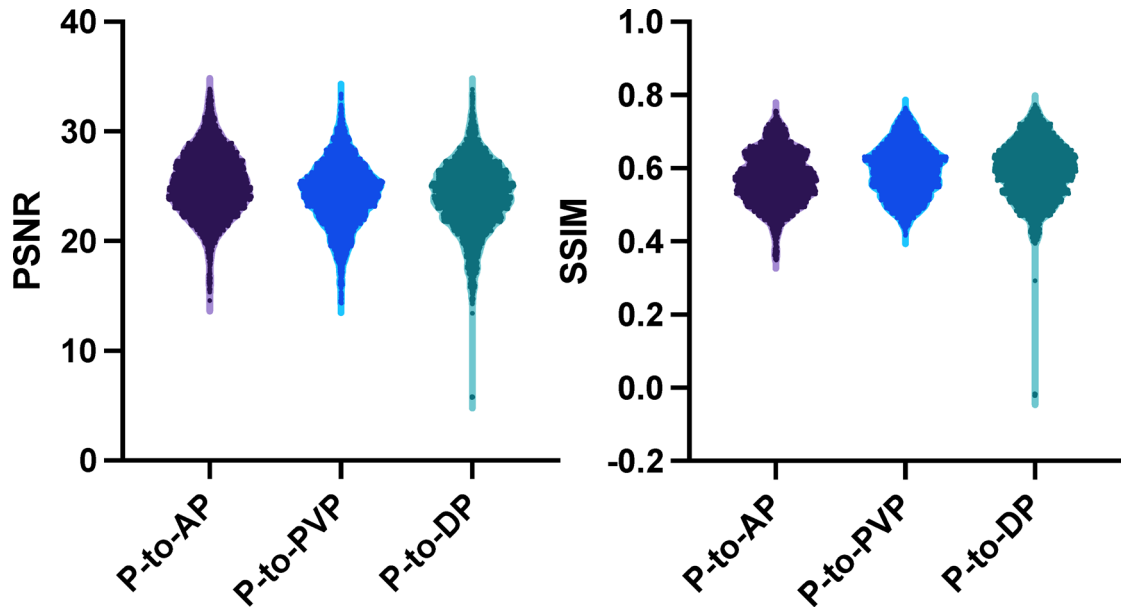


Fig. 3. The qualitative evaluation of the CT translation for three MPECT phases.

Task	Methods	Abnormal Dataset									
		FNH		HCC		HEM		CYST		Overall	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
P to AP	Baseline	26.47 ± 2.85	0.78 ± 0.06	25.56 ± 3.30	0.76 ± 0.07	25.50 ± 3.33	0.77 ± 0.07	26.58 ± 3.18	0.81 ± 0.06	26.00 ± 3.26	0.78 ± 0.07
	MAN-GAN	27.27 ± 2.91	0.81 ± 0.05	26.03 ± 3.45	0.78 ± 0.06	25.95 ± 3.47	0.79 ± 0.07	27.12 ± 3.30	0.83 ± 0.05	26.52 ± 3.40	0.80 ± 0.06
P to PVP	Baseline	25.82 ± 2.92	0.78 ± 0.06	24.99 ± 3.21	0.75 ± 0.07	24.27 ± 3.51	0.75 ± 0.08	26.32 ± 2.85	0.81 ± 0.06	25.38 ± 3.23	0.77 ± 0.07
	MAN-GAN	25.98 ± 2.89	0.78 ± 0.06	25.13 ± 3.27	0.76 ± 0.07	24.39 ± 3.57	0.75 ± 0.08	26.46 ± 2.87	0.82 ± 0.06	25.52 ± 3.28	0.78 ± 0.07
P to DP	Baseline	24.49 ± 2.98	0.75 ± 0.07	24.44 ± 3.17	0.74 ± 0.08	24.28 ± 3.60	0.74 ± 0.09	26.14 ± 3.01	0.80 ± 0.07	25.02 ± 3.31	0.76 ± 0.08
	MAN-GAN	24.91 ± 2.93	0.76 ± 0.07	24.68 ± 3.28	0.75 ± 0.08	24.57 ± 3.57	0.76 ± 0.08	26.41 ± 3.07	0.81 ± 0.06	25.30 ± 3.35	0.77 ± 0.08

Table 3. SR performance on dataset 2. Significant values are in bold. Values were presented as mean ± SD.

SSIM between the real and fake mask images. The distributions of the PSNR and SSIM for different phases are relatively concentrated.

The test performance of MAN-GAN on the disease dataset and external dataset

Dataset 2 was collected to test the performance of the images generated by MAN-GAN. As presented in Table 3, our proposed MAN-GAN also achieved promising performance on dataset 2, showing that both the PSNR and SSIM were greater for MAN-GAN than for the baseline. These results also indicated the generalizability of the MAN-GAN model.

To verify the generalization ability of the proposed MAN-GAN, we tested our method on dataset 3 for external validation. As listed in Table 4, we compared MAN-GAN with existing methods for three-phase translation. Figure 2b shows enhanced CT images generated by MAN-GAN and existing methods for three-phase translation. In P-to-DP translation, MAN-GAN can enhance kidney regions more precisely than other methods, as shown in the magnified patches, suggesting its superiority to existing methods.

Methods	Mode conversion					
	P to AP		P to PVP		P to DP	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
pix2pix ¹⁰	23.26 ± 2.90	0.74 ± 0.08	23.26 ± 2.93	0.74 ± 0.08	22.14 ± 3.50	0.70 ± 0.10
CycleGAN ¹³	23.20 ± 3.11	0.74 ± 0.08	22.76 ± 2.95	0.73 ± 0.08	22.01 ± 3.70	0.70 ± 0.10
CUT ¹²	22.65 ± 2.72	0.73 ± 0.08	21.65 ± 2.39	0.71 ± 0.08	21.48 ± 3.24	0.69 ± 0.10
F-LSeSim ¹¹	23.20 ± 2.81	0.73 ± 0.08	22.76 ± 2.76	0.72 ± 0.08	22.03 ± 3.50	0.70 ± 0.10
DINO ⁹	23.36 ± 2.81	0.73 ± 0.07	22.47 ± 2.64	0.71 ± 0.07	21.09 ± 3.17	0.67 ± 0.09
CHAN ⁸	23.14 ± 2.77	0.73 ± 0.08	22.74 ± 2.69	0.73 ± 0.08	21.72 ± 3.28	0.70 ± 0.10
Sun et al. ²¹	23.23 ± 2.83	0.73 ± 0.08	22.77 ± 2.72	0.72 ± 0.08	21.73 ± 3.30	0.69 ± 0.10
Chen et al. ²²	23.73 ± 2.90	0.75 ± 0.07	23.18 ± 2.80	0.74 ± 0.07	22.38 ± 3.54	0.71 ± 0.10
MAN-GAN	23.92 ± 3.11	0.76 ± 0.08	23.31 ± 3.0	0.74 ± 0.08	22.50 ± 3.70	0.72 ± 0.10

Table 4. Comparison with existing methods on dataset 3. Significant values are in bold. Values were presented as mean ± SD.

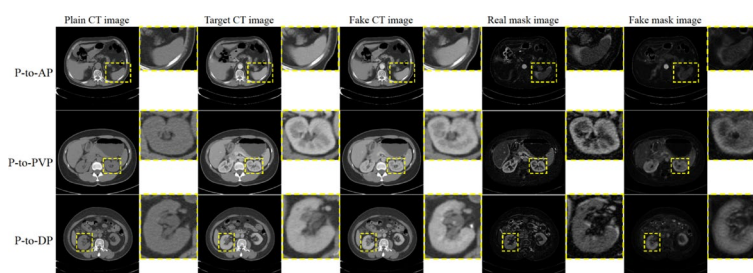


Fig. 4. The visual results of mask images for P-to-AP, P-to-PVP and P-to-DP translation on Dataset 1.

Visualization analysis

To visualize the generated mask images, Fig. 4 shows the mask images for different phases of CT translation. In P-to-AP translation, the fake mask image tends to enhance the liver regions as does the real mask image. When translating from plain to PVP CT images, the regions of the kidney are contrast-enhanced in both the real and fake mask images, indicating that MaskNet can capture contrast-enhanced regions as accurately as the subtraction technique can.

To prove the effectiveness of MAN-GAN in synthesizing realistic MPECT images, we constructed a gray distribution histogram and t-distributed stochastic neighbor embedding (t-SNE)³¹ to visualize the difference between real and fake MPECT images. Figure 5 shows the gray distribution histograms of non-contrast CT images and real and fake MPECT images from the main dataset. As shown in the first row, there is a large difference between the non-contrast CT images and real MPECT images. After translation from non-contrast CT images to fake MPECT images, the domain gap is significantly eliminated, indicating that the gray distribution of the fake MPECT images is remarkably close to that of the real images. To compare the differences in high-level space, Fig. 6 displays the t-SNE visualization for non-contrast CT images and real and fake MPECT images from the main dataset. Obviously, as depicted in the first row, the non-contrast CT images are far from the real MPECT images, indicating the marked difference between them. After the MPECT translation, the fake MPECT images are similar to the real images in different phases, suggesting the effectiveness of MAN-GAN in realistic MPECT translation.

Radiologists' evaluation analysis

According to the results of the quality analysis, the image quality of the generated three phase images was all above average for clinical diagnosis (> 4). Moreover, the similarities between real images and generated images in all three phases are all greater than 80. No significant differences were found in quality analysis or similarity analysis among the different generated phase images (P-to-AP, P-to-PVP, P-to-DP), with P values of 0.059 and 0.730, respectively. Upon further comparison between two different generated phase images (e.g., P-to-AP vs. P-to-PVP and P-to-DP vs. P-to-PVP), only the (P-to-AP vs. P-to-PVP) were significantly different in terms of quality assessment (Table 5).

Discussion

In this study, we first proposed a liver CT image generation network, MAN-GAN, which can automatically synthesize liver MPECT images using non-contrast CT images only. The proposed network novelly incorporated the subtraction strategy in clinical practice and devised a MaskNet model to predict the subtraction signals,

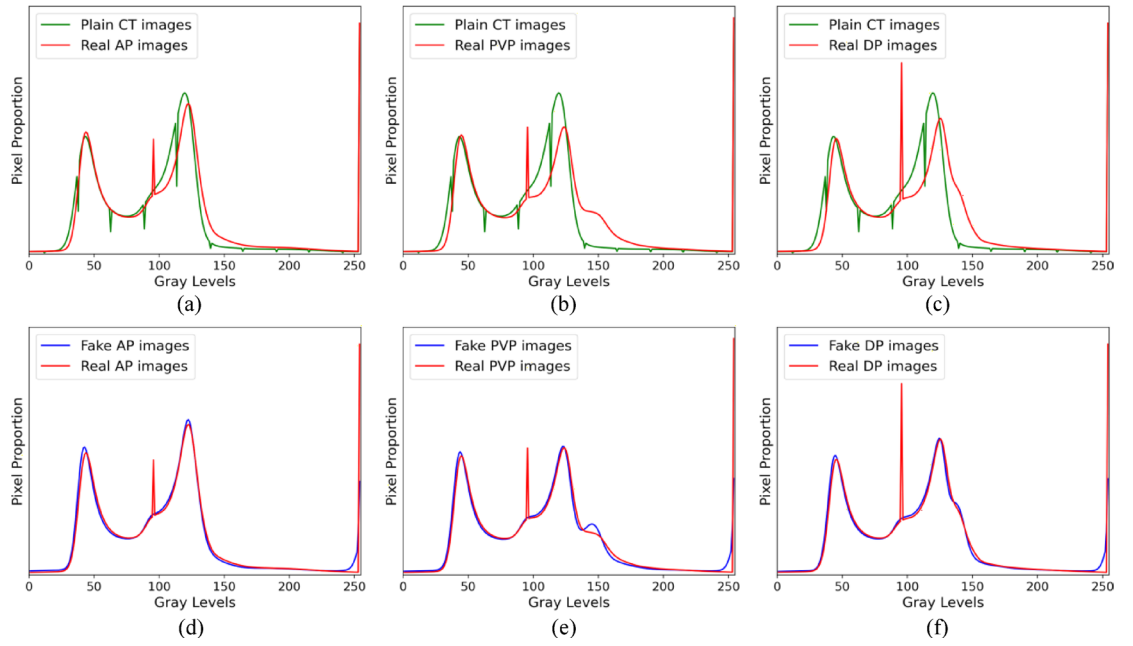


Fig. 5. The gray distribution histogram of non-contrast CT images, real and fake MPECT images from the main dataset. The first row displays the gray distribution histogram of non-contrast CT images and real MPECT images. The second row shows the gray distribution histogram of real and fake MPECT images.

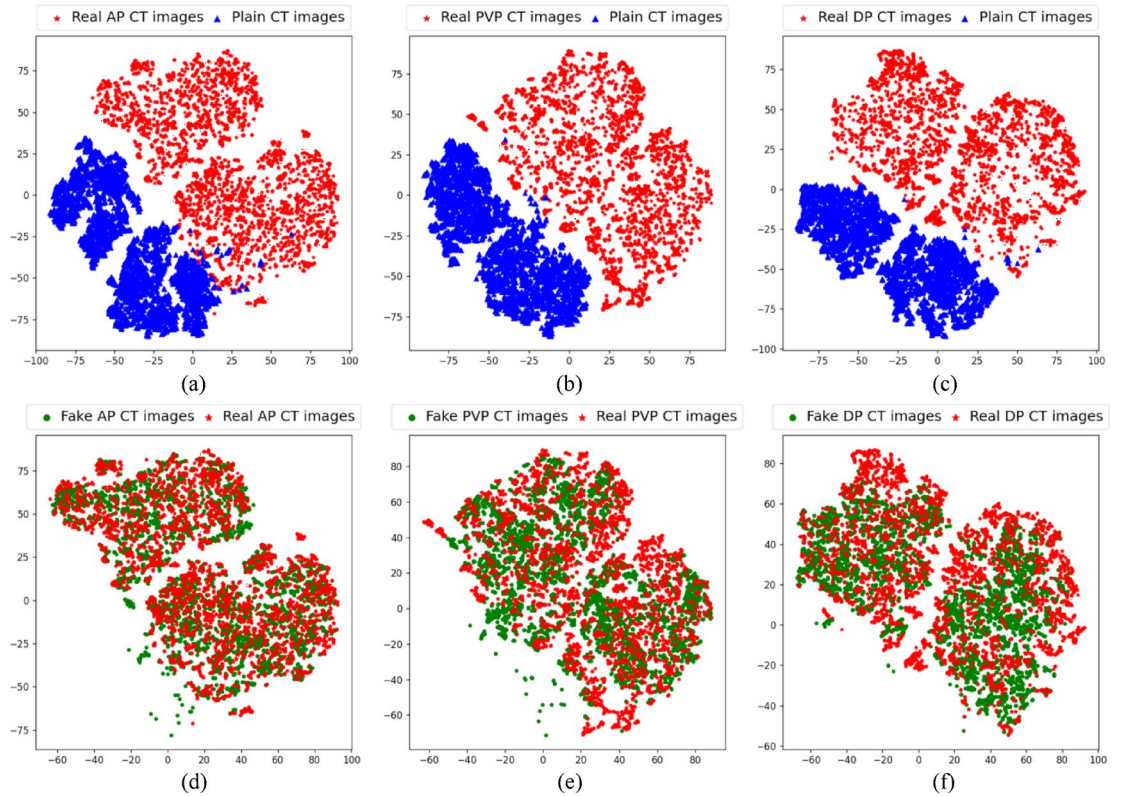


Fig. 6. The t-SNE visualization for the non-contrast CT images, real and fake MPECT images from the main dataset. The first row compares the results of non-contrast CT images and real MPECT images, while the second row shows the results of real and fake MEPCT images.

	Quality	Similarity
P-to-AP	4.0 (3.5, 4.5)	81.5 (78.0, 84.5)
P-to-PVP	4.0 (4.0, 4.5)	81.0 (78.0, 83.5)
P-to-DP	4.0 (3.5, 4.5)	81.5 (78.5, 84.5)
Statistical difference	P-to-AP vs. P-to-PVP: $p=0.03$	P-to-AP vs. P-to-PVP: $p=0.83$
	P-to-AP vs. P-to-DP: $p=0.96$	P-to-AP vs. P-to-DP: $p=0.58$
	P-to-PVP vs. P-to-DP: $p=0.05$	P-to-PVP vs. P-to-DP: $p=0.44$

Table 5. Radiologist's evaluation analysis. Values were presented as median (Q25, Q75).

which were subsequently used to guide contrasted CT synthesis. With this strategy, our proposed MAN-GAN outperformed the other state-of-the-art methods and achieved robust performance on an external dataset.

Liver inherently has a dual blood supply, making hemodynamics changes more complex to learn by the network. We introduce the strategy of subtraction (AP subtracting non-contrast CT images, PVP subtracting non-contrast CT images and DP subtracting non-contrast CT images) in clinical practice. The subtraction images can accurately reflect the difference between two phases and highlight the contrast-enhanced regions, which may aid in image translation. To leverage this domain knowledge effectively, the proposed MAN-GAN devises a MaskNet model to predict subtraction images and further uses this network to guide contrasted CT synthesis. Our proposed MAN-GAN outperformed the state-of-the-art methods and the baseline method in the current task, obtaining higher PSNR and SSIM values. The gray distribution histogram and t-SNE analysis further and more directly visualized the high similarity between real and fake MPECT images. Recently, Zhong et al. proposed a UMTL framework to utilize this domain knowledge to generate subtraction images and combine them with non-contrast CT images to obtain synthesized CT images³². While UMTL has shown significant success in CT synthesis, it requires both non-contrast CT and contrasted CT images as inputs during both the training and inference stages. This may contradict the initial intention of solving the problem of the unavailability of MPECT in clinical practice. In contrast, our proposed MAN-GAN addresses this concern effectively, as it needs only non-contrast CT images as input. This enables MAN-GAN to predict subtraction images for non-contrast CT scans and generate corresponding contrasted CT images. As a result, MAN-GAN represents a more practical and applicable solution for all patients in real clinical applications.

Our constructed MAN-GAN also achieved robust results in Dataset 2 and 3, indicating the generalizability of our model. In the P-to-AP translation, our model achieved the best performance in all the datasets. Specifically, it substantially outperforms the method proposed by Chen et al.²² by approximately 0.19 db. This indicated the superior generalization ability of MAN-GAN to existing methods. This finding also suggested that the most difficult translation is P-to-DP. This phenomenon is acceptable and may be attributed to the characteristic liver hemodynamic changes after injecting contrast media. The scanning of APs, PVPs and DPs is a continuous process; the shortest time interval between non-contrast CT and MPECT is AP, while the longest time interval is DP. The longer the time interval is, the more obvious the effect of individual heterogeneity is, and the more difficult the learning of hemodynamic changes is for the network.

The PSNR and SSIM are the most popular metrics for evaluating image translation performance^{33–36}. However, whether the image quality of synthesized images is sufficient for clinical diagnosis has not been well studied in prior studies. To address this puzzle, we further performed two radiologists' evaluation analyses, quality analysis and similarity analysis. In the quality analysis, we adopted an evaluation scale widely utilized in the quality assessment of radiological images³⁰. The evaluation results showed that the quality of the images was above average (greater than 4). This demonstrated that the quality of our images essentially met or even exceeded the quality of real images. Significant differences were observed in the comparison between the two distinct generated phase images (P-to-AP vs. P-to-PVP). We speculated that it might be attributed to the limited dataset, highlighting the necessity for a more extensive sample size to undergo further validation. Moreover, in similarity analysis, we made the first attempt to develop five criteria to simulate the experience of radiologists in actual clinical practice. The objective is to identify specific deficiencies in the generated images from a clinical application perspective. Similarly, our results showed that all the criteria (those with scores greater than 80) were satisfactory for clinical use. These encouraging results prompted us to accelerate the translation of our model to clinical practice. Notably, we observed relatively low scores on two of the five criteria, namely, the homogeneity of the enhancement and the degree of the enhancement. The complexity of hepatic hemodynamics, coupled with insufficient training data, may be the underlying cause. Recognizing these shortcomings is instrumental in guiding future improvements to our methodology.

Several limitations of the current study should be mentioned. First, the utilized dataset is relatively small, although larger than those in recently published papers³². A larger and external dataset should be collected to further improve the performance of our model in the future. Second, prior knowledge of liver hemodynamic changes was not embedded in the network. Guided by human knowledge, a DL network would be more capable of performing comprehensive tasks³⁷. The latest large language models may also give new solutions in this area³⁸. Subsequently, we only investigated the possibility of generating MPECT images from non-contrast CT images, not touching further directions for disease classification. Constructing a multitask network, such as one that combines image translation and disease diagnosis, will be our future direction. Lastly, we can also compare the region-wise or organ-wise difference between the generated and ground-truth CT images. In this paper, we mainly use the pixel-wise and perceptual metrics and cannot obtain the region-level or organ-level performance for further analysis. Our future work will leverage the organ segmentation models to segment out all the organs

first^{39,40}, implement 3D non-contrast and MPECT translation for each organ, and then evaluate the synthesis performance for each organ.

In conclusion, our proposed MAN-GAN demonstrated the feasibility of liver MPECT image translation based on non-contrast CT images only and achieves state-of-the-art performance via the subtraction strategy. Our method shows potential for solving the dilemma of liver CT contrast scanning and may provide valuable support for various clinical scenarios involving liver interactions. However, further validation and exploration are needed in the future to fully understand the potential and effectiveness of these methods.

Data availability

Specific inquiries for our data and analyses can be raised to corresponding author and will be shared based on reasonable request.

Received: 16 February 2025; Accepted: 7 July 2025

Published online: 22 July 2025

References

- Cao, G., Jing, W., Liu, J. & Liu, M. Countdown on hepatitis B elimination by 2030: The global burden of liver disease related to hepatitis B and association with socioeconomic status. *Hepatol. Int.* **16**, 1282–1296 (2022).
- Devarbhavi, H. et al. Global burden of liver disease: 2023 update. *J. Hepatol.* **79**, 516–537 (2023).
- Yoon, J., Park, S. H., Ahn, S. J. & Shim, Y. S. Atypical manifestation of primary hepatocellular carcinoma and hepatic malignancy mimicking lesions. *J. Korean Soc. Radiol.* **83**, 808–829 (2022).
- Ozaki, K., Higuchi, S., Kimura, H. & Gabata, T. Liver metastases: Correlation between imaging features and pathomolecular environments. *Radiographics* **42**, 1994–2013 (2022).
- Hinson, J. S. et al. Risk of acute kidney injury after intravenous contrast media administration. *Ann. Emerg. Med.* **69**, 577–586.e4 (2017).
- Sun, Z., Choo, G. H. & Ng, K. H. Coronary CT angiography: current status and continuing challenges. *Br. J. Radiol.* **85**, 495–510 (2012).
- Takita, H. et al. AI-based virtual synthesis of methionine PET from contrast-enhanced MRI: Development and external validation study. *Radiology* **308**, e223016 (2023).
- Gao, F. et al. Complementary, heterogeneous and adversarial networks for image-to-image translation. *IEEE Trans. Image Process.* **30**, 3487–3498 (2021).
- Vougioukas, K., Petridis, S. & Pantic, M. DINO: A conditional energy-based GAN for domain translation. Preprint at <https://doi.org/10.48550/arXiv.2102.09281> (2021).
- Isola, P., Zhu, J.-Y., Zhou, T. & Efros, A. A. Image-to-image translation with conditional adversarial networks. Preprint at <https://doi.org/10.48550/arXiv.1611.07004> (2018).
- Zheng, C., Cham, T.-J. & Cai, J. The Spatially-Correlative Loss for Various Image Translation Tasks. Preprint at <https://doi.org/10.48550/arXiv.2104.00854> (2021).
- Park, T., Efros, A. A., Zhang, R. & Zhu, J.-Y. Contrastive learning for unpaired image-to-image translation. In *Computer Vision—ECCV 2020* (eds Vedaldi, A. et al.) 319–345 (Springer International Publishing, Cham, 2020). https://doi.org/10.1007/978-3-030-58545-7_19.
- Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *2017 IEEE International Conference on Computer Vision (ICCV)* 2242–2251 (2017). <https://doi.org/10.1109/ICCV.2017.244>.
- Liu, Y. et al. CT synthesis from MRI using multi-cycle GAN for head-and-neck radiation therapy. *Comput. Med. Imaging Graph.* **91**, 101953 (2021).
- Huang, W. et al. Arterial spin labeling images synthesis from sMRI using unbalanced deep discriminant learning. *IEEE Trans. Med. Imaging* **38**, 2338–2351 (2019).
- Florkow, M. C. et al. Deep learning-based MR-to-CT synthesis: The influence of varying gradient echo-based MR images as input channels. *Magn. Reson. Med.* **83**, 1429–1441 (2020).
- Torrado-Carvajal, A. et al. Dixon-VIBE deep learning (DIVIDE) pseudo-CT synthesis for pelvis PET/MR attenuation correction. *J. Nucl. Med.* **60**, 429–435 (2019).
- Armanious, K. et al. MedGAN: medical image translation using GANs. *Comput. Med. Imaging Graph.* **79**, 101684 (2020).
- Armanious, K. et al. Unsupervised Medical Image Translation Using Cycle-MedGAN. In *2019 27th European Signal Processing Conference (EUSIPCO)* 1–5 (IEEE, A Coruna, Spain, 2019). <https://doi.org/10.23919/EUSIPCO.2019.8902799>.
- Upadhyay, U., Chen, Y., Hepp, T., Gatidis, S. & Akata, Z. Uncertainty-Guided Progressive GANs for Medical Image Translation. Preprint at <https://doi.org/10.48550/arXiv.2106.15542> (2021).
- Sun, Y., Wang, J., Shi, J. & Boppert, S. A. Synthetic polarization-sensitive optical coherence tomography by deep learning. *npj Digit. Med.* **4**, 1–7 (2021).
- Chen, L., Liang, X., Shen, C., Jiang, S. & Wang, J. Synthetic CT generation from CBCT images via deep learning. *Med. Phys.* **47**, 1115–1125 (2020).
- Chandrashekar, A. et al. A deep learning approach to generate contrast-enhanced computerised tomography angiograms without the use of intravenous contrast agents. *Eur. Heart J.* **41**, ehaa946-0156 (2020).
- Ouyang, C. et al. Causality-inspired single-source domain generalization for medical image segmentation. *IEEE Trans. Med. Imaging* **42**, 1095–1106 (2023).
- Chandrashekar, A. et al. A deep learning approach to generate contrast-enhanced computerised tomography angiograms without the use of intravenous contrast agents. *Eur. Heart J.* **41**(Supplement_2), ehaa946-0156 (2020).
- Huh, J. et al. Added value of CT arterial subtraction images in liver imaging reporting and data system treatment response categorization for transcatheter arterial chemoembolization-treated hepatocellular carcinoma. *Invest. Radiol.* **56**, 109–116 (2021).
- Bressem, K. K. et al. Instant outcome evaluation of microwave ablation with subtraction CT in an in vivo porcine model. *Invest. Radiol.* **54**, 333–339 (2019).
- Avants, B. B., Tustison, N. & Johnson, H. Advanced Normalization Tools (ANTs).
- Paszke, A. et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. Preprint at <https://doi.org/10.48550/arXiv.1912.01703> (2019).
- Kalra, M. K. et al. Clinical comparison of standard-dose and 50% reduced-dose abdominal CT: Effect on image quality. *Am. J. Roentgenol.* **179**, 1101–1106 (2002).
- Shi, S. Visualizing Data using t-SNE. ArXiv (2021).
- Zhong, L. et al. United multi-task learning for abdominal contrast-enhanced CT synthesis through joint deformable registration. *Comput. Methods Programs Biomed.* **231**, 107391 (2023).

33. Chang, Y., Li, Z., Saju, G., Mao, H. & Liu, T. Deep learning-based rigid motion correction for magnetic resonance imaging: A survey. *Meta-Radiology* **1**, 100001 (2023).
34. Huang, S. et al. Domain-scalable unpaired image translation via latent space anchoring. *IEEE Trans. Pattern Anal. Mach. Intell.* **45**(10), 11707–11719 (2023).
35. Wang, C. J., Rost, N. S. & Golland, P. Spatial-intensity transforms for medical image-to-image translation. *IEEE Trans. Med. Imaging* **42**, 3362–3373 (2023).
36. Suwanraksa, C., Bridhikitti, J., Liamsuwan, T. & Chaichulee, S. CBCT-to-CT translation using registration-based generative adversarial networks in patients with head and neck cancer. *Cancers* **15**, 2017 (2023).
37. When brain-inspired AI meets AGI. *Meta-Radiology* **1**, 100005 (2023).
38. Nazir, A. & Wang, Z. A comprehensive survey of ChatGPT: Advancements, applications, prospects, and challenges. *Meta-Radiology* **1**, 100022 (2023).
39. Salimi, Y., Shiri, I., Mansouri, Z. & Zaidi, H. Deep learning-assisted multiple organ segmentation from whole-body CT images. 2023.10.20.23297331 Preprint at <https://doi.org/10.1101/2023.10.20.23297331> (2023).
40. Wasserthal, J. et al. TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. *Radiol. Artif. Intell.* **5**, e230024 (2023).

Author contributions

Wei Zhao designed the study, performed data analysis, and wrote the manuscript. Wenting Chen contributed to data preprocessing, statistical analysis, and manuscript revision. Li Fan, Youlan Shang, Yisong Wang, and Weijun Situ assisted with data collection and image annotation. Wenzheng Li provided methodological guidance. Tianming Liu and Yixuan Yuan supervised the study and revised the manuscript. Jun Liu secured funding and finalized the manuscript. All authors approved the final version.

Funding

The study was supported by National Natural Science Foundation of China (62476291), Hunan Provincial Natural Science Foundation for Distinguished Young Scholars (2025JJ20097), Hunan Provincial Natural Science Foundation (2022JJ70139), the Research Foundation of Education Bureau of Hunan Province (24B0003).

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-10754-z>.

Correspondence and requests for materials should be addressed to Y.Y. or J.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025