

Virtual Contrast Enhancement for CT Scans of Abdomen and Pelvis

Jingya Liu^a, Yingli Tian^{*a}, Cihan Duzgol^b, Oguz Akin^{*b}, A. Muhteşem Ağildere^c, K. Murat Haberal^c, Mehmet Coşkun^c

^a*The City College of New York, New York, 10031, USA*

^b*Memorial Sloan Kettering Cancer Center, New York, 10065, USA*

^c*Baskent University, Ankara, 06810, Turkey*

Abstract

Contrast agents are commonly used to highlight blood vessels, organs, and other structures in magnetic resonance imaging (MRI) and computed tomography (CT) scans. However, these agents may cause allergic reactions or nephrotoxicity, limiting their use in patients with kidney dysfunctions. In this paper, we propose a generative adversarial network (GAN) based framework to automatically synthesize contrast-enhanced CTs directly from the non-contrast CTs in the abdomen and pelvis region. The respiratory and peristaltic motion can affect the pixel-level mapping of contrast-enhanced learning, which makes this task more challenging than other body parts. A perceptual loss is introduced to compare high-level semantic differences of the enhancement areas between the virtual contrast-enhanced and actual contrast-enhanced CT images. Furthermore, to accurately synthesize the intensity details as well as remain texture structures of CT images, a dual-path training schema is proposed to learn the texture and structure features simultaneously. Experiment results on three contrast phases (i.e. arterial, portal, and delayed phase) show the potential to synthesize virtual contrast-enhanced CTs directly from non-contrast CTs of the abdomen and pelvis for clinical evaluation.

Keywords: Contrast Enhanced CT, Image Synthesize, Deep Learning, Generative Adversarial Network

*Corresponding author. Email: ytian@ccny.cuny.edu, akino@mskcc.org.

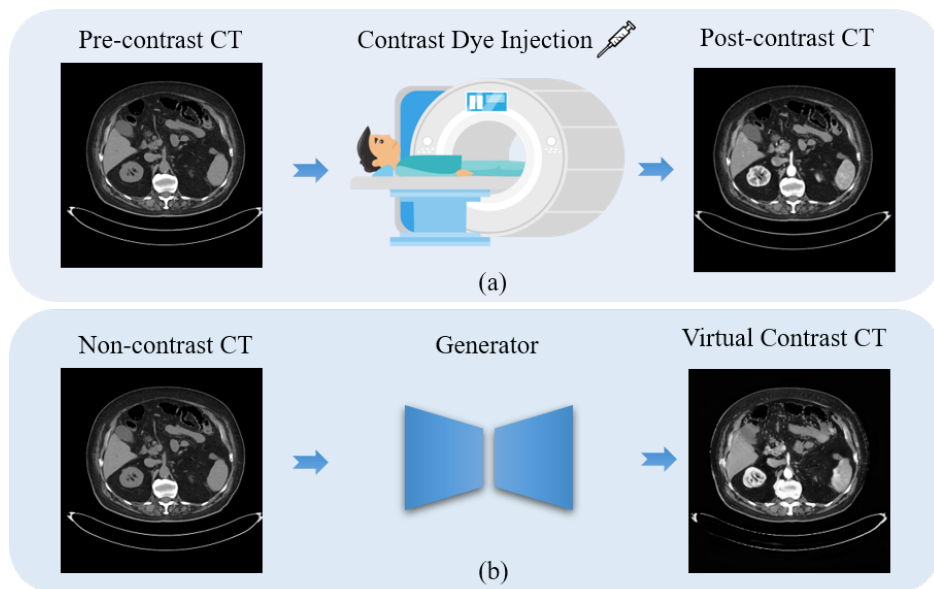


Figure 1: Illustration of contrast-enhanced CT generation. (a) Actual contrast-enhanced CT scan with contrast agent (dye) injection in patients. (b) Virtual contrast-enhanced CT by our proposed framework generated directly from the non-contrast CT.

1. Introduction

1.1. Motivation

Contrast agents are commonly applied in enhancing image-based diagnostic tests such as magnetic resonance imaging (MRI) and computed tomography (CT) to highlight blood vessels, organs, and other structures [1, 2]. Regarding the approximately 76 million computed tomography (CT) scans and 34 million magnetic resonance imaging (MRI) tests performed each year, half use intravenous contrast agents [3]. Recently, there are concerns about the safety of contrast dye. Ideally, the injected contrast agent should be eliminated from the body with no additional effects to patients [4]. However, these agents may potentially cause high risks for many patients, including allergic-like reactions at initial exposure, adverse reactions to pharmacologic toxicity, breakthrough reactions, contrast material-induced nephrotoxicity, and nephrogenic systemic fibrosis [5, 6]. These are side effects of contrast dye frequently observed in the clinic, each with a different underlying mechanism. In addition, the gadolinium to deposit in bone, brain, other tissues has been reported [7].

Researchers dedicately explored ways to reduce side effects or dye deposition, such as replacing with the macrocyclic agent, minimizing the concentration, and using natural D-glucose as an infusible biodegradable MRI agent [8]. However, repeated dye injections on insufficient contrast enhancement areas may cause further harms. Also, these methods require extensive experiments for safety purposes along with uncertainties due to individual differences. In contrast, we explore an alternative solution by generating contrast CT images virtually. A deep learning method enhanced image quality by synthesizing post-contrast brain MRI images from pre-contrast and post-contrast images with 10% low-dose (0.01 mmol/kg) of gadobenate dimeglumine [9]. Similarly, this paper attempted to automatically synthesize virtual contrast CT images for the abdomen and pelvis region to prevent the risks of patients with kidney dysfunction, as illustrated in Fig. 1. Unlike [9] with additional low-dose images, we directly generate virtual contrast-enhanced CTs from non-contrast CTs. This task is more challenging without multi-parametric images compared to brain MRIs [9, 10]. The CTs with no contrast enhancement that have the post-contrast CTs are denoted as the 'pre-contrast CTs', while 'non-contrast CT' refers to the CTs without post-contrast CTs. Similar to actual contrast-enhanced CTs, the virtual contrast-enhanced CTs highlight vascular and organ structures at multiple phases for better delineation of anatomical structures.

1.2. Challenges for Generating Virtual Contrast CTs

1. Misalignment between pre-contrast and post-contrast CT scans:

In clinical evaluation, post-contrast CT images are obtained after contrast agent injection at time intervals. The pose movement highly affect the alignment on repeated pre- and post-contrast enhanced CT pairs. Two main types of misalignments are observed between the pre- and post-contrast CT pairs. The first type is ridged movement which are related to the positional shifting of the patient's body in specific directions. This ridged movement can be aligned by affine transformation via enhanced correlation coefficient maximization. The second type belongs to the random respiratory difference caused by body movements (such as bowel peristalsis) and motion artifacts. The organ location and orientation shifts are hard to solve by image registration methods. As AI-based methods are data-driven, the misalignments make the task challenging, especially under per-pixel supervision.

2. The complexity of abdominal and pelvis CT scans: The abdominal and pelvic regions contain more complex organ structures than the

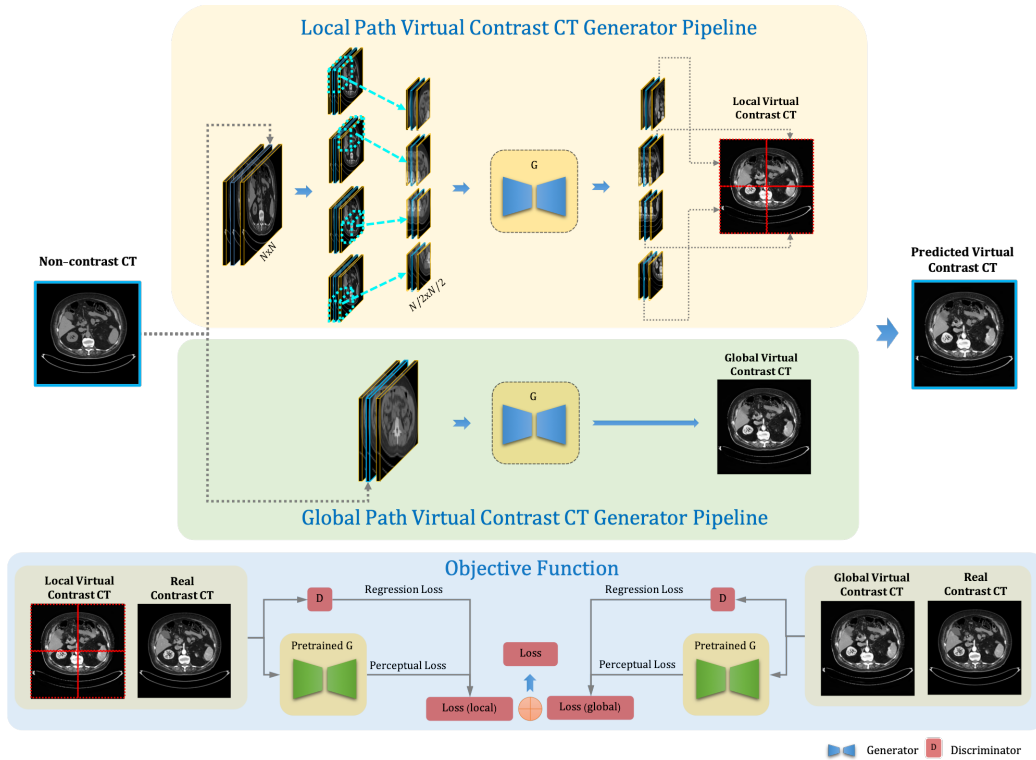


Figure 2: The proposed detail-aware dual-path framework synthesizes virtual contrast-enhanced CT images directly from non-contrast CT images. The global path takes three consecutive whole CT slices as input to extract global structure features. The local path divides the whole images into four patches to extract more detailed texture features and generates four corresponding virtual contrast patches. The four patches are integrated into a whole image for the objective function. In the training phase, a perceptual loss is employed to compare the virtual contrast-enhanced CT and actual contrast-enhanced CT. The total objective function combines the cost for both global and local paths for backpropagation.

brain, spine, and extremities [9, 11]. To synthesize from non-contrast to post-contrast domain, extracting rich features from limited data can easily lead to overfitting of the model.

1.3. Summary of Our Contributions

We propose a dual-path generative adversarial network (GAN) to synthesize virtual contrast-enhanced CTs directly from non-contrast CTs by preserving the texture and enhancing the pixel intensity. The main contributions of this paper are summarized in the following three aspects.

- We attempt to automatically generate virtual contrast-enhanced CT scans from non-contrast CT in the abdominal and pelvis region for the arterial, portal, and delayed phases. This is the first work to explore all three contrastive phases. We collected a dataset of CT scans of the abdomen and pelvis region for training and testing, including pre-and post-contrast CTs for the arterial, portal, and delayed phases commonly used in clinical practice.
- As shown in Fig. 2, we propose a dual-path GAN-based framework with a high-resolution generator. The high-resolution layers are applied in the generator and exchange information across multi-resolution representations for high-resolution image synthesis. To leverage the misalignment on repeated pre-and post-contrast enhanced CT pairs, the perceptual loss [12] is applied to better handle the misalignment compared to per-pixel loss, which can learn the semantic differences of the contrasting regions by comparing high-level feature representations between the actual and virtual contrast-enhanced CT images.
- To further improve the enhancement details, we design a dual-path training schema to simultaneously learn the correlation between the contrast-enhancement region from the global-path network and the fine-grained details from a local-path network. A self-supervised schema is proposed to generate a pre-trained model with rich feature extraction by employing the large public national lung screening trial (NLST) dataset [13] without additional annotations. The pre-trained model helps to overcome the overfitting problem to the model with limited data. The experimental results demonstrate the state-of-the-art performance in synthesizing virtual contrast-enhanced CTs directly from non-contrast CTs.

2. Related Work

Medical imaging, such as MRIs and CTs, plays an essential role in medical diagnosis. To better assist in-clinic diagnosis, several studies on medical image synthesis are reported, such as to synthesis of strawberry-like Fe₃O₄-Au hybrid nanoparticles at room temperature that simultaneously exhibited fluorescence, enhanced X-ray attenuation, and magnetic properties [14], generate photorealistic images of blood cells [15], reconstruct CT from two orthogonal X-rays [16], multi-contrast MRI synthesis [17], cross-modality MR image

synthesis [18], and brain MRI image synthesis [19, 20, 11]. These approaches show good diagnosis validity and may provide necessity of multiple data fusion to improve diagnosis performance [21, 22]. There are existing pixel intensity transform approaches for virtual medical image generation. Traditional handcraft samples and feature selection strategies, such as pair-wise matching of the patches [19] and maximizing the cross-correlation within the space of diffeomorphic maps [23]. Those methods are specially designed for certain types of images but are not robust to other imaging types. As convolutional neural networks (CNNs) have shown the advantage for accurate and effective learning, some methods are introduced for image synthesis. Zhao *et al.* [24] proposed a task-specific modification of an encoder-decoder network and generated corresponding MRI images from the original CT images. The proposed network is based on a well-known architecture, U-Net [25], which completes the semantic segmentation task by a skip connection from the down-sampled layers to up-sampled layers. However, batch normalization limits U-Net to preserve image intensity features. Moreover, the traditional *MSE* and *L1* loss functions can easily over-smooth the generated image. Better feature extraction and effective backpropagate mechanisms should be applied to synthesize CT images with complex features and image intensity variance.

In recent years, the generative adversarial networks (GANs) have been proved as effective methods for virtual image generation in computer vision and medical imaging fields [26, 27, 28, 29, 30, 31, 32, 33]. Tang *et al.* [34] aimed to classify the contrast stages, such as between portal and delayed phases and applied an unpaired GAN for data augmentation. Bayramoglu *et al.* [35] adopted conditional adversarial generative networks (cGANs) to virtually stain the specimens with a non-linear mapping. A few studies were conducted to generate contrast-enhanced brain MRI images [9, 10, 36]. Yurt *et al.* [37] proposed a GAN-based method to impute missing MRI sequences. Bahrami *et al.* [20] synthesized the 7T MRI images from 3T MRI for higher intensity by a CNN-based framework adopting the intensity and the tag features of the brain tissue as input. Dar *et al.* [17] proposed a GAN-based method to generate multi-contrast MRI brain exams. In order to reduce the potential risk of dye injection, Gong *et al.* [9] proposed a method to predict full-dose MRI images by using raw MRI images and gadolinium dose reduction MRI images. Compared to generating virtual contrast CT images of the abdomen and pelvis, brain MRI images have less misregistration and have multiparametric images to provide additional soft-tissue contrast. However,

synthesizing CT is more challenging without the help of multiparametric images like MRI [38]. Therefore, contrast agents are more common in CT than MRI to improve soft tissue differentiation and better depict various structures. Qian *et al.*[39] introduced an additional classification of CT and MRI to help the GAN-based model accurately capture the domain knowledge due to the intrinsic structure differences between the generation of MRI and CT. This method mainly focuses on synthesizing images of one modality from images of another modality. Unlike paper [39], our paper synthesizes images from the same modality but at different contrast enhancement stages.

Recently, Santini *et al.* [40] explored the synthetic enhancement on an attenuation in the cardiac chambers based on non-contrast cardiac CTs. However, compared to the CTs in cardiac region, the abdomen and pelvis CT scans are more challenging with more complex structure and features of organs and tissues. Recently, Kim *et al.* [41] reconstructed the synthetic contrast-enhanced CTs (DL-SCE-CT) at portal phase from nonenhanced CTs (NECT) in patients with acute abdominal pain (AAPa). Choi *et al.* [42] evaluated a deep learning model for generating synthetic contrast-enhanced CT (sCECT) from non-contrast chest CT (NCCT) via Pix2Pix GAN. Liu *et al.* [43] proposed a GAN-based model, DyeFreeNet, to reconstruct the post-contrast enhanced CT at arterial stage. These previous methods developed deep learning frameworks to enhance the contrasts for only one phase or required multi-stage training procedures. In this paper, we attempt to generate virtual contrast-enhanced CT scans of the abdomen and pelvis region for three contrast enhancement phases (arterial, portal, and delayed phases) and propose a one-stage end-to-end framework.

Synthesizing contrast CT images from non-contrast CT images is similar to the image translation task. The input image maps from the source domain to the target domain. Among the natural scene dataset, the features in the target domain are common among the data, such as the stripes and colors of a zebra and the color of cars. However, the contrast enhancement regions among slices of the same CT in the abdomen and pelvis region are different, where requires the model to learn various transferred features for a whole case. Recently, Pix2Pix [44] framework was proposed as an antagonistic network by calculating the distance between the predicted and the actual tag of pixels with tightly correlated images. In addition, many efforts have been performed in transfer learning with unpaired images. CycleGAN [45] learns the transfer from the input domain to the target domain by verifying the cycle consistency with inversed mapping. Contrastive learning GAN [46] proposed

an image-to-image transfer framework to maximize the mutual information between two domains by an unsupervised contrastive loss [47]. To accurately generated contrast-enhanced regions in CT scans, we apply the pixel-level mapping as the baseline model. We further compared the proposed pixel-level transfer learning framework with the unpaired image-to-image transfer learning methods: cycle-consistency and contrastive loss-based GAN.

3. Methods

As shown in Fig. 2, we propose a dual-path GAN-based framework to generate virtual contrast-enhanced CTs from non-contrast CTs. The global path takes three consecutive CT slices as input to extract global structure features and outputs one virtual contrast CT slice, as the local path takes four patches divided from CT images for detailed texture feature extraction. Section 3.1 introduces the framework for virtual contrast CT synthesis. Section 3.2 focuses on the perceptual loss acquisition by a contrast-aware pretrained model. Section 3.3 explains the dual-path training mechanism.

3.1. Framework for Virtual Contrast CT Synthesis

We follow the conditional generative adversarial network (cGAN) to learn the mapping from the input image to the output image. For a fine-grained contrast CT enhancement with slight pixel misalignment, we aim at tackling the problem by combining the pixel-level mapping and the domain adaptation. A generator G learns the mapping between the pre-contrast CT image x and the actual contrast-enhanced CT image c . The virtual contrast CT c' is synthesized by predicting the contrast enhancement level of the pixels from x , while a discriminator D is trained to distinguish the actual contrast-enhanced CT image c with virtual contrast CT image c' .

The objective of the proposed framework is shown as Eq. (1):

$$\mathcal{L}_{GAN}(G, D) = \mathbb{E}_{x,c}[LogD(x, c)] + \mathbb{E}_x[\log(1 - D(G(x)))], \quad (1)$$

where the generator G aims at minimizing the objective while the discriminator D maximizes it.

The Generator: The generated contrast-enhanced CTs require preserving the structures of non-contrast CTs and highlighting body structures the same as actual contrast-enhanced CTs. U-net [25] is an encoder-decoder network and can be used in image synthesis. However, the skip connections, downsamplings, and maximum pooling layers lead to a lack of details and

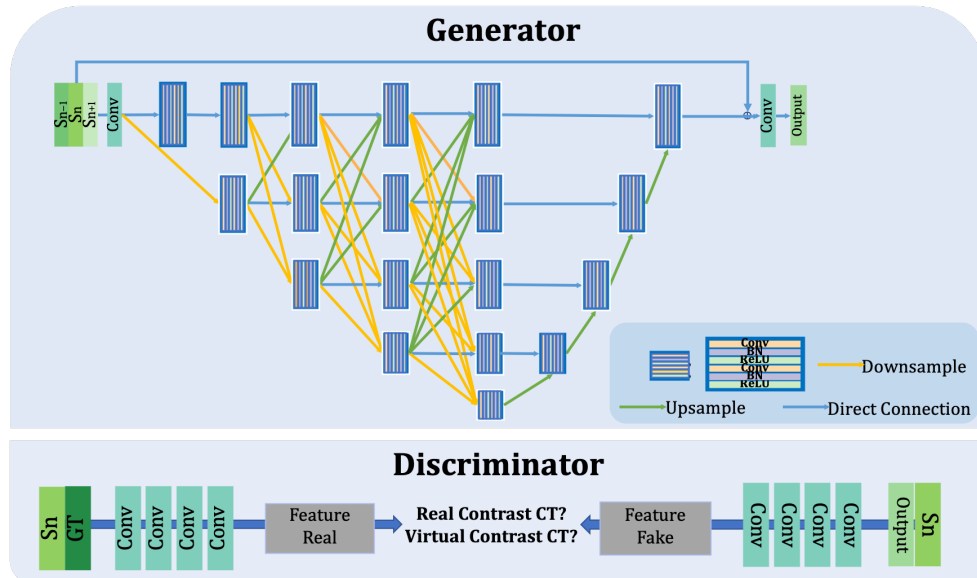


Figure 3: The architecture of the virtual contrast-enhanced CT predictor. 1) The generator composes five layers of the encoder-decoder network and takes three continuous CT images as the input. The parallel connections of the layers can preserve both the high-level and low-level features. 2) The discriminator distinguishes the actual contrast-enhanced CT and virtual contrast-enhanced CT by extracting the features through two networks with the input of actual contrast-enhanced CT and virtual contrast-enhanced CT respectively.

over-smooth on the virtual contrast-enhanced CTs. We apply the strong backbone High-Resolution Network (HRNet) [48] to maintain the high resolution by exchanging information across multi-resolution representations. Fig. 3 shows that the encoder network preserves the high-resolution features and high-semantic features by backpropagating these features through multiple layers. The convolutional block $L(i)$ is concatenated with all previous layers $L(i - k)$, where k is in the range of $(1, i - 1)$.

The encoder network contains convolutional block L_d as d refers the network depth (i.e. network layers) in each block of L_d . The current convolution layer $L_d(i)$ integrates the features from previous layers. Each convolution block is composed of a 2×2 max-pooling layer, 3×3 kernel size double convolution layer and batch normalization, with rectifying linear unit (ReLU) as the activation function. In this paper, five convolutional blocks are implemented. The decoder network adopts the skip connection strategy by gradually combining the features from high-level to low-level layers. To

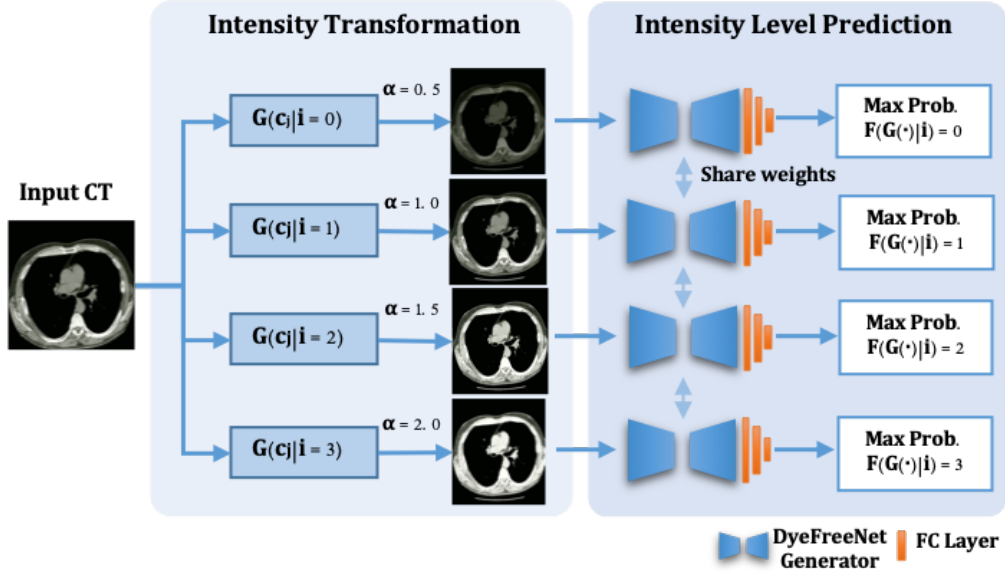


Figure 4: The framework to obtain the pretrained model on the NLST dataset. 1) Each input CT image is transformed into four image intensity levels with the intensity coefficients α for $[0.5, 1.0, 1.5, 2.0]$ of the original CT image, used as four image intensity categories for the pretext task without extra labeling. 2) The intensity-level classification network composes the generator network followed by three Fully Connected (FC) layers to obtain the maximum prediction probabilities for each intensity level. The trained generator is further employed as the pretrained model for the proposed generator.

retain the CT image structures of the non-contrast CT, the input image is concatenated with the output feature maps, followed by a convolutional layer to optimize the weight of the image intensity level.

The Discriminator: As shown in Fig. 1, the discriminator consists of four Convolutional (Conv) blocks with the channel sizes of (64, 128, 256, 512) and the feature sizes of $(256 \times 256, 128 \times 128, 64 \times 64, 32 \times 32)$. Each block contains a 2D Conv layer, followed by a BatchNorm layer and LeakyReLU with a slope of 0.2. A Conv layer flattens the features from the last Conv block to one-dimensional vectors. Mapping non-contrast CT domain to virtual contrast CT domain can be considered as an intensity regression task. *MSE* and *BCEwithlogits* losses are applied to optimize the discriminator by comparing following two feature maps: 1) the feature map extracted from input CT concatenated with the actual contrast-enhanced CT, which labeled as ones; 2) the feature map extracted from input CT concatenated with the

virtual CT, which labeled as zeros.

3.2. Perceptual Loss using Pretrained Network

Pretrained Model: A novel self-supervised training strategy is proposed to train a pretrained model for rich semantic and intensity-aware feature extraction. As shown in Fig. 4, first, the four-level image intensity transformation is applied. The non-contrast CT images are multiplied with the four intensity coefficients (α) levels [0.5, 1.0, 1.5, 2.0], denoted as four classes [0, 1, 2, 3]. Secondly, pretext task training is introduced to classify these four-intensity levels. The backbone network structure is the same as the generator, followed by three fully connected layers and a Softmax layer for the intensity-level classification. Finally, the last three fully connected layers are excluded. The CTs of abdomen and pelvic region from large public NLST dataset [13] are applied for training without requiring extra annotations. This pretrained model is sensitive to intensity variance and can extract rich contrast-aware features. The loss function is described as Eq. (2):

$$loss(c_j|i) = -\frac{1}{K} \sum_{I=0}^{(K-1)} \log(F(G(c_j, I)|i)), \quad (2)$$

where for the input CT slice c_j is transformed into K image intensity I levels by multiplying the coefficient i . The classification network $F(\cdot|I)$ learns the intensity features. The image intensity transformation $G(c_j|I)$ transforms the input CTs to the four image intensity levels.

Perceptual Loss: Perceptual loss [12] has been very popular in image style transformation, which provides accurate results by measuring perceptual differences in content and style between images at the feature level. The per-pixel objective functions (MSE and $L1$) may mislead the model to focus on structure biases between images. In contrast, the perceptual loss measures the differences on the high-level features and is more robust than the per-pixel losses by calculating the mean value on the sum of all squared errors. Inspired by the high sensitivity of human vision on image intensity variations, we apply the perceptual loss to learn the content information by comparing multi-layer feature maps via a pre-trained model.

The pretrained model extracts features from both virtual and actual contrast CTs. The input CTs are downsampled with the factors of $2\times$, $4\times$, $8\times$, and $16\times$ at the last four convolutional blocks. Perceptual heatmaps are generated by averaging the feature-level differences between virtual and actual

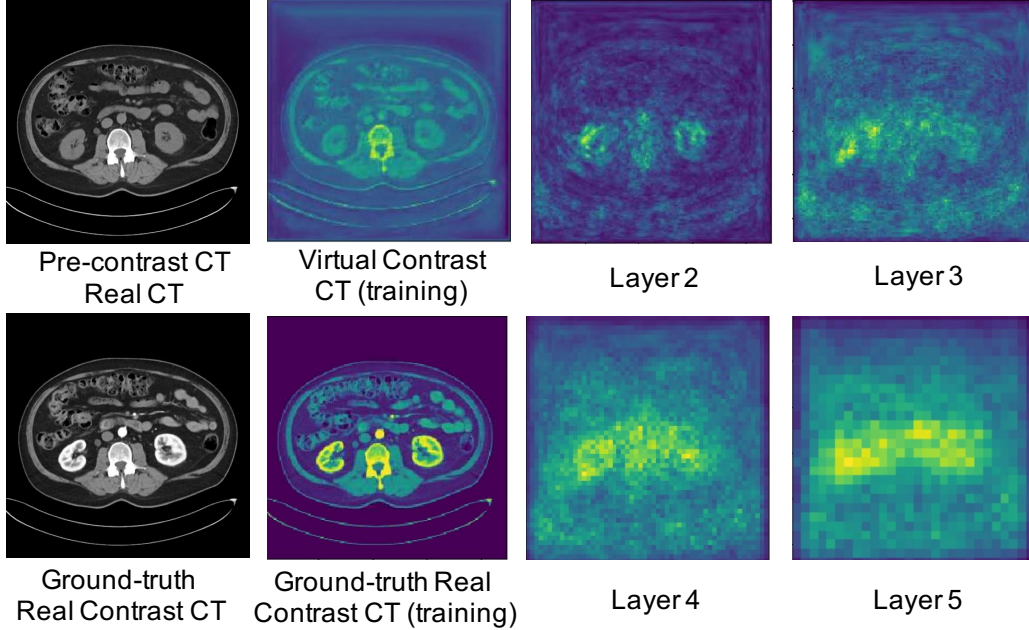


Figure 5: The illustration for heatmaps of perceptual loss computed by averaging the feature-level differences between virtual and actual contrast CTs at the last four convolution blocks.

contrast CTs, shown in Fig. 5. The contrasting regions between the virtual and actual contrast CT are highlighted on perceptual heatmaps in the kidney regions.

The objective function of the generator is described as Eq. (3):

$$G^* = \arg \min_G \max_D \mathcal{L} + \lambda_1 \mathcal{L}_{MSE}(G) + \sum^j \ell_{feat}^{\phi,j}(c', c), \quad (3)$$

where the λ_1 is hyperparameter to weight MSE loss (set as 1 in the training), j is referred to the j th convolutional block in range of (1, 4).

3.3. Dual-path Training Schema

It is essential to preserve both the global structures for organs and enhancement regions, as well as local contrast-enhanced details for blood vessels and soft tissue. A dual-path training schema is proposed to allow the network simultaneously to learn the correlation between the local and global feature representations. The input CT x for global path (with the size of $N \times N$) is

divided into four small patches $[x_1, x_2, x_3, x_4]$ with the size of $N/2 \times N/2$ for local path. The four generated virtual CT patches from the local path are combined as a virtual contrast CT c_{local} . The global path generates the virtual contrast CT image c_{global} . The final objective function of the generator combines the losses of the dual paths, shown as Eq. (4):

$$G^*_{final} = \lambda_2 G^*_{global} + \lambda_3 G^*_{local} . \quad (4)$$

where the λ_2 and λ_3 are hyperparameters to adjust the weights of the dual-path network (set as 0.6 and 0.4), G^* refers to Eq. (3).

4. Experiments

4.1. Dataset

4.1.1. Dataset Collection

CT examinations are obtained using dual-source multi-detector CT (Somatom Force Dual Source CT, Siemens Medical Solutions, Forchheim, Germany). Patients are positioned supine on the table. First, pre-contrast imaging of the abdomen was acquired from the dome of the liver to the iliac crest in an inspiratory breath hold by using a detector configuration of 192×0.6 mm, a tube current of 90 kVp, quality reference of 277 mAs. After intravenous injection of a non-ionic contrast agent 350 mg/ml (1.5 mL per kilogram of body weight at a flow rate of 4 ml/s), bolus tracking was started in the abdominal aorta at the level of the celiac trunk with a threshold of 100 HU. Scans were acquired using attenuation-based tube current modulation (CARE Dose 4D, Siemens).

4.1.2. Dataset Details

The dataset includes pre- and post-contrast CTs of arterial, portal, and delayed phases. A total of 65 patients are assigned to the study. In the training and validation, there are 48 cases (5,457 CT slices) for the arterial phase, 59 cases (6,269 CT slices) for the portal phase, and 55 cases (5,928 slices) for the delayed phase. Five cases (5 patients) with 590 CT slices are conducted for inference.

4.2. Data Processing

The data processing and pair selection strategy for the pre- and post-contrast CT scans are introduced as follows.

Corresponding CT Slice Pair Selection: The steps for pairing the pre- and post-contrast CT slices are as follows. 1) Select the corresponding CT images according to the table location; 2) Assign a unique ID to the CT slices of the same case and the same table location on the axial plane.

Pixel-wise Alignment: Due to the misalignment CT caused by body movements, it is necessary to calibrate the mismatch between CT image pairs, which greatly impacts the optimization process. Affine and translation transforms are employed to align the pre- and post-contrast CT images based on enhanced correlation coefficient maximization. Additionally, the cases with very large misalignment (larger than 6 pixels) are excluded. The pre- and post-contrast CTs with the largest overlapping structures are selected as pairs.

Window Size Selection: Hounsfield Unit (HU) is applied in CT images to represent the linear transformation of the measured attenuation coefficient. The HU value ranges from -1000 (air) to approximately 2000 (very dense bone). For a better analysis of the abdominal and pelvis areas, the window width and window level are set to 400 and 30 respectively.

4.3. Implementation Details

4.3.1. Pretrained Model

41,589 low-dose CT slices from a large public national lung screening test (NLST) dataset are employed for the pretrained model. The CT images are transformed at four CT image intensity levels by applying $[0.5, 1.0, 1.5, 2.0]$ coefficients. The learning rate of training was set to $1e^{-4}$ and decreased by 0.1 after 8 epochs, updated by Adam optimizer. Total training includes 10 epochs with a batch size of 2 on a GeForce GTX 1080 GPU using Pytorch 1.0 and Python 2.7.

4.3.2. Model Training

The CTs are resized to 256×256 . Three consecutive CT slices ($256 \times 256 \times 3$) mimic three RGB channels as the input of the global path network. For the CT scan with no adjacent section before or after the current slice, the current slice is duplicated to fill the blank. Fig. 2 shows that the CT image ($256 \times 256 \times 3$) is divided into four patches ($128 \times 128 \times 3$) as the input of the local path. The generator consists of five convolutional blocks with the output dimensions set to $[32, 64, 128, 256, 512]$. The decoder network follows the U-Net skip connection. 32 channels from the last layers are concatenated with the input, followed by two convolution layers. The generator is initialized

by the weights of the pretrained model. 5-fold validation is applied to the training and validation sets to select the hyper-parameter settings gave the best result. The learning rates are set to $5e^{-5}$ for generator and $2e^{-4}$ for discriminator during the training and decreasing by 0.1 after 40 epochs for a total of 50 epochs with a batch size of 2. The weight decay is set to $5e^{-4}$. Adam optimizer is applied for optimization.

4.4. Experimental Results

4.4.1. Radiologist Evaluation

As shown in Fig. 6, the virtual contrast-enhanced results of our framework and the existing state-of-the-art methods (i.e., unpaired image-to-image transfer method CycleGAN [45] and unsupervised contrastive loss based GAN [46]) are blindly assessed by two radiologists based on image quality and contrast enhancement for overall image quality, organs, and vessels. Pairs of pre- and actual contrast-enhanced CTs and the pre- and virtual contrast-enhanced CTs are applied for evaluation. The qualitative rating was based on a 5-point Likert scale, with the evaluation scores of poor (1), sub-optimal (2), acceptable (3), good (4), and excellent (5). Since the virtual contrast-enhanced CT images for different phases are required to mimic contrast enhancement of organs and vessels compared to the non-contrast CT images, the scores can be considered as “1” for no obvious changes from non-contrast CT image, “2” as mild virtual contrast enhancement, “3” for moderate virtual contrast enhancement, “4” for good virtual contrast enhancement, and “5” for very good virtual contrast enhancement.

The radiologists’ average assessment of our results is acceptable with the rate of “3” (see Fig. 6). Note that for the arterial phase, the framework is trained by CT images mainly highlighted arterial vascular structures. Therefore, the evaluation scores are slightly higher as more contrast enhancement is observed for the arteries. Our method outperforms CycleGAN [45] and Contrastive GAN [46] in all evaluation aspects. Although the contrastive loss-based GAN achieves comparable results to our proposed method in the arterial and portal phases, as detailed contrast enhancement urinary tract regions are required in the delayed phase, our approach surpasses the other two methods with a big margin in generating more precise detail enhancement and higher image quality.

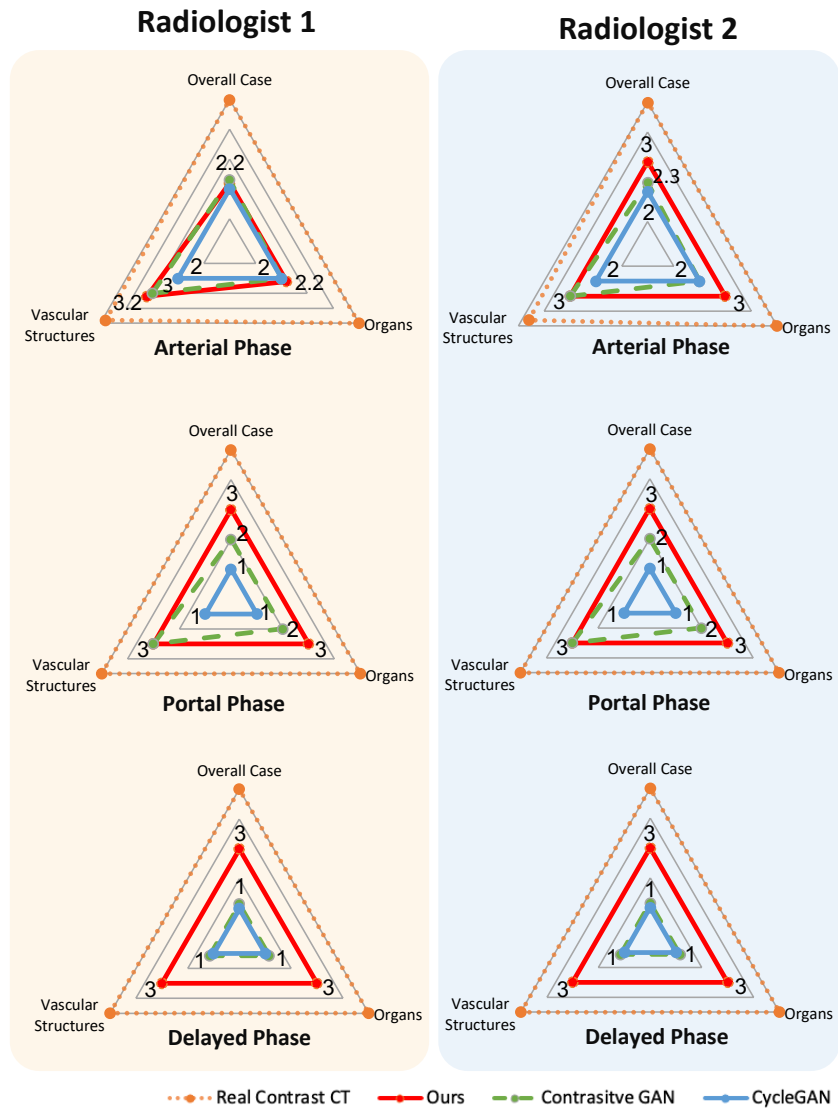


Figure 6: The assessments of two radiologists for the arterial, portal, and delayed phases in three aspects: overall case, organs, vascular structures by applying a 5-level score schema from 1 (poor) to 5 (excellent). The proposed method (red) outperforms the state-of-the-art methods: Contrastive GAN (green) [46] and CycleGAN (blue) [45] in all evaluation aspects.

Table 1: The quantitative comparisons between proposed framework with the state-of-the-art methods CycleGAN [45] and Contrastive GAN [46] for arterial phase (AP), portal phase (PP), and delayed phase (DP).

Phase	PSNR \uparrow	SSIM \uparrow	MSE \downarrow	SNU \downarrow
CycleGAN-AP [45]	19.16 \pm 1.89	0.66 \pm 0.05	215 \pm 16	0.36 \pm 0.08
CycleGAN-PP [45]	19.07 \pm 1.93	0.69\pm0.05	244 \pm 23	0.41 \pm 0.12
CycleGAN-DP [45]	17.36 \pm 1.81	0.65 \pm 0.07	257 \pm 38	0.58 \pm 0.17
Contrastive-AP [46]	19.85\pm3.04	0.674\pm0.07	198 \pm 22	0.33 \pm 0.25
Contrastive-PP [46]	19.47 \pm 2.74	0.66 \pm 0.07	228 \pm 36	0.49 \pm 0.35
Contrastive-DP [46]	18.42\pm2.36	0.60 \pm 0.08	246 \pm 14	0.52 \pm 0.41
Proposed-AP (ours)	19.09 \pm 1.88	0.63 \pm 0.05	194\pm19	0.25\pm0.09
Proposed-PP (ours)	19.79\pm1.17	0.65 \pm 0.05	226\pm29	0.34\pm0.16
Proposed-DP (ours)	17.23 \pm 1.59	0.65\pm0.05	244\pm22	0.46\pm0.14

4.4.2. Compared to the State-of-the-art Methods

Quantitative evaluations on four metrics are applied to compare virtual and actual contrast-enhanced CTs. Peak Signal to Noise Ratio (PSNR) measures voxel-wise differences. The Structural Similarity Index (SSIM) [49] evaluates the image quality based on non-local structural similarity. The Mean Absolute Error (MSE) and Spatial NonUniformity (SNU) [50] measure contrast enhancement accuracy. Table 1 shows the performance of proposed methods with state-of-the-art comparison for arterial phase (AP), portal phase (PP), and delayed phase (DP). The proposed method shows a slightly lower performance at *PSNR* and *SSIM* in some phases compared to CycleGAN [45] and Contrastive GAN [46]. However, *PSNR* and *SSIM* cannot solely represent the performance of contrast-enhancement regions. SNU and MSE calculate the contrast quality of the virtual images and demonstrate that our method outperforms other methods at all contrast-enhanced phases. The advantages of the proposed method are better visualized in Fig. 7.

Contrast enhancement details are shown in Fig. 7 for our proposed framework (Fig. 7(e)) compared to Contrastive GAN (Fig. 7(c)) and CycleGAN (Fig. 7(d)). Predominantly enhancement of arterial vascular structures is observed in the arterial phase. In the portal phase, organ parenchymal and venous vascular structures (such as portal veins and hepatic veins) are observed. In the delayed phase, while organs and vessels fade out, urinary excretion of contrast highlights the urinary tract.

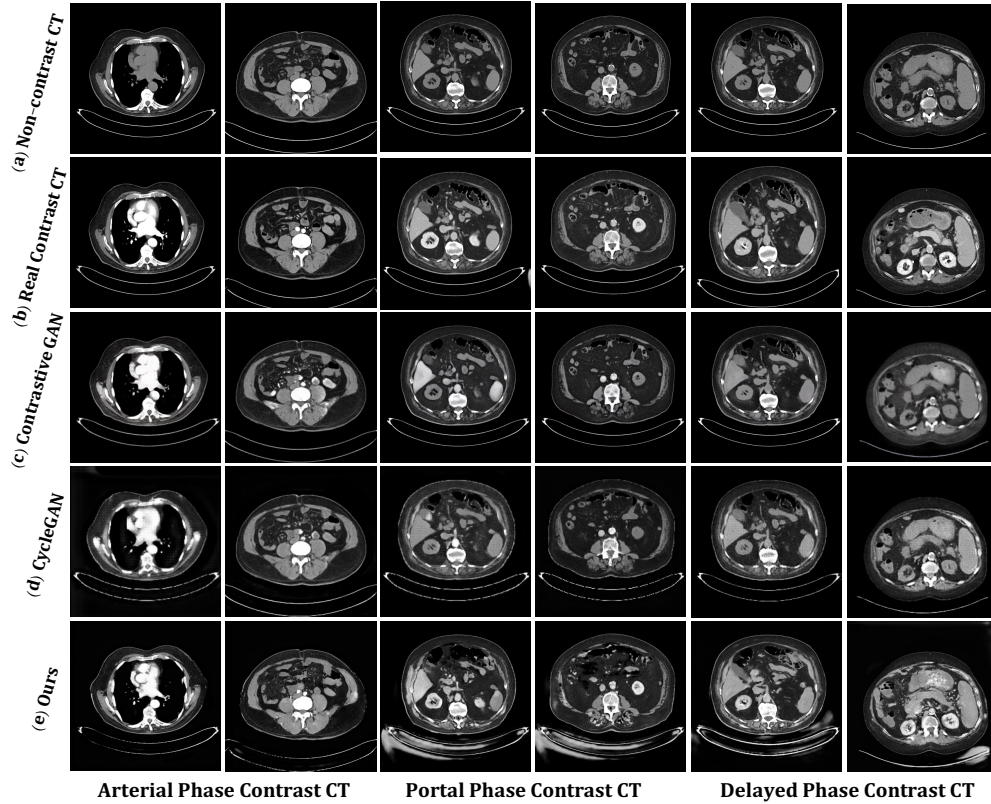


Figure 7: The visualizations of virtual contrast-enhanced CT synthesized by the proposed framework, compared with the existing state-of-the-art methods [45] [46] in the arterial, portal, and delayed phases. (a) The pre-contrast CT. (b) The actual contrast-enhanced CT. (c) The virtual contrast-enhanced CT by Contrastive GAN [46]. (d) The virtual contrast-enhanced CT by CycleGAN [45]. (e) The virtual contrast-enhanced CT by proposed framework.

Our result is sharper and has no pixelation to view compared to the other methods. Non-enhancing tissue is more clearly visualized. The proposed pixel-level mapping shows advantages to generating fine-grained enhancement regions. CycleGAN and Contrastive GAN show poor enhancements in details or fail to predict virtual contrast CT enhancement, especially at the portal and delayed phases.

Table 2: The quantitative evaluation for ablation study without discriminator (w/o Discr.), without perceptual loss (w/o Perc.), and without dual-path training schema (Dual) in the arterial phase (AP), portal phase (PP), and delayed phase (DP)

<i>Phase</i>	<i>PSNR</i> \uparrow	<i>SSIM</i>	<i>MSE</i> \downarrow	<i>SNU</i> \downarrow
w/o Discr.-AP	18.96 \pm 1.72	0.64 \pm 0.05	243 \pm 16	0.31 \pm 0.16
w/o Discr.-PP	19.44 \pm 1.69	0.69\pm0.05	232 \pm 23	0.37 \pm 0.35
w/o Discr.-DP	16.95 \pm 1.39	0.65 \pm 0.07	296 \pm 38	0.53 \pm 0.12
w/o Perc.-AP	18.36 \pm 1.50	0.63 \pm 0.06	234 \pm 44	0.42 \pm 0.34
w/o Perc.-PP	18.77 \pm 1.43	0.66 \pm 0.08	221 \pm 21	0.39 \pm 0.59
w/o Perc.-DP	16.86 \pm 1.32	0.64 \pm 0.06	257 \pm 16	0.52 \pm 0.38
w/o Dual.-AP	19.05 \pm 3.04	0.67\pm0.07	198 \pm 22	0.33 \pm 0.25
w/o Dual.-PP	19.47 \pm 2.74	0.66 \pm 0.07	238 \pm 36	0.49 \pm 0.35
w/o Dual.-DP	16.42 \pm 2.36	0.60 \pm 0.08	246 \pm 14	0.52 \pm 0.41
Proposed-AP	19.09\pm1.88	0.63 \pm 0.05	194\pm19	0.25\pm0.09
Proposed-PP	19.79\pm1.17	0.65 \pm 0.05	216\pm29	0.34\pm0.16
Proposed-DP	17.23\pm1.59	0.65\pm0.05	244\pm22	0.46\pm0.14

4.4.3. Ablation Study

This section describes the ablation study on the perceptual loss, dual-path training schema, and the discriminator. The quantitative and qualitative evaluation are shown in Table 2 and Fig. 8.

Effectiveness of the Adversarial Supervision: Per-pixel *MSE* and *L1* losses may omit high-frequency features and oversmooth texture details. Adversarial supervision outperforms the model without discriminator on all evaluation metrics shown in Table 2, especially for *MSE* and *SNU*. As illustrated in Fig. 8, our proposed method (Fig. 8(f)) shows more accurate intensity enhancement and higher resolution on the virtual contrast CT compared to the method without discriminator (Fig. 8(c)).

Effectiveness of the Perceptual Loss: Table 2 shows that the proposed model outperforms the model without perceptual loss on all evaluation metrics. Compared with the cases without perceptual loss (Fig. 8(d)), Fig. 8(f) shows better enhancement and fewer artifacts, demonstrating the benefits of content learning via feature-level differences between virtual and

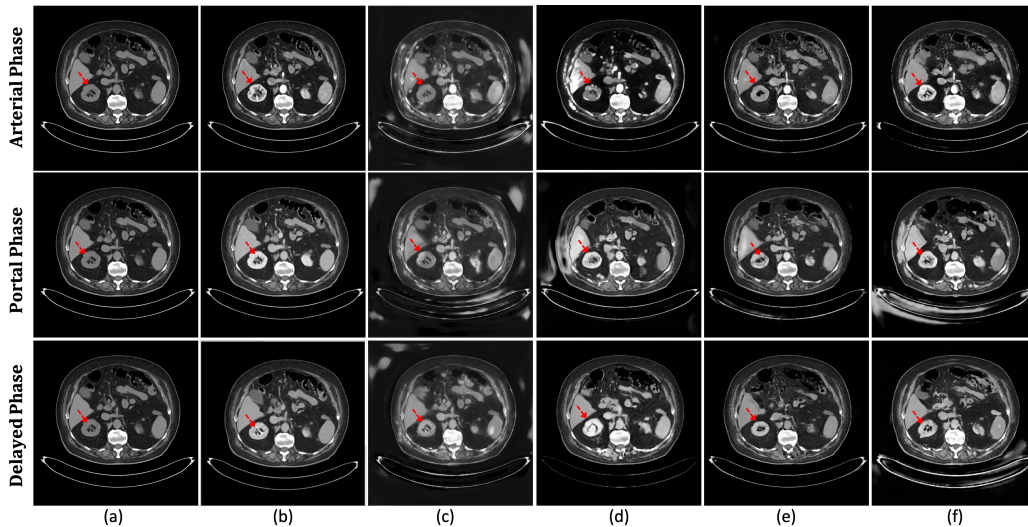


Figure 8: The illustration of the qualitative results for the ablation study by applying discriminator, perceptual loss, and dual-path training schema for the arterial, portal, and delayed phases. (a) The pre-contrast CT. (b) The actual contrast-enhanced CT. (c) The virtual contrast-enhanced CT without discriminator. (d) The virtual contrast-enhanced CT with only the global-path training. (e) The virtual contrast-enhanced CT without perceptual loss. (f) The proposed virtual contrast-enhanced CT.

actual contrast-enhanced CTs.

Effectiveness of the Dual-path Schema: Experiments evaluate the model with and without the dual-path training scheme. Fig. 8 shows that the aorta is well enhanced in the virtual contrast-enhanced CT generated by the models with and without dual-path schema. However, Fig. 8(e) shows insufficiently enhanced kidney region by the model trained with only the global-path network (highlighted by the red arrow). Fig. 8(f) demonstrates the virtual contrast CT with dual-path schema successfully highlights the detailed kidney region. Table 2 indicates that the model trained with dual-path training schema significantly outperforms the model with only the global-path network on all the evaluation scores with a large margin, especially for *SNU*.

Additionally, (Fig. 9 shows the loss curves of the baseline model (Fig. 9(i)), the proposed model with the pre-trained weights and the perceptual loss (Fig. 9(ii)), and the proposed model with the pre-trained weights, the perceptual loss, and dual-path strategy (Fig. 9(iii))). The training loss (Fig. 9(i)

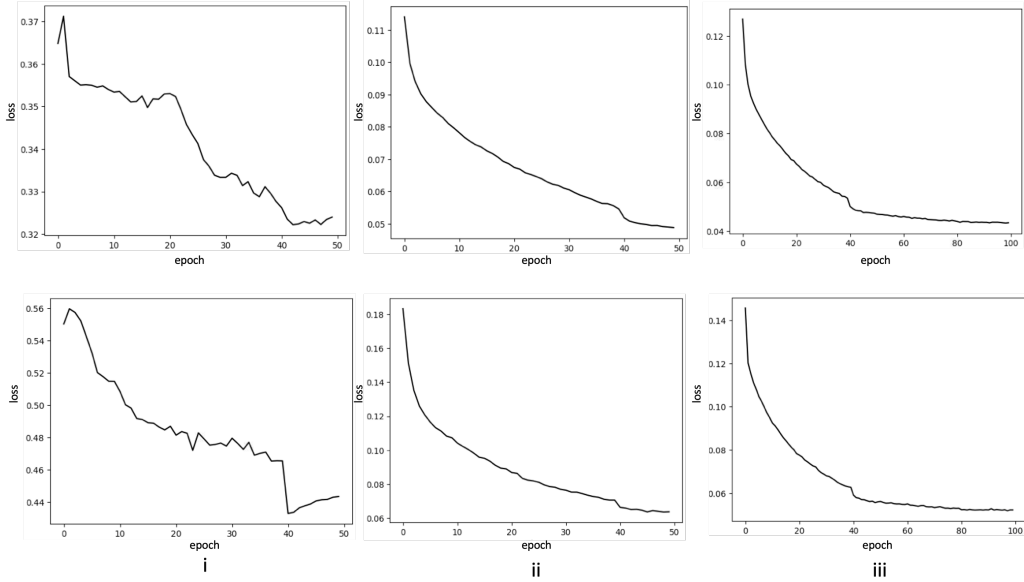


Figure 9: The losses in the training (upper) and validation (lower) for i) baseline model [44] ; ii) with pre-trained weights and perceptual loss; iii) with pre-trained weights, perceptual loss, and dual-path strategy (our proposed methods.)

upper image) shows the baseline network is hard to converge and needs an early stop. Fig. 9(ii) upper image confirms the model initialized with pre-trained weights and the perceptual loss is easily converged and effectively learns the contrasted enhanced features. By training with the pre-trained model, the perceptual loss, and dual-path strategy (Fig. 9(iii) upper image), the model shows better convergence. The validation loss of proposed model shows a small margin between validation and training (Fig. 9(ii) and (iii) lower image). Fig. 9(iii) shows the loss curve is nice converged by training 50 epochs.

4.4.4. Remaining Challenge

Fig. 10 summarizes the remaining limitations of virtual contrast CT synthesis as following three aspects: 1) poor, patchy, and insufficient enhancement, where the contrast is insufficient; 2) unnecessary enhancement, where artifacts appear; and 3) no enhancement for regions should be enhanced.

Two main challenges cause the limitations. Firstly, misalignments of CT pairs on physiologic motion highly affect the accuracy of the enhancement, especially in the delayed phase. Although the preprocessing step excludes

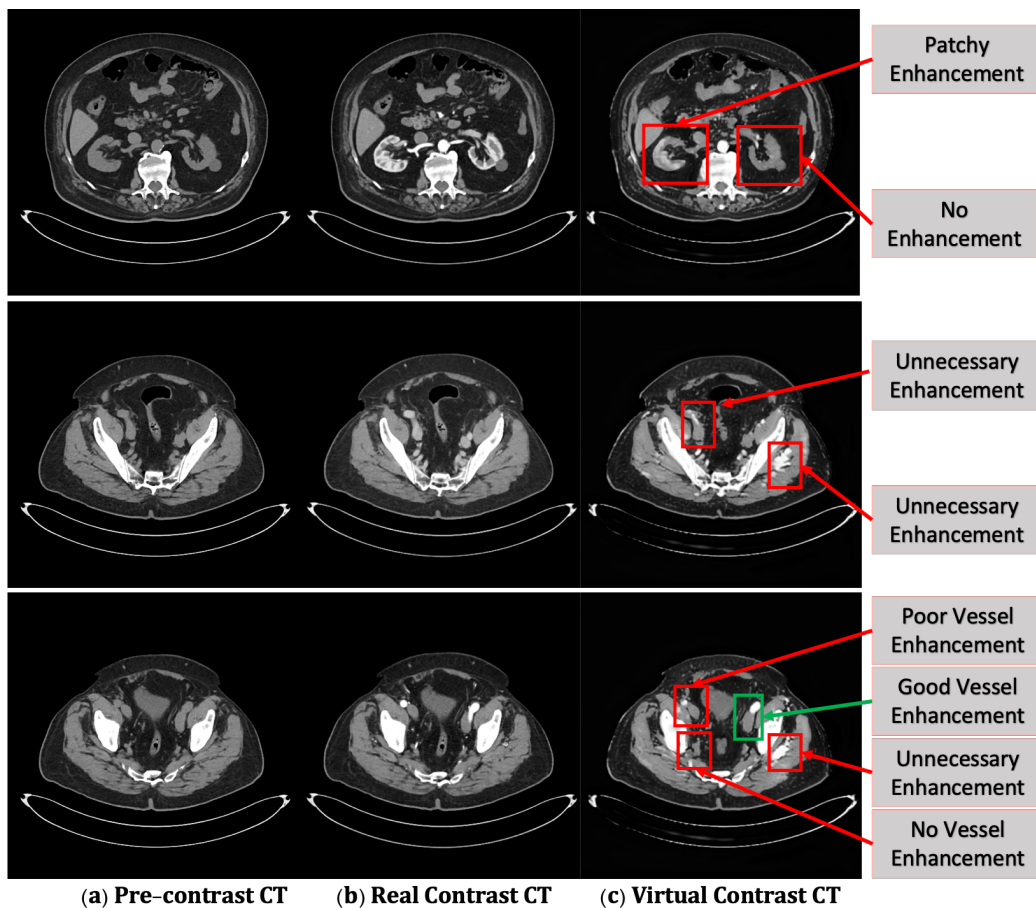


Figure 10: The failure example illustrations of virtual contrast-enhanced CT synthesized by the proposed framework. (a) The pre-contrast CT. (b) The actual contrast-enhanced CT. (c) The virtual contrast-enhanced CT.

CT pairs with large misalignment and applies the affine transformation to limit the misalignments within 3 pixels, few remaining misalignments are still associated with artifacts and unnecessary enhancement. Secondly, the abdominal and pelvic region consists of complex organs and anatomy. Lacking training samples leads to unsatisfied enhancements or no enhancement in details on a few virtual contrast CTs.

4.4.5. Future Work

This paper studies axial CTs as commonly used in routine radiology interpretations. We may explore the performances of coronal and sagittal CTs in the future. Additionally, we will explore the model that can deal with few data, such as active learning and few-shot learning for virtual contrast CT image generation with finer enhancement details. We will evaluate the virtual contrast CT image quality on downstream tasks, such as nodule detection and segmentation.

5. Conclusions

We have proved the concept of generating virtual contrast-enhanced CTs directly from non-contrast CT scans of the abdomen and pelvis via a GAN-based framework. A dual-path training strategy is proposed to maintain both the high-level structure features as well as the local texture features to produce realistic virtual contrast CTs with details. A perceptive loss is conducted to compare the feature level difference between virtual contrast CT and the actual contrast-enhanced CT together via a contrast-aware pre-trained model. The proposed framework has successfully predicted virtual contrast-enhanced CT density while maintaining the structural information of non-contrast images and has achieved good performance and demonstrated promising potential in clinical applications.

Acknowledgments

This work was supported in part by Memorial Sloan Kettering Cancer Center Support Grant/Core Grant P30 CA008748 (the collaborative research between CCNY and MSKCC), and the National Science Foundation under award number IIS-2041307.

References

- [1] J. Liu, M. Li, J. Wang, F. Wu, T. Liu, Y. Pan, A survey of MRI-based brain tumor segmentation methods, *Tsinghua Science and Technology* 19 (6) (2014) 578–595.
- [2] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, et al., A survey on deep learning in medical image analysis, *Medical image analysis* 42 (2017) 60–88.

- [3] D. J. Brenner, H. Hricak, Radiation exposure from medical imaging: time to regulate?, *Jama* 304 (2) (2010) 208–209.
- [4] K. R. Beckett, A. K. Moriarity, J. M. Langer, Safe use of contrast media: what the radiologist needs to know, *Radiographics* 35 (6) (2015) 1738–1750.
- [5] M. Andreucci, R. Solomon, A. Tasanarong, Side effects of radiographic contrast media: pathogenesis, risk factors, and prevention, *BioMed research international* 2014.
- [6] H.-J. Mentzel, J. Blume, A. Malich, C. Fitzek, J. R. Reichenbach, W. A. Kaiser, Cortical blindness after contrast-enhanced CT: complication in a patient with diabetes insipidus, *American journal of neuroradiology* 24 (6) (2003) 1114–1116.
- [7] S. Fox-Rawlings, D. Zuckerman, NCHR report: the health risks of MRIs with gadolinium-based contrast agents, *National Center for Health Research Q* 9.
- [8] J. W. Choi, W.-J. Moon, Gadolinium deposition in the brain: current updates, *Korean journal of radiology* 20 (1) (2019) 134–147.
- [9] E. Gong, J. M. Pauly, M. Wintermark, G. Zaharchuk, Deep learning enables reduced gadolinium dose for contrast-enhanced brain MRI, *Journal of Magnetic Resonance Imaging* 48 (2) (2018) 330–340.
- [10] J. Kleesiek, J. N. Morshuis, F. Isensee, K. Deike-Hofmann, D. Paech, P. Kickingereder, et al., Can Virtual Contrast Enhancement in Brain MRI Replace Gadolinium?: A Feasibility Study, *Investigative radiology* .
- [11] A. Bône, S. Ammari, J.-P. Lamarque, M. Elhaik, É. Chouzenoux, F. Nicolas, P. Robert, C. Balleyguier, N. Lassau, M.-M. Rohé, Contrast-enhanced brain MRI synthesis with deep learning: key input modalities and asymptotic performance, in: *International Symposium on Biomedical Imaging*, 2021.
- [12] J. Johnson, A. Alahi, L. Fei-Fei, Perceptual losses for real-time style transfer and super-resolution, in: *European conference on computer vision*, Springer, 694–711, 2016.

- [13] T. National Lung Screening Trial Research, The national lung screening trial: overview and study design, *Radiology* 258 (1) (2011) 243–253, [dataset].
- [14] H. Y. Zhao, S. Liu, J. He, C. C. Pan, H. Li, Z. Y. Zhou, et al., Synthesis and application of strawberry-like Fe₃O₄-Au nanoparticles as CT-MR dual-modality contrast agents in accurate detection of the progressive liver disease, *Biomaterials* 51 (2015) 194–207.
- [15] O. Bailo, D. Ham, Y. Min Shin, Red blood cell image generation for data augmentation using Conditional Generative Adversarial Networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 0–0, 2019.
- [16] X. Ying, H. Guo, K. Ma, J. Wu, Z. Weng, Y. Zheng, X2CT-GAN: Reconstructing CT from Biplanar X-Rays with Generative Adversarial Networks, in: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 10619–10628, 2019.
- [17] S. U. Dar, M. Yurt, L. Karacan, A. Erdem, E. Erdem, T. Çukur, Image synthesis in multi-contrast MRI with conditional generative adversarial networks, *IEEE transactions on medical imaging* 38 (10) (2019) 2375–2388.
- [18] B. Yu, L. Zhou, L. Wang, Y. Shi, J. Fripp, P. Bourgeat, Ea-GANs: edge-aware generative adversarial networks for cross-modality MR image synthesis, *IEEE transactions on medical imaging* 38 (7) (2019) 1750–1762.
- [19] X. Cao, J. Yang, Y. Gao, Y. Guo, G. Wu, D. Shen, Dual-core steered non-rigid registration for multi-modal images via bi-directional image synthesis, *Medical image analysis* 41 (2017) 18–31.
- [20] K. Bahrami, F. Shi, I. Rekik, D. Shen, Convolutional neural network for reconstruction of 7T-like images from 3T MRI using appearance and anatomical features, in: *Deep Learning and Data Labeling for Medical Applications*, Springer, 39–47, 2016.
- [21] S. Wang, M. E. Celebi, Y.-D. Zhang, X. Yu, S. Lu, X. Yao, Q. Zhou, M.-G. Miguel, Y. Tian, J. M. Gorriz, et al., *Advances in data preprocessing*

- for biomedical data fusion: an overview of the methods, challenges, and prospects, *Information Fusion* 76 (2021) 376–421.
- [22] Y.-D. Zhang, Z. Zhang, X. Zhang, S.-H. Wang, MIDCAN: A multiple input deep convolutional attention network for Covid-19 diagnosis based on chest CT and chest X-ray, *Pattern recognition letters* 150 (2021) 8–16.
 - [23] S. Roy, A. Carass, A. Jog, J. L. Prince, J. Lee, MR to CT registration of brains using image synthesis, in: *Medical Imaging 2014: Image Processing*, vol. 9034, International Society for Optics and Photonics, 903419, 2014.
 - [24] C. Zhao, A. Carass, J. Lee, Y. He, J. L. Prince, Whole brain segmentation and labeling from CT using synthetic MR images, in: *International Workshop on Machine Learning in Medical Imaging*, Springer, 291–298, 2017.
 - [25] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: *International Conference on Medical image computing and computer-assisted intervention*, Springer, 234–241, 2015.
 - [26] X. Yi, E. Walia, P. Babyn, Generative adversarial network in medical imaging: A review, *Medical image analysis* (2019) 101552.
 - [27] J. M. Wolterink, T. Leiner, M. A. Viergever, I. Išgum, Generative adversarial networks for noise reduction in low-dose CT, *IEEE transactions on medical imaging* 36 (12) (2017) 2536–2545.
 - [28] P. Costa, A. Galdran, M. I. Meyer, M. Niemeijer, M. Abràmoff, A. M. Mendonça, et al., End-to-end adversarial retinal image synthesis, *IEEE transactions on medical imaging* 37 (3) (2017) 781–791.
 - [29] A. BenTaieb, G. Hamarneh, Adversarial stain transfer for histopathology image analysis, *IEEE transactions on medical imaging* 37 (3) (2017) 792–802.
 - [30] Y. Huo, Z. Xu, H. Moon, S. Bao, A. Assad, T. K. Moyo, et al., Synsegnet: Synthetic segmentation without target modality ground truth, *IEEE transactions on medical imaging* 38 (4) (2018) 1016–1025.

- [31] T. M. Quan, T. Nguyen-Duc, W.-K. Jeong, Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss, *IEEE transactions on medical imaging* 37 (6) (2018) 1488–1497.
- [32] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, et al., DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction, *IEEE transactions on medical imaging* 37 (6) (2017) 1310–1321.
- [33] D. Nie, R. Trullo, J. Lian, L. Wang, C. Petitjean, S. Ruan, Q. Wang, D. Shen, Medical image synthesis with deep convolutional adversarial networks, *IEEE Transactions on Biomedical Engineering* 65 (12) (2018) 2720–2730.
- [34] Y. Tang, H. H. Lee, Y. Xu, O. Tang, Y. Chen, D. Gao, S. Han, R. Gao, C. Bermudez, M. R. Savona, et al., Contrast phase classification with a generative adversarial network, in: *Medical Imaging 2020: Image Processing*, vol. 11313, International Society for Optics and Photonics, 1131310, 2020.
- [35] N. Bayramoglu, M. Kaakinen, L. Eklund, J. Heikkila, Towards virtual h&e staining of hyperspectral lung histology images using conditional generative adversarial networks, in: *Proceedings of the IEEE International Conference on Computer Vision*, 64–71, 2017.
- [36] X. Yang, Y. Lin, Z. Wang, X. Li, K.-T. Cheng, Bi-modality Medical Image Synthesis using Semi-supervised Sequential Generative Adversarial Networks, *IEEE Journal of Biomedical and Health Informatics* .
- [37] M. Yurt, S. U. Dar, A. Erdem, E. Erdem, K. K. Oguz, T. Çukur, must-GAN: Multi-stream generative adversarial networks for MR image synthesis, *Medical image analysis* 70 (2021) 101944.
- [38] J. M. Pereira, C. B. Sirlin, P. S. Pinto, G. Casola, CT and MR imaging of extrahepatic fatty masses of the abdomen and pelvis: techniques, diagnosis, differential diagnosis, and pitfalls, *Radiographics* 25 (1) (2005) 69–85.
- [39] P. Qian, K. Xu, T. Wang, Q. Zheng, H. Yang, A. Baydoun, J. Zhu, B. Traughber, R. F. Muzic, Estimating CT from MR abdominal images

- using novel generative adversarial networks, *Journal of Grid Computing* 18 (2) (2020) 211–226.
- [40] G. Santini, L. M. Zumbo, N. Martini, G. Valvano, A. Leo, A. Ripoli, F. Avogliero, D. Chiappino, D. Della Latta, Synthetic contrast enhancement in cardiac CT with Deep Learning, arXiv preprint arXiv:1807.01779 .
- [41] S. W. Kim, J. H. Kim, S. Kwak, M. Seo, C. Ryoo, C.-I. Shin, S. Jang, J. Cho, Y.-H. Kim, K. Jeon, The feasibility of deep learning-based synthetic contrast-enhanced CT from nonenhanced CT in emergency department patients with acute abdominal pain, *Scientific reports* 11 (1) (2021) 1–10.
- [42] J. W. Choi, Y. J. Cho, J. Y. Ha, S. B. Lee, S. Lee, Y. H. Choi, J.-E. Cheon, W. S. Kim, Generating synthetic contrast enhancement from non-contrast chest computed tomography using a generative adversarial network, *Scientific reports* 11 (1) (2021) 1–10.
- [43] J. Liu, Y. Tian, A. M. Ağildere, K. M. Haberal, M. Coşkun, C. Duzgol, O. Akin, DyeFreeNet: Deep Virtual Contrast CT Synthesis, in: *International Workshop on Simulation and Synthesis in Medical Imaging*, Springer, 80–89, 2020.
- [44] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, Image-to-image translation with conditional adversarial networks, in: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1125–1134, 2017.
- [45] J.-Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: *Proceedings of the IEEE international conference on computer vision*, 2223–2232, 2017.
- [46] T. Park, A. A. Efros, R. Zhang, J.-Y. Zhu, Contrastive learning for unpaired image-to-image translation, in: *European Conference on Computer Vision*, Springer, 319–345, 2020.
- [47] T. Chen, S. Kornblith, M. Norouzi, G. Hinton, A simple framework for contrastive learning of visual representations, in: *International conference on machine learning*, PMLR, 1597–1607, 2020.

- [48] K. Sun, B. Xiao, D. Liu, J. Wang, Deep high-resolution representation learning for human pose estimation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5693–5703, 2019.
- [49] Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli, Image quality assessment: from error visibility to structural similarity, IEEE transactions on image processing 13 (4) (2004) 600–612.
- [50] J. Harms, Y. Lei, T. Wang, R. Zhang, J. Zhou, X. Tang, W. J. Curran, T. Liu, X. Yang, Paired cycle-GAN-based image correction for quantitative cone-beam computed tomography, Medical physics 46 (9) (2019) 3998–4009.