# Contrast Flow Pattern and Cross-Phase Specificity-Aware Diffusion Model for NCCT-to-Multiphase CECT Synthesis

Kaiyi Zheng[1,2,3], Mu Huang[1,2,3], Xinming Li[4], Jianhua Ma[5], Qianjin Feng[1,2,3], Wei Yang[1,2,3]*, and Liming Zhong[1,2,3]*

[1] School of Biomedical Engineering, Southern Medical University, Guangzhou, 510515, China
[2] Guangdong Provincial Key Laboratory of Medical Image Processing, Guangzhou, 510515, China
[3] Guangdong Province Engineering Laboratory for Medical Imaging and Diagnostic Technology, Guangzhou, 510515, China
{weiyanggm, limingzhongmindy}@gmail.com
[4] Department of Radiology, Zhujiang Hospital, Southern Medical University, Guangzhou, China
[5] School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, China

**Abstract.** Multiphase contrast-enhanced computed tomography (CT) is clinically significant in providing vascular structure and lesion phase-specific enhancements. Yet, its clinical utility is constrained by intrinsic contrast agent-associated risks (e.g., nephrotoxicity, allergic reactions) and multiphase cumulative radiation exposure. To tackle this, synthesizing contrast-enhanced CT (CECT) using non-contrast CT (NCCT) offers a potential alternative. However, achieving a high-quality synthesis of multiphase CECT remains challenging due to the contrast agent (CA)-induced complex contrast flow dynamics and the specific variations across phases. Therefore, this paper proposes a contrast flow pattern and cross-phase specificity-aware diffusion model for NCCT-to-multiphase CECT synthesis. Specifically, a contrast flow pattern learning mechanism is integrated into the conditional diffusion model, which enables orderly phase transitions while ensuring anatomically and temporally coherent enhancement synthesis. Furthermore, a phase distinction network is introduced to align cross-phase specificity features with the contrast features in synthesized CECT images. Experimental results on multicenter abdomen CT datasets have demonstrated the superiority of our method compared to state-of-the-art methods.

**Keywords:** Multiphase CECT · Image synthesis · Contrast agent-free · Diffusion model

## 1 Introduction

Multiphase contrast-enhanced computed tomography (CECT) provides critical diagnostic value through vascular characterization and lesion-specific enhance-

ment patterns[1, 2]. Unfortunately, the clinical application of CECT faces limitations due to contraindications and patient-specific factors. The use of iodinated contrast agents (CA) for CECT scans carries potential risks such as severe allergic reactions and nephrotoxicity[3]. Moreover, multiphase CECT protocols lead to cumulative radiation exposure, raising concerns for patients requiring repeated imaging. On the contrary, non-contrast computed tomography (NCCT) has the advantage of lower radiation exposure and avoids the potential adverse effects of CA. Nevertheless, the poor visibility of abnormalities and organ contours in NCCT scans increases the diagnostic difficulty for radiologists. Therefore, developing a CA-free synthesis framework for generating diagnostic-quality images is of significant promise.

With recent advances in deep learning, contrast-enhanced synthesis from NCCT images has emerged as a promising alternative to real CECT scans [4–8]. Pang et al. [9] developed a generative adversarial network (GAN)-based framework for chest NCCT-to-CECT synthesis and applied it in pulmonary vessel segmentation. Zhong et al. [10] presented CKAP-Net to synthesize late arterial phase CECT images for multi-organ segmentation in NCCT images. Unlike these specific single-phase synthesis methods, Uhm et al. [11] proposed DiagnosisGAN for completing any missing CT phase based on other phases. It is worth noting that numerous current studies focus on synthesizing specific phases, while only a few can synthesize multiphase. These methods fail without exception in CA-free situations, in which all contrast-enhanced phases are absent. To address this, Liu et al. [12] pioneered multiphase CECT synthesis, including arterial, portal, and delayed phases. The proposed dual-path GAN enables NCCT-to-multiphase CECT synthesis by preserving the texture and enhancing the pixel intensity. Zhong et al. [13] introduce united multi-task learning for joint synthesis and deformable registration of abdominal CECT, which synthesizes CECT images with acceptable quality on multiphase data. However, these methods treat each phase independently, neglecting the inherent temporal dependencies and inter-phase relationships. This isolated approach can lead to inconsistencies in synthesized images, such as anatomical misalignments or unrealistic enhancement patterns, thereby compromising diagnostic reliability.

The primary challenge in synthesizing multiphase CECT images lies in accurately modeling the dynamic distribution of contrast agents (CA) across different phases. Such a CA flow pattern contains abundant diagnostic relative information. Existing methods struggle to model the temporal coherence of CA flow across phases, leading to anatomically inconsistent enhancement (e.g., arterial-phase hyperenhancement misaligned with portal-phase washout). Moreover, another challenge of multiphase CECT synthesis is the lack of cross-phase difference CA knowledge learning. In the clinic, each phase captures distinct temporal and spatial patterns of CA enhancement, which are critical for precise diagnosis.

To tackle the aforementioned issues and utilize the CA knowledge, we propose a contrast flow pattern and cross-phase specificity-aware diffusion model (CFPS-Diff) for NCCT-to-multiphase CECT synthesis. Specifically, contrast flow pattern learning (CFPL) is integrated into the conditional diffusion model, which en-

ables any-to-follow phase orderly transitions. The CFPL facilitates the dynamic distribution of CA in the synthesis of multiphase CECT images, thereby ensuring anatomically and temporally coherent enhancement synthesis. Furthermore, cross-phase specificity learning (CPSL) was realized through the designed phase distinction network to align cross-phase specificity features with the contrast features in synthesized CECT images, facilitating cross-phase enhancement synthesis. We conduct experiments on multi-center clinical datasets, comprising NCCT and triple-phase CECT, demonstrating the superior performance of the proposed CFPS-Diff in synthesizing CECT images compared with state-of-the-art (SOTA) methods. Our code is available at https://github.com/Kindyz/CFPS-Diff.git.

## 2 Method

**Problem Formulation.** In the training stage, a set of multiphase data, namely NCCT, arterial phase (AP), venous phase (VP), and delayed phase (DP) CECT volumes, are given, respectively. The entire dataset is denoted by $S = \{(X_i, Y_i^A, Y_i^V, Y_i^D) \mid i = 1, 2, \ldots, N\}$ where $i$ denotes the $i-th$ volume. Correspondingly, we assign phase labels to all volumes in $S$ according to their index, forming a set $\{\xi_i\} \in \{0, 1, 2, 3\}$. In the testing stage, solely the NCCT volume $X_i$ is provided, and the objective is to synthesize other triple phase CECT volumes $\{(Y_i^A, Y_i^V, Y_i^D)\}$.

**Conditional Diffusion Model.** Diffusion models have demonstrated exceptional capabilities in generating high-quality medical images by modeling complex data distributions through iterative denoising processes [14]. This iterative denoising mechanism intrinsically shares inherent similarities with CA-induced enhancement evolution. Therefore, we construct CFPS-Diff based on the denoising diffusion probabilistic model (DDPM) [15], aiming to effectively model the temporal dynamics of CA distribution in multiphase CECT synthesis. The forward process starts from the initial target phase CT volume $x_0 \sim S$ and progressively adds Gaussian noise in $t$ timesteps according to a predefined variance schedule $\{\beta\}$ and produces noise target phase CT volume $x_t$. Letting $\alpha_t := 1-\beta_t$ and $\bar{\alpha}_t = \prod_{i=1}^{t}$, each $x_t$ can be directly sampled via the closed-form expression:

$$\begin{aligned} \text{(Forward):} \quad & x_t = \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1-\bar{\alpha}_t}, \quad \epsilon \sim \mathcal{N}(0, \sigma^2 \mathbf{I}), \\ \text{(Reverse):} \quad & x_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left( \left( 1 - \frac{\beta_t}{1-\bar{\alpha}_t} \right) x_t + \frac{\sqrt{\bar{\alpha}_t}\beta}{1-\bar{\alpha}_t} u_\theta \left( x_t, x_c, w_{flow}, t \right) \right) \\ & \quad + \sigma_t \epsilon, \quad x_c \sim S \end{aligned} \tag{1}$$

where $x_c$ represents the CT volume of the conditional phase, $u_\theta$ represents the image generator to predict the target phase CT volume directly, and $w_{flow}$ denotes the phase sequential embedding vector.

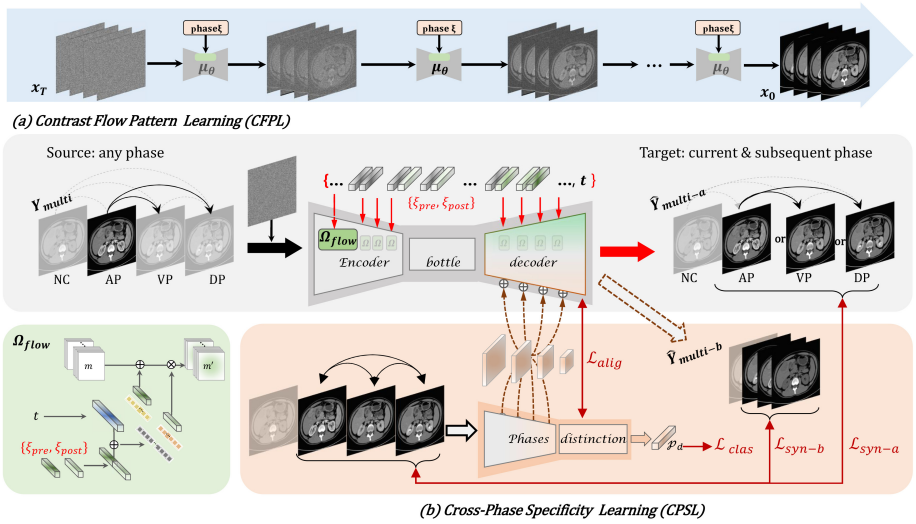


**Fig. 1.** The overview of our proposed CFPS-Diff. (a) Contrast Flow Pattern (CFP)-Aware Learning. (b) Cross-Phase Specificity (CPS)-Aware Learning.

**Contrast Flow Pattern Learning (CFPL).** Our proposed CFPL is designed to tackle the challenge of modeling dynamic contrast flow in multiphase CT synthesis. As shown in Fig. 1(a), the CFPL introduces the guidance of order phase embedding to synthesize the target phase CT volume. In particular, we randomly generate a sample pair $(\xi_{pre}, \xi_{post})$ using a time-constrained sequential sampling formula:

$$\xi_{pre} \sim \mathcal{U}(0,3), \quad \xi_{post} \sim \mathcal{U}\{\xi_{pre}, \xi_{pre}+1, \ldots, 3\} \tag{2}$$

where $\xi_{pre}$ and $\xi_{post}$ are the labels of the source phase and target phase, respectively. This sequential sampling strategy enforces $\xi_{pre} \leq \xi_{post}$ to reflect the temporal monotonicity of the contrast flow pattern. Following, the phase sequential embedding module $\Omega_{flow}$ is used to embed the phase flow and fusion with timestep-conditional representation $e$ in the diffusion model and modulate feature maps $m$ in each encoder and decoder layer of $u_\theta$ to $m_{flow}$. The process is formulated as:

$$\begin{aligned} w_{flow} &= \mathbf{\Omega}_{flow}\left(\xi_{pre}, \xi_{post}, e_t\right) \\ m\prime &= \left(m \oplus \mathcal{M}_{\text{shift}}\left(w_{flow}\right)\right) \odot \mathcal{M}_{\text{scale}}\left(w_{flow}\right) \end{aligned} \tag{3}$$

where $e$ is a timestep-conditional representation in the diffusion model, $\mathcal{M}_{\text{shift}}$ and $\mathcal{M}_{\text{scale}}$ denote the embedding layers for obtaining shift and scale vectors, respectively, and $\oplus$ and $\odot$ present element-wise addition and multiplication operations, respectively. The loss function is a weighted combination of mean absolute error (MAE) loss and wavelet loss[16] and formalized as:

$$\mathcal{L}_{syn-a} = \lambda_1 |\hat{x}_0 - x_0| + \lambda_2 \mathcal{L}_{wavelet}(\hat{x}_0, x_0) \tag{4}$$

where $\lambda_1$ and $\lambda_2$ are the weights for the MAE loss and the wavelet loss.

**Cross-Phase Specificity Learning (CPSL).** Radiologists major rely on the differential enhancement observed across multiple phases of contrast-enhanced CT to detect abnormalities and differentiate disease entities. To reinforce this critical specificity synthesis, we proposed the CPSL to align the synthesis enhancement features with phase distinction features (shown in Fig.1. (b)). The phase distinction network, denoted as $D$ is tasked with contrast difference extraction and phase classification on real multiphase volumes. In each training iteration, the synthesis and classification networks are optimized simultaneously and the multi-scale enhancement features learned by the extractor $F_d$ in $D$ were used to constrain decoder $F_u$ in $u_\theta$. The loss functions are depicted as:

$$\mathcal{L}_{alig} = \sum_{l=1}^{N} \left| F_{u,l}\left(x_t, x_c, w_{flow}, t\right) - F_{d,l}\left(x_0\right) \right|, \tag{5}$$

where $l$ denotes the $l-th$ feature layer. Inspired by [17], a class distance weighted cross-entropy loss (CDW-Loss) is introduced to boost the cross-phase various awareness of CFPS-Diff.

$$\mathcal{L}_{clas} = -\sum_{n=0}^{N-1} log(1 - D(x_0)) \times |n - c|^{\lambda_3}, \tag{6}$$

where $l$ denotes $l-th$ feature layer, $n$ is the index class in the output layer, $c$ is the index of the ground-truth phase label $\xi_{post}$ and $\lambda_3$ is a hyperparameter that determines the strength of the coefficient. The classification loss $\mathcal{L}_{clas}$ boosts the phase various aware of CFPS-Diff of various through introducing distance constraints. Furthermore, the phase difference features extracted by the distinction network can be used to reconstruct the specific CECT image, optimized by:

$$\mathcal{L}_{syn-b} = \left| u_\theta\left(x_t, x_c, F_d\left(x_0\right), t\right) - x_0 \right| \tag{7}$$

**Losses.** The total loss $\mathcal{L}_{total}$ of the CFPS-Diff is defined as:

$$\mathcal{L}_{total} = \mathcal{L}_{syn-a} + \mathcal{L}_{syn-b} + \lambda_4 \mathcal{L}_{clas} + \lambda_5 \mathcal{L}_{alig}, \tag{8}$$

where $\lambda_4$ and $\lambda_5$ are the weights for the $\mathcal{L}_{clas}$ loss and $\mathcal{L}_{alig}$ loss, respectively.

## 3 Experiments

### 3.1 Experimental setup

**Dataset.** The following three medical image datasets are employed in this study. We constructed the CT-1 dataset by collecting CT images from 547 patients

**Table 1.** Ablation study for the proposed components on the CT-1 dataset.

| CFPL | CPSL | CDW-Loss | MAE | PSNR | SSIM |
|---|---|---|---|---|---|
| ✗ | ✗ | ✗ | 38.73 | 28.07 | 0.787 |
| ✓ | ✗ | ✗ | 34.78 | 28.78 | 0.801 |
| ✓ | ✓ | ✗ | 30.70 | 29.51 | 0.824 |
| ✓ | ✓ | ✓ | 30.69 | 29.51 | 0.825 |

across AP, VP, and DP, along with corresponding body segmentation annotations, from a local hospital, which was then used for model training and validation. Additionally, CT images from 116 cases across two other local hospitals were collected to form the CT-2 set, which was designated as the test set. Among these, 28 cases featuring rare pathological types were selected from the three centers to constitute the external test set. The image preprocessing involved both registration and grayscale normalization. We performed affine spatial normalization using FSL and symmetric normalization (SyN) via ANTS to align the multiphase CECT images with NCCT images. The CT images were clipped within the range of [-1000, 1000] Hounsfield units (HU) and normalized to [-1, 1] using a piecewise linear mapping approach.

**Implementation.** All experiments were conducted in PyTorch on a server with an NVIDIA GeForce RTX 3090 GPU. The input size for the synthesis models was $256 \times 256 \times 8$. Data augmentation, including random translation and flipping, was applied to enhance image diversity. The timestep was set to 1000. We used the AdamW optimizer with a polynomial decay learning rate strategy, starting at $2 \times 10^{-4}$ for up to 50 epochs. $\lambda_1, \lambda_2, \lambda_4$, and $\lambda_5$ are set to 10, 0.1, 0.1 and 0.1, respectively.

**Evaluation Metric.** To evaluate synthesis performance, we used the MAE, structural similarity index measurement (SSIM), and peak signal-to-noise ratio (PSNR) to assess pixel-level similarity and image quality, as well as Frechet Inception Distance (FID)[18] and Learned Perceptual Image Patch Similarity (LPIPS)[19] to measure the similarity of data distributions between the synthetic and real images.

### 3.2 Experimental results

**Ablation Study.** To assess the contribution of each key component in CFPS-Diff, we conducted an ablation study on the validation set from the CT-1 dataset (as shown in Table 1), evaluating different model variants. We used DDPM as the baseline to synthesize multiphase CECT from NCCT. We then introduced CFPL, CPSL, and CDW-Loss, resulting in three ablation models to analyze the impact of each component. Compared with the baseline, the whole CFPS-Diff framework led to a reduction in mean MAE from 38.73 to 30.69, an increase in mean SSIM from 0.787 to 0.82, and an elevation in mean PSNR from 28.07

**Table 2.** Quantitative comparisons of different models for NCCT-to-multiphase CECT synthesis on multi-enter datasets.

| Dataset | Model | Arterial Phase | | | Venous Phase | | | Delay Phase | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | MAE | PSNR | SSIM | MAE | PSNR | SSIM | MAE | PSNR | SSIM |
| CT-2 Dataset | U-Net | 35.79 | 28.81 | 0.802 | 36.80 | 28.29 | 0.788 | 36.02 | 28.21 | 0.795 |
| | Pix2pix | 45.43 | 27.57 | 0.749 | 61.16 | 26.15 | 0.616 | 49.38 | 27.18 | 0.695 |
| | Pix2pixHD | 47.10 | 27.25 | 0.715 | 38.28 | 28.12 | 0.774 | 36.97 | 28.13 | 0.788 |
| | CycleGAN | 42.52 | 27.89 | 0.751 | 48.84 | 27.24 | 0.704 | 44.04 | 27.52 | 0.739 |
| | SwinUNETR | 32.66 | 29.24 | 0.819 | 34.51 | 28.50 | 0.802 | 34.43 | 28.38 | 0.805 |
| | TransUNet | 33.48 | 29.12 | 0.815 | 36.04 | 28.26 | 0.793 | 34.90 | 28.34 | 0.802 |
| | ResViT | 34.09 | 29.03 | 0.809 | 34.47 | 28.61 | 0.800 | 36.23 | 28.19 | 0.796 |
| | DDPM | 39.47 | 28.21 | 0.791 | 39.99 | 27.84 | 0.780 | 38.72 | 27.79 | 0.794 |
| | **CFPS-Diff** | **30.44** | **29.79** | **0.832** | **31.63** | **29.35** | **0.820** | **31.52** | **29.02** | **0.822** |
| External Set | U-Net | 35.95 | 28.53 | 0.798 | 35.44 | 28.54 | 0.791 | 35.29 | 28.39 | 0.797 |
| | Pix2pix | 44.96 | 27.38 | 0.750 | 58.94 | 26.35 | 0.625 | 47.42 | 27.46 | 0.704 |
| | Pix2pixHD | 47.51 | 27.15 | 0.702 | 37.63 | 28.24 | 0.773 | 36.42 | 28.33 | 0.787 |
| | CycleGAN | 42.42 | 27.64 | 0.749 | 46.73 | 27.48 | 0.711 | 42.33 | 27.73 | 0.746 |
| | SwinUNETR | 32.94 | 28.92 | 0.816 | 33.18 | 28.75 | 0.805 | 33.78 | 28.57 | 0.807 |
| | TransUNet | 33.50 | 28.84 | 0.813 | 34.96 | 28.44 | 0.797 | 33.83 | 28.63 | 0.806 |
| | ResViT | 34.59 | 28.64 | 0.804 | 33.37 | 28.78 | 0.803 | 35.69 | 28.32 | 0.798 |
| | DDPM | 39.49 | 28.03 | 0.790 | 38.70 | 28.10 | 0.782 | 37.87 | 28.01 | 0.797 |
| | **CFPS-Diff** | **30.87** | **29.41** | **0.830** | **30.46** | **29.61** | **0.823** | **30.91** | **29.24** | **0.825** |

to 29.51. In particular, the sequential addition of the CFPL and CPSL modules resulted in improvements of 10.20% and 11.73% in MAE, as well as 1.78% and 2.87% in SSIM, respectively. These results demonstrate that the proposed modules significantly enhance the synthesis of multiphase CECT from NCCT.

**Comparison with Literature.** We conducted a series of experiments to compare the performance of our proposed CFPS-Diff with other SOTA synthesis methods, including U-Net[20], Pix2pix [21], Pix2pixHD[22], CycleGAN [23], SwinUNETR[24], TransUNet [25], ResViT[26] and DDPM.The quantitative results of all synthesis methods on our CT-2 Dataset and external set are presented in Table 2. Compared with these eight methods, our CFPS-Diff achieves the lowest MAE, highest SSIM, and highest PSNR, with a mean MAE of 31.20, a mean SSIM of 0.825, and a mean PSNR of 29.39 on the CT-2 Dataset, and 30.75, 0.826, and 29.42 on the external set. Specifically, compared to the suboptimal SwinUNETR, our model achieves notable improvements of 6.8%, 8.35%, and 8.45% in MAE across AP, VP, and DP, respectively, on CT-2 Dataset, and 6.28%, 8.2%, and 8.5% on external set. Fig. 2. visualizes the synthetic triple-phase CECT images from different methods. It is observed that the synthetic CECT images generated by our CFPS-Diff show richer contrast-enhanced details, especially in the arterial arteries, than those generated by other SOTA synthesis methods. To more comprehensively evaluate the quality of synthetic CECT images by different methods, we employed two commonly used metrics

**Table 3.** Evaluation of different models for NCCT-to-multiphase CECT synthesis based on data distribution similarity on External Set

| Model | Arterial Phase | | Venous Phase | | Delay Phase | |
|---|---|---|---|---|---|---|
| | FID | LPIPS$_{\times 10}$ | FID | LPIPS$_{\times 10}$ | FID | LPIPS$_{\times 10}$ |
| U-Net | 31.77 | 0.95±0.15 | 37.89 | 1.13±0.18 | 31.30 | 0.96±0.14 |
| Pix2pix | 28.53 | 0.92±0.14 | 35.53 | 1.12±0.17 | 30.39 | 0.95±0.14 |
| Pix2pixHD | 27.65 | 1.03±0.16 | 26.97 | 1.00±0.17 | 29.58 | 0.88±0.14 |
| CycleGAN | 29.95 | 0.99±0.14 | 31.30 | 1.17±0.18 | 25.77 | 0.95±0.14 |
| SwinUNETR | 23.50 | 0.75±0.13 | 28.17 | 0.89±0.15 | 23.41 | 0.77±0.13 |
| TransUNet | 27.50 | 0.79±0.13 | 29.50 | 0.91±0.15 | 27.35 | 0.79±0.12 |
| ResViT | 27.74 | 0.81±0.14 | 26.68 | 0.92±0.17 | 27.61 | 0.85±0.15 |
| DDPM | 29.64 | 0.86±0.14 | 32.56 | 1.02±0.17 | 27.45 | 0.85±0.13 |
| **CFPS-Diff** | **21.46** | **0.70±0.14** | **26.43** | **0.86±0.16** | **20.56** | **0.71±0.13** |

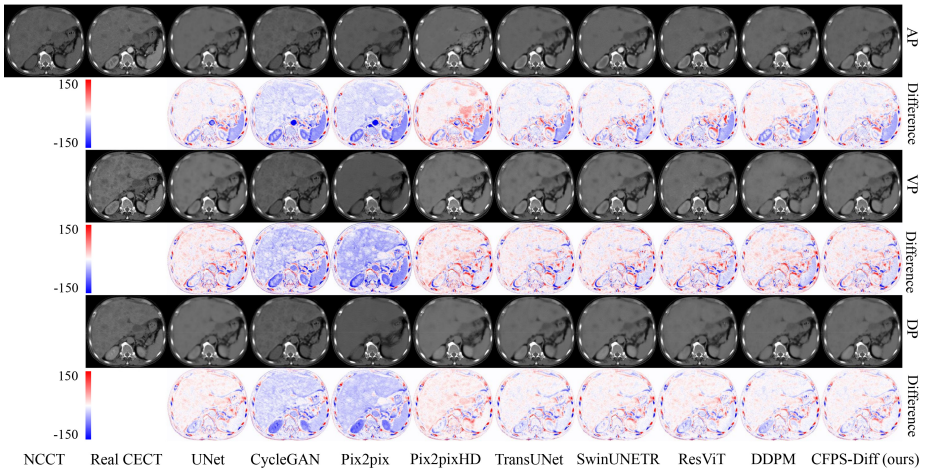


**Fig. 2.** Visual results of synthetic multiphase CECT images generated by different synthesis methods on external set. The color maps are the difference between the synthetic CECT and the real CECT.

for assessing visual quality, namely FID and LPIPS, with the results presented in Table 3. Our CFPS-Diff achieved improvements of 32.45%, 30.25%, and 34.31% in FID for the three phases, and enhancements of 26.32%, 23.89%, and 26.04% in LPIPS, respectively, compared to the lower limit.

## 4 Conclusion

In this paper, we introduce CFPS-Diff, a united framework for synthesizing multiphase CECT images from NCCT images. Unlike existing approaches that treat each phase independently, our proposed CFPS-Diff integrates CFPL within a conditional diffusion model. This integration enables smooth and orderly transitions between phases, ensuring both anatomical and temporal consistency in

the synthesized enhancement patterns by effectively modeling the dynamic distribution of contrast agents. Furthermore, we introduce CPSL through a phase distinction network, which aligns cross-phase specificity features with contrast information, thereby enhancing the synthesized CECT images and improving quantitative synthesis metrics. Extensive experiments conducted on multi-center datasets demonstrate the superior performance of CFPS-Diff compared to state-of-the-art medical image synthesis methods.

**Acknowledgments** This work was partially supported by grants from the National Natural Science Foundation of China (No. 82472050, No. 62101239) and Guangdong Provincial Key Laboratory of Medical Image Processing (No. 2020B1212060039).

**Disclosure of Interests** The authors have no competing interests to declare that are relevant to the content of this article.

## References

1. T. J. Van Oostenbrugge, J. J. Fütterer, and P. F. Mulders, "Diagnostic imaging for solid renal tumors: a pictorial review," *Kidney Cancer*, vol. 2, no. 2, pp. 79–93, 2018.
2. B. Gulpinar, E. Peker, M. Kul, A. H. Elhan, and N. Haliloglu, "Liver metastases of neuroendocrine tumors: is it possible to diagnose different histologic subtypes depending on multiphasic ct features?" *Abdominal Radiology*, vol. 44, pp. 2147–2155, 2019.
3. W. Bottinor, P. Polkampally, and I. Jovin, "Adverse reactions to iodinated contrast media," *International Journal of Angiology*, vol. 22, no. 03, pp. 149–154, 2013.
4. J. Peng, Y. Liu, D. Jiang, X. Wang, P. Peng, S. He, W. Zhang, and F. Zhou, "Deep learning and gan-synthesis for auto-segmentation of pancreatic cancer by non-enhanced ct for adaptive radiotherapy," *International Journal of Radiation Oncology, Biology, Physics*, vol. 117, no. 2, pp. e499–e500, 2023.
5. A. Chandrashekar, N. Shivakumar, P. Lapolla, A. Handa, V. Grau, and R. Lee, "A deep learning approach to generate contrast-enhanced computerised tomography angiograms without the use of intravenous contrast agents," *European Heart Journal*, vol. 41, no. Supplement_2, pp. ehaa946–0156, 2020.
6. J. W. Choi, Y. J. Cho, J. Y. Ha, S. B. Lee, S. Lee, Y. H. Choi, J.-E. Cheon, and W. S. Kim, "Generating synthetic contrast enhancement from non-contrast chest computed tomography using a generative adversarial network," *Scientific reports*, vol. 11, no. 1, p. 20403, 2021.
7. M. Seo, D. Kim, K. Lee, S. Hong, J. S. Bae, J. H. Kim, and S. Kwak, "Neural contrast enhancement of ct image," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2021, pp. 3973–3982.
8. J. Lyu, Y. Fu, M. Yang, Y. Xiong, Q. Duan, C. Duan, X. Wang, X. Xing, D. Zhang, J. Lin *et al.*, "Generative adversarial network–based noncontrast ct angiography for aorta and carotid arteries," *Radiology*, vol. 309, no. 2, p. e230681, 2023.

9. H. Pang, S. Qi, Y. Wu, M. Wang, C. Li, Y. Sun, W. Qian, G. Tang, J. Xu, Z. Liang *et al.*, "Ncct-cect image synthesizers and their application to pulmonary vessel segmentation," *Computer Methods and Programs in Biomedicine*, vol. 231, p. 107389, 2023.
10. L. Zhong, R. Xiao, H. Shu, K. Zheng, X. Li, Y. Wu, J. Ma, Q. Feng, and W. Yang, "Ncct-to-cect synthesis with contrast-enhanced knowledge and anatomical perception for multi-organ segmentation in non-contrast ct images," *Medical Image Analysis*, vol. 100, p. 103397, 2025.
11. K.-H. Uhm, S.-W. Jung, M. H. Choi, S.-H. Hong, and S.-J. Ko, "A unified multiphase ct synthesis and classification framework for kidney cancer diagnosis with incomplete data," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 12, pp. 6093–6104, 2022.
12. J. Liu, Y. Tian, C. Duzgol, O. Akin, A. M. Ağıldere, K. M. Haberal, and M. Coşkun, "Virtual contrast enhancement for ct scans of abdomen and pelvis," *Computerized Medical Imaging and Graphics*, vol. 100, p. 102094, 2022.
13. L. Zhong, P. Huang, H. Shu, Y. Li, Y. Zhang, Q. Feng, Y. Wu, and W. Yang, "United multi-task learning for abdominal contrast-enhanced ct synthesis through joint deformable registration," *Computer Methods and Programs in Biomedicine*, vol. 231, p. 107391, 2023.
14. A. Kazerouni, E. K. Aghdam, M. Heidari, R. Azad, M. Fayyaz, I. Hacihaliloglu, and D. Merhof, "Diffusion models in medical imaging: A comprehensive survey," *Medical image analysis*, vol. 88, p. 102846, 2023.
15. J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in Neural Information Processing Systems*, vol. 33, pp. 6840–6851, 2020.
16. H. Huang, R. He, Z. Sun, and T. Tan, "Wavelet-srnet: A wavelet-based cnn for multi-scale face super resolution," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 1689–1697.
17. G. Polat, I. Ergenc, H. T. Kani, Y. O. Alahdab, O. Atug, and A. Temizel, "Class distance weighted cross-entropy loss for ulcerative colitis severity estimation," in *Annual Conference on Medical Image Understanding and Analysis*. Springer, 2022, pp. 157–171.
18. M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "Gans trained by a two time-scale update rule converge to a local nash equilibrium," *Advances in neural information processing systems*, vol. 30, 2017.
19. R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 586–595.
20. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 2015, pp. 234–241.
21. P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134.
22. T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional gans," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 8798–8807.
23. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2223–2232.

24. A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth, and D. Xu, "Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images," in *International MICCAI brainlesion workshop*. Springer, 2021, pp. 272–284.
25. J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, "Transunet: Transformers make strong encoders for medical image segmentation," *arXiv preprint arXiv:2102.04306*, 2021.
26. O. Dalmaz, M. Yurt, and T. Çukur, "Resvit: residual vision transformers for multimodal medical image synthesis," *IEEE Transactions on Medical Imaging*, vol. 41, no. 10, pp. 2598–2614, 2022.