Masked Residual Diffusion Probabilistic Model with Regional Asymmetry Prior for Generating Perfusion Maps from Multi-phase CTA

Yuxin Cai[1], Jianhai Zhang[2], Lei He[1], and Aravind Ganesh[2], and Wu Qiu[1⋆]

[1] College of Life Science and Technology, Huazhong University of Science and Technology, China
[2]Departments of Clinical Neuroscience, Cumming School of Medicine,University of Calgary, Canada.

Abstract. Multiphase CT angiography (mCTA) has become an important diagnostic tool for acute ischemic stroke (AIS), offering insights into occlusion sites and collateral circulation. However, its broader application is hindered by the need for specialized interpretation, contrasting with the intuitive nature of CT perfusion (CTP). In this work, we propose a novel diffusion based generative model to generate CTP-like perfusion maps, enhancing AIS diagnosis in resource-limited settings. Unlike traditional diffusion models that restore images by predicting the added noise, our approach uses a masked residual diffusion probabilistic model (MRDPM) to recover the residuals between the predicted and target image within brain regions of interests for more detailed generation. To target denoising efforts on relevant regions, noise is selectively added into the brain area only during diffusion. Furthermore, a Multi-scale Asymmetry Prior module and a Brain Region-Aware Network are proposed to incorporate anatomical prior information into the MRDPM to generate finer details while ensuring consistency. Experimental evaluations with 514 patient images demonstrate that our proposed method is able to generate high quality CTP-like perfusion maps, outperforming several other generative models regarding the metrics of MAE, LPIPS, SSIM, and PSNR. The code is publicly available at `https://github.com/UniversalCAI/MRDPM-with-RAP`.

Keywords: Acute ischemic stroke · CT perfusion imaging · Multi-phase CT angiography· Diffusion models · Medical image synthesis

1 Introduction

Multiphase CT angiography (mCTA) emerges as a critical advancement in the diagnostic landscape for acute ischemic stroke (AIS), primarily utilized to assess occlusions and collateral status for patients with AIS [10,16]. However, its clinical adoption has been hampered by the need for specialized interpretative expertise,

⋆ Corresponding Author: wuqiu@hust.edu.cn

contrasting with CTP's more intuitive evaluation process. This challenge has spurred efforts to leverage mCTA in generating perfusion maps, *e.g.,* CBF, CBV, and Tmax, similar to those provided by CTP, aiming to optimize stroke care in the settings constrained by imaging resources.

To tackle the challenges in interpreting mCTA, several methods have been introduced, including traditional deconvolution algorithms [4,5], random forest techniques [11], and convolutional neural network (CNN)-based approaches [17]. These existing methods, however, suffer from noise and the low temporal resolution characteristic of mCTA, resulting in generated maps of diminished quality and missing the intricate patterns typically observed in CTP images. Despite these advancements, a notable gap persists in rendering mCTA as intuitive and detail-rich as CTP, accentuating the need for continued innovation in this area.

Image generation models, *e.g.*, generative adversarial networks (GANs) [18,7], have revolutionized the field of cross-modality medical image synthesis [6,22,1], offering new avenues for enhancing mCTA's interpretability. More recently, diffusion model has established new performance benchmarks in image synthesis [14,13]. Compared to GANs, diffusion model demonstrates more stable training processes and a reduced likelihood of losing data distribution [2,21]. However, the application of diffusion model to medical imaging faces unique challenges, including the complex structure of medical images and the critical need for precision in image translation [8,12,9]. In order to address these challenges, we propose a novel approach that harnesses the ability of diffusion models for the synthesis of CTP perfusion images from mCTA data.

Our key technical contributions are summarized as follows: 1) a novel masked residual diffusion model architecture is proposed, which surpasses current diffusion models in generating finer image details while ensuring consistency; 2) we introduce a multi-scale asymmetry prior module, which enhances the guidance for the diffusion model by integrating and fusing latent space features derived from the initial predictor; 3) additionally, a brain region-aware network is designed for improving feature extraction from anatomical areas within the brain.

To our knowledge, this paper represents the first study to introduce diffusion model to synthesize CTP-like maps from multiphase CTA. This integrated approach is able to bridge the gap in interpretability between mCTA and CTP, providing a more intuitive and detailed method for diagnosing AIS.

2 Methodology

Given a dataset $\{(x, y)\}_{i=1}^{N}$, where $x \in \mathbb{R}^{1\times H\times W}$ and $y \in \mathbb{R}^{4\times H\times W}$ denote the output and input samples, respectively. The output x has one channel for CTP images, and the input y has 4 channels composed of one non-contrast CT (NCCT, typically acquired prior to CTA at baseline) and three temporal phases of mCTA. Our aim is to efficiently generate CTP-like maps x , *i.e.*, CBF, CBV, and Tmax, based on mCTA images y.

To overcome the challenges posed by the inherent randomness of Gaussian noise in existing diffusion models for medical image generation, we introduce

an integrated framework. This framework implements a residual-based diffusion model to ensure stable outputs with two tailor-made modules: the Multi-scale Asymmetry Prior (MAP) module and the Brain Region-Aware Network (BRAN). Our objective is to refine the precision of noise prediction while concentrating on the brain and its pertinent areas, thus elevating both the quality and efficiency of image generation. The framework of our proposed method is illustrated in Fig. 1.

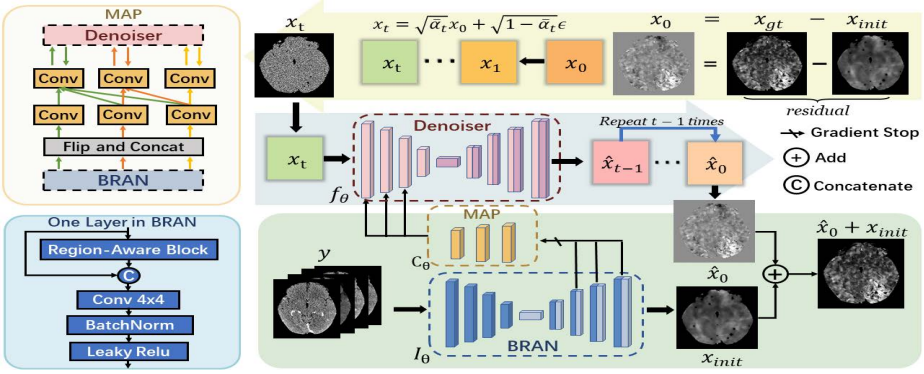

Fig. 1. Overview of our proposed Masked Residual Diffusion Model framework

2.1 Masked Residual Diffusion Model

The randomness increases in the reverse process of diffusion models, even with a good initial condition, due to the stochastic nature of Gaussian noise. Therefore, a residual-based diffusion model [19] is proposed to stabilize the outputs in the deterministic inference. Unlike conventional diffusion models, the proposed residual diffusion model restores the residual between the ground truth x_{gt} and an initial mapping $x_{\text{init}} = I_\theta(y)$ (Fig. 1). The predictor $I_\theta(y)$ is pre-trained to ensure the next diffusion model can be steadily trained. The pre-training objective $\mathcal{L}_{\text{init}}$ is defined as the mean absolute error between the x_{init} and the x_{gt} :

$$\mathcal{L}_{\text{init}} = \mathbb{E} \left\| x_{\text{gt}} - x_{\text{init}} \right\|_1 \tag{1}$$

During the training phase, $I_\theta(y)$ is fine-tuned within the diffusion model to improve adaptability and ensure consistent output quality. The diffusion model, represented as f_θ, aims to estimate noise $\hat{\epsilon}$ in the reverse process:

$$\begin{aligned} \hat{\boldsymbol{\epsilon}} &= f_\theta(x_t, \bar{\alpha}, y, \Omega) \\ \texttt{Sampling:}\ x_t &= \sqrt{\bar{\alpha}_t} x_0 + \sqrt{1-\bar{\alpha}_t}\epsilon, \quad \epsilon \sim \mathcal{N}(0, \mathbf{I}) \\ \texttt{Residual:}\ x_0 &= x_{\text{gt}} - x_{\text{init}} \end{aligned} \tag{2}$$

where $\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i, \alpha_i$ denotes variance schedule, Ω denotes conditions.

In contrast to the traditional diffusion process, our approach introduces noise selectively into the foreground regions during the forward diffusion phase, specifically designed for brain images where the background outside of the brain is typically non-essential. This technique refines the denoising process by focusing on the vital task of restoring information in areas of interest, rather than indiscriminately applying noise across the entire image. We employ a binary foreground mask, M_{FG}, representing the brain, which can be readily derived using morphological operations from mCTA images. Subsequently, we refine the diffusion loss function $\mathcal{L}_{\text{diff}}$ by computing the mean square error (MSE) between the actual noise ϵ and its estimated counterpart $\hat{\epsilon}$ with this computation being exclusively restricted to the regions delineated by M_{FG} (Eq. (3)). This modification ensures that only the foreground areas are considered in the loss calculation, enhancing the specificity and efficiency of our denoising process.

$$\mathcal{L}_{\text{diff}} = \mathbb{E} \left\| M_{FG} \cdot (\boldsymbol{\epsilon} - \hat{\boldsymbol{\epsilon}}) \right\|_2^2 \tag{3}$$

Despite the fact that fine-tuning I_θ during training does not require the computation of gradients [19], it is essential to preserve the stability of I_θ. This preservation helps in preventing the deterioration of the pre-trained I_θ and ensures the stability of the diffusion model's objectives. Consequently, we introduce an additional loss function weighted by λ, similar to the approach outlined in Eq.(1), to provide supervision for I_θ. λ is empirically set as 0.1. This adjustment redefines the training phase objective as follows:

$$\mathcal{L} = \mathcal{L}_{\text{diff}} + \lambda \mathcal{L}_{\text{init}}, \quad \lambda > 0 \tag{4}$$

2.2 Multi-scale Asymmetry Prior

Radiologists' interpretations heavily rely on their prior knowledge, encompassing anatomy and contextual understanding of medical images. Experts can spot abnormalities by comparisons between the ischemic and contralateral hemispheres within specific anatomical regions. Inspired by this fact, a functional module C_θ is proposed to extract the latent space features Z_i and Z_f from the initial predictors I_θ and the denoiser f_θ, respectively. The fused outputs, called Multi-scale Asymmetry Prior (MAP), are subsequently used as additional condition for predicting the noise in the diffusion model. The module structure C_θ is illustrated in Fig. 2, and the process is defined as below:

$$\Omega := C_\theta(Z_i, Z_f) \tag{5}$$

Specifically, $\hat{z}_I^l \in Z_I$ represents a replicated feature from the l-th layer of the initial predictor. We employ a sampling module δ^l to process $\hat{z}_I^l$ at each layer. To augment the model's awareness of asymmetrical horizontal patterns, we apply a mirroring operation to $\hat{z}_I^l$, generating a horizontally flipped counterpart, labeled as z_i^l. Utilizing a pyramid feature fusion approach, we ensure these features contribute not just within their original layers but are also integrated into

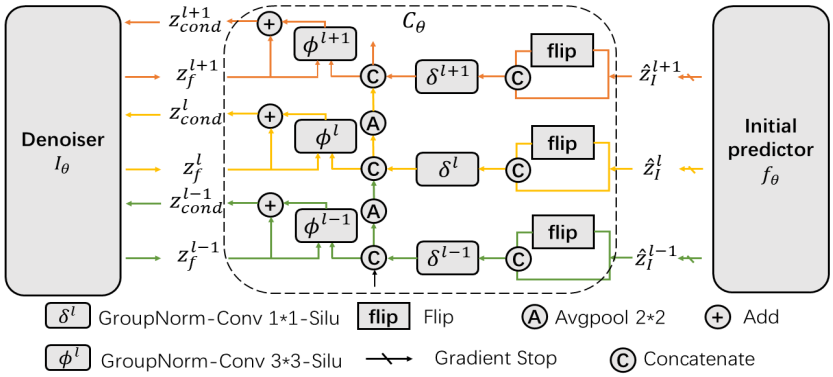

Fig. 2. Structure of C_θ

higher layers through average pooling. This strategy bolsters the model's ability to refine features across levels, ensuring a comprehensive multi-scale feature representation. This process can be formulated as follows:

$$z_i^l = \delta^l \left(\text{flip} \left(\hat{z}_I^l \right), \hat{z}_I^l \right) \tag{6}$$

$$z_i^{1 \to l} = \text{concat} \left(z_i^l, \text{avgpool} \left(z_i^{1 \to l-1} \right) \right) \tag{7}$$

The conditional sampling module, denoted as ϕ^l, integrates the diffusion model's layer-specific feature $z_f^l \in Z_f$ with the aggregated multi-layer feature $z_i^{1 \to l}$ through ϕ^l sampling. This feature is then added back to the original input feature, forming the enhanced feature z_{cond}^l returned to the diffusion model:

$$z_{\text{cond}}^l = z_f^l + \phi^l \left(z_f^l, z_i^{1 \to l} \right) \tag{8}$$

$$Z_{\text{cond}} = \{ z_{\text{cond}}^1, z_{\text{cond}}^2, \ldots, z_{\text{cond}}^l \} \tag{9}$$

This methodology ensures that each layer's features are both refined and expanded upon, providing a rich, multi-layered feature set for the model.

2.3 Brain Region-Aware network

Brain imaging highlights unique cerebral structures, yet traditional convolutional modules, designed for direct mapping, struggle to leverage this anatomical structure information. Addressing this, we introduce a Brain Region-Aware Network (BRAN) module as the initial predictor I_θ [3] illustrated in Fig. 3. BRAN improves mCTA to CTP mapping accuracy by integrating features more effectively across brain sub-regions through an improved U-shape architecture [7].

Specifically, our method roughly segments brain images into four regions, considered as background, gray matter, white matter, and ventricle. It is implemented by a guided Mask GM which divides the spatial domain into $D = 4$

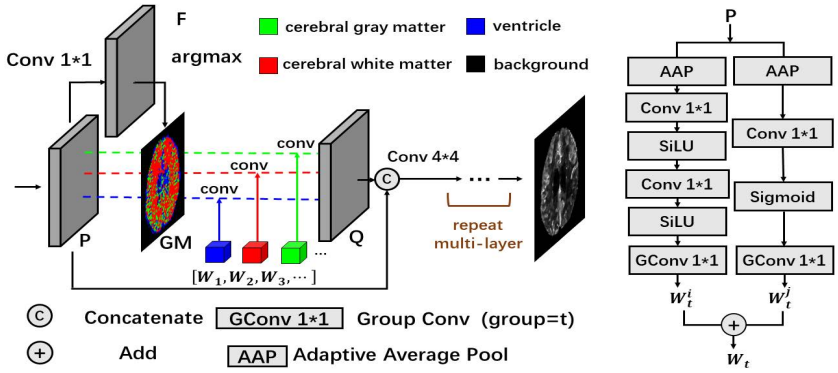

Fig. 3. Architecture of a Layer in BRAN

subregions. Each subregion is with a unique set of filters W. The GM is dynamically generated from the input feature map $P \in \mathbb{R}^{U\times V\times C}$ via a 1×1 convolution, producing a D-channel feature map $F = \mathrm{Conv}_{1\times 1}(P)$, $F \in \mathbb{R}^{U\times V\times D}$, where U, V, and C represent the height, width, and number of channels, respectively [3]. Each channel of F represents one sub-region, and the GM at each point (u, v) is determined by the channel with the highest value. This process is described as follows:

$$GM_{u,v} = \mathrm{argmax}\left(F^0_{u,v}, F^1_{u,v}, \cdots, F^{D-1}_{u,v}\right) \tag{10}$$

$$S_t = \{(u, v)|GM(u, v) = d\}, \quad d \in [0, \ldots, D-1] \tag{11}$$

To enhance the mapping capability of the BRAN module, we augment the original filter strategy by incorporating two sets of filters with varying depths for each sub-region, denoted as $W^i = [W^i_0, \cdots, W^i_{D-1}]$ and $W^j = [W^j_0, \cdots, W^j_{D-1}]$. The filters W^i_d and W^j_d, corresponding to region S_d, are convolved with the input feature P within their specific sub-regions. This is followed by an aggregation step to generate the final feature map Q, as illustrated by:

$$W_d = W^i_d + W^j_d, \quad d \in [0, \ldots, D-1] \tag{12}$$

$$Q = \sum_{\substack{d=0 \\ (u,v)\in S_d}}^{D-1} P_{u,v} * W_d \tag{13}$$

where W_d represents the sum of W^i_d and W^j_d. The process of acquiring W^i_d and W^j_d is depicted in Fig. 3.

3 Experiments and Results

Datasets and Pre-processing. Our study involved imaging data from 514 AIS patients, each having undergone mCTA and CTP images as reference standard

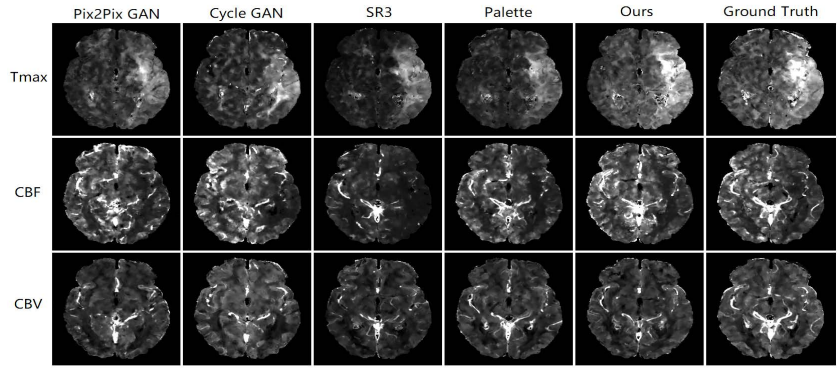

Fig. 4. Comparative Results of CTP Map Generation

prior to treatment. We split the dataset into 411 images for training and 103 for testing. To prepare for neural network analysis, we standardized preprocessing: this included aligning modalities, axial slicing, and adjusting images (cropping, rotation, translation) for brain centralization and skull stripping to remove non-brain elements. Images were resized to 256×256 pixels at $0.625mm \times 0.625mm$ spacing. We normalized CBF to [0,200], CBV to [0,20], and Tmax to [0,30].

Hyper-parameters Settings. The learning rate remains fixed for the initial 100 epochs and then linearly declines from 5×10^{-5} to 5×10^{-7} over the subsequent 100 epochs, with a batch size of 4 throughout. Noise variances linearly spanned from $\beta_1 = 10^{-6}$ to $\beta_T = 0.01$, with $T = 1000$ diffusion steps. The model was trained on a single NVIDIA RTX A100 GPU with 40 GB memory using the Adam optimizer in PyTorch. An Exponential Moving Average (EMA) over the model parameters with a decay rate of 0.9999 was employed.

Evaluation Metrics. Four quantitative metrics, *i.e.*, Mean Absolute Error (MAE), Learned Perceptual Image Patch Similarity (LPIPS), Structural Similarity Index Measure(SSIM), and Peak Signal-to-Noise Ratio (PSNR), were used for evaluations. Notably, LPIPS[20], as a deep learning derived perceptual metric, canevaluate the generated images from a perspective of human visual perception.

Results. The quantitative and qualitative results are shown in Tab. 1 and Fig.4, respectively. Our model demonstrated great performance in both structural consistency and perceptual quality. We compared our approach with five advanced image synthesis techniques: two GAN based methods of Pix2Pix [7] and CycleGAN [23], and two diffusion model based methods of SR3 [15] and Palette [14], with each method trained using its official settings for optimal performance. For comparisons, the masked diffusion strategy, as elaborated in Section 2.1, was applied to both SR3 and Palette. Results in the Tab.1 show that our method surpasses other methods in both consistency and perceptual metrics.

An ablation study further dissected the impact of key model components on the image generation process. Incrementally removing elements such as mask, residual diffusion (Res), BRAN, and MAP allowed us to gauge their contribution

to the model's performance. An initial predictor based on the generator proposed by Isola et al.[7] is applied for baseline comparison with BRAN. As evidenced by the results in Tab. 2, each component played a vital role in enhancing the accuracy and quality of the synthesized images, with the full configuration yielding the best performance.

Table 1. Quantitative results on different perfusion maps.

Parameters	Methods	MAE ↓	LPIPS ↓	SSIM ↑	PSNR ↑
Tmax	Pix2Pix [7]	0.051 ± 0.024	0.198 ± 0.050	0.640 ± 0.133	20.2 ± 3.30
	CycleGAN [23]	0.065 ± 0.032	0.218 ± 0.055	0.598 ± 0.146	18.2 ± 3.34
	SR3 [15]	0.058 ± 0.036	0.207 ± 0.071	0.641 ± 0.143	19.8 ± 4.25
	Palette [14]	0.051 ± 0.028	0.193 ± 0.057	0.648 ± 0.135	20.5 ± 3.85
	Ours	$\mathbf{0.047 \pm 0.023}$	$\mathbf{0.179 \pm 0.051}$	$\mathbf{0.649 \pm 0.135}$	$\mathbf{21.1 \pm 3.57}$
CBF	Pix2Pix [7]	0.044 ± 0.019	0.208 ± 0.051	0.655 ± 0.120	20.2 ± 2.84
	CycleGAN [23]	0.051 ± 0.024	0.220 ± 0.055	0.646 ± 0.123	19.4 ± 2.97
	SR3 [15]	0.044 ± 0.022	0.208 ± 0.059	0.667 ± 0.124	20.6 ± 3.20
	Palette [14]	0.043 ± 0.020	0.203 ± 0.057	0.670 ± 0.121	20.7 ± 3.10
	Ours	$\mathbf{0.041 \pm 0.018}$	$\mathbf{0.192 \pm 0.049}$	$\mathbf{0.673 \pm 0.116}$	$\mathbf{21.0 \pm 2.98}$
CBV	Pix2Pix [7]	0.050 ± 0.022	0.196 ± 0.051	0.655 ± 0.126	20.0 ± 2.86
	CycleGAN [23]	0.057 ± 0.026	0.203 ± 0.056	0.640 ± 0.130	19.2 ± 2.90
	SR3 [15]	0.057 ± 0.035	0.207 ± 0.088	0.649 ± 0.149	19.7 ± 3.60
	Palette [14]	0.053 ± 0.031	0.192 ± 0.071	0.660 ± 0.138	20.1 ± 3.42
	Ours	$\mathbf{0.046 \pm 0.021}$	$\mathbf{0.171 \pm 0.049}$	$\mathbf{0.668 \pm 0.126}$	$\mathbf{20.6 \pm 2.88}$

Table 2. Ablation Study Results

Mask	Res	BRAN	MAP	MAE ↓	LPIPS ↓	SSIM ↑	PSNR ↑
×	×	×	×	0.065 ± 0.035	0.268 ± 0.102	0.585 ± 0.142	18.9 ± 3.64
✓	×	×	×	0.051 ± 0.028	0.188 ± 0.057	0.640 ± 0.135	20.2 ± 3.85
✓	✓	×	×	0.050 ± 0.025	0.191 ± 0.052	0.641 ± 0.138	20.4 ± 3.51
✓	✓	✓	×	0.049 ± 0.023	0.189 ± 0.052	0.647 ± 0.136	20.7 ± 3.45
✓	✓	✓	✓	$\mathbf{0.047 \pm 0.023}$	$\mathbf{0.180 \pm 0.051}$	$\mathbf{0.649 \pm 0.135}$	$\mathbf{21.1 \pm 3.57}$

4 Discussion and Conclusion

Our work presents a clinically-applicable approach in medical imaging, specifically for CTP-like map generation from mCTA images. The proposed masked residual diffusion probabilistic model, enhanced by foreground restriction masks and residual diffusion, offering a blend of precision and detail. In addition, MAP and BRAN are designed to address the specific challenges of medical image generation, ensuring the accurate reproduction of intricate cerebral structures and perfusion patterns. The application of this model in clinical settings could greatly augment diagnostic capabilities for AIS, providing high-quality CTP-like maps from readily accessible mCTA images. This advancement is especially promising for resource-limited settings, where access to advanced imaging technologies may

be constrained. By delivering detailed perfusion maps that closely mimic those produced by traditional CTP, our model not only bridges a critical gap in stroke diagnostics but also enhances the potential for early and accurate intervention, underscoring its clinical relevance and potential impact on patient care.

Acknowledgments. This work was supported in part by the National Key Research and Development Program of China (2023YFC2410802), the Hubei Provincial Key Research and Development Program (2023BCB007), the Fundamental Research Funds for the Central Universities (HUST:2024JYCXJJ035), the High-Performance Computing platform of Huazhong University of Science and Technology and computer power at Wuhan Seekmore Intelligent Imaging Inc.

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

References

1. Armanious, K., Jiang, C., Fischer, M., Küstner, T., Hepp, T., Nikolaou, K., Gatidis, S., Yang, B.: Medgan: Medical image translation using gans. Computerized medical imaging and graphics **79**, 101684 (2020)
2. Bau, D., Zhu, J.Y., Wulff, J., Peebles, W., Strobelt, H., Zhou, B., Torralba, A.: Seeing what a gan cannot generate. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 4502–4511 (2019)
3. Chen, J., Wang, X., Guo, Z., Zhang, X., Sun, J.: Dynamic region-aware convolution. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 8064–8073 (2021)
4. Chung, K.J., Khaw, A.V., Pandey, S.K., Lee, D.H., Mandzia, J.L., Lee, T.Y.: Feasibility of deconvolution-based multiphase ct angiography perfusion maps in acute ischemic stroke: Simulation and concordance with ct perfusion. Journal of Stroke and Cerebrovascular Diseases **31**(12), 106844 (2022)
5. Chung, K.J., Pandey, S.K., Khaw, A.V., Lee, T.Y.: Multiphase ct angiography perfusion maps for predicting target mismatch and ischemic lesion volumes. Scientific Reports **13**(1), 21976 (2023)
6. Dalmaz, O., Yurt, M., Çukur, T.: Resvit: Residual vision transformers for multimodal medical image synthesis. IEEE Transactions on Medical Imaging **41**(10), 2598–2614 (2022)
7. Isola, P., Zhu, J.Y., Zhou, T., Efros, A.A.: Image-to-image translation with conditional adversarial networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1125–1134 (2017)
8. Kazerouni, A., Aghdam, E.K., Heidari, M., Azad, R., Fayyaz, M., Hacihaliloglu, I., Merhof, D.: Diffusion models for medical image analysis: A comprehensive survey. arXiv preprint arXiv:2211.07804 (2022)
9. Khader, F., Müller-Franzes, G., Tayebi Arasteh, S., Han, T., Haarburger, C., Schulze-Hagen, M., Schad, P., Engelhardt, S., Baeßler, B., Foersch, S., et al.: Denoising diffusion probabilistic models for 3d medical image generation. Scientific Reports **13**(1), 7303 (2023)
10. Menon, B.K., d'Esterre, C.D., Qazi, E.M., Almekhlafi, M., Hahn, L., Demchuk, A.M., Goyal, M.: Multiphase ct angiography: a new tool for the imaging triage of patients with acute ischemic stroke. Radiology **275**(2), 510–520 (2015)

11. Qiu, W., Kuang, H., Ospel, J.M., Hill, M.D., Demchuk, A.M., Goyal, M., Menon, B.K.: Automated prediction of ischemic brain tissue fate from multiphase computed tomographic angiography in patients with acute ischemic stroke using machine learning. Journal of stroke **23**(2), 234–243 (2021)
12. Rguibi, Z., Hajami, A., Zitouni, D., Elqaraoui, A., Zourane, R., Bouajaj, Z.: Improving medical imaging with medical variation diffusion model: An analysis and evaluation. Journal of Imaging **9**(9), 171 (2023)
13. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 10684–10695 (2022)
14. Saharia, C., Chan, W., Chang, H., Lee, C., Ho, J., Salimans, T., Fleet, D., Norouzi, M.: Palette: Image-to-image diffusion models. In: ACM SIGGRAPH 2022 Conference Proceedings. pp. 1–10 (2022)
15. Saharia, C., Ho, J., Chan, W., Salimans, T., Fleet, D.J., Norouzi, M.: Image super-resolution via iterative refinement. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(4), 4713–4726 (2022)
16. Tan, Z., Parsons, M., Bivard, A., Sharma, G., Mitchell, P., Dowling, R., Bush, S., Churilov, L., Xu, A., Yan, B.: Comparison of computed tomography perfusion and multiphase computed tomography angiogram in predicting clinical outcomes in endovascular thrombectomy. Stroke **53**(9), 2926–2934 (2022)
17. Wang, C., Shi, Z., Yang, M., Huang, L., Fang, W., Jiang, L., Ding, J., Wang, H.: Deep learning-based identification of acute ischemic core and deficit from non-contrast ct and cta. Journal of Cerebral Blood Flow & Metabolism **41**(11), 3028–3038 (2021)
18. Wang, T., Lei, Y., Fu, Y., Wynne, J.F., Curran, W.J., Liu, T., Yang, X.: A review on medical imaging synthesis using deep learning and its clinical applications. Journal of applied clinical medical physics **22**(1), 11–36 (2021)
19. Whang, J., Delbracio, M., Talebi, H., Saharia, C., Dimakis, A.G., Milanfar, P.: Deblurring via stochastic refinement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 16293–16303 (2022)
20. Zhang, R., Isola, P., Efros, A.A., Shechtman, E., Wang, O.: The unreasonable effectiveness of deep features as a perceptual metric. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 586–595 (2018)
21. Zhao, S., Ren, H., Yuan, A., Song, J., Goodman, N., Ermon, S.: Bias and generalization in deep generative models: An empirical study. Advances in Neural Information Processing Systems **31** (2018)
22. Zhou, T., Fu, H., Chen, G., Shen, J., Shao, L.: Hi-net: hybrid-fusion network for multi-modal mr image synthesis. IEEE transactions on medical imaging **39**(9), 2772–2781 (2020)
23. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE international conference on computer vision. pp. 2223–2232 (2017)