HF-ResDiff: High-Frequency-guided Residual Diffusion for Multi-dose PET Reconstruction

Zixin Tang[1], Caiwen Jiang[1], Zhiming Cui[1], and Dinggang Shen[1,2,3,4(✉)]

[1] School of Biomedical Engineering
[2] State Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai 201210, China
dgshen@shanghaitech.edu.cn
[3] Shanghai United Imaging Intelligence Co. Ltd., Shanghai 200030, China
[4] Shanghai Clinical Research and Trial Center, Shanghai 201210, China

Abstract. Positron Emission Tomography (PET), an advanced nuclear imaging technology capable of visualizing human biological processes, plays an irreplaceable role in diagnosing various diseases. Nonetheless, PET imaging necessitates the administration of radionuclides into the human body, inevitably leading to radiation exposure. To mitigate the risk, many studies seek to reconstruct high-quality standard-dose PET from low-dose PET to reduce the required dosage of radionuclides. However, these efforts perform poorly in capturing high-frequency details in images. Meanwhile, they are limited to single-dose PET reconstruction, overlooking a clinical fact: *due to inherent individual variations among patients, the actual dose level of PET images obtained can exhibit considerable discrepancies.* In this paper, we propose a multi-dose PET reconstruction framework that aligns closely with clinical requirements and effectively preserves high-frequency information. Specifically, we design a **H**igh-**F**requency-guided **Res**idual **Diff**usion Model (**HF-ResDiff**) that enhances traditional diffusion models by 1) employing a simple CNN to predict low-frequency content, allowing the diffusion model to focus more on high-frequency counterparts while significantly promoting the training efficiency, 2) incorporating a Frequency Domain Information Separator and a High-frequency-guided Cross-attention to further assist the diffusion model in accurately recovering high-frequency details, and 3) embedding a Dose Control module to enable the diffusion model to accommodate PET reconstruction at different dose levels. Through extensive experiments, our HF-ResDiff outperforms the state-of-the-art methods in PET reconstruction across multiple doses.

Keywords: PET Reconstruction · High-frequency Restoration· Residual Diffusion Model · Discrete Wavelet Transform.

1 Introduction

Positron Emission Tomography (PET) is a sophisticated functional imaging modality, distinguished by its sensitive visualization of biological processes within

the human body [16]. Capitalizing on this advantage, PET plays a crucial role in diagnosing and early intervening in many diseases, such as Alzheimer's disease, cancers, and others [18]. Nevertheless, PET imaging requires the injection of a certain dosage of radioactive isotopes into the human body to obtain images that meet clinical diagnostic needs, which inevitably introduces radiation hazards [2]. To mitigate this predicament, a trade-off approach focusing on reconstructing high-quality standard-dose PET (SPET) from low-dose PET (LPET) is proposed to balance diagnostic efficacy with patient safety.

Over the past few decades, numerous PET enhancement algorithms have been explored. The earliest are based on filtering methods, such as non-local means [1], bilateral filters [10], and guided filters [25]. While these methods are notably straightforward, they are limited in performance and often result in overly smoothed images. With the advent of deep learning, a new generation of neural network-based PET enhancement algorithms are introduced [6, 19, 12]. Benefiting from the powerful neural networks, these methods often achieve commendable results. However, it is still challenging to reconstruct high-frequency details in PET images [13, 15]. More importantly, these end-to-end approaches generally focus only on single-dose PET reconstruction, which contradicts the clinical fact: *due to inherent individual variations among patients, the actual dose level of PET images obtained can exhibit considerable discrepancies.* These limitations restrict their application in real-world clinical settings.

Recently, the diffusion models have demonstrated immense success in image generation [9]. By deconstructing complex image mapping tasks into a series of denoising steps, diffusion models surpass other methods in generative capability. Nonetheless, directly applying diffusion models to PET reconstruction poses several significant challenges. First, unlike 2D photographic images, PET scans are three-dimensional, which markedly amplifies the already considerable computational demands of diffusion models. Second, the iterative noise addition and removal process may result in the erosion of high-frequency details, which is unacceptable in the context of PET imaging as the preservation of detailed information is imperative for accurate diagnostics [7, 14, 5].

To this end, inspired from [21], we propose the High-Frequency-guided Residual Diffusion Model (HF-ResDiff) adaptable for multi-dose PET reconstruction with an emphasis on high-frequency details. Specifically, we use a less computationally intensive CNN to recover low-frequency components with optimized efficiency. Consequently, the diffusion model is solely employed to predict the residual between the initial output of the CNN and the target SPET, i.e., the high-frequency counterparts. Meanwhile, we embed a Frequency Domain Information Separator (FD-Info Separator) and a High-frequency-guided Cross-attention (HF-guided CA) within the diffusion model, which isolate high-frequency information and thus better recover intricate details. Furthermore, we introduce a Dose Control module, enabling our HF-ResDiff to accommodate multi-dose PET reconstruction. Extensive experiments conducted on public datasets demonstrate that our method outperforms state-of-the-art approaches

both quantitatively and qualitatively, underscoring its promising potential for real-world clinical applications.

2 Methodology

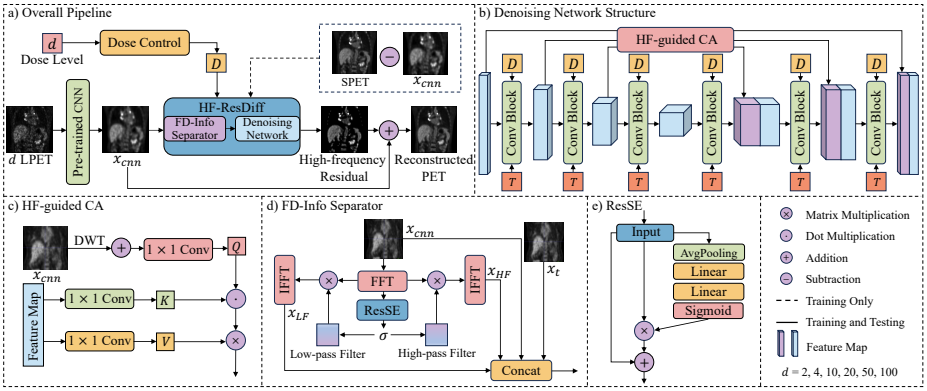

Fig. 1. Overview of the proposed method. a) Pipeline for both training and testing phases; b) Architecture of the denoising network of HF-ResDiff, with details of HF-guided CA depicted in c); and d) Mechanism of FD-Info Separator, with the structure of ResSE depicted in e).

The pipeline of our HF-ResDiff is illustrated in Fig. 1. Given an input LPET image x with a dose level of d, it is first passed through a pre-trained CNN to generate a low-frequency coarse prediction x_{cnn}. Subsequently, x_{cnn} is fed into the diffusion model, which is integrated with FD-Info Separator and HF-guided CA for further high-frequency restoration. Meanwhile, the Dose Control module extracts dosage-specific embedding D from the dose level d to control the diffusion model for tailored reconstruction. The final reconstructed PET is obtained by adding the low-frequency prediction and the high-frequency residual together. Below, we introduce the details of each component within HF-ResDiff.

2.1 Pre-train CNN for Predicting Low-Frequency Information

Predicting low-frequency information from images is much simpler than capturing high-frequency details. Deploying powerful diffusion models to recover low-frequency image information is inefficient, resulting in unnecessary computational expenditure. In light of this, in our HF-ResDiff framework, the diffusion model is solely employed to reconstruct the high-frequency content in PET images, while the initial prediction of low-frequency information is entrusted to the less computationally demanding CNN.

To maximize the capability of CNN in predicting both low-frequency and an extended range of high-frequency information, thereby alleviating the burden on the diffusion model, we design two specialized loss functions $\mathcal{L}_{FFT}$ and $\mathcal{L}_{DWT}$ based on common frequency processing operations, i.e., the Fast Fourier Transform (FFT) [4] and the Discrete Wavelet Transform (DWT) [17]. $\mathcal{L}_{FFT}$ calculates the mean squared error (MSE) between the magnitudes of the FFT coefficients derived from the ground truth (GT) and the prediction. This constraint targets at the consistency of overall frequency distribution. In comparison, $\mathcal{L}_{DWT}$ exclusively concentrates on the high-frequency spectrum. The DWT decomposes the GT and the prediction into eight distinct sub-bands (one low-frequency sub-band and seven high-frequency sub-bands). $\mathcal{L}_{DWT}$ only calculates the MSE between the high-frequency sub-bands of GT and prediction correspondingly. The final training loss $\mathcal{L}_{CNN}$ for CNN is formulated as follows:

$$\mathcal{L}_{CNN} = \mathcal{L}_{GT} + \alpha\mathcal{L}_{FFT} + \beta\mathcal{L}_{DWT}, \tag{1}$$

where α and β are hyperparameters used to adjust the weights. $\mathcal{L}_{GT}$ directly calculates the MSE between the GT and the prediction in the spatial domain. In this manner, we can pre-train a CNN to generate preliminary prediction x_{cnn} dominated by low-frequency contents but also containing partial high-frequency elements.

2.2 HF-ResDiff Model

Subsequently, x_{cnn} is input into the diffusion model to predict high-frequency residuals. Herein, we design the FD-Info Separator and HF-guided CA based on FFT and DWT to further assist the diffusion model in capturing high-frequency information.

FD-Info Separator After x_{cnn} is input into the diffusion model, the FD-Info Separator initially segregates it into high-frequency x_{HF} and low-frequency x_{LF} components, allowing the network to distinctly focus on reconstructing information across varying frequency bands. The detailed structure of the FD-Info Separator is shown in Fig. 1 d). For $x_{cnn} \in \mathbb{R}^{H\times W\times D\times C}$, we first perform FFT to obtain the frequency domain feature map M and then construct a high-pass filter and a low-pass filter, with their standard deviation σ calculated as:

$$\sigma = min(|ResSE(M) + \frac{1}{2}l|, l), \tag{2}$$

where $l = min(H, W, D)$. The ResSE (Residual Squeeze and Excitation) module, depicted in Fig. 1 e), is designed based on [8] and [11] to adaptively calculate the range of σ. With σ, the high-pass $H(u, v, k)$ and low-pass $L(u, v, k)$ filters are given as:

$$H(u, v, k) = 1 - e^{-P^2(u,v,k)/2\sigma^2}, L(u, v, k) = e^{-P^2(u,v,k)/2\sigma^2}, \tag{3}$$

where $P(u, v, k)$ is the distance from any point (u, v, k) in the frequency domain to the center point. We then obtain the filtered feature maps by multiplying M with $H(u, v, k)$ and $L(u, v, k)$, and perform inverse FFT (IFFT) on them to acquire x_{HF} and x_{LF} in the spatial domain.

FD-Info Separator allows for an explicit filtering of high and low frequency components based on FFT. The two resulted feature maps x_{HF} and x_{LF}, along with original x_{cnn} and current noisy image x_t, are all concatenated across the channel dimension and fed into the denoising network.

HF-guide CA The classic UNet architecture employs skip connections to fuse features across varying scopes [20], yet it falls short in adequately capturing high-frequency components. Moreover, due to noise interference, the boundary between high and low frequency components is not constant during the diffusion process. In such context, DWT is more suitable for frequency analysis as it serves as an adaptive segregation. Therefore, we integrate a HF-guided CA based on DWT to enhance the network's ability to utilize high-frequency information and synthesize fine-grained details.

The details of the HF-guided CA are shown in Fig. 1 c). Specifically, we first apply DWT on the input x_{cnn} and segregate the high-frequency sub-bands to generate Q through a 1×1 convolution. Concurrently, within each layer of the UNet, the downsampled feature map undergoes distinct linear transformations to produce K and V. The output M_{hf} can be obtained as follows:

$$M_{hf} = \text{Softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V, \tag{4}$$

where d_k is the number of columns of matrix Q. M_{hf} is then concatenated with the corresponding upsampled feature map on the decoder side, as depicted in Fig. 1 b).

2.3 Multi-dose PET Reconstruction

To ensure our HF-ResDiff model is adaptable for multi-dose PET reconstruction scenarios, we devise a Dose Control module to steer and regulate the diffusion model for varying doses. Similar to the time step embedding approach, we utilize the sinusoidal positional encoding [22] to encode the dose level d associated with the input LPET into a dose embedding D:

$$D = \begin{cases} \sin\left(d \times \exp\left(\frac{-\log(10000)\times 2i}{dim}\right)\right) & \text{if } i \text{ is even} \\ \cos\left(d \times \exp\left(\frac{-\log(10000)\times (2i-1)}{dim}\right)\right) & \text{if } i \text{ is odd} \end{cases}, \tag{5}$$

where dim is the dimension of the model, and i is the dimension index (ranging from 0 to dim - 1). The sine and cosine functions provide a unique encoding for each dose level, allowing the model to discern intricate differences in different dose levels. The resulting dose embedding D is then fed into each residual block of the denoising network along with the time step embedding T.

Table 1. Quantitative comparison of our HF-ResDiff with five state-of-the-art generation methods, in terms of PSNR [dB] and SSIM [%].

Method	DRF=50		DRF=10		DRF=2	
	SSIM	**PSNR**	**SSIM**	**PSNR**	**SSIM**	**PSNR**
3D-UNet [3]	65.02(7.17)	28.29(0.42)	78.35(3.89)	32.24(0.32)	81.51(3.39)	33.57(0.23)
3D-cGAN [24]	76.08(5.58)	30.66(0.54)	89.32(2.03)	35.88(0.54)	92.01(1.23)	37.86(0.55)
PET-Diff [5]	87.56(2.94)	32.77(0.55)	90.18(1.14)	36.97(0.75)	95.28(0.29)	40.58(0.58)
DISGAN [23]	83.08(5.51)	33.21(0.41)	93.49(0.82)	38.66(0.40)	94.94(0.78)	40.94(0.45)
Con-Diff [7]	88.19(1.96)	33.17(0.36)	90.60(2.29)	36.63(0.59)	96.57(0.33)	41.28(0.57)
Proposed	**89.83(1.92)**	**34.24(0.34)**	**93.71(0.90)**	**38.75(0.62)**	**97.20(0.21)**	**41.69(0.53)**

2.4 Implementation Details

Our HF-ResDiff is implemented using the PyTorch framework and trained on a Nvidia Tesla V100 GPU with 40GB memory. α and β in Equation 1 are initialized as 0.2 and 0.1 respectively. The number of time steps for the diffusion model is set to 2000 during training and 400 for testing. The training process lasts 40 epochs, using Adam optimizer with a learning rate of 0.0001. The quantitative results are evaluated based on two metrics, i.e., Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM).

3 Experiments

3.1 Dataset

Our dataset originates from Ultra Low Dose PET Imaging Challenge Dataset[5]. This dataset contains whole-body 18F-FDG PET imaging samples acquired from Siemens Biograph Vision Quadra and United Imaging uEXPLORER. We use a subset of 197 subjects to develop our method, with 157 subjects for training and the rest 40 for testing. LPET with dose reduction factors (DRFs) at 2, 4, 10, 20, 50, and 100 are reconstructed from the counts of a time window resampled at the middle of the acquisition, with time reduced correspondingly. In this study, we focus on the abdominal PET only, so the whole-body scans are cropped to the abdominal region, with a resolution of 256 $\times$ 256 $\times$ 128. To reduce the dependence on GPU memory, we extract the overlapped patches of size 64 $\times$ 64 $\times$ 64 from each cropped abdominal PET.

3.2 Comparative Experiments

We compare our proposed method with five state-of-the-art methods, which can be divided into three classes: 1) CNN-based method, including 3D U-Net [3]; 2) GAN-based method, including 3D-cGAN [24] and DISGAN [23]; and 3) Diffusion model-based method, including PET Diffusion (PET-Diff) [5] and

[5] https://ultra-low-dose-pet.grand-challenge.org/Dataset/

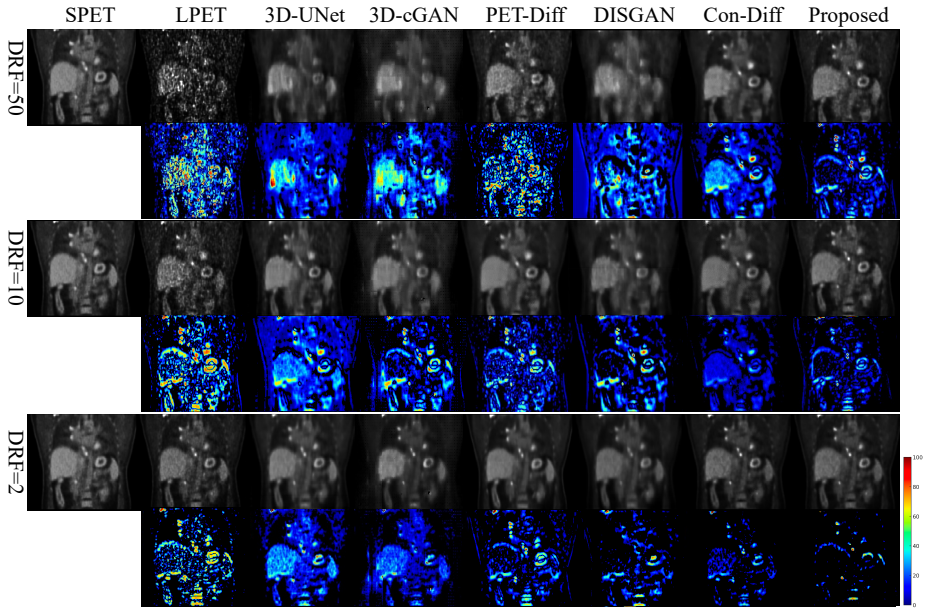

Fig. 2. Visual comparison of reconstructed PET images produced by six different methods. From left to right are the SPET, LPET, results of five other comparison methods (3rd-7th columns) and our HF-ResDiff (8th column). The corresponding error maps between the generated results and SPET are shown in the 2nd, 4th, and 6th rows.

Contrastive Diffusion (Con-Diff) [7]. All comparison methods, as well as our proposed method, are implemented using a consistent data processing approach. The quantitative and qualitative results are provided in Table 1 and Fig. 2, respectively.

Quantitative Comparison: The results in Table 1 show that our HF-ResDiff achieves the best overall performance and surpasses other diffusion model-based methods. It validates the efficacy of our enhancements to the traditional diffusion model. Moreover, an important observation is the relative stability of our HF-ResDiff model when the DRF increases. In contrast, competing methods tend to experience a pronounced decline in performance under similar conditions. This demonstrates that our HF-ResDiff is robust to PET reconstruction across varying doses, highlighting its potential for clinical application.

Qualitative Comparison: The PET images reconstructed under three different imaging doses are shown in Fig. 2. It is suggested that, across all doses, our HF-ResDiff can reconstruct PET images with the slightest noise but the clearest detailed texture. Further, the PET images reconstructed by our HF-ResDiff have the lightest color in the error maps, demonstrating their closest similarity to the SPET. All these observations validate the effectiveness of our proposed HF-ResDiff and its superior performance over the state-of-the-art approaches.

Table 2. Quantitative results of ablation study, in terms of PSNR [dB] and SSIM [%].

Method	DRF=50		DRF=10		DRF=2	
	SSIM	**PSNR**	**SSIM**	**PSNR**	**SSIM**	**PSNR**
Std-Diff	87.56(2.94)	32.77(0.55)	90.18(1.14)	36.97(0.75)	95.28(0.29)	40.58(0.58)
CNN-Diff	88.42(2.03)	33.28(0.38)	91.66(0.97)	37.60(0.67)	95.87(0.29)	40.61(0.61)
CNN-Diff-FD	89.19(1.96)	34.09(0.37)	92.96(0.93)	38.41(0.66)	96.68(0.23)	41.44(0.57)
Proposed	**89.83(1.92)**	**34.24(0.34)**	**93.71(0.90)**	**38.75(0.62)**	**97.20(0.21)**	**41.69(0.53)**

3.3 Ablation Study

To verify the effectiveness of each proposed strategy, we conduct a series of ablation studies, including: 1) Std-Diff: standard diffusion model; 2) CNN-Diff: adopting CNN for coarse prediction; 3) CNN-Diff-FD: CNN-Diff with frequency domain processing, i.e., FD-Info Separator and HF-guided CA; and 4) HF-ResDiff (proposed): CNN-Diff-FD with the Dose Control module. All methods use the same experimental settings, and their quantitative results are given in Table 2.

Statistics in Table 2 reveals several key insights into our study. First, the CNN-Diff model outperforms Std-Diff, highlighting the advantage of integrating a pre-trained CNN for low-frequency content prediction, which effectively alleviates the workload on the diffusion model and leads to improved outcomes. Second, the CNN-Diff-FD model, enriched with frequency domain processing, exceeds the performance of the CNN-Diff model and underscores the benefit of incorporating the FD-Info Separator and HF-guided CA. It proves that explicitly separating high-frequency components bolsters the diffusion model's focus on generating high-frequency details. Additionally, the HF-ResDiff model, equipped with the Dose Control module, demonstrates superior performance in scenarios of dose variation (e.g., when DRF is 10 and 2), validating the effectiveness of the Dose Control module. These observations collectively affirm the efficacy of the strategies implemented in our HF-ResDiff, each contributing positively to PET reconstruction.

4 Conclusion

In this study, we propose an innovative multi-dose universal PET reconstruction method to reduce radiation risks associated with PET imaging. We present the High-Frequency-guided Residual Diffusion Model (HF-ResDiff), which enhances traditional diffusion models through a series of targeted augmentations. First, it employs a pre-trained CNN to predict low-frequency components, thereby reducing computational burden on the diffusion model. Moreover, it leverages specialized frequency-domain processing approaches, namely the FD-Info Separator and HF-guided CA, to sharpen the model's focus on the intricate high-frequency details in PET images. Last, a Dose Control module is integrated, utilizing specific dose-derived embeddings to inform and guide the reconstruction process.

Extensive experiments on public datasets have demonstrated that our method outperforms existing state-of-the-art techniques, in both quantitative measurements and qualitative assessments.

Acknowledgments. This work was supported in part by National Natural Science Foundation of China (grant numbers 62131015, 62250710165, U23A20295), the STI 2030-Major Projects (No. 2022ZD0209000), Shanghai Municipal Central Guided Local Science and Technology Development Fund (grant number YDZX20233100001001), Science and Technology Commission of Shanghai Municipality (STCSM) (grant number 21010502600), and The Key R&D Program of Guangdong Province, China (grant numbers 2023B0303040001, 2021B0101420006).

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

References

1. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. In: IEEE Computer Society Conference on Computer Vision and Pattern Recognition. pp. 60–65. IEEE (2005)
2. Chawla, S.C., Federman, N., Zhang, D., Nagata, K., Nuthakki, S., McNitt-Gray, M., Boechat, M.I.: Estimated cumulative radiation dose from PET/CT in children with malignancies: a 5-year retrospective review. Pediatric Radiology **40**, 681–686 (2010)
3. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: Medical Image Computing and Computer-Assisted Intervention. pp. 424–432. Springer (2016)
4. Cooley, J.W., Tukey, J.W.: An algorithm for the machine calculation of complex Fourier series. Mathematics of Computation **19**, 297–301 (1965)
5. Gong, K., Johnson, K., El Fakhri, G., Li, Q., Pan, T.: PET image denoising based on denoising diffusion probabilistic model. European Journal of Nuclear Medicine and Molecular Imaging **51**, 358–368 (2024)
6. Häggström, I., Schmidtlein, C.R., Campanella, G., Fuchs, T.J.: DeepPET: A deep encoder–decoder network for directly solving the PET image reconstruction inverse problem. Medical Image Analysis **54**, 253–262 (2019)
7. Han, Z., Wang, Y., Zhou, L., Wang, P., Yan, B., Zhou, J., Wang, Y., Shen, D.: Contrastive diffusion model with auxiliary guidance for coarse-to-fine PET reconstruction. In: Medical Image Computing and Computer-Assisted Intervention. pp. 239–249. Springer (2023)
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778. IEEE (2016)
9. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in Neural Information Processing Systems **33**, 6840–6851 (2020)
10. Hofheinz, F., Langner, J., Beuthien-Baumann, B., Oehme, L., Steinbach, J., Kotzerke, J., van den Hoff, J.: Suitability of bilateral filtering for edge-preserving noise reduction in PET. EJNMMI Research **1**(1), 23 (2011)

11. Hu, J., Shen, L., Sun, G.: Squeeze-and-excitation networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 7132–7141. IEEE (2018)
12. Jiang, C., Pan, Y., Cui, Z., Nie, D., Shen, D.: Semi-supervised standard-dose PET image generation via region-adaptive normalization and structural consistency constraint. IEEE Transactions on Medical Imaging **42**(10), 2974–2987 (2023)
13. Jiang, C., Pan, Y., Cui, Z., Shen, D.: Reconstruction of standard-dose PET from low-dose PET via dual-frequency supervision and global aggregation module. In: 2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2022)
14. Jiang, C., Pan, Y., Liu, M., Ma, L., Zhang, X., Liu, J., Xiong, X., Shen, D.: PET-diffusion: Unsupervised PET enhancement based on the latent diffusion model. In: Medical Image Computing and Computer-Assisted Intervention. pp. 3–12. Springer (2023)
15. Jiang, C., Pan, Y., Shen, D.: TriDoRnet: Reconstruction of standard-dose PET from low-dose PET in triple (projection, image, and frequency) domains. In: 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI). pp. 1–5. IEEE (2023)
16. Karnabi, E.: Positron emission tomography. In: Cardiology Procedures: A Clinical Primer, pp. 77–88. Springer (2022)
17. Mallat, S., Hwang, W.L.: Singularity detection and processing with wavelets. IEEE Transactions on Information Theory **38**(2), 617–643 (1992)
18. Nordberg, A., Rinne, J.O., Kadir, A., Långström, B.: The use of PET in Alzheimer disease. Nature Reviews Neurology **6**(2), 78–87 (2010)
19. Reader, A.J., Corda, G., Mehranian, A., da Costa-Luis, C., Ellis, S., Schnabel, J.A.: Deep learning for PET image reconstruction. IEEE Transactions on Radiation and Plasma Medical Sciences **5**(1), 1–25 (2020)
20. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention. pp. 234–241. Springer (2015)
21. Shang, S., Shan, Z., Liu, G., Zhang, J.: Resdiff: Combining cnn and diffusion model for image super-resolution. arXiv preprint arXiv:2303.08714 (2023)
22. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in Neural Information Processing Systems **30** (2017)
23. Wang, Q., Mahler, L., Steiglechner, J., Birk, F., Scheffler, K., Lohmann, G.: DIS-GAN: Wavelet-informed discriminator guides GAN to MRI super-resolution with noise cleaning. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 2452–2461. IEEE (2023)
24. Wang, Y., Yu, B., Wang, L., Zu, C., Lalush, D.S., Lin, W., Wu, X., Zhou, J., Shen, D., Zhou, L.: 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. Neuroimage **174**, 550–562 (2018)
25. Yan, J., Lim, J.C.S., Townsend, D.W.: Mri-guided brain PET image filtering and partial volume correction. Physics in Medicine & Biology **60**(3), 961 (2015)