PET Image Denoising Based on 3D Denoising Diffusion Probabilistic Model: Evaluations on Total-Body Datasets

Boxiao Yu[1], Savas Ozdemir[2], Yafei Dong[3], Wei Shao[4], Kuangyu Shi[5], and Kuang Gong[1⋆]

[1] J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida, Gainesville, FL, USA
[2] Department of Radiology, University of Florida, Jacksonville, FL, USA
[3] Yale School of Medicine, Yale University, New Haven, CT, USA
[4] Department of Medicine, University of Florida, Gainesville, FL, United States
[5] Department of Nuclear Medicine, University of Bern, Bern, Switzerland

Abstract. Due to various physical degradation factors and limited photon counts detected, obtaining high-quality images from low-dose Positron emission tomography (PET) scans is challenging. The Denoising Diffusion Probabilistic Model (DDPM), an advanced distribution learning-based generative model, has shown promising performance across various computer-vision tasks. However, currently DDPM is mainly investigated in 2D mode, which has limitations for PET image denoising, as PET is usually acquired, reconstructed, and analyzed in 3D mode. In this work, we proposed a 3D DDPM method for PET image denoising, which employed a 3D convolutional network to train the score function, enabling the network to learn 3D distribution. The total-body ^{18}F-FDG PET datasets acquired from the Siemens Biograph Vision Quadra scanner (axial field of view >1m) were employed to evaluate the 3D DDPM method, as these total-body datasets needed 3D operations the most to leverage the rich information from different axial slices. All models were trained on 1/20 low-dose images and then evaluated on 1/4, 1/20, and 1/50 low-dose images, respectively. Experimental results indicated that 3D DDPM significantly outperformed 2D DDPM and 3D UNet in qualitative and quantitative assessments, capable of recovering finer structures and more accurate edge contours from low-quality PET images. Moreover, 3D DDPM revealed greater robustness when there were noise level mismatches between training and testing data. Finally, comparing 3D DDPM with 2D DDPM in terms of uncertainty revealed 3D DDPM's higher confidence in reproducibility.

Keywords: PET image denoising · Diffusion models · Low-dose PET

⋆ Corresponding author: kgong@bme.ufl.edu

1 Introduction

Positron Emission Tomography (PET) is highly valued for disease management of cancer [20], neurodegenerative diseases [1], and cardiac diseases [5] due to its high sensitivity and precise quantification capabilities. Concerns regarding radiation exposure and potential cancer risk [19, 10] necessitate the reduction of PET injection dose. This reduction, however, leads to a deterioration of PET image quality, which can comprise its quantitative accuracy and lesion detectability [24]. Therefore, restoring high-quality PET images from low-dose PET datasets emerges as a crucial challenge.

With the availability of extensive training data and substantial computational resources, large-scale neural network training becomes feasible, leading to rapid developments in deep learning-based PET image denoising. Convolutional Neural Networks (CNNs) show their superiority in various PET image denoising evaluations [12, 3, 14, 21]. The integration of skip connections [9], residual learning [30, 18], and perceptual loss [7] has further improved CNN's performance for PET image denoising. However, CNNs can sometimes produce overly smoothed results that can overlook lesions or pathology changes. Generative Adversarial Networks (GANs), with their novel adversarial loss item, can generate results with less spatial blurring and better visual quality [27, 27, 6], even at scenarios without utilizing paired training data [17, 31]. However, they may suffer from instability during the adversarial training process, which needs careful hyperparameter tuning.

In recent years, diffusion models have surpassed GANs as the leading image generation model [4]. The Denoising Diffusion Probabilistic Model (DDPM) [11], an advanced distribution learning-based model, has shown promising performance across various medical image processing tasks [15, 23, 32]. In the forward diffusion phase of DDPM, random Gaussian noise is gradually injected into the image to perturb the input data, and then the model learns to reverse this diffusion process to construct the desired data samples from the noise. Current diffusion model-based denoising methods are mainly focusing on 2D mode[13, 8, 16], limiting their application in PET imaging as PET is inherently a 3D imaging modality. Although some works proposed to include adjacent slices into the network training to address the issue of 3D inconsistency[29], the network used to generate the score function still relies on 2D convolutional networks, hindering the ability to perceive 3D structural information from PET volumes.

In this work, we proposed a 3D DDPM method for total-body PET image denoising. A 3D convolutional network with residual blocks was implemented to generate the score function, allowing the model to learn the 3D distribution information of PET. Experimental results based on total-body ^{18}F-FDG PET datasets of different dose levels indicated that the proposed 3D DDPM method significantly outperformed 2D DDPM and 3D UNet in both qualitative and quantitative aspects, demonstrating its superior robustness. Moreover, we explored the advantage of 3D DDPM over 2D DDPM in terms of uncertainty, where 3D DDPM model exhibited higher confidence in reproducibility.

2 Methodology

2.1 Diffusion Models

Given a noise-free data point $\mathbf{x}_0$ sampled from the real data distribution $q(\mathbf{x})$, the forward process of the diffusion model [25, 11] is defined as a Markov process, where Gaussian noise is gradually added to $\mathbf{x}_0$ in accordance with a variance schedule $\{\beta_t \in (0,1)\}_{t=1}^{T}$, over a large enough number of time steps T, which can be expressed as:

$$q(\mathbf{x}_{1:T}|\mathbf{x}_0) = \prod_{t=1}^{T} q(\mathbf{x}_t|\mathbf{x}_{t-1}), \qquad q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}). \quad (1)$$

Based on the property that $\mathbf{x}_t$ can be sampled at any arbitrary time step in a closed form, the forward process can be further expressed with the notation $\alpha_t = 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{s=1}^{t} \alpha_s$:

$$q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I}) \quad (2)$$

If the above process can be reversed, we can sample from $q(\mathbf{x}_{t-1}|\mathbf{x}_t)$ and reconstruct the clean data from a Gaussian noise $\mathbf{x}_T \sim \mathcal{N}(0, \mathbf{I})$. When β_t is small enough, $q(\mathbf{x}_{t-1}|\mathbf{x}_t)$ also follows Gaussian distribution and is tractable when conditioned on $\mathbf{x}_0$:

$$q(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{x}_0) = \mathcal{N}(\mathbf{x}_{t-1}; \tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0), \frac{1-\bar{\alpha}_{t-1}}{1-\bar{\alpha}_t}\beta_t\mathbf{I}), \quad (3)$$

$$\text{where} \quad \tilde{\boldsymbol{\mu}}_t(\mathbf{x}_t, \mathbf{x}_0) = \frac{\sqrt{\bar{\alpha}_{t-1}}\beta_t}{1-\bar{\alpha}_t}\mathbf{x}_0 + \frac{\sqrt{\alpha_t}(1-\bar{\alpha}_{t-1})}{1-\bar{\alpha}_t}\mathbf{x}_t. \quad (4)$$

However, as $\mathbf{x}_0$ remains unknown during the sampling process, it is requisite to learn a model p_θ to approximate these conditional probabilities for executing the reverse diffusion process:

$$p_\theta(\mathbf{x}_{0:T}) = p(\mathbf{x}_T) \prod_{t=1}^{T} p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t), \quad (5)$$

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, t), \boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t)), \quad (6)$$

where $\boldsymbol{\mu}_\theta$ and $\boldsymbol{\Sigma}_\theta$ denote the predicted mean and variance, respectively, with the input $\mathbf{x}_t$ and t through a network with parameter θ. Following the prior works [11, 22], we can then parameterize $\boldsymbol{\mu}_\theta$ to predict the noise $\boldsymbol{\epsilon}_\theta$ instead of the mean value:

$$\boldsymbol{\mu}_\theta(\mathbf{x}_t, t) = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right) \quad (7)$$

Hence, $\mathbf{x}_{t-1}$ can be computed as follows:

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)\right) + \sigma_t\mathbf{z}, \quad \text{where } \mathbf{z} \sim \mathcal{N}(0, \mathbf{I}) \quad (8)$$

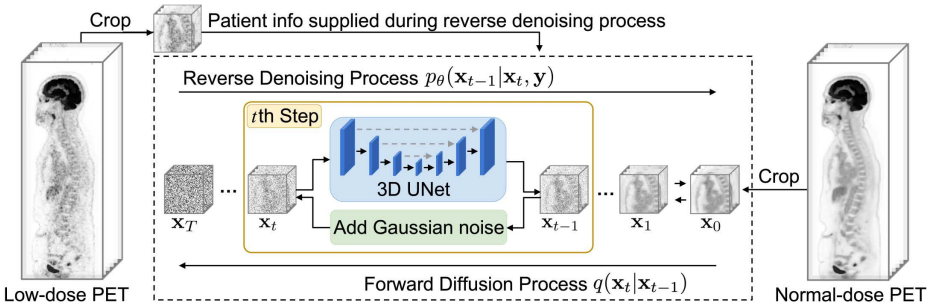

Fig. 1. Diagram of the proposed 3D DDPM framework.

2.2 Conditional PET Image Denoising Based on 3D DDPM

The above DDPM framework is employed for unconditional image generation. To facilitate PET image denoising, a noisy PET image is required to be input into the network to guide the image generation, rather than the random generation of new samples. Fig. 1 shows the diagram of the proposed 3D DDPM. The network's input is transformed to paired of predicted high-quality PET image $\mathbf{x}$ and original low-quality PET image $\mathbf{y}$, with the score function $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, t)$ modified to $\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{y}, t)$. Additionally, learnable parameter $\boldsymbol{\Sigma}_\theta(\mathbf{x}_t, t)$ can be set to untrained time dependent constants $\sigma_t^2\mathbf{I}$ [11]. Consequently, Eq. 6, 7, and 8 can be reformulated as:

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t, \mathbf{y}) = \mathcal{N}(\mathbf{x}_{t-1}; \boldsymbol{\mu}_\theta(\mathbf{x}_t, \mathbf{y}, t), \sigma_t^2\mathbf{I}) \tag{9}$$

$$\boldsymbol{\mu}_\theta(\mathbf{x}_t, \mathbf{y}, t) = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{y}, t)\right) \tag{10}$$

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\mathbf{x}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{y}, t)\right) + \sigma_t\mathbf{z}, \tag{11}$$

For the score-function optimization, the objective utilized in the training processing is therefore expressed as:

$$L(\theta) = \mathbb{E}_{t,\mathbf{x}_0,\boldsymbol{\epsilon}}\left[\|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta(\mathbf{x}_t, \mathbf{y}, t)\|^2\right]. \tag{12}$$

As this is the feasibility study testing 3D DDPM for PET image denoising, we adopted the most widely used 3D UNet architecture as the 3D DDPM's backbone network to recover clean data from noisy inputs. We expect better performance can be achieved if more advanced 3D network is employed as the backbone network.

3 Experiments and Results

3.1 Dataset and Implementation details

The Siemens Biograph Vision Quadra data from the Ultra-low Dose PET Imaging Challenge was utilized in our experiments, which comprised 377 total-body ^{18}F-FDG PET datasets. Fig.2 illustrates details of the data employed in our experiments along with visual examples of different dose levels. To enhance data processing efficiency, parts of background areas and the patients' legs were cropped, resulting in an image matrix size of $192 \times 288 \times 520$, corresponding to dimensions in the coronal, sagittal, and axial directions, respectively. The voxel size was $1.65 \times 1.65 \times 1.65$ mm^3. The image intensity was expressed in Standardized Uptake Value (SUV). The 1/20 low-dose and normal-dose PET data were selected as the low- and high-quality paired data for the model training and evaluation, which were randomly divided into training data (302 datasets), validation data (15 datasets), and test data (60 datasets). Additionally, the 1/4 and 1/50 low-dose PET data from the test data pool were further evaluated by the model to assess the robustness of different methods to discrepancies in noise levels between the training and test datasets.

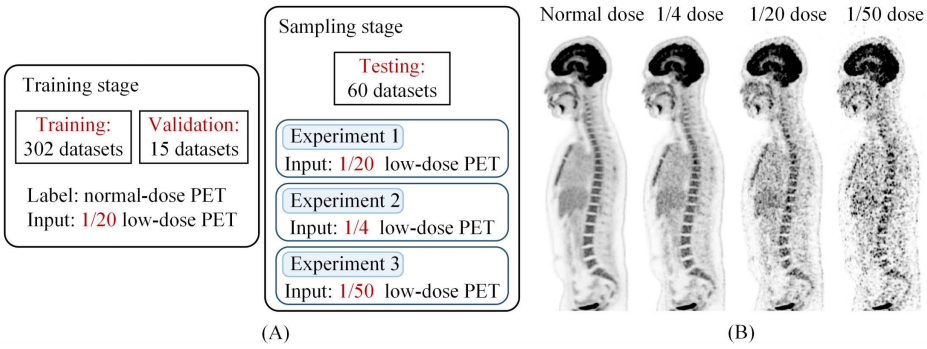

Fig. 2. (A) Details about the datasets used in the training, validation, and test scenarios. (B) Sagittal views of one example normal-dose, 1/4, 1/20, and 1/50 low-dose PET data from the Siemens Biograph Vision Quadra total-body ^{18}F-FDG PET datasets.

Considering the available GPU memory, the total-body PET data were randomly cropped into patches of size $96 \times 96 \times 96$ for network input during training. During the sampling stage, the total-body PET was first chunked into six segments along the axial direction with an overlap of 10 voxels, then fed sequentially into the 3D model, and finally combined together to generate the final total-body PET image with a weighted arithmetic mean applied to the overlapping parts. The time step T was set to 1000 throughout the forward and reverse diffusion processes, and a linear noise schedule was implemented. The training batch size was set to 8, utilizing 8 NVIDIA A100 GPUs for distributed

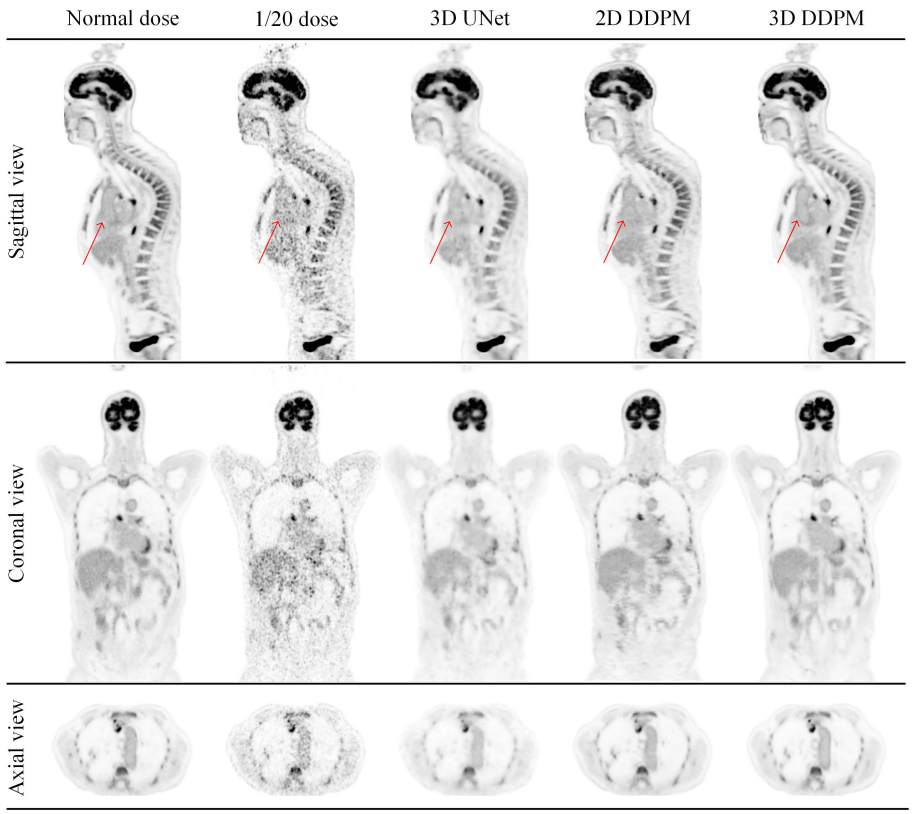

Fig. 3. Three views of one test dataset denoised using different methods, along with the normal-dose PET image and 1/20 low-dose PET image.

training. The training time was approximately nine days, and the testing time was on average 28 minutes per test dataset.

The 2D DDPM [4] and 3D UNet [2]-based image denoising were employed as reference methods. The 2D DDPM network took an additional two neighboring axial slices as the input to avoid axial artifacts, inspired by the work of Gong et al. [8] The input patch size for the 3D UNet was aligned with 3D DDPM to ensure a fair comparison. All methods were implemented using PyTorch; the learning rate was set to 1×10^{-4}, and mixed precision training techniques were applied to enhance computational efficiency.

3.2 Evaluation Metrics

To quantitatively evaluate the denoising results of various methods, the Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) [28] were adopted, utilizing normal-dose PET images $\mathbf{y}$ as the ground truth. PSNR

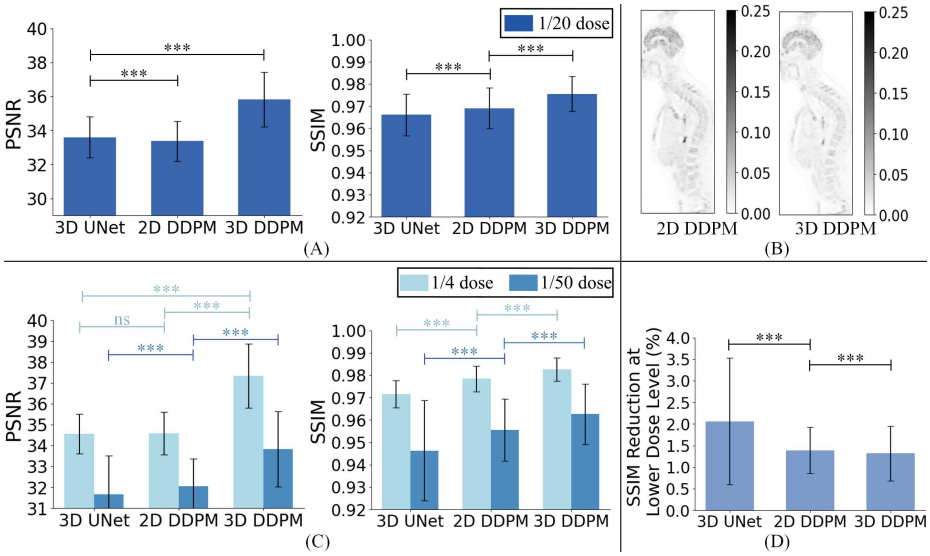

Fig. 4. (A) The PSNR and SSIM values calculated based on 60 1/20 low-dose test datasets. (B) Sagittal view of the uncertainty maps of one test data calculated from 20 realizations for 2D DDPM and 3D DDPM.(C) The PSNR and SSIM values calculated based on 60 1/4 and 1/50 low-dose test datasets, respectively. (D) The SSIM percentage reduction when the dose level of 60 test datasets decreased from 1/20 to 1/50. ***, ns located at the top of the bar plot represents p-value <0.001, p-value >0.05, respectively.

approximated the quality of reconstruction for denoised low-dose PET images $\mathbf{x}$ in comparison to $\mathbf{y}$ based on their Mean Squared Error (MSE) calculated as :

$$PSNR(\mathbf{x}, \mathbf{y}) = 10 \cdot \log_{10}[MAX_I / MSE(\mathbf{x}, \mathbf{y})], \tag{13}$$

where MAX_I represented the maximum possible pixel value of the images. SSIM compared $\mathbf{x}$ and $\mathbf{y}$ based on three properties: luminance, contrast, and structure, and was defined as:

$$SSIM(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_{\mathbf{x}}\mu_{\mathbf{y}} + C_1)(2\sigma_{\mathbf{xy}} + C_2)}{(\mu_{\mathbf{x}}^2 + \mu_{\mathbf{y}}^2 + C_1)(\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{y}}^2 + C_2)}, \tag{14}$$

where $\mu_{\mathbf{x}}(\mu_{\mathbf{y}})$ denoted the mean, $\sigma_{\mathbf{x}}^2(\sigma_{\mathbf{y}}^2)$ was the variance of $\mathbf{x}(\mathbf{y})$, and $\sigma_{\mathbf{xy}}$ denoted the cross-correlation between $\mathbf{x}$ and $\mathbf{y}$. C_1, C_2, and C_3 were small constants that ensure numeric stability, which were set based on the default settings in scikit-image [26] metrics. The Wilcoxon signed-rank tests for PSNR and SSIM were performed to ascertain whether there were statistically significant differences between the different methods.

3.3 Results

The qualitative evaluation results of the proposed 3D DDPM and other reference methods are shown in Fig. 3, which indicated that the results produced by 3D

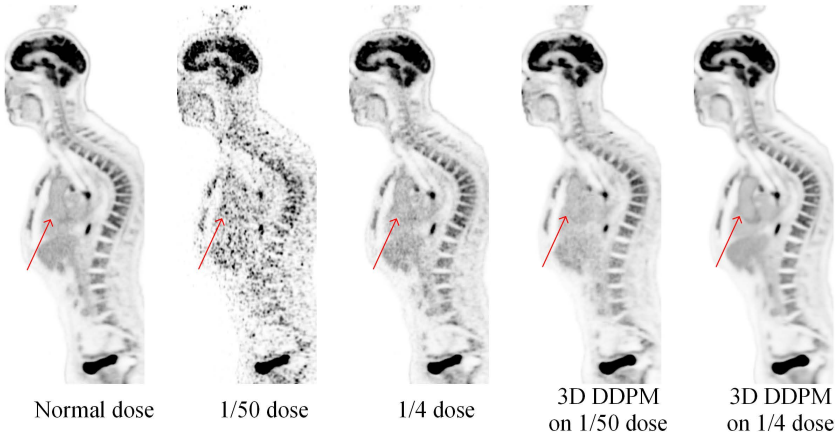

Fig. 5. Sagittal view of 1/4 and 1/50 low-dose PET images and the corresponding denoising results using the proposed 3D DDPM, along with the normal-dose PET.

UNet were overly smooth. In contrast, 2D DDPM and 3D DDPM generated more realistic denoising results. 3D DDPM significantly surpassed both 2D DDPM and 3D UNet across various body organs, revealing better structural details and more precise edge contours. The quantitative results shown in Fig. 4(A) matched with our qualitative observations. 3D DDPM achieved significantly better performance in both PSNR and SSIM than the other two methods. One advantage of DDPM is the ability to generate the uncertainty map. Fig. 4(B) presents the uncertainty map of the 3D DDPM results, generated using 20 different random seeds, alongside the 2D DDPM for comparison. The results indicate that the 3D DDPM exhibits lower uncertainty, especially in the lesion regions, compared to 2D DDPM. This demonstrates the higher reliability of 3D DDPM in quantitative characteristics, reflecting its high confidence in reproducibility.

Furthermore, all methods were directly applied to data with varying noise levels without any fine-tuning. The 3D DDPM achieved the best outcomes across all evaluation metrics (as shown in Fig. 4(C)), showcasing the superiority on data with higher or lower doses than the training dataset, thus demonstrating its excellent generalization performance. Notably, when the test data dose was reduced from 1/20 to 1/50, the 3D DDPM exhibited the minimal decline in SSIM (as shown in Fig. 4(D)), indicating the robustness of the 3D DDPM approach. The qualitative results (as shown in Fig. 5) revealed that the 3D DDPM consistently produced promising denoising results regardless of the PET image dose level. As the dose increased, the denoising results were further improved; when the dose decreased, features nearly invisible in low-dose images were effectively recovered.

4 Conclusion

In this work, we proposed a 3D DDPM framework to enhance the quality of low-dose total-body PET images. Our qualitative and quantitative results indicated the superiority of the proposed 3D DDPM over 2D DDPM and 3D UNet methods, demonstrating the advantage of extending diffusion models to 3D mode for total-body PET image denoising. Furthermore, our method exhibits strong robustness, which consistently conducted promising denoising results across PET images with varying noise levels, indicating its substantial practical significance and potential for clinical applications.

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

References

1. Barthel, H., Schroeter, M.L., Hoffmann, K.T., Sabri, O.: Pet/mr in dementia and other neurodegenerative diseases. In: Seminars in nuclear medicine. vol. 45, pp. 224–233. Elsevier (2015)
2. 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–MICCAI 2016: 19th International Conference, Athens, Greece, October 17-21, 2016, Proceedings, Part II 19. pp. 424–432. Springer (2016)
3. Cui, J., Gong, K., Guo, N., Wu, C., Meng, X., Kim, K., Zheng, K., Wu, Z., Fu, L., Xu, B., et al.: Pet image denoising using unsupervised deep learning. European journal of nuclear medicine and molecular imaging **46**, 2780–2789 (2019)
4. Dhariwal, P., Nichol, A.: Diffusion models beat gans on image synthesis. Advances in neural information processing systems **34**, 8780–8794 (2021)
5. Dorbala, S., Di Carli, M.F.: Cardiac pet perfusion: prognosis, risk stratification, and clinical management. In: Seminars in nuclear medicine. vol. 44, pp. 344–357. Elsevier (2014)
6. Fu, Y., Dong, S., Niu, M., Xue, L., Guo, H., Huang, Y., Xu, Y., Yu, T., Shi, K., Yang, Q., et al.: Aigan: Attention–encoding integrated generative adversarial network for the reconstruction of low-dose ct and low-dose pet images. Medical Image Analysis **86**, 102787 (2023)
7. Gong, K., Guan, J., Liu, C.C., Qi, J.: Pet image denoising using a deep neural network through fine tuning. IEEE Transactions on Radiation and Plasma Medical Sciences **3**(2), 153–161 (2018)
8. 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 pp. 1–11 (2023)
9. Hashimoto, F., Ohba, H., Ote, K., Teramoto, A., Tsukada, H.: Dynamic pet image denoising using deep convolutional neural networks without prior training datasets. IEEE access **7**, 96594–96603 (2019)
10. Hashimoto, F., Onishi, Y., Ote, K., Tashima, H., Reader, A.J., Yamaya, T.: Deep learning-based pet image denoising and reconstruction: a review. Radiological Physics and Technology pp. 1–23 (2024)

11. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. Advances in neural information processing systems **33**, 6840–6851 (2020)
12. Jaudet, C., Weyts, K., Lechervy, A., Batalla, A., Bardet, S., Corroyer-Dulmont, A.: The impact of artificial intelligence cnn based denoising on fdg pet radiomics. Frontiers in oncology p. 3136 (2021)
13. 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: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 3–12. Springer (2023)
14. Kaplan, S., Zhu, Y.M.: Full-dose pet image estimation from low-dose pet image using deep learning: a pilot study. Journal of digital imaging **32**(5), 773–778 (2019)
15. 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)
16. Lee, S., Chung, H., Park, M., Park, J., Ryu, W.S., Ye, J.C.: Improving 3d imaging with pre-trained perpendicular 2d diffusion models. arXiv preprint arXiv:2303.08440 (2023)
17. Lei, Y., Dong, X., Wang, T., Higgins, K., Liu, T., Curran, W.J., Mao, H., Nye, J.A., Yang, X.: Whole-body pet estimation from low count statistics using cycle-consistent generative adversarial networks. Physics in Medicine & Biology **64**(21), 215017 (2019)
18. Lv, Y., Xi, C.: Pet image reconstruction with deep progressive learning. Physics in Medicine & Biology **66**(10), 105016 (2021)
19. Miglioretti, D.L., Johnson, E., Williams, A., Greenlee, R.T., Weinmann, S., Solberg, L.I., Feigelson, H.S., Roblin, D., Flynn, M.J., Vanneman, N., et al.: The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. JAMA pediatrics **167**(8), 700–707 (2013)
20. Ming, Y., Wu, N., Qian, T., Li, X., Wan, D.Q., Li, C., Li, Y., Wu, Z., Wang, X., Liu, J., et al.: Progress and future trends in pet/ct and pet/mri molecular imaging approaches for breast cancer. Frontiers in oncology **10**, 1301 (2020)
21. Nerella, S., Bandyopadhyay, S., Zhang, J., Contreras, M., Siegel, S., Bumin, A., Silva, B., Sena, J., Shickel, B., Bihorac, A., et al.: Transformers in healthcare: A survey. arXiv preprint arXiv:2307.00067 (2023)
22. Nichol, A.Q., Dhariwal, P.: Improved denoising diffusion probabilistic models. In: International Conference on Machine Learning. pp. 8162–8171. PMLR (2021)
23. Özbey, M., Dalmaz, O., Dar, S.U., Bedel, H.A., Özturk, Ş., Güngör, A., Çukur, T.: Unsupervised medical image translation with adversarial diffusion models. IEEE Transactions on Medical Imaging (2023)
24. Schaefferkoetter, J.D., Yan, J., Sjöholm, T., Townsend, D.W., Conti, M., Tam, J.K.C., Soo, R.A., Tham, I.: Quantitative accuracy and lesion detectability of low-dose 18f-fdg pet for lung cancer screening. Journal of Nuclear Medicine **58**(3), 399–405 (2017)
25. Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., Ganguli, S.: Deep unsupervised learning using nonequilibrium thermodynamics. In: International conference on machine learning. pp. 2256–2265. PMLR (2015)
26. Van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T.: scikit-image: image processing in python. PeerJ **2**, e453 (2014)
27. 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)

28. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing **13**(4), 600–612 (2004)
29. Xie, H., Gan, W., Zhou, B., Chen, X., Liu, Q., Guo, X., Guo, L., An, H., Kamilov, U.S., Wang, G., et al.: Dose-aware diffusion model for 3d ultra low-dose pet imaging. arXiv preprint arXiv:2311.04248 (2023)
30. Xue, H., Teng, Y., Tie, C., Wan, Q., Wu, J., Li, M., Liang, G., Liang, D., Liu, X., Zheng, H., et al.: A 3d attention residual encoder–decoder least-square gan for low-count pet denoising. Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **983**, 164638 (2020)
31. Zhou, L., Schaefferkoetter, J.D., Tham, I.W., Huang, G., Yan, J.: Supervised learning with cyclegan for low-dose fdg pet image denoising. Medical image analysis **65**, 101770 (2020)
32. Zhu, R., Li, X., Zhang, X., Ma, M.: Mri and ct medical image fusion based on synchronized-anisotropic diffusion model. IEEE Access **8**, 91336–91350 (2020)