

Supervised Diffusion-Model-Based PET Image Reconstruction

George Webber (✉)¹[0009–0003–2573–7690], Alexander Hammers^{1,2}[0000–0001–9530–4848], Andrew P. King¹[0000–0002–9965–7015], and Andrew J. Reader¹[0000–0002–2726–3383]

¹ School of Biomedical Engineering and Imaging Sciences, King’s College London, London, UK

{firstname}.{lastname}@kcl.ac.uk

² Guy’s & St Thomas’ PET Centre, St Thomas’ Hospital, London, UK

Abstract. Diffusion models (DMs) have recently been introduced as a regularizing prior for PET image reconstruction, integrating DMs trained on high-quality PET images with unsupervised schemes that condition on measured data. While these approaches have potential generalization advantages due to their independence from the scanner geometry and the injected activity level, they forgo the opportunity to explicitly model the interaction between the DM prior and noisy measurement data, potentially limiting reconstruction accuracy. To address this, we propose a supervised DM-based algorithm for PET reconstruction. Our method enforces the non-negativity of PET’s Poisson likelihood model and accommodates the wide intensity range of PET images. Through experiments on realistic brain PET phantoms, we demonstrate that our approach outperforms or matches state-of-the-art deep learning-based methods quantitatively across a range of dose levels. We further conduct ablation studies to demonstrate the benefits of the proposed components in our model, as well as its dependence on training data, parameter count, and number of diffusion steps. Additionally, we show that our approach enables more accurate posterior sampling than unsupervised DM-based methods, suggesting improved uncertainty estimation. Finally, we extend our methodology to a practical approach for fully 3D PET and present example results from real [¹⁸F]FDG brain PET data.

Keywords: PET Image Reconstruction · Diffusion Models · Inverse Problems · Score-Based Generative Modeling.

1 Introduction

Positron emission tomography (PET) is a widely used nuclear medicine technique for quantitatively imaging functional processes within the body [19]. Clinicians and researchers rely on PET imaging to measure disease biomarkers, guide treatment decisions, and study physiological processes [1].

Reducing the radioactive counts used for PET images is desirable to minimize radiation exposure and/or reduce scan times [12]. However, reconstructing PET

images from limited radioactive counts is an ill-posed inverse problem due to the high-variance Poisson noise inherent in PET measurement data. State-of-the-art reconstruction algorithms address this issue by conditioning on image-based priors, which can be learned [11, 6] or hand-crafted [25]. These priors, combined with measured data, define a posterior distribution of possible reconstructed images [5]. The posterior distribution reflects the uncertainty of the observed data, which is important in medical image reconstruction to quantify the reliability of image features and structures that may impact diagnosis [4]. However, most PET reconstruction methods only estimate the posterior mean [15].

Recent work has leveraged diffusion models (DMs) [21, 7] for PET reconstruction [20, 27, 8, 26, 9, 28]. As the current state-of-the-art in generative image modeling [2], DM-based approaches offer two key advantages: 1) the potential to surpass the image quality of other deep-learned (DL) reconstruction approaches and 2) to fully model the posterior distribution of possible reconstructions from noisy data.

Previous approaches have integrated DMs trained on high-quality PET images with *unsupervised* schemes that condition on measured data [20, 27, 8]. This enables these algorithms to remain agnostic to injected dose and scanner configuration, but they do not learn how the data consistency updates interact with the DM-based denoising updates, thereby compromising both quantitative accuracy and fidelity to the DM’s learned manifold.

In this work, we propose and investigate an efficient *supervised* conditioning framework tailored to PET reconstruction. Building on the DEFT (Doob’s h-transform Efficient Fine Tuning) framework proposed by Denker *et al.* [3], we pre-train a DM on high-quality PET images, which we then freeze while learning a measurement-conditional module with sinogram data. The design of our measurement-conditional module integrates solutions to several PET-specific challenges, such as the non-negativity of PET’s Poisson likelihood model and the wide intensity range of PET images.

The contributions of our paper are therefore:

- We present the first supervised learning DM-based approach to PET image reconstruction.
- We propose and validate PET-specific modifications to existing DM-based frameworks for solving inverse problems.
- We match the performance of state-of-the-art supervised reconstruction approaches in our *in-silico* experiments while also enabling posterior sampling.
- We demonstrate applicability to real-world clinical fully 3D ^{18}F FDG brain PET data.

2 Methodology

2.1 Preliminaries

PET reconstruction PET reconstruction from measured sinograms is an inverse problem [14]. If $\mathbf{x}$ is the true radiotracer distribution, $\mathbf{A}$ is the system model

(incorporating attenuation and normalization), and $\mathbf{b}$ models scatter and randoms components, the true mean $\mathbf{q}$ of noisy measurements $\mathbf{m}$ (e.g. a sinogram) may be modeled as

$$\mathbf{q} = \mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{m} \sim \text{Poisson}(\mathbf{q}). \quad (1)$$

For an image estimate $\mathbf{x}$, let the Poisson log-likelihood (PLL) of noisy measurements $\mathbf{m}$ with respect to $\mathbf{x}$ be $L(\mathbf{x}|\mathbf{m}) := PLL(\max(\mathbf{x}, 0)|\mathbf{m}, \mathbf{A}, \mathbf{b})$, and let $p(\mathbf{x})$ be the prior probability density of $\mathbf{x}$.

Diffusion models The central construct in a DM is a diffusion process mapping from high-quality images to unstructured Gaussian noise, often modeled as a stochastic differential equation (SDE) (over time t from 0 to 1) [23]. This SDE defines transition probabilities p_t ; the score vector $\nabla_{\mathbf{x}} \log p_t(\mathbf{x}_t)$ then represents the direction of removing artificial Gaussian noise from high-quality images. One may train a score network $\mathbf{s}_\theta(\mathbf{x}_t, t)$ to estimate $\nabla_{\mathbf{x}} \log p_t(\mathbf{x}_t)$ by optimization with respect to the Denoising Score Matching (DSM) objective [24]:

$$\min_{\theta} \left\{ \mathbb{E}_{t \sim U[0,1]} \mathbb{E}_{\mathbf{x}_0 \sim p_0} \mathbb{E}_{\mathbf{x}_t \sim p_t(\mathbf{x}_t|\mathbf{x}_0)} \left[\|\mathbf{s}_\theta(\mathbf{x}_t, t) - \nabla_{\mathbf{x}} \log p_t(\mathbf{x}_t|\mathbf{x}_0)\|_2^2 \right] \right\} \quad (2)$$

For image generation, we sample the initial image estimate $\mathbf{x}_1$ as Gaussian noise, before iteratively denoising by approximating the reverse-time SDE using $\mathbf{s}_\theta(\mathbf{x}_t, t)$. By interleaving denoising steps with unsupervised conditioning steps, or by training $\mathbf{s}_\theta$ with conditioning information, we may perform reconstruction.

These processes usually involve Tweedie’s estimate $\hat{\mathbf{x}}_0(\mathbf{x}_t, \mathbf{s}_\theta) = \frac{\mathbf{x}_t + \sigma_t \mathbf{s}_\theta(\mathbf{x}_t, t)}{\mu_t}$, which estimates the conditional expectation $\mathbb{E}(\mathbf{x}_0|\mathbf{x}_t)$, given image iterate $\mathbf{x}_t$ and trained score network $\mathbf{s}_\theta(\mathbf{x}_t, t)$, where μ_t and σ_t are terms analytically derived from p_t (see [20]). In this work, we use the variance-preserving SDE (see [27]).

2.2 Approach

Figure 1 presents an overview of our approach, PET-DEFT, named for its adaptation of the DEFT framework [3] to PET reconstruction.

To deal with PET images’ wide dynamic range, we work in a normalized space for diffusion updates and un-normalize for data fidelity updates. Training on image data, we calculate scale factor $c = c(\mathbf{x})$; training or reconstructing from measurements, we estimate $c \approx \hat{c}(\mathbf{m})$, where M is the μ -map for dataset i and $\mathbf{x}_{\text{MLEM}_{10}}$ is the image estimate after 10 iterations of MLEM (Maximum-Likelihood Expectation-Maximization) [18]:

$$c = c(\mathbf{x}) := \|\mathbf{x}\|_1 / \|M\|_0 \quad ; \quad \hat{c} = \hat{c}(\mathbf{m}) := \|\mathbf{1}_{M>0} \cdot \mathbf{x}_{\text{MLEM}_{10}}\|_1 / \|M\|_0. \quad (3)$$

We firstly pre-train $\mathbf{s}_\theta(\mathbf{x}_t, t)$ on c -normalized PET images (see [20, 27]). We then introduce paired datasets $(\mathbf{x}^{(i)}, \mathbf{m}^{(i)})_{i=1}^N$ from which to learn the conditional score network $\mathbf{s}_\phi$. Given noisy image $\mathbf{x}_t$, we calculate Tweedie’s estimate using the pre-trained score $\hat{\mathbf{x}}_0(\mathbf{x}_t, \mathbf{s}_\theta)$. This estimate is in normalized diffusion space; we transform it to calibrated measurement space via multiplication with $\hat{c}$. We

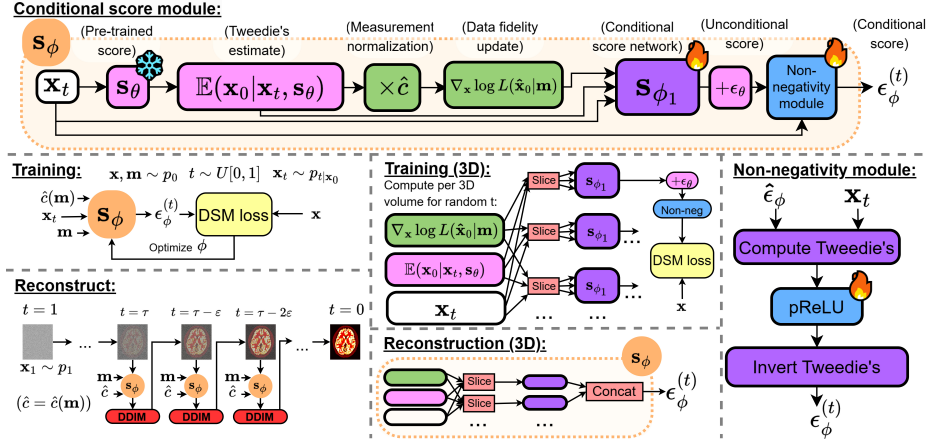


Fig. 1: Methodology of proposed approach. The conditional score module $\mathbf{s}_\phi$ (top) learns to estimate a noisy image iterate $\mathbf{x}_t$'s conditional score $\epsilon_\phi^{(t)}$ (akin to a denoising vector). In $\mathbf{s}_\phi$, the unconditional Tweedie's estimate (from pre-trained $\mathbf{s}_\theta$) is measurement-normalized and used to compute a data fidelity update. Then $\mathbf{s}_\phi$ processes these terms, before adding $\mathbf{s}_\theta$ and encouraging non-negativity of the conditional Tweedie's estimate, to output $\epsilon_\phi^{(t)}$. Reconstruction uses $\epsilon_\phi^{(t)}$ as a denoising direction to guide the DDIM scheme for reverse diffusion. In 3D, we precompute memory-intensive components of $\mathbf{s}_\phi$ and train $\mathbf{s}_{\phi_1}$ as a 2.5D network on tranverse slices. We then reconstruct as in 2D, except with $\epsilon_\phi^{(t)}$ formed by concatenating slice-wise outputs. Fire (ice) indicates trainable (frozen) modules.

then calculate a data fidelity update on this estimate as the gradient of the log-likelihood $\nabla_{\mathbf{x}} \log L(\hat{\mathbf{x}}_0(\mathbf{x}_t, \mathbf{s}_\theta) | \mathbf{m})$.

The noisy iterate $\mathbf{x}_t$, Tweedie's estimate $\hat{\mathbf{x}}_0(\mathbf{x}_t, \mathbf{s}_\theta)$, and the data fidelity update are then concatenated and inputted to the network $\mathbf{s}_\phi$. The unconditional score $\mathbf{s}_\theta(\mathbf{x}_t, t)$ is added, so that $\mathbf{s}_\phi$ is trained to predict the residual.

Lastly, we apply a novel mechanism to encourage the non-negativity of our conditional Tweedie's estimate. This is motivated by a desire to keep the image iterate values $\mathbf{x}_t$ positive so they are incorporated into later data fidelity updates (as the Poisson likelihood has the domain $\mathbb{R}^{\geq 0}$), thereby improving the conditional score estimation during test-time reconstruction. We calculate Tweedie's estimate from the conditional score $\hat{\mathbf{x}}_0(\mathbf{x}_t, \mathbf{s}_\theta + \mathbf{s}_\phi)$, apply a pReLU and invert Tweedie's estimate to get a score estimate corresponding to reduced negative values in $\hat{\mathbf{x}}_0(\mathbf{x}_t, \mathbf{s}_\theta + \mathbf{s}_\phi)$:

$$\epsilon_\phi^{(t)} = (\mu_t \cdot \text{pReLU}(\hat{\mathbf{x}}_0(\mathbf{x}_t, \mathbf{s}_\theta + \mathbf{s}_\phi)) - \mathbf{x}_t) / \sigma_t. \quad (4)$$

We then freeze $\mathbf{s}_\theta$ and train the output $\epsilon_\phi^{(t)}$ on the conditional DSM objective. To perform reconstruction, we use $\mathbf{s}_\phi$ with the DDIM (Denoising Diffusion Implicit Model) [22] scheme (100 steps) for simulating the reverse diffusion SDE.

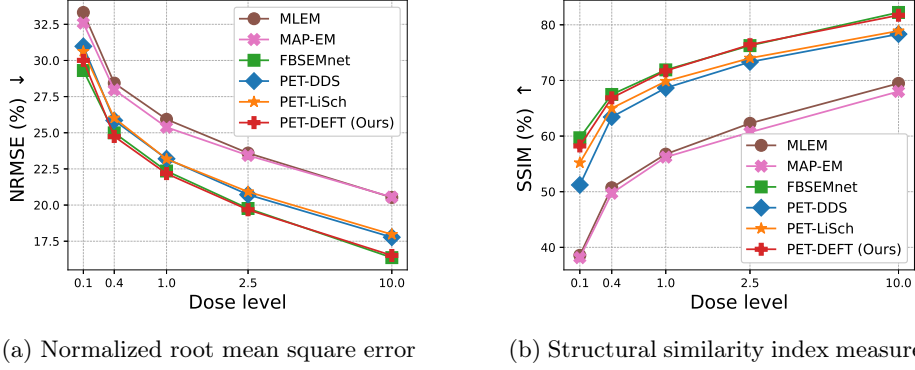


Fig. 2: Quantitative accuracy of PET reconstruction algorithms at 5 dose levels.

Fully 3D reconstruction Our approach extends efficiently to 3D, as we may decouple $\mathbf{A}$, $\mathbf{s}_\theta$ and $\mathbf{s}_\phi$. For a given 3D dataset, we first compute the pre-trained update components; we then train our conditional network on randomly-shuffled pre-computed slices of Tweedie’s estimate, $\mathbf{x}_t$ and the data fidelity update. We train a 2.5D network for $\mathbf{s}_\theta$ and $\mathbf{s}_{\phi_1}$, with $3\times$ as many input channels to accommodate information on 3 consecutive transverse slices (while retaining one output channel for predicting the score for the central slice). The 3D score estimate is then the concatenation of transverse slice-wise score estimates.

3 Experiments

3.1 Experimental details

Dataset 39 3D ^{18}F FDG PET brain scans were simulated by segmenting real MR volumes and assigning realistic activity values to different tissues (following Mehranian *et al.* [11]). 30 central 2D transverse slices were taken from each of these simulated scans to form the ground truth images used in this study. Measured data was generated by forward modeling each 2D image and applying Poisson noise to the resulting sinogram. Corrective factors were modeled as purely attenuation and contamination was modeled as a constant background of 30% of simulated prompts. We define dose level 1.0 as when the average count rate was 1.7×10^6 . For real data, we used 25 (2) previously acquired real 3D ^{18}F FDG datasets for training (validation) (see [11] for dataset details). Real training images were MR-assisted reconstructions with the Bowsher prior [16].

Implementation details The forward model $\mathbf{A}$ (common to all methods) was implemented with ParallelProj [17], with dimensions approximating the Siemens Biograph mMR. 900 datasets were used for model training, 150 for validation and 10 for testing. Evaluations were made from three independent realizations of Poisson noisy sinograms. SSIM (structural similarity index measure) and

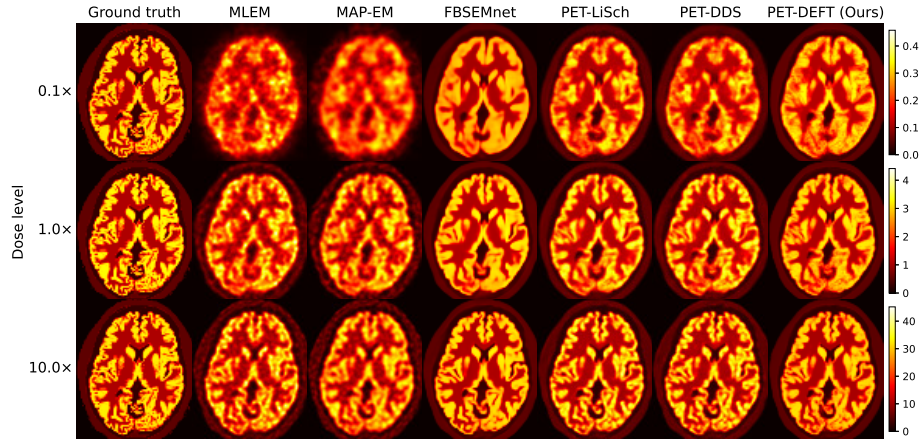


Fig. 3: PET brain phantom reconstructions with increasing dose level.

LPIPS (learned perceptual image patch similarity) are perceptual metrics, while NRMSE (normalized root mean square error) is a distortion metric. NRMSE and SSIM values for DM-based methods were computed on the mean of 12 reconstructed samples (from the same noisy sinogram). We implemented the model architecture in [2] for unconditional and conditional score modules, varying the numbers of input channels as necessary. PyTorch [13] and an Nvidia RTX 3090 GPU were used for experiments. Adam [10] with learning rate 1×10^{-4} (8×10^{-7}) was used for 2D (3D) training. Supervised models were trained to minimize validation loss. Hyperparameters for MLEM, MAP-EM, PET-LiSch, and PET-DDS were tuned to minimize test-time NRMSE. In 2D, we set DDIM’s stochasticity hyperparameter $\eta = 0.1$ (following [20, 27]), while in 3D, we increased η to 0.9, as this yielded results more visually consistent with the OSEM image.

3.2 Comparison against other methods

MLEM is a clinical standard model-based algorithm [18], while MAP-EM with a patch-based prior [25] is a state-of-the-art model-based approach. PET-DDS [20] and PET-LiSch [27] are unsupervised DM-based PET reconstruction approaches. FBSEMnet is a supervised deep-learned unrolled iterative approach, representing the current state-of-the-art for deep-learned PET reconstruction [11].

Quantitative analysis Figure 2 shows that our proposed method PET-DEFT performs on par with FBSEMnet, outperforming unconditional diffusion and model-based approaches across 5 dose levels. In general, the methods are grouped by performance into supervised DL, unsupervised DL and model-based methods.

Qualitative analysis Figure 3 shows that our proposed approaches yield perceptually strong reconstruction accuracy at multiple dose levels. As counts de-

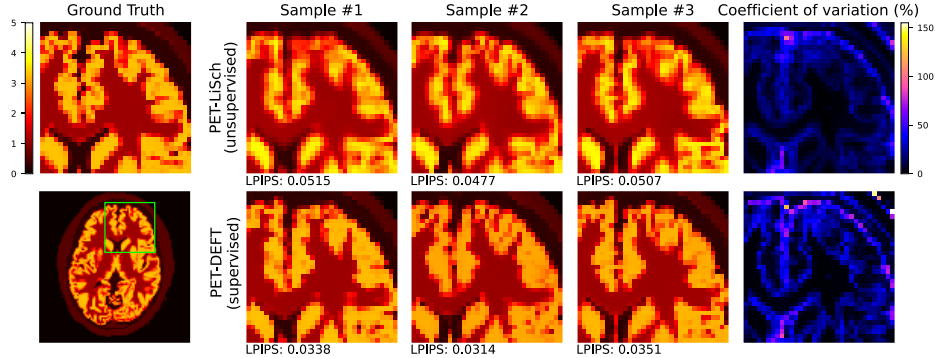


Fig. 4: Example samples from unsupervised (top) and supervised (bottom) DM-based PET reconstruction (ours) at dose level 1.0.

crease, PET-DEFT images resemble blurred scans (due to the averaging of multiple samples), retaining some fine-grained features, while FBSEMnet resembles a lower resolution blocky image (as it doesn’t model the posterior distribution and is trained with an L2 loss).

3.3 Posterior sampling and uncertainty

Figure 4 shows single reconstruction samples from our supervised approach PET-DEFT and the unsupervised approach PET-LiSch. As is clear visually and also indicated by the LPIPS score, samples from PET-DEFT yield superior adherence to the real manifold of ground truth images.

3.4 Ablation studies

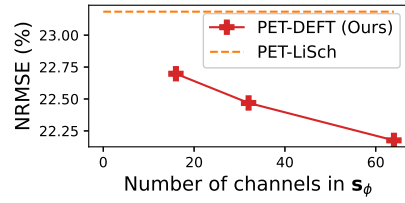
Figure 5(a) shows that the measurement normalization mechanism is essential for PET-DEFT’s performance. Figure 5(b) shows that with the non-negativity module, we may perform as few as 5 reverse diffusion steps and retain our peak quantitative accuracy, compared to 8 steps without. In Figure 5(c) and (d), we show relative resilience of PET-DEFT to reducing the number of parameters or training sinogram-image pairs; the trend in Fig. 5(c) suggests possible further performance improvements with a larger conditioning network.

3.5 Reconstruction from 3D Real $[^{18}\text{F}]$ FDG data

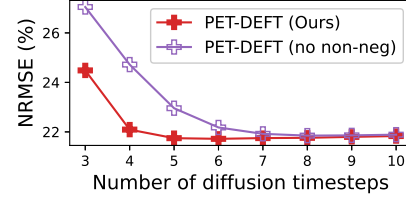
Compared to unrolled iterative methods such as FBSEMnet, our DM-based approach trains more efficiently (as the neural network and forward operator are accessed once each and decoupled for backpropagation), and also enables posterior sampling. We showcase these abilities by demonstrating reconstruction from real 3D $[^{18}\text{F}]$ FDG sinogram data (using a commercial 24GB GPU) in Figure 6.

Approach	NRMSE (%)
No normalization step	30.0
Measurement-normalized	22.1

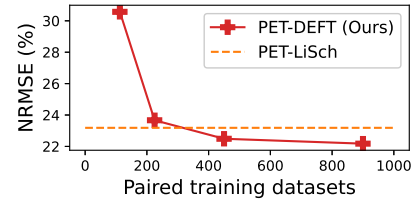
(a) Effect of measurement normalization



(c) Number of parameters



(b) Effect of non-negativity module with few diffusion timesteps



(d) Number of training examples

Fig. 5: Ablation studies into PET-DEFT's conditional score network (dose 1.0).

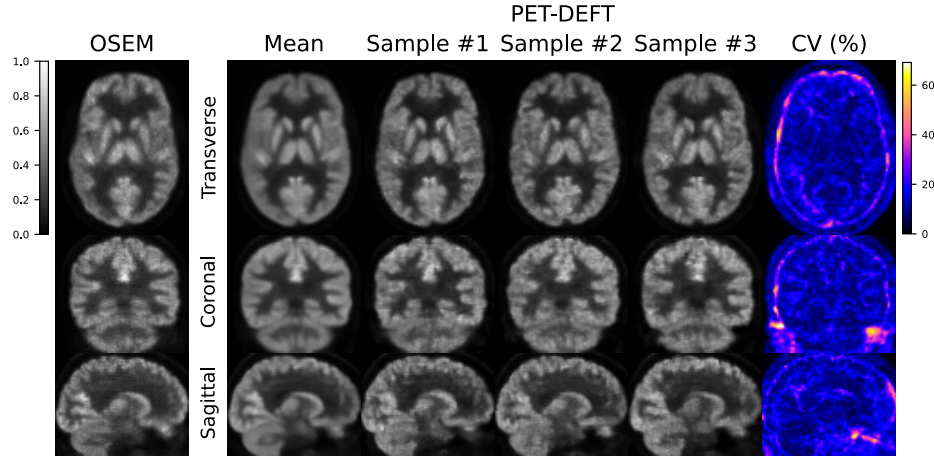


Fig. 6: Slices from image reconstructed with PET-DEFT from 3D real data. Mean and coefficient of variation (CV) are calculated over 12 samples.

4 Conclusion

In this paper, we propose novel and efficient methodology for supervised DM-based reconstruction of PET images. We introduce PET-specific mechanisms for training a conditional score network, including measurement normalization and a non-negativity module to address the wide dynamic range of PET images and the non-negativity of the Poisson likelihood. We found that our approach PET-DEFT outperforms unsupervised DM-based reconstruction and matches state-of-the-art PET reconstruction on experiments on simulated PET datasets. Our approach also enables more accurate posterior sampling for PET reconstruction. Lastly, we showed the feasibility and efficiency of our approach by reconstructing from real ^{18}F FDG PET data, with an innovative training strategy that decouples the 3D PET physics operator from a 2.5D conditional score network.

Acknowledgments. GW acknowledges funding from the EPSRC CDT in Smart Medical Imaging [EP/S022104/1] and a GSK studentship. This work was also supported in part by EPSRC grant [EP/S032789/1]. The authors have applied a Creative Commons Attribution (CC BY) licence to any Accepted Author Manuscript version arising, in accordance with King’s College London’s Rights Retention policy.

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

References

1. Anand, S., Singh, H., Dash, A.: Clinical applications of PET and PET-CT. *Medical Journal Armed Forces India* **65**(4), 353–358 (2009)
2. Dhariwal, P., Nichol, A.: Diffusion models beat GANs on image synthesis. *Advances in Neural Information Processing Systems* **34**, 8780–8794 (2021)
3. Denker, A., Vargas, F., Padhy, S., Didi, K., Mathis, S.V., Barbano, R., Dutordoir, V., Mathieu, E., Komorowska, U.J., Lio, P.: DEFT: Efficient fine-tuning of diffusion models by learning the generalised h -transform. *Advances in Neural Information Processing Systems* **37**, 19636–19682 (2024)
4. Filipović, M., Barat, E., Dautremer, T., Comtat, C., Stute, S.: PET Reconstruction of the Posterior Image Probability, Including Multimodal Images. *IEEE Transactions on Medical Imaging* **38**(7), 1643–1654 (2019)
5. Filipović, M., Dautremer, T., Comtat, C., Stute, S., Barat, E.: Reconstruction, analysis and interpretation of posterior probability distributions of PET images, using the posterior bootstrap. *Physics in Medicine and Biology* **66**(12), 125014 (2021)
6. Guazzo, A., Colarieti-Tosti, M.: Learned primal dual reconstruction for PET. *Journal of Imaging* **7**(12), 248 (2021)
7. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. *Advances in Neural Information Processing Systems* **33**, 6840–6851 (2020)
8. Hu, J., Song, B., Fessler, J.A., Shen, L.: Patch-based diffusion models beat whole-image models for mismatched distribution inverse problems. *arXiv preprint arXiv:2410.11730* (2024). <https://doi.org/10.48550/arXiv.2410.11730>

9. Hu, R., Wu, D., Tivnan, M., Guo, N., Yoon, S., Chen, Z., Wang, Y., Luo, J., Xue, S., Cui, J., Liu, H., Li, Q.: Unsupervised low-dose PET image reconstruction based on pre-trained denoising diffusion probabilistic prior. *Journal of Nuclear Medicine* **65**(supplement 2), 241109–241109 (2024)
10. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. In: *International Conference on Learning Representations (ICLR)* (2015).
11. Mehranian, A., Reader, A.J.: Model-based deep learning PET image reconstruction using forward-backward splitting expectation-maximization. *IEEE Transactions on Radiation and Plasma Medical Sciences* **5**(1), 54–64 (2020)
12. Nievelstein, R.A.J., Quarles van Ufford, H.M.E., Kwee, T.C., Bierings, M.B., Ludwig, I., Beek, F.J.A., de Klerk, J.M.H., Mali, W.P.T.M., de Bruin, P.W., Geleijns, J.: Radiation exposure and mortality risk from CT and PET imaging of patients with malignant lymphoma. *European Radiology* **22**(9), 1946–1954 (2012)
13. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Köpf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S.: PyTorch: An imperative style, high-performance deep learning library. In: *Advances in Neural Information Processing Systems 32 (NeurIPS 2019)*, pp. 8026–8037 (2019)
14. Reader, A.J., Corda, G., Mehranian, A., Costa-Luis, C.d., Ellis, S., Schnabel, J.A.: Deep learning for PET image reconstruction. *IEEE Transactions on Radiation and Plasma Medical Sciences* **5**(1), 1–25 (2021)
15. Reader, A.J., Pan, B.: AI for PET image reconstruction. *British Journal of Radiology* **96**(1150), 20230292 (2023)
16. Schramm, G., Holler, M., Rezaei, A., Vunckx, K., Knoll, F., Bredies, K., Boada, F., Nuyts, J.: Evaluation of parallel level sets and Bowsher’s method as segmentation-free anatomical priors for time-of-flight PET reconstruction. *IEEE Transactions on Medical Imaging* **37**(2), 590–603 (2018)
17. Schramm, G., Thielemans, K.: PARALLELPROJ—an open-source framework for fast calculation of projections in tomography. *Frontiers in Nuclear Medicine* **3** (2024)
18. Shepp, L.A., Vardi, Y.: Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging* **1**(2), 113–122 (1982)
19. Shukla, A.K., Kumar, U.: Positron emission tomography: An overview. *Journal of Medical Physics* **31**(1), 13–21 (2006)
20. Singh, I.R.D., Denker, A., Barbano, R., Kereta, Ž., Jin, B., Thielemans, K., Maass, P., Arridge, S.: Score-based generative models for PET image reconstruction. *Machine Learning for Biomedical Imaging 2*(Generative Models), 547–585 (2024)
21. Sohl-Dickstein, J., Weiss, E., Maheswaranathan, N., Ganguli, S.: Deep unsupervised learning using nonequilibrium thermodynamics. In: *Proceedings of the 32nd International Conference on Machine Learning*, 2256–2265 (2015)
22. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. In: *9th International Conference on Learning Representations (ICLR)* (2021)
23. Song, Y., Sohl-Dickstein, J., Kingma, D.P., Kumar, A., Ermon, S., Poole, B.: Score-based generative modeling through stochastic differential equations. In: *International Conference on Learning Representations (ICLR)* (2021)
24. Vincent, P.: A connection between score matching and denoising autoencoders. *Neural Computation* **23**(7), 1661–1674 (2011)
25. Wang, G., Qi, J.: Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization. *IEEE Transactions on Medical Imaging* **31**(12), 2194–2204 (2012)

26. Webber, G., Mizuno, Y., Howes, O.D., Hammers, A., King, A.P., Reader, A.J.: Generative-model-based fully 3D PET image reconstruction by conditional diffusion sampling. In: IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), 1–2 (2024)
27. Webber, G., Mizuno, Y., Howes, O.D., Hammers, A., King, A.P., Reader, A.J.: Likelihood-scheduled score-based generative modeling for fully 3D PET image reconstruction. *IEEE Transactions on Medical Imaging*, Early Access (2025). <https://doi.org/10.1109/TMI.2025.3576483>
28. Xie, T., Cui, Z.X., Luo, C., Wang, H., Liu, C., Zhang, Y., Wang, X., Zhu, Y., Chen, G., Liang, D., Jin, Q., Zhou, Y., Wang, H.: Joint diffusion: Mutual consistency-driven diffusion model for PET-MRI co-reconstruction. *Physics in Medicine & Biology* **69**(15), 155019 (2024)