Pseudo-MRI-Guided PET Image Reconstruction Method Based on a Diffusion Probabilistic Model

Weijie Gan, Huidong Xie, Carl von Gall, Günther Platsch, Michael T. Jurkiewicz, Andrea Andrade, Udunna C. Anazodo, Ulugbek S. Kamilov, Hongyu An, and Jorge Cabello

Abstract

Anatomically guided PET reconstruction using MRI information has been shown to have the potential to improve PET image quality. However, these improvements are limited to PET scans with paired MRI information. In this work we employed a diffusion probabilistic model (DPM) to infer T1-weighted-MRI (deep-MRI) images from FDG-PET brain images. We then use the DPM-generated T1w-MRI to guide the PET reconstruction. The model was trained with brain FDG scans, and tested in datasets containing multiple levels of counts. Deep-MRI images appeared somewhat degraded than the acquired MRI images. Regarding PET image quality, volume of interest analysis in different brain regions showed that both PET reconstructed images using the acquired and the deep-MRI images improved image quality compared to OSEM. Same conclusions were found analysing the decimated datasets. A subjective evaluation performed by two physicians confirmed that OSEM scored consistently worse than the MRI-guided PET images and no significant differences were observed between the MRI-guided PET images. This proof of concept shows that it is possible to infer DPM-based MRI imagery to guide the PET reconstruction, enabling the possibility of changing reconstruction parameters such as the strength of the prior on anatomically guided PET reconstruction in the absence of MRI.

1 Introduction

The combination of anatomical information with positron emission tomography (PET) measured data has been long explored as a way to regularize the PET inverse reconstruction problem [1]. Magnetic resonance imaging (MRI) can provide exquisite anatomical high detail between different soft tissues and low noise, especially in brain imaging where 3D MRI scans have isotropic spatial resolution. Therefore, MRI has the

This work was supported in part by the Academic Medical Organization of Southwestern Ontario (AMOSO; Grant ID: INN20-023) and Physician's Services Incorporated (PSI) Foundation - LRI7767306.

This work involved human subjects in its research. Approval of all ethical and experimental procedures and protocols was granted by the Western University Research Ethics Board (HSREB 104900).

Weijie Gan is with the Department of Computer Science & Engineering, Washington University in St. Louis, St. Louis, MO, USA.

Huidong Xie is Department of Radiology and Biomedical Imaging, Yale University, New Haven, CT, USA.

Michael T. Jurkiewicz is with Department of Medical Imaging and Department of Medical Biophysics, Western University, 1151 Richmond St, London, ON N6A 5C1, Canada.

Andrea Andrade is with Department of Pediatrics and Department of Clinical Neurological Sciences, Western University, 1151 Richmond St, London, ON N6A 5C1, Canada.

Udunna C. Anazodo is with Lawson Health Research Institute, 268 Grosvenor St, London, ON N6A 4V2, Canada and Neurology and Neurosurgery, Montreal Neurological Institute, McGill University, Montreal, QC, H3A 2B4 Canada.

Günther Platsch is with Siemens Healthineers AG, Erlangen, Germany.

Ulugbek S. Kamilov is with the Department of Computer Science & Engineering and Department of Electrical & System Engineering, Washington University in St. Louis, St. Louis, MO, USA.

Hongyu An is with the Mallinckrodt Institute of Radiology, the Departments of Neurology, Biomedical Engineering, Electrical & System Engineering, and with the Division of Biology and Biomedical Sciences, Washington University in St. Louis, MO, USA

Jorge Cabello and Carl von Gall are with Siemens Medical Solutions USA, Inc., 810 Innovation Dr, Knoxville, TN 37932, USA.

potential to improve PET image quality by reducing noise, improving spatial resolution and quantification. It also has the potential of reducing the PET scan time or injected activity while retaining the PET image quality. However, access to PET/MRI technology is not as extended as to PET/CT, hence the application of MRI-guided PET (MRIg-PET) reconstruction algorithms is strongly limited.

Anatomically-guided PET reconstruction relies on the premise that there exists a correlation between the anatomical information and functional information. Brain gray matter consists of neurons and dendrites connecting surrounding neurons, while white matter consists of axons (surrounded by fatty myelin) connecting neurons from different regions of the brain. Due to their composition there is 2.5-4.1 higher glucose consumption in the grey matter than in the white matter [2]. On the other hand, T1-weighted MRI shows high intensity in fatty tissues like the white matter, and darker intensity in the gray matter. This direct correlation between tissue intensity in MRI and glucose consumption makes FDG brain PET/MRI the usual case where anatomically-guided PET reconstruction methods are tested.

PET image reconstruction using prior information is often formulated using a *maximum-a-posteriori* approach [3,4], where the data and prior information are combined in the cost function using a hyperparameter to control the weight of the prior information over the acquired PET data. Even though there have been multiple efforts to develop algorithms to adaptively calculate the hyperparameter based on the measured data, it still often requires manual tuning [5].

Inference of MRI from PET imagery is not novel. The estimation of MRI from PET, using florbetapir for amyloid imaging, was explored using generative adversarial networks for quantitative PET analysis [6]. The resulting volumetric PET analysis using the generated MRI images compared to the acquired MRI images was similar, even though the generated MRI images were blurry, the soft tissue contrast was low and the overall image quality was poor and unrealistic. Inversely, deep generated FDG-PET images have been synthesized using multiple convolution U-net generators from MRI using the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset [7].

The use of artificial intelligence (AI) models for anatomically guided PET reconstruction has already been employed to create a model that takes as input an OSEM-PET reconstructed image and a paired T1w-MRI image, to produce an anatomically guided PET reconstructed image [8]. A downside of this approach is that a model is created using as input reconstructed PET images generated with certain reconstruction parameters (i.e. with or without point spread function [PSF] modeling), hence, if different reconstruction parameters are employed the output of the network is uncertain. In addition, this approach uses a specific MRI sequence, T1w-MRI in the case of [8], preventing its application to other MRI sequences which may be more relevant for other disorders.

Diffusion probabilistic models (DPM) refer to a new class of deep generative models which have recently gained popularity due to the excellent performance (see recent reviews in [9–11]). The key concept behind DPM is to iteratively add noise to an original data point until it becomes a pure noise sample, and then reverse this process by denoising, thereby generating new samples. During this process, DPM learn the data distribution, enabling the synthesis of unseen, high-quality data samples. This approach has demonstrated remarkable results, particularly in generating high-resolution natural images.

Recent studies have shown the potential of DPM in synthesis of PET or MR images [12–17]. For example, DPM has been applied to generate pseudo PET images from the low-dose variants [16,17] or corresponding MR images [14,15].

In this work we trained a DPM model with 25 datasets of paired T1w MR images and FDG-PET images reconstructed using a standard iterative parameter (ordered subset expected maximum, OSEM), to generate a pseudo-T1w-MRI (deep-MRI) image that is employed to guide the PET reconstruction (deep-PET) using a maximum-a-posteriori (MAP) reconstruction algorithm. The model was then tested in eight datasets from the same cohort not included in the training. This approach allows the parameters of the reconstruction, including the strength of the influence of the anatomical information, to be modified according to the needs of the application, as opposed to prior attempts [8], such as modifying the hyper-parameter or the number of iterations. A detailed description of the study cohort, PET acquisition and reconstruction parameters are summarized in [20].

Furthermore, the model generated with the full dose PET reconstructed images, was evaluated providing



Figure 1: An illustration of the proposed deep-MRI-guided PET image reconstruction. Firstly, we generate a deep-MRI image from the OSEM PET image by using a pre-trained diffusion model (see its training process in Section 2.1 and Section 2.2). Then, the deep-MRI-guided PET image is obtained by solving an optimization of the anatomically-guided reconstruction given the generated deep-MRI images and the raw PET data.

as input reconstructed PET images obtained from decimated datasets at 25% and 5% of the original datasets. Resulting deep-MRI images were compared with the real MRI images, and more importantly, the resulting deep-PET images were compared to the PET images obtained using the real MRI images. The comparison was performed objectively using a volume of interest analysis, and validated clinically based on observations from experienced physicians.

The methodology presented in this work can be applied to PET/CT scans and emerging dedicated brain PET scanners, in absence of a paired MRI scan, or also in PET/MRI scans where the MRI contains artifacts, such as patient motion, artifacts coming from coils failures or patients with metallic implants.

2 Methods

2.1 Diffusion Probabilistic Model

DPM consists of two Markov processes: the fixed forward process and the learning-based reverse process. The forward process starts from a sample of a T1w-MRI image $\boldsymbol{x}_0 \sim q(\boldsymbol{x})$ and gradually adds Gaussian noise according to the following transition probability:

$$q(\boldsymbol{x}_t | \boldsymbol{x}_{t-1}) := \mathcal{N}(\boldsymbol{x}_t; \sqrt{1 - \beta_t \boldsymbol{x}_{t-1}}, \beta_t \mathbf{I}) , \qquad (1)$$

where $\mathcal{N}(\cdot)$ represents the Gaussian probability density function, and $\beta_{1:T}$ refers to a variance schedule subject to $\beta_t \in (0, 1)$ for all t = 1, ..., T. An important feature of this forward process is that it allows us to sample at any arbitrary step t directly conditioned on \mathbf{x}_0

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

where $\alpha_t := 1 - \beta_t$ and $\bar{\alpha}_t = \prod_{s=1}^t \alpha_s$. The latent variables $\boldsymbol{x}_{1:T}$ have the same dimensionality as the T1w-MRI image. The latent \boldsymbol{x}_T is nearly an isotropic Gaussian distribution for a properly designed β_t schedule. Since the latent \boldsymbol{x}_T is a known isotropic Gaussian distribution, one can generate a new \boldsymbol{x}_T and then obtain a T1w-MRI image \boldsymbol{x}_0 by progressively sampling from the reverse posterior $q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)$. However, this reverse posterior is tractable only if \boldsymbol{x}_0 is known

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

where

$$\mu_q(\boldsymbol{x}_t, \boldsymbol{x}_0) = \frac{\sqrt{\alpha_t}(1 - \bar{\alpha}_{t-1})\boldsymbol{x}_t + \sqrt{\bar{\alpha}_{t-1}}(1 - \alpha_t)\boldsymbol{x}_0}{1 - \bar{\alpha}_t} .$$
(4)

Noted that $q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t) := q(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, \boldsymbol{x}_0)$, where the extra conditioning term \boldsymbol{x}_0 is superfluous due to the Markov property. DPM proposes to learn a parameterized Gaussian transitions $p_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t)$ to approximate (3)

$$p_{\boldsymbol{\theta}}(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t) = \mathcal{N}\left(\boldsymbol{x}_{t-1}; \boldsymbol{\mu}_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t), \sigma_t^2 \mathbf{I}\right) , \qquad (5)$$

where

$$\mu_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t) = \frac{1}{\sqrt{\alpha_t}} \left(\boldsymbol{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \boldsymbol{\epsilon}_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t) \right) \,. \tag{6}$$

Here, $\epsilon_{\theta}(\boldsymbol{x}_t, t)$ denotes a neural network. The training objective of $\epsilon_{\theta}(\boldsymbol{x}_t, t)$ can be formulated as follow by considering a training set of $\{\boldsymbol{x}_i\}_i$

$$\mathbb{E}_{\boldsymbol{x},\epsilon,t\sim[1,T]}\left[\left\|\boldsymbol{\epsilon}-\boldsymbol{\epsilon}_{\boldsymbol{\theta}}(\boldsymbol{x}_{t},t)\right\|^{2}\right],\tag{7}$$

where \mathbf{x}_t denotes a noisy latent sampled from eq. (2) as $\mathbf{x}_t = \sqrt{\overline{\alpha}_t}\mathbf{x} + \sqrt{(1 - \overline{\alpha}_t)}\epsilon$ for $\epsilon \sim \mathcal{N}(0, \mathbf{I})$. Recent studies [22,23] have shown the improved performance by using the learned variance $\sigma_t^2 \coloneqq \sigma_{\theta}^2(\mathbf{x}_t, t)$. We also adopted this approach. To be specific, we have $\sigma_{\theta}(\mathbf{x}_t, t) \coloneqq \exp(v \log \beta_t + (1 - v) \log \tilde{\beta}_t)$, where $\tilde{\beta}_t$ refers to the lower bounds for the reverse diffusion posterior variances [24], and v denotes the network output. We used a single neural network with two separate output channels to estimate the mean and the variance of (5) jointly. Based on the learned reverse posterior $p_{\theta}(\mathbf{x}_{t-1}|\mathbf{x}_t)$, the iteration of obtaining a \mathbf{x}_0 from a \mathbf{x}_T can be formulated as

$$\boldsymbol{x}_{t-1} = \mu_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t) + \sigma_t \boldsymbol{z}, \text{ where } \boldsymbol{z} \sim \mathcal{N}(0, \mathbf{I}) .$$
 (8)

2.2 PET-conditioned MRI Synthesis Diffusion Model

The conventional diffusion models discussed in Section 2.1 are for unconditional MR image synthesis, namely generating arbitrary realistic MR image. In this study, we reconstructed a set of PET images using the OSEM algorithm, and then trained a diffusion model using paired OSEM-PET reconstructed images and T1w-MRI images, in order to generate pseudo MR images that are consistent with the OSEM-PET images. To be specific, let c denote the OSEM-PET images. Our goal is to learn a new parameterized Gaussian transitions $p_{\theta}(x_{t-1}|x_t, c)$ to approximate eq. (3) instead of eq. (5) in the original DPM. This approach can be achieved by simply supplying c as the additional network input. The new training objective can then be formulated as follows by considering a training set MRI-PET image pairs $\{(x_i, c_i)\}_i$

$$\mathbb{E}_{(\boldsymbol{x},\boldsymbol{c}),\epsilon,t\sim[1,T]}\left[\left\|\boldsymbol{\epsilon}-\boldsymbol{\epsilon}_{\boldsymbol{\theta}}(\boldsymbol{x}_{t},t,\boldsymbol{c})\right\|^{2}\right].$$
(9)

Based on $p_{\theta}(\boldsymbol{x}_{t-1}|\boldsymbol{x}_t, \boldsymbol{c})$, the iteration of the PET-conditioned DPM becomes

$$\boldsymbol{x}_{t-1} = \mu_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t, \boldsymbol{c}) + \sigma_t \boldsymbol{z}, \text{ where } \boldsymbol{z} \sim \mathcal{N}(0, \mathbf{I}) , \qquad (10)$$

where

$$\mu_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t, \boldsymbol{c}) = \frac{1}{\sqrt{\alpha_t}} \left(\boldsymbol{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_{\boldsymbol{\theta}}(\boldsymbol{x}_t, t, \boldsymbol{c}) \right) \,. \tag{11}$$

2.3 3D Image Synthesis using 2D Diffusion Model

PET and MRI are 3D imaging modalities, yet directly synthesizing an entire 3D MR volume with a diffusion model poses practical challenges due to the extremely high memory and computational cost. An alternative approach involves training a diffusion model for 2D PET-conditioned MR image synthesis and subsequently generating the entire 3D MR volume in a *slice-by-slice* manner. Although effective, this simple approach may lead to shaggy imaging artifacts along the slice dimension [25–27]. This issue arises because direct slice-byslice synthesis fails to consider the spatial correlation of the 3D volume along the slice dimension. To mitigate such artifacts, we adapted [25] to incorporate an optimization strategy into each diffusion sampling step, leveraging a total variation (TV) prior exclusively imposed along the slice direction. This enforces spatial consistency in the 3D volume, particularly in the slice direction. To specify, let C(x, y, z) represent a 3D PET volume, and $X_T(x, y, z)$ denote a 3D latent variable sampled from an isotropic Gaussian distribution, where x, y, and z denote physical coordinates. The diffusion sampling process iteratively updates the entire $X_t(x, y, z)$ volume for t ranging from T to 0. In each iteration, each slice of $X_t(x, y, z)$ is independently updated as follows

$$\boldsymbol{Y}_{t-1}(x, y, z_i) = \mu_{\boldsymbol{\theta}}(\boldsymbol{X}_t(x, y, z_i), t, \boldsymbol{C}(x, y, z_i)) + \sigma_t \boldsymbol{n},$$
(12)

where $\boldsymbol{n} \sim \mathcal{N}(0, \boldsymbol{I})$, and $z_i = 1, \dots, N_z$ with N_z being the total number of slices. Note that $\boldsymbol{X}_t(x, y, z_i)$ and $\boldsymbol{C}(x, y, z_i)$ are 2D images, allowing the use of standard 2D diffusion models. Subsequently, we solve the optimization problem below to obtain $\boldsymbol{X}_{t-1}(x, y, z)$

$$\begin{aligned} \mathbf{X}_{t-1}(x, y, z) &= \\ &\arg\min_{\mathbf{X}(x, y, z)} \|\mathbf{X}(x, y, z) - \mathbf{Y}_{t-1}(x, y, z)\|_{2}^{2} + \tau \|\mathbf{D}_{z}\mathbf{X}(x, y, z)\|_{1} \end{aligned} \tag{13}$$

where D_z denotes a gradient operator along the z dimension, and τ is a trade-off parameter. In practice, we use gradient descent to solve this optimization problem and optimize τ through grid-search on a validation set.

2.4 Anatomically-guided Reconstruction Method

The MRIg-PET image reconstruction method used in this work was the Bayesian *maximum-a-posteriori*, including the MRI information modeled as a Markov random field. The prior model was formulated as

$$R(u) = \sum_{j}^{N} \phi(\sum_{b \in N_j} \xi_{jb} \omega_{jb} \psi(u_j - u_b)), \qquad (14)$$

where ϕ is the potential function operating in N_j around voxel index j, ξ_{jb} is a coefficient defined as $1/2N_j\sigma_u^2$ (σ_u is the standard deviation in the neighborhood N_j in the PET image u), ω_{jb} is a similarity coefficient, and ψ is the quadratic similarity relation. ψ is weighted by ω_{jb} based on the Burg joint entropy between the PET and MRI images [3], defined as

$$\omega_{jb} = \frac{1}{p(u_j, v_j)} \exp(-\frac{(u_j - u_b)^2}{2\sigma_u^2}) \exp(-\frac{(v_j - v_b)^2}{2\sigma_v^2}).$$
(15)

where \boldsymbol{v} is the MRI image, σ is the standard deviation in N_j in \boldsymbol{u} and \boldsymbol{v} , and $p(u_j, v_j)$ is defined as a non-parametric Parzen window. The weights are constrained to only those voxels which fulfill a certain level of similarity [18]. The optimization was written as a preconditioned gradient ascend [19].

The derivative of R(u) is included in the reconstruction with a hyper-parameter which balances the influence of the prior and the measured data. The hyper-parameter was individually recalculated at each iteration (k) for each voxel (j), depending on the image noise following

$$\beta_j^k = \alpha \sqrt{z_j^{k-1} \sum_{i=1}^M g_{ij} n_i a_i},\tag{16}$$

where α is a constant (1 in this work), g is the system model, n is the normalization and a is the attenuation. The weight z_j^{k-1} is calculated based on the MRI and PET information as follows: First, the MRI image intensity is scaled to the dynamic range of the PET image at each iteration. Secondly, each voxel is filtered by calculating the mean intensity inside the region of interest determined by similar intensities in the anatomical image [20].

2.5 Effect of Decimated Datasets

One of the potential applications of anatomically-guided PET reconstruction is the possibility of reducing the scan time or injected activity in the subject. Therefore, low counts applications render to investigate the performance of the AI model with reduced counts datasets. For that purpose, we randomly decimated the acquired listmode files, from 100% of the counts, down to 25% and 5% of the original counts in the eight subjects used for validation. The validation was performed visually, as well as quantitatively via volume of interest (VOI) analysis using Statistics Parametric Mapping (SPM12) [21]. The analysed VOIs comprise a range of cortical and sub-cortical regions of different sizes.

2.6 Subjective Image Quality Evaluation

Reconstructed PET images using OSEM and MRIg-PET (with the acquired and the deep-MRI images) of the eight datasets used for evaluation were inspected by two board-certified nuclear medicine physicians with >10 years of experience. All the images were evaluated using a 4-point scale (excellent, good, poor, nondiagnostic) on noise, sharpness, general image quality, and confidence. The evaluation was performed on the three dose levels: 100%, 25% and 5%. A paired t-test was performed between patients scored by both physicians reconstructed with the three methods to determine the statistical significance.

3 Results

3.1 MRI Analysis

Figure 2 shows an example of an acquired T1w-MRI image and its paired result from the DPM, together with the PET reconstructed image using OSEM as input to the model, the relative difference between both MRI images, and the relative difference masked by the gray and white matter masks. The masks of the segmented gray and white matter of the acquired T1w-MRI is shown in figure 2, together with the histograms of the gray and white matter from the acquired and deep-MRI images.



Figure 2: MRI acquired from scanner (top left), deep-MRI (top center), relative difference (top right), reconstructed PET image using OSEM (bottom left), gray and white matter masks (bottom center) and histogram of gray and white matter from acquired and deep-MRI images (bottom right) of an exemplar subject.

Visually, both MRI images show high structural resemblance, but the acquired MRI shows better soft tissue contrast and spatial resolution than the deep-MRI. The relative difference shows that the highest error is visible outside the brain area, where voxels with low intensity values are predominant. Overall, the mean relative differences in the gray and white matter between the acquired MRI and the deep-MRI are -3.4% and 1.4% respectively. The histograms show similar shapes but the histogram corresponding to the deep-MRI image confirms the aforementioned reduced contrast between gray and white matter.

3.2 Qualitative PET Image Quality Comparison

Figure 3 shows the axial slices of the PET reconstructed images obtained with OSEM and MRIg-PET using the acquired and deep-MRI images, together with the relative difference measured between the two PET reconstructed images using the acquired and deep-MRI images. Both MRIg-PET images show high similarities, and both show reduced noise and higher sharpness than the OSEM image. Visually, relative differences do not exceed $\sim \pm 20\%$, showing the largest differences in areas with low PET signal.



Figure 3: Axial slices at different levels of reconstructed PET image using OSEM (first column), MRIq-PET image obtained using the acquired MRI (second column), MRIg-PET image obtained using the deep-MRI (third column), and relative difference between the MRIg-PET images obtained with the acquired and deep-MRI images.

3.3 PET Volume of Interest Analysis

We measured the mean and standard deviation in each VOI of the eight subjects used for validation, normalized to the Hammers atlas in MNI space. Figure 4 shows the mean relative difference between the mean values measured in each ROI between the MRIg-PET images using the acquired and the deep-MRI images, and between the MRIg-PET image and OSEM images. Figure 4 also shows the mean coefficient of variation (CV), measured in each VOI for OSEM and MRIg-PET using the acquired and the deep-MRI images. The VOIs used in the analysis are shown on the right side of figure 4.

The VOIs in figure 4 are presented in order of increasing volume from the amygdala ($\sim 2.3 \text{ mL}$) up to the cerebellum ($\sim 114 \text{ mL}$). The relative differences show a correlation with volume size, showing large differences between mean intensities obtained with the different reconstructions for the smallest VOIs, and small differences for the larger regions. The differences between MRIg-PET using the different MRI images show smaller differences than between MRIg-PET images with the acquired MRI and OSEM. The CVs obtained with MRIg-PET, independently of the MRI image used, were consistently lower than with OSEM.



Figure 4: Mean relative difference between MRIg-PET reconstruction using the acquired and deep-MRI images and between MRIg-PET reconstruction using the the acquired MRI and OSEM for each ROI, averaged between eight subjects (top left). CV measured in each ROI obtained with MRIg-PET reconstruction using the acquired and deep-MRI images, and OSEM, averaged between eight subjects (bottom left).

3.4 Performance with Decimated Datasets

Figure 5 shows the reconstructed PET axial slices of the same patient as figure 3, using 100% of the acquired counts, and the corresponding images obtained with 25% and 5% of the acquired counts.

Visually the MRIg-PET reconstructed images using the acquired and deep-MRI show high resemblance. The relative differences show errors up to 20% in the cold regions, mainly where the CSF is located, and lower in the hot and warm regions, corresponding to the gray and white matter. The MRIg-PET reconstructed images show overall similar improvements compared to the OSEM-reconstructed image in terms of noise and image sharpness.

Figure 6 shows the relative difference between mean intensities measured in multiple VOIs, between the PET reconstructed images obtained with the acquired MRI and the deep-MRI. In addition, the improvement in noise of each MRI-guided PET reconstructed image compared to OSEM is also shown in figure 6. Comparing the mean intensity in the VOIs of the PET reconstructed images obtained with the measured and deep-MRI images, the largest differences are correlated with the VOI volume (with the exception of the gray and white matter), showing the smallest regions the largest differences. The largest differences were measured in the amygdala ($\sim 2\%$), followed by the gray and white matter. We did not observe systematic differences between the different levels of counts in the datasets, suggesting that the quality of the deep-MRI is relatively independent on the PET image noise.

The noise measured between OSEM and the MRIg-PET images (with both MRIs) was similar for all the VOIs, showing the largest difference in the amygdala, and consistently having higher improvements between OSEM and either MRIg-PET method for the lowest level of counts and lower improvement for the highest level of counts.

3.5 Subjective Image Quality Evaluation

In order to confirm if the quantitative measurements from Sections 3.3 and 3.4 can be transferred to visual improvement from a clinical perspective, the images of the eight subjects used for validation were inspected by two physicians. The results obtained from the two physicians and the eight patients were combined to



Figure 5: Axial slices of PET reconstructed images from the full dataset (first row), 25% of the counts (second row), 5% of the counts (third row), reconstructed with OSEM (left column), and with MRIg-PET using the acquired MRI (second column) and the deep-MRI (third column). The relative difference between the PET reconstructed images using MRIg-PET with the measured and deep-MRI are shown in the fourth column.

extract the mean and standard deviation for each reconstruction algorithm and each dose level. Figure 7 shows the resulting combined scores for noise, sharpness, image quality and confidence, for the three dose levels: 100%, 25% and 5%. The statistical significance is indicated as ns (no significance), * (p<0.05), ** (p<0.01) and *** (p<0.001).

PET images reconstructed with OSEM were scored as good (2) for all the figures of merit with 100% and 25% of the counts, while MRIg-PET images were scored as excellent (1) for noise and as good (2) for sharpness, image quality and confidence, for 100% and 25% of the counts. However, MRIg-PET images were overall scored consistently slightly higher (lower score) than OSEM images. MRIg-PET shows overall superior performance compared to OSEM in noise, sharpness and image quality as shown in figure 7. However, the three algorithms show similar confidence for the 100% and 50% dose levels. With respect to the two MRIg-PET algorithms, the physicians rated the images similarly, showing in some cases better scores for the MRIg-PET reconstruction using the acquired MRI, but statistically not significant. Most differences were found between OSEM and MRIg-PET in the case of noise for all the dose levels. Additionally, significant differences were found between OSEM and MRIg-PET for all the figures of merit for the lowest dose level of 5%.

3.6 Results with Defective MRI

The poor quality in MRI images may happen due to subject motion or hardware problems. For instance, the MRI-T1w sequence takes 5-min to be acquired. During that time, the subject may move the head causing motion artifacts and resulting in poor image quality. Such images may potentially introduce artifacts in the PET images, when combined during the MRI-guided PET reconstruction process.

Figure 8 shows an example of a 12-year old male which shows strong artifacts in the MRI-T1w images due to motion, which are transferred to the MRIg-PET image. The deep-MRI image shows lower image contrast but no motion artifacts. Therefore, the MRIg-PET reconstructed image obtained with the deep-MRI image does not show artifacts, and shows higher resemblance to the OSEM images, compared to the MRIg-PET image based on the acquired MRI image. Some regions with visible artifacts in the MRIg-PET images.



Figure 6: Mean intensity difference between PET images obtained with the acquired MRI and deep-MRI for different regions, obtained for 100%, 25% and 5% of the measured counts (top). Noise difference between OSEM and the MRIg-PET reconstructed image obtained with the acquired MRI (bottom left) and deep-MRI (bottom right) for different regions, obtained for 100%, 25% and 5% of the measured counts.

4 Discussion

The use of low noise and 3D high spatial resolution MRI to guide the PET reconstruction has extensively been investigated to regularize the ill-posed inversion problem of PET image reconstruction. There are a considerable amount of algorithms that formulate the problem in different ways. However, all those methods require a paired MRI with a high level of correlation between the functional or metabolic information from PET and anatomical information from MRI, which greatly limits the application of MRI-guided PET reconstruction.

Diffusion probabilistic models (DPM) are emerging as a method to produce natural realistic images with high spatial resolution. Therefore, we employed in this work DPM to generate highly realistic MRI images from nothing else than PET images as input, in order to use as reference to guide the PET image reconstruction, rather than for any clinical purpose. The premise is that there needs to be be a high correlation between the patterns in the PET and the MRI images, as is usually the case in FDG-PET and MRI-T1w.

In MRI-guided PET reconstruction, a certain level of information is transferred from the MRI image to the PET image. If too much information is transferred, the PET image may resemble an MRI image while if too little information is transferred, the PET image will be similar to the OSEM-PET image. Usually, the level of transferred information is limited in order to keep the image as PET-like as possible, while improving the image quality. The purpose of this work was to generate an MRI image of enough realism and quality to be able to guide the PET image reconstruction without introducing hallucinations, not as accurate as the acquired MRI images, but still improving the PET image quality.

The resulting MRI-T1w images obtained from DPM highly resembled the acquired MRI-T1w images, showing in general similar noise properties. Spatial resolution was slightly degraded and contrast between soft tissues decreased in the deep-MRI images compared to the acquired MRI images. However, by inspecting the MRIg-PET images obtained with the acquired MRI and the deep-MRI, small differences were observed, suggesting that the level of degradation in the MRI images was not enough to be transferred to the PET images, while still improving the PET image quality. This was confirmed by a VOI analysis performed in multiple cortical and subcortical brain regions.



Figure 7: 4-point scale on noise, sharpness, image quality and confidence evaluated for OSEM, MRIg-PET using the acquired MRI, and MRIg-PET using the deep-MRI, for three dose levels. The statistical significance between OSEM and MRIg-PET with the acquired MRI, and between the two MRIg-PET images is shown for each plot.

Similar findings were observed in decimated datasets down to 25% and 5% from the original number of counts. The largest differences in mean intensity between MRIg-PET images with the acquired and deep-MRI were observed to be $\sim 2\%$ in the amygdala and white matter. Comparing the noise measured in all the regions between OSEM and both MRIg-PET, the improvement introduced by MRIg-PET was similar in all the regions.

To confirm if the quantitative observations could be transferred to actual clinical readings, two physicians blindly evaluated all the images using a 4-point score on noise, sharpness, image quality and confidence on the reading. Their observations confirmed that the MRIg-PET images obtained with the acquired and the deep-MRI showed high resemblance in the four figures of merit, with no statistical significance between images. Comparing to OSEM, the only consistent observations were statistically significant improvement of MRIg-PET compared to OSEM in noise, and for the four figures of merit measured on the images reconstructed with 5% of the original counts. Besides statistical measurements, MRIg-PET images in general scored better (lower score - better image quality) than OSEM images. The more similar scores between OSEM and MRIg-PET were obtained in the reading confidence.

Finally we also highlighted the benefits of using a deep-MRI in PET/MRI cases where the MRI was defective. While hardware problems are not common, patient motion is often encountered, especially in patients who undergo neurological disorders. The MRIg-PET obtained with deep-MRI showed higher resemblance to the OSEM image compared to the acquired MRI in terms of hallucinations, while preserving the expected improved image quality from MRIg-PET compared to OSEM. It is important to highlight that patient motion during the PET measurement was not addressed in this work, and remains future work. It is also important to note that subject motion is a more critical problem to address when using anatomical information in the reconstruction compared to reconstruction algorithms without anatomical guidance, since misregistrations can potentially introduce strong artifacts in the PET reconstructed images.

This study considers a relatively small PET/MRI training dataset acquired from 25 subjects. We observed that this small training dataset is sufficient for the diffusion model to generate realistic MRI-T1w images and perform well on the downstream task of deep-MRI-guided PET reconstruction. However, despite the high perceptual quality, we identified quantitative errors in the synthesized MRI images when compared to the acquired MRI-T1w images. We hypothesize that this discrepancy may be attributed to the breath of the data. Our future work involves applying this study to a dataset acquired from a larger number of subjects, with the potential to mitigate such quantitative errors in MRI synthesis.



Figure 8: Axial slices of an exemplar subject who moved during the acquisition of the MRI-T1w sequence. The first column is the originally acquired MRI-T1w images, the second column is the deep-MRI-T1w images, the third column is the PET reconstructed images using OSEM, the fourth column is the MRIg-PET reconstructed image with the acquired MRI-T1w image, and the fifth column is the MRIg-PET reconstructed image with the deep-MRI-T1w image

The work presented here was only explored for FDG scans, but it could be in principle extended to other radio-tracers, as long as the PET and MRI images present high correlation. A limitation of this approach is cases where PET and MRI images show different patterns or structures, such as cases with PET-only or MRI-only lesions.

5 Conclusion

We showed in this work a method to produce pseudo-MRI-T1w images from PET-FDG reconstructed images using a diffusion probabilistic model for MRI-guided PET image reconstruction. Resulting images were compared to reconstructed images using the acquired MRI-T1w, reaching the same image quality, and always improving the image quality achieved with OSEM. The method was shown to work successfully for different levels of PET counts down to 25% of the counts. Improvements for the 5% datasets were also observed, but the images were overall rated as poor.

Acknowledgments

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests with the work reported in this paper: Jorge Cabello and Carl von Gall are full-time employees, Siemens Medical Solutions USA, Inc. Günther Platsch is full-time employee, Siemens Healthineers AG, Germany.

References

- C. Comtat, P. E. Kinahan, J. A. Fessler, T. Beyer, D. W. Townsend, M. Defrise and C. Michel, Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels, Physics in Medicine and Biology, vol. 47, no. 1, 2002.
- [2] Y. J. Jeong, H. J. Yoon, and D. Y. Kang, Assessment of change in glucose metabolism in white matter of amyloid-positive patients with Alzheimer disease using F-18 FDG PET, Medicine, vol. 96, no. 48, 2017.

- [3] A. Mehranian, M. Belzunce, F. Niccolini, M. Politis, C. Prieto, F. Turkheimer, A. Hammers, A. Reader, "PET image reconstruction using multi-parametric anato-functional priors," in Physics in Medicine and Biology, 2017, 62, pp. 5975.
- [4] G. Schramm, M. Holler, A. Rezai, K. Vunckx, F. Knoll, K. Bredies, F. Boada, J. Nuyts, "Evaluation of Parallel Level Sets and Bowsher's Method as Segmentation-Free Anatomical Priors for Time-of-Flight PET Reconstruction," in IEEE Trans Med Imaging. 2018 Feb; 37(2): 590–603.
- [5] Y. J. Tsai, G. Schramm, S. Ahn, A. Bousse, S. Arridge, J. Nuyts, B. F. Hutton, C. W. Stearns, K. Thielemans, "Benefits of Using a Spatially-Variant Penalty Strength With Anatomical Priors in PET Reconstruction," in IEEE Trans Med Imaging. 2020 Jan; 39(1): 11–22.
- [6] H. Choi and D. S. Lee, Generation of Structural MR Images from Amyloid PET: Application to MR-Less Quantification, Journal of Nuclear Medicine, vol. 59, no. 7, 2018.
- [7] J. Zhang, X. He, L. Qing, F. Gao, and B. Wang. BPGAN: Brain PET synthesis from MRI using generative adversarial network for multi-modal Alzheimer's disease diagnosis, Computer Methods and Programs in Biomedicine, vol. 217, 2022.
- [8] G. Schramm, D. Rigie, T. Vahle, A. Rezaei, K. Van Laere, T. Shepherd, J. Nuyts, F. Boada. Approximating anatomically-guided PET reconstruction in image space using a convolutional neural network, Neuroimage, 224:117399, 2021.
- [9] F. Croitoru, V. Hondru, R. T. Ionescu, and M. Shah, "Diffusion models in vision: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 9, pp. 10850–10869, 2023.
- [10] H. Cao et al., "A Survey on Generative Diffusion Model," arXiv:2209.02646, Jul. 2023.
- [11] A. Kazerouni et al., "Diffusion Models for Medical Image Analysis: A Comprehensive Survey," Medical Image Analysis, 2023.
- [12] L. Jiang, Y. Mao, X. Chen, X. Wang, and C. Li, "CoLa-Diff: Conditional Latent Diffusion Model for Multi-Modal MRI Synthesis," arXiv:2303.14081, Mar. 2023.
- [13] S. Pan et al., "Cycle-guided Denoising Diffusion Probability Model for 3D Cross-modality MRI Synthesis," arXiv:2305.00042, 2023.
- [14] J. Ouyang et al., "Predicting FDG-PET images from multi-contrast MRI using deep learning in patients with brain neoplasms," Journal of Magnetic Resonance Imaging, 2023.
- [15] T. Xie et al., "Synthesizing PET images from High-field and Ultra-high-field MR images Using Joint Diffusion Attention Model," arXiv:2305.03901, 2023.
- [16] S. Pan et al., "Full-dose PET Synthesis from Low-dose PET Using High-efficiency Diffusion Denoising Probabilistic Model," arXiv:2308.13072, 2023.
- [17] K. Gong, K. A. Johnson, G. E. Fakhri, Q. Li, and T. Pan, "PET image denoising based on denoising diffusion probabilistic models," European Journal of Nuclear Medicine and Molecular Imaging, 2023.
- [18] J. E. Bowsher *et al.*, "Utilizing MRI information to estimate F18-FDG distributions in rat flank tumors," IEEE Symposium Conference Record Nuclear Science 2004., Rome, Italy, 2004, pp. 2488-2492 Vol. 4, doi: 10.1109/NSSMIC.2004.1462760.
- [19] J. Nuyts, D. Beque, P. Dupont and L. Mortelmans, "A concave prior penalizing relative differences for *maximum-a-posteriori* reconstruction in emission tomography," in IEEE Transactions on Nuclear Science, vol. 49, no. 1, pp. 56-60, Feb. 2002, doi: 10.1109/TNS.2002.998681.

- [20] J. Cabello, M. T. Jurkiewicz, A. Andrade, T. L. S. Benzinger, H. An and U. C. Anazodo, "Evaluation of an MRI-Guided PET Image Reconstruction Approach With Adaptive Penalization Strength," in IEEE Transactions on Radiation and Plasma Medical Sciences, doi: 10.1109/TRPMS.2024.3352983.
- [21] W. Penny, K. Friston, J. Ashburner, S. Kiebel and T. Nichols, "Statistical Parametric Mapping: The Analysis of Functional Brain Images", Elsevier, 2007.
- [22] P. Dhariwal, and A. Nichol. "Diffusion models beat gans on image synthesis." Advances in neural information processing systems 34 (2021): 8780-8794.
- [23] A. Q. Nichol, and P. Dhariwal. "Improved denoising diffusion probabilistic models." International conference on machine learning. PMLR, 2021.
- [24] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," in Proc. Adv. Neural Inf. Process. Syst., 2020, pp. 6840–6851.
- [25] H. Chung, D. Ryu, M. T. McCann, M. L. Klasky, J. C. Ye. "Solving 3d inverse problems using pretrained 2d diffusion models," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023.
- [26] H. Xie, W. Gan, B. Zhou, X. Chen, Q. Liu, X. Guo, L. Guo, H. An, U. S. Kamilov, G. Wang, C. Liu. "DDPET-3D: Dose-aware Diffusion Model for 3D Ultra Low-dose PET Imaging," arXiv:2311.04248, 2023.
- [27] Y. Hu, S.V. Kothapalli, W. Gan, A.L. Sukstanskii, G.F. Wu, M. Goyal, D.A. Yablonskiy and U.S. Kamilov "DiffGEPCI: 3D MRI Synthesis from mGRE Signals using 2.5 D Diffusion Model," arXiv:2311.18073, 2023.