List-Mode PET Image Reconstruction Using Deep Image Prior

Kibo Ote, Fumio Hashimoto, Yuya Onishi, Takashi Isobe, and Yasuomi Ouchi

Abstract—List-mode positron emission tomography (PET) image reconstruction is an important tool for PET scanners with many lines-of-response and additional information such as time-of-flight and depth-of-interaction. Deep learning is one possible solution to enhance the quality of PET image reconstruction. However, the application of deep learning techniques to list-mode PET image reconstruction has not been progressed because list data is a sequence of bit codes and unsuitable for processing by convolutional neural networks (CNN). In this study, we propose a novel list-mode PET image reconstruction method using an unsupervised CNN called deep image prior (DIP) which is the first trial to integrate list-mode PET image reconstruction and CNN. The proposed list-mode DIP reconstruction (LM-DIPRecon) method alternatively iterates the regularized list-mode dynamic row action maximum likelihood algorithm (LM-DRAMA) and magnetic resonance imaging conditioned DIP (MR-DIP) using an alternating direction method of multipliers. We evaluated LM-DIPRecon using both simulation and clinical data, and it achieved sharper images and better tradeoff curves between contrast and noise than the LM-DRAMA, MR-DIP and sinogram-based DIPRecon methods. These results indicated that the LM-DIPRecon is useful for quantitative PET imaging with limited events while keeping accurate raw data information. In addition, as list data has finer temporal information than dynamic sinograms, list-mode deep image prior reconstruction is expected to be useful for 4D PET imaging and motion correction.

Index Terms—Deep neural network, image reconstruction, list-mode, positron emission tomography, unsupervised learning.

I. INTRODUCTION

POSITRON emission tomography (PET) is an in-vivo imaging tool with a wide range of clinical applications such as oncology, cardiology, and neurology [1]. In PET scan, firstly a radiotracer is injected into a subject. Then, the radiotracer emits a positron, and an annihilation between the positron and electron generates two 511 keV gamma rays. The PET detector ring catches two gamma rays as a coincidence event (i.e., the two annihilation photons (or gamma rays) are detected within a time coincidence window). The coincidence event indicates that an activity is on the line-of-response (LOR), but it cannot be determined where the activity is on the LOR. Hence, an image reconstruction process is required for tomographic imaging.

Positron emission tomography image reconstruction is a challenging task because the number of events is limited, and it is also an ill-posed inverse problem. Regularization using prior information about PET images is often used to suppress the ill-posedness and statistical noise. For example, the spatial smoothness of an image is used to regularize the PET image reconstruction through Gibbs priors [2], [3]. In addition, anatomical information from X-ray computational tomography (CT) or magnetic resonance imaging (MRI) can be used for edge-preserving regularization of PET image reconstruction [4], [5], [6]. Besides these conventional algorithms, some deep learning-based approaches have recently been proposed for PET image reconstruction and provide excellent performance [7], [8], [9], [10].

One strategy of deep learning-based PET image reconstruction is to learn the image reconstruction from sinograms through large amounts of training datasets [11], [12]. For example, DeepPET has been proposed [13], which performs a direct mapping from a sinogram space into reconstructed image space, using a fully convolutional neural network (CNN). It is easy to handle with CNN because both the sinogram and reconstructed image are represented by the tensor. In another strategy, [14] proposed a regularization method that uses a CNN representation as prior information. However, these deep learning-based methods still require large amounts of datasets for network training. In recent years, unsupervised learning has attracted attention for PET image reconstruction, particularly deep image prior (DIP) [15], which uses the CNN architecture itself as a regularization, is used as a new regularization tool for PET image reconstruction because
it requires no prior training data [16], [17], [18]. However, the number of bins of a sinogram is dependent on the number of LORs, and further increases in orders of magnitude if the PET scanner has the ability to measure additional information, such as time-of-flight (TOF) and depth-of-interaction (DOI). This is because the TOF and DOI increase the dimension of the sinogram and increase the number of LORs. Hence, the sinogram-mode PET image reconstruction becomes memory intensive and sometimes infeasible, when the PET scanner has many LORs and additional information.

List-mode PET image reconstruction is the promising solution for TOF- and DOI-PET image reconstruction when the number of events per frame becomes low [19] and is suitable for motion correction because the list data holds fine temporal information. Hence, improving the list-mode PET image reconstruction is important for advanced PET scanners. However, the incorporation of deep learning techniques to list-mode PET image reconstruction has not advanced because list data is a sequence of bit codes and is not suitable for processing by CNNs.

To surpass the above barriers, in this study, we propose list-mode PET image reconstruction using DIP which is the first trial to integrate list-mode PET image reconstruction and CNN. The proposed method, called LM-DIPRecon, alternatively iterates list-mode PET image reconstruction and image processing by DIP according to an alternating direction method of multipliers (ADMM) [20]. The proposed LM-DIPRecon is a strong algorithm for PET image reconstruction as keeping accurate raw data information while a sinogram-based reconstruction loses some information to calculate practical and manageable sinograms. We evaluated the proposed LM-DIPRecon method using Monte-Carlo simulation and clinical data. The LM-DIPRecon method reduces noise and maintains the contrast compared with post-processing by the DIP and sinogram-based DIPRecon methods.

II. RELATED WORKS

In this section, we introduce list-mode PET image reconstruction and DIP.

A. List-Mode PET Image Reconstruction

The list data is formulated as,

\[ U = \{ i (t) \mid t = 1, 2, \ldots, T \}, \]

where \( i \) is an index of a coincidence event, \( T \) is the number of events, and \( i (t) \) is an index of LOR measuring the \( t \)-th event.

The list-mode log-likelihood function [21] is defined as,

\[ L(U \mid x) = \sum_i \sum_j a_{i(t)} j x_j - \sum_i \sum_j a_{ij} x_j, \]

where \( x \) is an unknown image, \( j \) is an index of a voxel, and \( a_{ij} \) represents the contribution of the voxel \( j \) to the LOR \( i \).

We can reconstruct the image from the list data by maximizing the list-mode log-likelihood function with an iterative algorithm. For fast global optimization, the list-mode dynamic row action maximum likelihood algorithm (LM-DRAMA) has been developed [22], [23]. The reconstructed image \( x_{EM} \) is calculated using LM-DRAMA updates, as follows:

\[ x^{(k,q+1)}_{j,EM} = x^{(k,q)}_{j} + x^{(k,q)}_{j} \lambda (k,q) \left( \frac{N_{sub}}{S} \sum_{r \in \text{Sub}_q} a_{r(t)} j \right) - 1, \]

\[ S_j = \sum_i a_{ij}, \]

\[ y^{(k,q)}_{j} = \sum_j a_j x^{(k,q)}_j, \]

\[ \lambda^{(k,q)} = \frac{\beta}{\beta + q + \gamma k N_{sub}}, \]

where \( k \) is the number of main iteration, \( q \) is the number of sub-iteration, \( N_{sub} \) is the number of subsets of list data, \( S \) is the sensitivity image [24], \( y \) is the forward projection of the image, \( \lambda \) is a subset-dependent relaxation coefficient, \( \text{Sub}_q \) is a subset of list data accessed at the \( q \)-th sub-iteration, and \( \beta \) and \( \gamma \) are the parameters of the relaxation coefficient.

B. Supervised Deep Learning

The supervised deep learning approach trains a CNN to map a degraded image to a clean image, and requires large amounts of pairs of degraded and clean images as the dataset.

Supervised deep learning is formulated as,

\[ \arg\min_\theta \sum_{x \in D} \| f_\theta (x^s_0) - x^s_{\text{ref}} \|^2, \]

where \( f \) is a CNN, \( \theta \) is a parameter of the CNN such as the connection weight, and \( x^s_0 \) and \( x^s_{\text{ref}} \) are \( s \)-th pair of measured degraded image and clean image in the dataset \( D \).

C. DIP

In contrast to supervised deep learning, an unsupervised deep learning technique requires only the degraded image as the dataset without clean images [25]. Further, the deep image prior requires no prior training data other than the measured degraded image itself [15]. The DIP method optimizes a CNN to map a prior distribution image to a measured degraded image. When there is no prior information, DIP inputs a random image noise to the CNN as the prior distribution image. CNNs have the property of restoring signals faster than noise [15]. By stopping the optimization early, we can obtain a restored image as an output of the CNN.

The DIP method is formulated as,

\[ \arg\min_\theta \| f_\theta (z) - x^s_0 \|^2, \]

where \( z \) is a prior distribution image.

In addition to the network architecture, the prior distribution image is the key factor that decides the performance of the DIP. For example, if we use the CT or MRI as the prior information [26], the recovery of PET images by DIP becomes faster and more accurate than the case of no prior information. For dynamic PET image denoising, we can also use the static PET image as the prior information [27].
III. METHODOLOGY

In this section, we integrate list-mode PET image reconstruction and DIP into one algorithm using ADMM.

A. Integration of LM-DRAMA and DIP

To regularize list-mode PET image reconstruction using DIP, we consider the following constrained optimization problem.  
\[
\max_{x,\theta} L(U \mid x) \quad \text{s.t.} \quad x = f_\theta(z),
\]
where the unknown image \( x \) is constrained to the output of the CNN.

The constrained optimization problem in (9) can be rewritten by the augmented Lagrangian method.  
\[
\min_{x,\theta} \max_{\mu} -L(U \mid x) + g^T (x - f_\theta(z)) + \frac{\rho}{2} \| x - f_\theta(z) \|^2
\]
where \( \rho \) is a positive constant and \( g \) is a dual variable or Lagrange multipliers. The optimization problem of (10) can be written in a slightly different form [20].

\[
\min_{x,\theta} \max_{\mu} -L(U \mid x) + \frac{\rho}{2} \| x - f_\theta(z) \|^2 + \frac{\rho}{2} \| \mu \|^2
\]
where \( \mu = g/\rho \) is a scaled dual variable.

The ADMM splits the optimization problem of (11) into the two subproblems and the update of scaled dual variable [20].

\[
x^{(n+1)} = \arg\max_x L(U \mid x) - \frac{\rho}{2} \| x - f_{\theta^{(n)}}(z) + \mu^{(n)} \|^2,
\]
\[
\theta^{(n+1)} = \arg\min_{\theta} \| f_\theta(z) - x^{(n+1)} - \mu^{(n)} \|^2,
\]
\[
\mu^{(n+1)} = \mu^{(n)} + x^{(n+1)} - f_{\theta^{(n+1)}}(z),
\]
where \( n \) is the number of iterations of three steps of the ADMM.

B. Solving Subproblem (12)

The subproblem (12) can be solved by the optimization transfer method [16], [28], [29]. Firstly, to optimize each voxel independently, we construct the surrogate function of \( L(U \mid x) \) as,

\[
Q_L(x \mid x^{(n)}) = \sum_{j} S_j \left( x_{j,EM}^{(n+1)} \log x_j - x_j \right).
\]
where \( x_{j,EM}^{(n+1)} \) is an updated image of \( x^{(n)} \) by (3).

The \( Q_L(x \mid x^{(n)}) \) will suffice the following two conditions from the convergence property of the DIP [30].

\[
Q_L(x \mid x^{(n)}) - Q_L(x^{(n)} \mid x^{(n)}) \leq L(U \mid x) - L(U \mid x^{(n)}),
\]
\[
\nabla Q_L(x^{(n)} \mid x^{(n)}) = \nabla L(U \mid x^{(n)}).
\]

By replacing the \( L(U \mid x) \) with the surrogate function \( Q_L(x \mid x^{(n)}) \), we obtain the surrogate objective function of subproblem (12) as,

\[
P \left( x_j \mid x^{(n)} \right) = S_j \left( x_{j,EM}^{(n+1)} \log x_j - x_j \right) - \frac{\rho}{2} \left( x_j - x_{j,base}^{(n)} \right)^2,
\]
\[
x_{j,base}^{(n)} = f_{\theta^{(n)}}(z) - \mu^{(n)}_j,
\]
where \( x_{j,base} \) is a base image of the regularization.

The surrogate objective function (18) can be optimized voxel by voxel. Hence, by setting the derivative of equation (18) to be zero, we can obtain the following update equation.

\[
x_{j}^{(n+1)} = x_{j,base}^{(n)} - \frac{S_j}{\rho} + \sqrt{\left( \frac{S_j}{\rho} \right)^2 + 4 \left( x_{j,EM}^{(n+1)} \right)^2} / 2.
\]

As a consequence, the algorithm of subproblem (12) becomes a combination of LM-DRAMA and regularization by equation (20).

In this study, we set \( N_{\text{sub}} = 40, \beta = 30 \), and \( \gamma = 0.1 \) for the LM-DRAMA. These are the default settings of LM-DRAMA. In addition, we set \( \rho = 0.5 \) for the regularization based on the experimental results (Fig. 5).

C. Solving Subproblem (13)

Subproblem (13) can be solved by the DIP with \( x_{\text{label}} = x_{j,base}^{(n+1)} + \mu^{(n)}_j \). In this study, we employ the MR-DIP which uses the MR image as the prior distribution image. In addition, we use the 3D U-net architecture [31] for the MR-DIP because it is suitable for medical image processing, and it may have a good inductive bias for the PET image. Fig 1 shows the network architecture of the 3D U-net in this study. The 3D U-net consists of encoder, decoder, and skip connections.

The encoder extracts the feature maps to generate the PET image from the MR image. The combination of 3D convolution with a kernel size of three and a leaky rectified linear unit (LReLU) is repeated twice before down-sampling. Down-sampling is performed by the combination of 3D convolution with kernel size of four and stride of two and LReLU. At each down-sampling, the number of voxels along \( x, y, \) and \( z \) directions are halved and the number of channels of feature map is doubled.

The decoder reconstructs the PET image from the extracted feature maps. The combination of 3D convolution with kernel
size of three and LReLU is repeated twice before up-sampling. Up-sampling is performed by the 3D convolution with kernel size of one and LReLU and trilinear interpolation. At each up-sampling, the number of voxels along x, y, and z directions are doubled, and the number of channels of feature map is halved. The output image is reconstructed by 3D convolution with kernel size of three with a single channel output.

The skip connection adds the feature maps before down-sampling to the feature maps before up-sampling by bypassing as shown in Fig. 1.

To stabilize the training, we clip the gradient norm at 1.0 [32] and take an exponential moving average (EMA) of the network output [15] as,

\[ f_{\theta(n,m+1)}(z) = \eta f_{\theta(n,m)}(z) + (1 - \eta) f_{\theta(n,m+1)}(z), \]

where \( \eta \) is a smoothing constant, \( f \) is an EMA of DIP outputs, \( m \) and \( M_2 \) are the number of sub-iteration and total sub-iteration of DIP for solving subproblem (13), respectively. In this study, we set \( \eta = 0.99 \).

**D. Overall Algorithm**

The overall algorithm of the LM-DIPRecon method is shown in Algorithm 1.

To ready the initial network \( \theta^{(0)} \), we take a two-step approach [16]. At first, we reconstruct the image with one main-iteration of the LM-DRAMA and perform the MR-DIP as,

\[
\arg\min_{\theta} \theta(n) = \left\| f_{\theta}(z) - \frac{x^{(1)}}{EM} \right\|^2. \tag{24}
\]

In this step, we update the network parameters by the Adam method [33] with 1000 epochs. We set one main-iteration of LM-DRAMA and 1000 epochs of MR-DIP as a relatively small number of iterations in the warmup of CNN because an initial image is preferred to be smooth.

In the second step, we iterate the three steps of the ADMM as shown in Algorithm 1. In each step of the ADMM, there is no need to solve the subproblems until convergence. For subproblem (12), we set \( M_1 = 2 \) and perform two sub-iterations of the regularized LM-DRAMA. For subproblem (13), we set \( M_2 = 10 \) and update the network parameters using the limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) method with 10 epochs and a learning rate of 1.0. As the L-BFGS method has a fast convergence rate when the parameters are near a local optimum, we use the Adam method in the first step and the L-BFGS method in the second step. We followed the previous study [16] for the settings of \( M_1 \) and \( M_2 \).

Last, after \( N \) iterations, the algorithm outputs the image as,

\[ \hat{x} = f_{\theta(N-1)}(z), \tag{25} \]

where \( \hat{x} \) is a final image.

**Algorithm 1 Algorithm of the LM-DIPRecon Method**

**First step**

1: Reconstruct the image with one main iteration of the LM-DRAMA.

2: Running Adam method with 1000 epochs to ready the initial network. MR image \( z \) is input.

\[ \theta^{(0)} = \left\| f_{\theta}(z) - \frac{x^{(1)}}{EM} \right\|^2. \]

**Second step**

Input: Iteration number \( N \), Sub-iteration numbers \( M_1 \) and \( M_2 \). Initial network \( \theta^{(0)} \), MR image \( z \), positive constant \( \rho \)

1: \( x^{(0)} = f_{\theta^{(0)}}(z) \)

2: \( \mu^{(0)} = 0 \)

3: for \( n = 0 \) to \( N - 1 \) do

4: \( x^{(n,0)} = x^{(n)} \)

5: \( x_{\text{base}} = f_{\theta^{(0)}}(z) - \mu^{(n)} \)

6: for \( m = 0 \) to \( M_1 - 1 \) do

7: \( u = nM_1 + m \)

8: \( k = \lceil u/N_{\text{sub}} \rceil \) where \( \lceil \cdot \rceil \) is round down.

9: \( q = u \% N_{\text{sub}} \)

10: \( \lambda = \frac{\rho + q + k}{N_{\text{sub}}} \)

11: \( x_{j,EM} = x_j^{(n,m)} + x_j^{(n,m)} \left( \frac{N_{\text{sub}}}{3} \sum_{t \in \text{Sub}_{q}} (x_t^{(n,m)}/\gamma^{(n,m)} - 1) \right) \)

12: \( x_j^{(n,m+1)} = x_j^{(n,m)} - \frac{S_j}{2} + \gamma x_j^{(n,m)} \left( x_j^{(n,m+1)}/\gamma^{(n,m)} - 1 \right) \)

13: end for

14: \( x^{(n+1)} = x^{(n,M_1-1)} \)

15: \( x_{\text{label}} = x^{(n,+1)} + \mu^{(n)} \)

16: Running L-BFGS method with \( M_2 \) iterations to get \( \theta^{(n+1)} = \| f_{\theta}(z) - x_{\text{label}} \|^2 \)

17: \( \mu^{(n+1)} = \mu^{(n)} + x^{(n,M_1-1)} - f_{\theta^{(n+1)}}(z) \)

18: end for

19: return \( \hat{x} = f_{\theta(N-1)}(z) \)

**IV. Experimental Setup**

We evaluated the proposed methods using the simulation data and clinical data. We compared the proposed method relative to the LM-DRAMA and the MR-DIP. Image size was \( 128 \times 128 \times 80 \) and voxel size was \( 2.6 \text{ mm} \times 2.6 \text{ mm} \times 2.4 \text{ mm} \). The details of the data correction can be seen in the Appendix.

**A. Simulation Data**

We performed Monte-Carlo simulation using our simulator. The simulation included the attenuation and scatter effects based on the delta-tracking method [34], [35]. The positron range, angular deviation and random events were not included in the simulation. We made a digital brain phantom from a segmented MR image downloaded from the BrainWeb [36]. The digital brain phantom was set to a contrast of \( 1.00:0.25:0.05 \) for gray matter (GM), white matter (WM) and CSF. In addition, we inserted the three tumor regions with a contrast of 1.0, 1.1, and 1.2. Attenuation coefficients were set to 0.00958 mm\(^{-1}\) for the soft tissue and CSF, and 0.0151 mm\(^{-1}\) for the bone tissue.
The detector arrangement was the same as our brain PET scanner (Hamamatsu HITS-655000 [37]) with four-layer DOI. The detector unit consists of $32 \times 32$ lutetium-yttrium orthosilicate (LYSO) scintillator array with a 1.2 mm pitch and 64 multi-pixel photon counters (MPPCs) in an $8 \times 8$ arrangement. To ensure that the load for the count rate in every layer is equally sensitive for gamma rays, the scintillator thickness was designed to be 3 mm, 4 mm, 5 mm, and 8 mm toward the outside of the field-of-view [37]. The scanner thickness was designed to be 3 mm, 4 mm, 5 mm, and 8 mm every layer is equally sensitive for gamma rays, the scintillator arrangement. To ensure that the load for the count rate in the denoising performance, we thinned the list data into $1/20$, $7 \times 10^7$ events. Furthermore, to evaluate the uptake of the QAA in the thalamus. The list data of 42-62 min had $4.7 \times 10^7$ events. To evaluate the uptake on the GM and tumor, $\bar{b}$ is the average of multiple ROI values set to WM, $K_a$ is the number of ROIs on the GM and tumor, $K_b$ is the number of ROIs on WM, $a$ is an ROI uptake on the GM and tumor, and $\bar{b}$ is an ROI value on the WM.

$$\bar{a} = \frac{1}{K_a} \sum_{k=1}^{K_a} a_k,$$

$$\bar{b} = \frac{1}{K_b} \sum_{k=1}^{K_b} b_k,$$

where $\bar{a}$ is the average of multiple region of interest (ROI) uptakes set to GM and tumor, $\bar{b}$ is the average of multiple ROI values set to WM, $K_a$ is the number of ROIs on the GM and tumor, $K_b$ is the number of ROIs on WM, $a$ is an ROI uptake on the GM and tumor, and $\bar{b}$ is an ROI value on the WM.

The MRI scan was performed using a 1.5 T MR scanner (Signa HDxt, GE, USA) with the following acquisition parameters: 3D mode sampling, Repetition Time/Time to Echo (TR/TE) (25/Minimum, T1-weighted), 30° flip angle, 1.5 mm slice thickness with no gap, and 256 $\times$ 256 matrices. The MR image was interpolated and aligned to the PET image using PMOD software (version 3.0, PMOD Technologies Ltd., Zurich, Switzerland) [39].

C. Evaluation Metrics

In the simulation study, we evaluated the signal-to-noise ratio (SNR) for quantitative evaluation.

$$\text{SNR} = 10 \log_{10} \left( \frac{1}{R} \sum_{j \in R} x_j \right)^2 \left( \frac{1}{R} \sum_{j \in R} (x_j - x_{j, \text{ref}})^2 \right),$$

where $x_{\text{ref}}$ is a phantom image, $R$ is a region of the whole brain, and $N_R$ is a number of voxels inside the whole brain. For tumor quantification, the tumor value ratio (TR) was calculated as

$$\text{TR} = \frac{1}{N_{\text{tumor}}} \sum_{j \in R_{\text{tumor}}} \frac{x_j}{x_{j, \text{ref}}},$$

where $R_{\text{tumor}}$ is a union of the three tumor regions and $N_{\text{tumor}}$ is a number of voxels inside the tumor regions.

In addition, we evaluated the curves of contrast recovery coefficient (CRC) versus normalized standard deviation (NSTD) to evaluate the tradeoff between contrast and noise. The CRC was calculated by the following equations.

$$\text{CRC} = \frac{\left( \frac{\bar{a}}{\bar{b}} - 1 \right)}{\left( \frac{\bar{a}_{\text{ref}}}{\bar{b}_{\text{ref}}} - 1 \right)},$$

$$\bar{a} = \frac{1}{N_{a,k}} \sum_{j \in R_{a,k}} x_j,$$

$$\bar{b} = \frac{1}{N_{b,k}} \sum_{j \in R_{b,k}} x_j,$$

where $N_{a,k}$ is the number of voxels inside the $k$-th ROI on the GM and tumor, $N_{b,k}$ is the number of voxels inside the $k$-th ROI on the WM, $R_{a,k}$ is the $k$-th ROI on the GM and tumor, and $R_{b,k}$ is the $k$-th ROI on the WM.

The NSTD can be calculated by the following equation.

$$\text{NSTD} = \frac{1}{\bar{b}} \sqrt{\frac{1}{K_b} \sum_{k=1}^{K_b} (b_k - \bar{b})^2},$$

where $K_a = 20$ for the GM, $K_a = 18$ for the tumor, and $K_b = 25$ for the WM.

In the clinical study, for ROI uptake quantification, the uptakes were calculated for the ROI on the thalamus. In addition, STD was calculated for the reconstructed images of clinical data.

In both simulation and clinical studies, we compared the following methods and parameters.

- **LM-DRAMA**: We performed the LM-DRAMA with 1, 2, 3, and 4 main-iterations. We applied the 3D Gaussian filter with full-width-at-half-maximum of 3 mm as a post smoothing. For display, we chose the LM-DRAMA with two main-iterations because it is a default setting of the HITS-655000.

- **MR-DIP**: We performed the MR-DIP using the reconstructed image of LM-DRAMA with two main-iteration,
Fig. 2. Images of simulation data reconstructed by the proposed method and the other methods. From left to right, MRI and phantom image, LM-DRAMA with full-count, LM-DRAMA, MR-DIP, LOR-DIPRecon and LM-DIPRecon with 1/20 count. Each image is tagged with its SNR and TR. White dotted lines indicate the position of profiles in Fig. 3.

Fig. 3. Profiles on images of simulation data reconstructed by the proposed method and the other method. Left is a profile with six-pixel width across two small tumors, as indicated by the white dotted lines on the first row of Fig. 2. Right is a profile with seven-pixel width across one large tumor, as indicated by the white dotted lines on the second row of Fig. 2. Arrows indicate the position of the tumors.

as the label. We used the Adam method and performed 2000 epochs at maximum.

- **LOR-DIPRecon**: We performed the sinogram-based DIPRecon [16] as LOR-DIPRecon. The events were accumulated in the LOR sinogram which is a sinogram not interpolated to evenly space and maintains a raw sampling pitch of LOR in a radial direction. To reduce the number of LORs, the deep pairs of DOI layer were combined to shallow pairs of DOI layer by the DOI compression (DOIC) method [40], using a redundant characteristic of DOI-PET to preserve DOI information as exactly as possible. We compressed the 16 DOI pairs to one pair to facilitate an implementation of LOR-based image reconstruction. In addition, we bundled the crystals into $4 \times 4$ in radial and axial directions to reduce the number of LORs. This crystal bundling is a setting for high-speed (HS) mode reconstruction of HITS-655000 [36]. In the clinical study with HITS-655000, the HS-mode reconstruction is routinely used. The ring difference was sampled with a maximum ring difference of $\pm 40$ and span of three. In the results, the size of the LOR sinogram was $128 \times 128 \times 80 \times 27$ (slice) (ring difference). Even after the DOIC and crystal bundling, the number of LORs was 5 and 7.5 times greater than the number of events in the 1/20 count simulation and 1/10 count clinical data, respectively. The image reconstruction from LOR sinograms was performed by the LOR-DRAMA which is the LOR-OSEM method [41].
with the subset-dependent relaxation [23]. We set the $N_{\text{sub}} = 32$ and $\beta = 50$ for the LOR-DRAMA. For the LOR-DIPRecon, we performed the two-step approach. In the first step, we performed the MR-DIP using Adam method with 1000 epochs, where the label was the reconstructed image by one main-iteration of LOR-DRAMA. The trained network by MR-DIP was used as the initial network parameter $\theta^{(0)}$. In the second step, we performed 320 iterations of LOR-DIPRecon at maximum. The other hyper parameters were the same as the LM-DIPRecon.

- **LM-DIPRecon**: We performed the two-step approach as explained above. In the first step, we performed the MR-DIP using the Adam method with 1000 epochs, where the label was the reconstructed image by one main-iteration of LM-DRAMA. The trained network by MR-DIP was used as the initial network parameter $\theta^{(0)}$. In the second step, we performed 200 iterations of Algorithm 1 at maximum with sub-iterations of $M_1 = 2$ and $M_2 = 10$ for subproblem (12) and (13), respectively. We used the L-BFGS method for subproblem (13) and set $\rho = 0.5$ for the regularization.

V. RESULTS

Fig. 2 shows the images of simulation data reconstructed by the proposed method and the other methods. The images were tagged with their SNRs and TRs. The reconstructed image of LM-DRAMA with 1/20 count was noisy because the post smoothing was limited to 3 mm. Applying MR-DIP reduced the noise at a slight cost of the contrast. The LM-DIPRecon method maintained the contrast and provided sharper images than those of MR-DIP and LOR-DIPRecon.

Fig. 3 shows the profiles across the tumor regions. The peaks of profile on MR-DIP and LOR-DIPRecon at tumor positions tended to be lower than the phantom as shown in Fig. 3. The LM-DIPRecon obtained closer peaks to the phantom than the MR-DIP and LOR-DIPRecon.

Fig. 4 shows the tradeoff curves between contrast and noise in the simulation study: TR versus SNR (left) and CRC versus NSTD (right). Fill markers correspond to the images, as shown in Fig. 2. TR was calculated for the union of three tumor regions. CRC was calculated for the ROIs on the gray matter and tumor regions. NSTD was calculated for the ROIs on the white matter. The plots of LM-DRAMA correspond to 1, 2, 3, and 4 main-iterations. The plots of MR-DIP correspond to 100, 200, . . . , 2000 epochs. The plots of LOR-DIPRecon correspond to 32, 64, . . . , 320 iterations. The plots of LM-DIPRecon correspond to 20, 40, . . . , 200 iterations.

Fig. 5. Effects of hyper parameter $\rho$ on the tradeoff curves between TR and SNR in the simulation study with 1/20 count.
Fig. 6. Images of simulation data with 1/20 count when there is misregistration between MR and PET images. From left to right, MRIs with zero, two and four voxel shift in x direction, images of LM-DIPRecon using MRIs with zero, two and four voxel shift in x direction. The images of LM-DIPRecon are tagged with TR and SNR. dx is the amount of shift in voxel. Fourth row shows the zoomed-in images on the red square region. The red arrow indicates the difference between the images at dx of zero and four.

The LM-DIPRecon provided the better SNR than that of LOR-DIPRecon. In addition, it showed the better CRC than that of LOR-DIPRecon in the case of the same NSTD.

To assess how sensitive the proposed method is to hyperparameters, we performed the LM-DIPRecon with ρ of 0.1, 0.3, 0.5 and 1.0 on the simulation data with 1/20 count. Among the hyperparameters, primarily ρ determines the tradeoff between contrast and noise by controlling the strength of regularization by MR-DIP. Fig. 5 shows the tradeoff curves between TR and SNR for different values of ρ. ρ = 0.5 provided the best tradeoff curve in terms of the SNR.

To assess how sensitive the proposed method is to mismatch between MR and PET images, we deliberately shifted the MR image in the x direction with two or four voxels in simulation data, and performed the LM-DIPRecon. Fig. 6 and 7 shows the images and the tradeoff curves between contrast and noise, respectively, on the simulation data with 1/20 count when the misregistration between the MR and PET images was zero, two, and four voxels in x direction. When the shift of MRI in the x direction was four voxels, the SNR decreased about 0.7 dB on average.

Fig. 8 shows the images of clinical data with 1/10 count reconstructed by the proposed method and the other methods. The images showed that the $^{11}$C-MeQAA accumulated dominantly in the thalamus and striatum [38]. Although the reconstructed images at 1/10 count were very noisy, MR-DIP, LOR-DIP and LM-DIPRecon reduced the noises. In addition, LM-DIPRecon restored the gray matter structure more clearly than MR-DIP. The LOR-DIPRecon provided the smoother image than that of LM-DIPRecon.
Fig. 8. Images of clinical data with 1/10 count reconstructed by the proposed method and the other methods. From left to right, MRI, LM-DRAMA, MR-DIP, LOR-DIPRecon and LM-DIPRecon. The images are tagged with the NSTD and the uptake. The fourth row shows the zoom-in images on the red square region.

Fig. 9 shows the tradeoff curves between contrast and noise in the clinical study with 1/10 count: Uptake vs. NSTD. The LM-DIPRecon provided a higher uptake than those of the other methods.

Fig. 10 shows the images of clinical data with full-count reconstructed by the proposed method and the other methods. The LM-DIPRecon imaged the clearer structure of gray matter than the other methods and reduced the noise relative to LM-DRAMA.

VI. DISCUSSION

In this study, we incorporated DIP into list-mode PET image reconstruction using an ADMM framework. As the computation efficiency of list-mode PET image reconstruction depends on the number of events and not on the number of LORs, the proposed LM-DIPRecon method will be suitable for a PET scanner with many LORs and additional information such as TOF and DOI.

In the simulation study, LM-DIPRecon provided sharper images visually than those of MR-DIP (Fig 2). This is because the number of iterations of list-mode PET image reconstruction in LM-DIPRecon is virtually larger than that of MR-DIP. The MR-DIP train a CNN to map an MR image to a PET image reconstructed using LM-DRAMA with two main-iterations. Therefore, the degree of convergence of the MR-DIP image to the maximum likelihood solution is limited to the level of two main-iterations. On the other hand, LM-DIPRecon performs more iterations of regularized LM-DRAMA through solving subproblem (10). For example, $N = 200$ and $M_1 = 2$ correspond to 10 main-iterations with 40 subsets. Hence, the LM-DIPRecon method could reconstruct a more convergent image than the MR-DIP within the constraint that the image is representable by the CNN. LM-DIPRecon started optimization with CNN prepared by the first step, but MR-DIP is also different in that it performed optimization from random initialization. This is a reason why the images of LM-DIPRecon were sharper than those of MR-DIP. In addition, LM-DIPRecon provided sharper image than that of LOR-DIPRecon because it enables full use of DOI information.

The LM-DIPRecon method provided better tradeoff curves between contrast and noise than those of the other methods (Fig. 4). However, in some cases, the tradeoff curves of LM-DIPRecon and MR-DIP overlapped (Fig 4. Left).
the computation time is prioritized, MR-DIP may become the preferred option instead of LM-DIPRecon. However, the merit of LM-DIPRecon is consistency with measured data. The MR-DIP method can be consistent with a label image, but the consistency with the measured data is limited by the label image in principle. Hence, the risk of eliminating the signal such as tumor by the non-linear image processing such as MR-DIP may be reduced by the LM-DIPRecon. When the data consistency is prioritized, the LM-DIPRecon is recommended. The LM-DIPRecon provided better contrast than that of LOR-DIPRecon in the case of the same noise level. This is because the list-mode approaches can fully use the DOI information.

The LM-DIPRecon provided the close tradeoff curves when ρ is within 0.3 to 1.0 (Fig. 5). When ρ is 0.1, the strength of regularization was not sufficient to suppress the noise. We expect that the proposed method is stable for the choice of ρ within 0.3 to 1.0.

The TR seemed more robust to misregistration between MRI and PET than SNR (Fig. 7). This is because the tumors were not included in MRI from the beginning, and TR was not affected by the misregistration. The SNR was relatively sensitive to misregistration between MRI and PET (See the tradeoff curve with dx = 4 in Fig. 7). This is likely because the misregistration of piecewise constant regions with large volume such as gray matter and white matter weaken the noise suppression performance of MR-DIP. From these results, although the proposed method is considered to be robust to the absence of signals such as tumors in MRI, it is considered to be relatively sensitive to misregistration between MRI and PET in terms of noise.

In the simulation study, a tumor on the image of the proposed method seemed to have higher contrast than the ground truth (Fig. 2). While an exact cause will be difficult to identify, it can be inferred that this is also because of the mismatch between MRI and PET. It can be considered that the noise suppression by the MR-DIP is weak on the tumor regions because the tumors were not included in the MR images. In addition, it was reported that the maximum likelihood (ML) reconstructions cause edge and noise artifacts as progress to convergence [42]. It may be concluded that ML reconstruction-caused artifacts on tumor regions could not be suppressed by MR-DIP because of tumor absence in MRI. These two factors are considered to have caused the uncertainty of tumor uptake.

In the clinical study with 1/10 count, LOR-DIPRecon provided a smoother image than that of LM-DIPRecon (Fig. 8). This can be considered as the smoothing effect of the DOI compression and crystal bundling. However, this smoothing may weaken the contrast. For example, the fourth row of Fig. 10 zoomed in the red square region in the first row of Fig. 10, where these are the results of clinical study with full-count. The red arrow indicated the hot spot which can be seen by the LM-DIPRecon, but not by the LOR-DIPRecon. This indicated that the smoothing effect by the DOI compression and crystal bundling has a risk to obscure the hot spot. In contrast, the LM-DIPRecon may be helpful to find the hot spot. Although the LOR-DIPRecon provided a smoother image than that of MR-DIP, this is because MR-DIP was applied to the reconstructed images of LM-DRAMA. It is worth noticing that LOR-DIPRecon provided a sharper image than that of MR-DIP with LOR-DRAMA as reported in the previous study [16]. In the clinical study, LM-DIPRecon visualized a clearer structure of GM than MR-DIP (Fig. 8 and 10). The image quality of real data depends on factors such as an imperfection of data correction or a subject motion. Note that these factors degrade the image quality and may make the visual difference between the reconstruction methods small. The LM-DIPRecon method provided a better tradeoff curve of uptake vs. NSTD than the other methods, similar to the simulation study (Fig. 9).

Currently, the number of LORs of research tomographs have achieved the order of 10^{9}. For example, the number of LORs of uEXPLORER total body PET/CT scanner is 92 × 10^{9} [43]. In addition, the number of LORs of brain PET scanner with four-layer DOI detectors which was used in this study is 107 × 10^{9} [37]. To full-use of the measured information in these scanners, the list-mode PET image reconstruction is still needed. The simulation and clinical results of the brain PET scanner with DOI detectors demonstrated the feasibility of the LM-DIPRecon for the PET scanners with a large number of LORs.

The limitation of the LM-DIPRecon is a long computation time. In this study, the regularized LM-DRAMA was computed using six cores of Intel i7-5930k 3.5GHz through the message passing interface, and the MR-DIP was computed using NVIDIA Quadro P6000 graphics card. In this condition, the two sub-iterations of regularized LM-DRAMA took 2-3 seconds, and the 10 sub-iterations of L-BFGS method for the MR-DIP took about 96 seconds. Hence, 200 iterations

![Fig. 9. Tradeoff curves between contrast and noise on the clinical study with 1/10 count: Uptake versus NSTD. Fill markers correspond to the images as shown in Fig. 5. Uptake was calculated for the ROI on the thalamus. NSTD was calculated for the ROIs on the white matter. The plots of LM-DRAMA correspond to 1, 2, 3, and 4 main-iterations. The plots of MR-DIP correspond to 100, 200, . . . , 2000 epochs. The plots of LOR-DIPRecon correspond to 32, 64, . . . , 320 iterations. The plots of LM-DIPRecon correspond to 20, 40, . . . , 200 iterations.](image-url)
Fig. 10. Images of clinical data with full-count reconstructed by the proposed method and the other methods. From left to right, MRI, LM-DRAMA, MR-DIP, LOR-DIPRecon and LM-DIPRecon. The images are tagged with the NSTD and the uptake. The fourth row shows the zoom-in images on the red square region. The red arrow indicates a convincing accumulation of $^{11}$C-MeQAA in the striatum on the LM-DIPRecon, but not on the LOR-DIPRecon.

of LM-DIPRecon took about $2 \times 10^4$ seconds in our environment. As the computational cost of MR-DIP was higher than regularized LM-DRAMA, we could reduce the total computational cost by increasing the number of sub-iterations of the regularized LM-DRAMA and decreasing the number of iterations of the LM-DIPRecon.

Another limitation of this study is a mismatch between MR and PET images. In simulation data, the structures of MR and PET images had agreed except for tumor regions, but not for the clinical data. This is because MRI and PET were taken separately in this study. This mismatch can be considered as one reason why the performance of the proposed method degraded in clinical data compared to simulation data. We expect that the MR-PET scanner easily eliminates this problem.

In this study, we demonstrated the proposed method on the four-layer DOI-PET scanner [37]. Another demonstration of the proposed method on TOF-PET scanners is also an important future study, and it is expected to strongly support the usefulness of the proposed method. In near future, we plan to evaluate the proposed method using clinical data of the dedicated brain TOF-PET scanner [44].

In the future, an extension of LM-DIPRecon to 4D image reconstruction [45], [46] can be considered. As list-mode data has finer temporal information than dynamic sinograms, LM-4D-DIPRecon may be useful for dynamic PET imaging or direct parametric imaging [47]. For example, 4D image reconstruction models the time activity curves as a linear combination of temporal basis functions, and estimates the coefficients of linear combination directly from dynamic sinogram or list data [45], [48]. By choosing temporal basis functions with correlation, noise can be suppressed by 4D PET image reconstruction. For list data, we can compute the contribution of event to the temporal basis function in millisecond order, but in dynamic sinogram we restricted to frame time unit and, which may result in a loss of some information. Hence, the proposed method is also promising as a framework for 4D PET image reconstruction using DIP [17]. The high temporal resolution of list data is also useful for motion correction [49], [50]. From these ideas, the LM-DIPRecon can
be expected to be useful for the 4D and 5D problems involving dynamic imaging of moving organs [51]. In this study, the MR image was used as a prior distribution image of the DIP, but an MR-guided deep decoder (MR-GDD) [52], which uses the MR image through an attention mechanism, has already been proposed. As the MR-GDD works to improve the contrast of the tumor which does not exist on the MR image, the list-mode PET image reconstruction using MR-GDD may improve the contrast of the tumor relative to the LM-DIPRecon.

VII. CONCLUSION

We proposed list-mode PET image reconstruction using DIP which is the first trial to integrate list-mode PET image reconstruction and CNN. We incorporated DIP into list-mode PET image reconstruction using an ADMM framework. We evaluated the proposed method using simulation and clinical data. The proposed method provided a sharper image, higher contrast, and lower noise than the other methods. These results indicated that the proposed method is promising for the next generation of brain PET scanners.

APPENDIX

In the experiments, data corrections [53] and the shift-invariant image-space point-spread-function (PSF) [54] were incorporated in the LM-DRAMA as,

$$\tilde{x}_{EM}^{(k+1)} = x_{EM}^{(k)},$$

$$x_{EM}^{(k+1)} = x_{EM}^{(k)} \chi_{EM}(k) + x_{EM}^{(k)} \chi_{EM}(k),$$

$$\chi_{EM}(k) = \left( \frac{N_{sub}}{\tilde{S}_j} \sum_{j} P_{ij} \sum_{t \in \text{Sub}_t} a_{ij(t)} \delta(t) - 1 \right),$$

$$x_{j}^{(k+1)} = \max \left( x_{j}^{(k+1)}, 0 \right),$$

$$y_{ij}^{(k)} = A_j \left( \sum_{j} P_{ij} x_{EM}^{(k)} \right) + C_i \frac{B_j}{B_i},$$

$$\tilde{S}_j = \sum_{i} P_{ij} B_i a_{ij},$$

$$A_j = e^{-\sum_{i} a_{ij} \mu_j},$$

$$\delta(t) = \begin{cases} 1 & \text{t is prompt} \\ -1 & \text{t is delayed} \end{cases},$$

where A is an attenuation probability, B is an efficiency of LOR, C is a scatter component, P is a PSF, $\tilde{S}$ is a sensitivity image including the LOR efficiency and the PSF, $\mu$ is an image of linear attenuation coefficient (µ-map), and $\delta$ is a sign of event for the random correction of the delayed subtraction.

The efficiency of LOR was estimated by the component-based normalization technique [55]. The scatter component for data correction was estimated by the single scatter simulation method [56]. In the simulation study, the $\mu$-map of the phantom image was used for the attenuation correction (AC) and the scatter estimation. In the clinical study, the $\mu$-map was estimated by the segmentation of the reconstructed image without AC. The bone tissues were inserted in the surface of segmented image using dilation and erosion filters. The PSF was modeled by the 3-D Gaussian function with full-width at half-maximum of one voxel because the voxel size was larger than the spatial resolution of HITS-655000 [37].

ACKNOWLEDGMENT

The authors would like to thank the members of the Fifth Research Group with the Central Research Laboratory, Hamamatsu Photonics K. K., for their kind support. This work was supported by JSPS KAKENHI Grant Number JP22K07762.

REFERENCES

[1] M. E. Phelps, PET: Molecular Imaging and its Biological Applications. New York, NY, USA: Springer-Verlag, 2004. [Online]. Available: https://www.springer.com/lb/book/9780387403595
[2] J. Qi et al., “High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner,” Phys. Med. Biol., vol. 43, no. 4, pp. 1001–1014, 1998.
[3] J. Nuyts et al., “A concave prior penalizing relative differences for maximum-a-posterior reconstruction in emission tomography,” IEEE Trans. Nucl. Sci., vol. 49, no. 1, pp. 56–60, Aug. 2002.
[4] G. Wang and J. Qi, “PET image reconstruction using kernel methods,” IEEE Trans. Med. Imag., vol. 34, no. 1, pp. 61–71, Jan. 2015.
[5] B. Bai, Q. Li, and R. M. Leahy, “Magnetic resonance-guided positron emission tomography image reconstruction,” Seminars Nucl. Med., vol. 43, no. 1, pp. 30–44, Jan. 2013.
[6] C. Comtat et al., “Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels,” Phys. Med. Biol., vol. 47, no. 1, pp. 1–20, 2002.
[7] A. J. Reader et al., “Deep learning for PET image reconstruction,” IEEE Trans. Radiat. Plasma Med. Sci., vol. 5, no. 1, pp. 1–25, Jan. 2021.
[8] K. Gong et al., “The evolution of image reconstruction in PET: From filtered back-projection to artificial intelligence,” PET Clin., vol. 16, no. 4, pp. 533–542, 2021.
[9] A. Mehranian and A. J. Reader, “Model-based deep learning PET image reconstruction using forward-backward splitting expectation-maximization,” IEEE Trans. Radiat. Plasma Med. Sci., vol. 5, no. 1, pp. 54–64, Jun. 2020.
[10] K. Ote and F. Hashimoto, “Deep-learning-based fast TOF-PET image reconstruction using direction information,” Radiolog. Phys. Technol., vol. 15, no. 1, pp. 72–82, Mar. 2022.
[11] B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, and M. S. Rosen, “Image reconstruction by domain-transform manifold learning,” Nature, vol. 555, pp. 487–492, Mar. 2018.
[12] W. Whiteley, W. K. Luk, and J. Gregor, “DirectPET: Full-size neural network PET reconstruction from sinogram data,” J. Med. Imag., vol. 7, no. 3, Feb. 2020, Art. no. 032503.
[13] I. Högström, C. R. Schmidtlein, G. Campanella, and T. J. Fuchs, “DeepPET: A deep encoder–decoder network for directly solving the PET image reconstruction inverse problem,” Med. Image Anal., vol. 54, pp. 253–262, May 2019.
[14] K. Gong et al., “Iterative PET image reconstruction using convolutional neural network representation,” IEEE Trans. Med. Imag., vol. 38, no. 3, pp. 675–685, Mar. 2019.
[15] A. Vedaldi, V. Lempitsky, and D. Ulyanov, “Deep image prior,” Int. J. Comput. Vis., vol. 128, no. 7, pp. 1867–1888, Jul. 2020.
[16] K. Gong, C. Catana, J. Qi, and Q. Li, “PET image reconstruction using deep image prior,” IEEE Trans. Med. Imag., vol. 38, no. 7, pp. 1655–1665, Jul. 2019.
[17] T. Yokota, K. Kawai, M. Sakata, Y. Kimura, and H. Hontani, “Dynamic PET image reconstruction using nonnegative matrix factorization incorporated with deep image prior,” in Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV), Oct. 2019, pp. 3126–3135.
[18] F. Hashimoto, K. Ote, and Y. Onishi, “PET image reconstruction incorporating deep image prior and a forward projection model,” IEEE Trans. Radiat. Plasma Med. Sci., vol. 6, no. 8, pp. 841–846, Nov. 2022, doi: 10.1109/TRPMS.2022.3161569.
[19] M. G. Spangler-Bickell, T. W. Deller, V. Bettinardi, and F. Jansen, “Ultra-fast list-mode reconstruction of short PET frames and example applications,” J. Nucl. Med., vol. 62, no. 2, pp. 287–292, Feb. 2021.
[20] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” Found. Trends Mach. Learn., vol. 3, no. 1, pp. 1–122, 2011.

[21] X. Cao, Q. Xie, and P. Xiao, “A regularized relaxed ordered subset list-mode reconstruction algorithm and its preliminary application to undersampling PET imaging,” Phys. Med. Biol., vol. 60, no. 1, pp. 49–66, Jan. 2015.

[22] T. Nakayama and H. Kudo, “Derivation and implementation of ordered-subsets algorithms for list-mode PET data,” in Proc. IEEE Nucl. Sci. Symp. Conf. Rec., Oct. 2005, pp. 1950–1954.

[23] E. Tanaka and H. Kudo, “Subset-dependent relaxation in block-iterative algorithms for image reconstruction in emission tomography,” Phys. Med. Biol., vol. 48, no. 10, pp. 1405–1422, May 2003.

[24] J. Qi, “Calculation of the sensitivity image in list-mode reconstruction for PET,” IEEE Trans. Nucl. Sci., vol. 53, no. 5, pp. 2746–2751, Oct. 2006.

[25] J. Lehtinen et al., “Noisy2Noise: Learning image restoration without clean data,” 2018, arXiv:1803.04189.

[26] J. Cui et al., “PET image denoising using unsupervised deep learning,” Eur. J. Nucl. Med. Mol. Imag., vol. 46, no. 13, pp. 2780–2789, Dec. 2019.

[27] F. Hashimoto, H. Ohba, K. Ote, A. Teramoto, and H. Tsukada, “Dynamic PET image denoising using deep convolutional neural networks without prior training datasets,” IEEE Access, vol. 7, pp. 96594–96603, 2019.

[28] K. Lange, D. R. Hunter, and I. Yang, “Optimization transfer using surrogate objective functions,” J. Comput. Graph. Statist., vol. 9, no. 1, pp. 1–20, Mar. 2000.

[29] G. Wang and J. Qi, “ Penalized likelihood PET image reconstruction using patch-based edge-preserving regularization,” IEEE Trans. Med. Imag., vol. 31, no. 12, pp. 2194–2204, Dec. 2012.

[30] E. S. H. Neto and Á. R. D. Pierro, “Convergence results for scaled gradient algorithms in positron emission tomography,” Inverse Problems, vol. 21, no. 6, pp. 1905–1914, Dec. 2005.

[31] Ö. Çiçek, “3D U-Net: Learning dense, volumetric segmentation from sparse annotation,” in Proc. IEEE Int. Conf. Med. Image Comput. and Computer-Assisted Intervention MICCAI 2016, vol. 9901, Cham, Switzerland: Springer, 2016, pp. 424–432.

[32] R. Pascancu, T. Mikolov, and Y. Bengio, “On the difficulty of training recurrent neural networks,” 2012, arXiv:1211.5063.

[33] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2014, arXiv:1412.6980.

[34] M. Ljungberg, A. Larsson, and L. Johansson, “A new collimator simulation in SIMIND based on the delta-scattering technique,” IEEE Trans. Nucl. Sci., vol. 52, no. 5, pp. 1370–1375, Oct. 2005.

[35] E. R. Woodcock et al., “Techniques used in the GEM code for Monte Carlo neutronics calculations in reactors and other systems of complex geometry,” in Proc. Conf. Appl. Comput. Methods React. Problems, May 1965, pp. 557–579.

[36] D. L. Collins et al., “Design and construction of a realistic digital brain phantom,” IEEE Trans. Med. Imag., vol. 17, no. 3, pp. 463–468, Jun. 1998.

[37] M. Watanabe et al., “Performance evaluation of a high-resolution brain PET scanner using four-layer MPAC DOI detectors,” Phys. Med. Biol., vol. 62, no. 17, pp. 7148–7166, 2017.

[38] K. Nakaizumi et al., “In vivo depiction of α7 nicotinic receptor loss for cognitive decline in Alzheimer’s disease,” J. Alzheimer’s Disease, vol. 61, no. 4, pp. 1355–1365, Jan. 2018.