Neural KEM: A Kernel Method With Deep Coefficient Prior for PET Image Reconstruction

Siqi Li, Kuang Gong, *Member, IEEE*, Ramsey D. Badawi, *Senior Member, IEEE*, Edward J. Kim, Jinyi Qi, *Fellow, IEEE*, and Guobao Wang, *Senior Member, IEEE*

Abstract—Image reconstruction of low-count positron emission tomography (PET) data is challenging. Kernel methods address the challenge by incorporating image prior information in the forward model of iterative PET image reconstruction. The kernelized expectation-maximization (KEM) algorithm has been developed and demonstrated to be effective and easy to implement. A common approach for a further improvement of the kernel method would be adding an explicit regularization, which however leads to a complex optimization problem. In this paper, we propose an implicit regularization for the kernel method by using a deep coefficient prior, which represents the kernel coefficient image in the PET forward model using a convolutional neural-network. To solve the maximum-likelihood neural network-based reconstruction problem, we apply the principle of optimization transfer to derive a neural KEM algorithm. Each iteration of the algorithm consists of two separate steps: a KEM step for image update from the projection data and a deep-learning step in the image domain for updating the kernel coefficient image using the neural network. This optimization algorithm is guaranteed to monotonically increase the data likelihood. The results from computer simulations and real patient data have demonstrated that the neural KEM can outperform existing KEM and deep image prior methods.

Index Terms—Dynamic PET, image reconstruction, kernel methods, optimization transfer (OT), deep image prior (DIP).

I. INTRODUCTION

TOMOGRAPHIC image reconstruction for positron emission tomography (PET) is challenging because of the

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Institutional Review Board at the University of California, Davis.

Siqi Li, Ramsey D. Badawi, and Guobao Wang are with the Department of Radiology, University of California Davis Health, Sacramento, CA 95817 USA (e-mail: sqlli@ucdavis.edu; rdbadawi@ucdavis.edu; gbwang@ucdavis.edu).

Kuang Gong is with the Gordon Center for Medical Imaging, Massachusetts General Hospital and Harvard Medical School, Boston, MA 02114 USA (e-mail: kgong@mgh.harvard.edu).

Edward J. Kim is with the Comprehensive Cancer Center, University of California Davis Health, Sacramento, CA 95817 USA (e-mail: jhkim@ucdavis.edu).

Jinyi Qi is with the Department of Biomedical Engineering, University of California at Davis, Davis, CA 95616 USA (e-mail: qi@ucdavis.edu).

Digital Object Identifier 10.1109/TMI.2022.3217543

ill-conditioned problem and low counting statistics [1]. The kernel method addresses this challenge by integrating image prior information into the forward model of PET image reconstruction (e.g., [2], [3], [4], [5], [6], [7], [8], [9]). The *a priori* information can come from composite time frames in a dynamic PET scan [2], [3], or from the co-registered anatomical images, e.g., magnetic resonance (MR) images [4], [5]. The derived kernel expectation-maximization (KEM) algorithm [2] has been demonstrated to be effective and is easy to implement [2], [3], [4], [5], [10].

To further improve the kernel method such as for higher temporal resolution dynamic PET imaging or for low-dose PET imaging, a straightforward approach would be adding an explicit regularization form on the kernel coefficient image to stabilize the solution [2]. This can be achieved using either conventional penalty functions (e.g., [12], [13], [14], [15]) or convolutional neural network (CNN) based penalties (e.g., [16], [17]). However, such regularization-based methods generally require a complex optimization algorithm, involve one or more hard-to-tune hyper-parameters, and need to run for many iterations for a convergent solution.

In this paper, we propose an implicit regularization for the kernel method by using CNN to represent the kernel coefficient image in the kernelized model. The use of CNN representation shares the same spirit of the work of Gong et al. [18], [19], [20] and others [21], [22], [23] that employs deep image prior (DIP) [24] for PET image reconstruction. Differently the CNN representation in this work is applied in the kernel coefficient space instead of the original PET activity image space, resulting in a modified kernel method with deep coefficient prior.

One challenge with solving the corresponding optimization problem is that the neural network is involved in the projection domain, resulting in a large-scale, nonlinear reconstruction problem. The alternating direction method of multipliers (ADMM) is a popular optimization approach to solving this kind of problems, e.g., in [18] and [21]. However, the hyper parameters associated with an ADMM algorithm are challenging to tune in practice. In this work, we derive an easy-to-implement iterative algorithm by using the principle of optimization transfer [26], [27], [28] for the neural network-based reconstruction. We call the new algorithm *neural* KEM to differentiate it from the original KEM algorithm.

There are also other ways to explore deep learning for PET reconstruction [29], [30], [31], [32], [33], such as the direct end-to-end mapping of PET image from projection [34] and unrolled model-based deep-learning reconstruction [35], [36]. All these approaches require pre-training using a large population-based database, which is not always available. Similar to the original kernel method [2] and the DIP method [18], the proposed method does not require population-based pre-training but is solely based on the data of single subjects.

The remaining of this paper is organized as follows. Section II introduces the background materials of the kernel method and DIP method for PET image reconstruction. Section III describes the proposed neural KEM method that combines the kernel method with deep coefficient prior. We then present 2D and 3D computer simulation studies in Sections IV and V and a real patient data study in Section VI to demonstrate the improvement of the proposed method over existing methods. Finally discussions and conclusions are drawn in Sections VII and VIII.

II. BACKGROUND

A. PET Image Reconstruction

PET projection measurement $y = \{y_i\}_{i=1}^N$ can be well modeled as independent Poisson random variables using the log-likelihood function [1],

$$L(y|x) = \sum_{i=1}^N y_i \log \overline{y}_i - \overline{y}_i - \log y_i!, \tag{1}$$

where the expectation of the projection data, $\overline{y} = \{\overline{y}_i\}_{i=1}^N$, is related to the unknown activity image $x = \{x_j\}_{j=1}^J$ through

$$\overline{y} = Px + r, \tag{2}$$

where x_j denotes the PET image intensity value in pixel j. N is the total number of detector pairs and J is the number of image pixels. P is the detection probability matrix and includes normalization factors for scanner sensitivity, scan duration, deadtime correction and attenuation correction. r is the expectation of random and scattered events [1].

The maximum likelihood estimate of the activity image x is found by maximizing the Poisson log-likelihood,

$$\hat{x} = \arg \max_{x \geq 0} L(y|x). \tag{3}$$

A common way of seeking the solution of (3) is the maximum likelihood expectation maximization (ML-EM) algorithm [37].

B. Kernel EM for PET Reconstruction

The image estimate by standard ML-EM is commonly noisy due to the limited counting statistics of PET emission data. To suppress noise, the kernel method [2] incorporates an image prior into the forward projection of PET reconstruction by describing the image intensity x_j using kernels,

$$x_j = \sum_{l \in \mathcal{N}_j} \alpha_l \kappa(f_j, f_l), \tag{4}$$

where $\mathcal{N}_j$ defines the neighborhood of pixel j, e.g., by the k-nearest neighbors (kNN, [38]). $\kappa(\cdot, \cdot)$ is the kernel function (e.g., radial Gaussian) and f denotes the low-dimensional feature vector that is extracted at each pixel from the image prior z (e.g. the composite images in dynamic PET or anatomical image in PET/MR or PET/CT). The equivalent matrix-vector form of (4) is

$$x = K\alpha, \tag{5}$$

where K is a sparse square kernel matrix with its (j, l)th element being $\kappa(f_j, f_l)$. α denotes the corresponding kernel coefficient image.

Substituting the kernelized image model (5) into the standard PET forward projection model in (2) gives the following kernelized forward projection model for PET image reconstruction,

$$\overline{y} = PK\alpha + r. \tag{6}$$

The maximum-likelihood estimate of α can be found by the kernel EM algorithm [2],

$$\alpha^{n+1} = \frac{\alpha^n}{w} \cdot \left(K^T P^T \frac{y}{PK\alpha^n + r} \right), \tag{7}$$

where

$$w = K^T P^T \mathbf{1}_N, \tag{8}$$

and $\mathbf{1}_N$ is a vector with all elements being 1. n denotes the iteration number and the superscript "T" denotes matrix transpose. The vector multiplication and division are element-wise operations. Note that the KEM update becomes the standard EM update if K is an identity matrix. Once α is estimated, the final reconstructed PET image is given by

$$\hat{x} = K\hat{\alpha}. \tag{9}$$

Note that the same update equation (7) is also used by the sieves method [11]. The difference is that the sieves method uses a stationary Gaussian kernel [11], while the kernel method here uses data-driven spatially variant kernels that are derived from the image prior z.

The estimated kernel coefficient image α by the standard kernel method [2] may still suffer from noise, as demonstrated later in this paper. One possible way for improvement is to add an explicit penalty function to stabilize the estimation of α as indicated in the original kernel method [2], which however may result in a more challenging optimization problem and involves at least one more regularization parameter to tune.

C. PET Reconstruction Using DIP

The DIP method for PET reconstruction is proposed in [18] based on the representation ability of CNNs. Instead of using the linear kernel representation in (6), the PET image x can be also described by a nonlinear representation using neural networks and the image prior data z,

$$x = \beta(\theta|z), \tag{10}$$
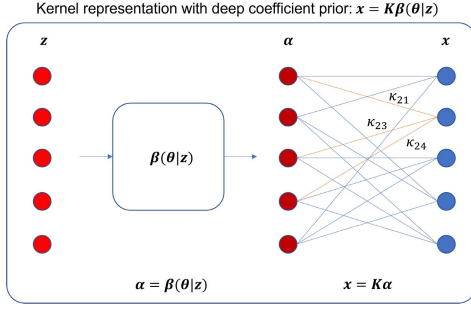
Fig. 1. Graphical illustration of the kernel representation with deep coefficient prior.
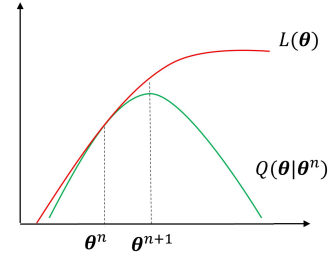

Fig. 2. Illustration of optimization transfer used in this work. The surrogate function $Q(\theta|\theta^n)$ minorizes the original likelihood function $L(\theta)$. $Q(\theta|\theta^n)$ is designed to be easier to optimize. The solution θ^{n+1} guarantees a monotonic increase in L.

where β is a neural netwok model with z the input images and θ the network weights. After substituting the DIP model (10) into (2), the PET forward projection model becomes

$$\overline{y} = P\beta(\theta|z) + r. \tag{11}$$

The maximum-likelihood estimate of the unknown θ is obtained by

$$\hat{\theta} = \arg\max_{\theta} L\Big(y|\beta(\theta|z)\Big). \tag{12}$$

Once θ is estimated, the PET activity image is calculated as

$$\hat{x} = \beta(\hat{\theta}|z). \tag{13}$$

To solve the resulting nonlinear optimization problem, Gong et al. [18] use the ADMM algorithm,

$$x^{n+1} = \arg\max_{x} L(y|x) - \frac{\rho}{2}||x - \beta(\theta^n|z) + \mu||^2, \tag{14}$$

$$\theta^{n+1} = \arg\min_{\theta} ||\beta(\theta|z) - (x^{n+1} + \mu^n)||^2, \tag{15}$$

$$\mu^{n+1} = \mu^n + x^{n+1} - \beta(\theta^{n+1}|z), \tag{16}$$

where the subproblem (14) is a penalized-likelihood image reconstruction problem and the subproblem (15) is an image-domain DIP learning using a mean-square error (MSE) loss function. ρ is a hyper-parameter. One well-known weakness of the ADMM algorithm is that ρ is usually difficult to tune.

III. PROPOSED NEURAL KEM

A. Kernel Method With Deep Coefficient Prior

In this work, we propose to describe the kernel coefficient image α in the kernel method as a function of neural networks,

$$\alpha = \beta(\theta|z), \tag{17}$$

where z and θ are again the input (e.g., the composite image prior in dynamic PET [2]) and weights of the neural network β, sharing the same spirit of the DIP model [18]. This provides a kernel representation with deep coefficient prior for PET image,

$$x = K\beta(\theta|z). \tag{18}$$

Fig. 1 shows a graphical illustration of the proposed model using neural network layers, of which the last layer is linear and has fixed network weights as determined by the kernel matrix K.

The proposed model (18) becomes the DIP model in [18] if the kernel matrix K is an identity matrix; the model is also equivalent to the standard kernel model [2] if the neural network β is an identity mapping. When a more complex neural network model (e.g., U-net) is used, the deep coefficient prior then introduces an implicit regularization to stabilize the estimation of the kernel coefficient image α.

By substituting the proposed image model in (18) into the standard PET forward projection model in (2), we obtain the following forward model for PET image reconstruction,

$$\overline{y} = PK\beta(\theta|z) + r. \tag{19}$$

The unknown θ of the neural network is estimated from the projection data by maximizing the Poisson log-likelihood,

$$\hat{\theta} = \arg\max_{\theta} L\Big(y|K\beta(\theta|z)\Big). \tag{20}$$

Once $\hat{\theta}$ is estimated, the PET image is obtained by

$$\hat{x} = K\beta(\hat{\theta}|z). \tag{21}$$

B. Tomographic Reconstruction of Neural Networks Using Optimization Transfer

The optimization problem in (20) is challenging to solve because the unknown θ is non-linearly involved in the projection domain. One possible solution is the ADMM algorithm as used in [18] but tuning the hyperparameter ρ is nontrivial in practice. Here we develop an easy-to-implement optimization transfer algorithm using a similar concept as used for nonlinear parametric PET image reconstruction of tracer kinetics [25], [26] and for joint image registration and reconstruction [27].

The basic idea of optimization transfer [28] is to construct a surrogate function $Q(\theta|\theta^n)$ at iteration n, which minorizes the original objective function $L(\theta)$, as illustrated in Fig. 2. Following the concave property of the log function [28] (also see Eq. 9 in [26]) and treating PK as a single matrix A with its (i, j)th element being a_{ij}, we have the following inequality,

$$\log \overline{y}_i = \log \Big(\sum_{j=1}^{J} a_{ij}\beta_j(\theta|z) + r_i \Big)$$

$$\geq \sum_{j=1}^{J} \frac{a_{ij}\beta_j(\theta^n|z)}{\overline{y}_i^n} \log \beta_j(\theta|z) + c_i^n, \tag{22}$$

where $\overline{y}_i^n = [A\boldsymbol{\beta}(\boldsymbol{\theta}^n|\boldsymbol{z})+\boldsymbol{r}]_i$. The iteration-dependent constant,

$$c_i^n = \left(\sum_{j=1}^{J} \frac{a_{ij}\beta_j(\boldsymbol{\theta}^n|\boldsymbol{z})}{\overline{y}_i^n} \log \frac{\overline{y}_i^n}{\beta_j(\boldsymbol{\theta}^n|\boldsymbol{z})} \right) + \frac{r_i}{\overline{y}_i^n} \log \overline{y}_i^n, \quad (23)$$

is independent of the unknown parameter $\boldsymbol{\theta}$ and is thus omitted hereafter.

Based on (22), an EM-type surrogate function $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^n)$ can be built for the original likelihood function L in a similar way as used in [26],

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^n) = \sum_{j=1}^{J} w_j \left(\hat{\alpha}_j^{n+1} \log \beta_j(\boldsymbol{\theta}|\boldsymbol{z}) - \beta_j(\boldsymbol{\theta}|\boldsymbol{z}) \right), \quad (24)$$

where w_j corresponds to the jth pixel of $\boldsymbol{w}$ defined in (8). $\hat{\boldsymbol{\alpha}}^{n+1}$ is an intermediate kernel coefficient image updated with

$$\hat{\boldsymbol{\alpha}}^{n+1} = \frac{\boldsymbol{\alpha}^n}{\boldsymbol{w}} \cdot \left(\boldsymbol{K}^T \boldsymbol{P}^T \frac{\boldsymbol{y}}{\overline{\boldsymbol{y}}^n} \right), \quad (25)$$

which is one iteration of KEM defined by (7) with $\boldsymbol{\alpha}^n \triangleq \boldsymbol{\beta}(\boldsymbol{\theta}^n|\boldsymbol{z})$.

The surrogate $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^n)$ resembles an image-domain Poisson log-likelihood function (with a pixel-wise weight $\boldsymbol{w}$). Using (22), it is straightforward to prove that the surrogate satisfies the following two conditions for optimization transfer [28],

$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^n) - Q(\boldsymbol{\theta}^n|\boldsymbol{\theta}^n) \leq L\left(\boldsymbol{y}|\boldsymbol{K}\boldsymbol{\beta}(\boldsymbol{\theta}|\boldsymbol{z})\right) - L\left(\boldsymbol{y}|\boldsymbol{K}\boldsymbol{\beta}(\boldsymbol{\theta}^n|\boldsymbol{z})\right),$$
$$(26)$$

$$\nabla Q(\boldsymbol{\theta}^n|\boldsymbol{\theta}^n) = \nabla L\left(\boldsymbol{y}|\boldsymbol{K}\boldsymbol{\beta}(\boldsymbol{\theta}^n|\boldsymbol{z})\right), \quad (27)$$

where ∇ denotes the gradient with respect to $\boldsymbol{\theta}$.

The original optimization problem in (20) is now equivalently transferred into the maximization of the surrogate function (24) at each iteration n,

$$\boldsymbol{\theta}^{n+1} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\boldsymbol{\theta}^n), \quad (28)$$

which performs an image-domain neural-network learning for seeking a $\boldsymbol{\beta}$ approximation to the intermediate kernel coefficient image $\hat{\boldsymbol{\alpha}}^{n+1}$. The learning can be implemented using any existing optimization algorithm (e.g., the Adams optimizer) that is available in a deep-learning software library such as PyTorch or TensorFlow. Because of $Q(\boldsymbol{\theta}^{n+1}|\boldsymbol{\theta}^n) \geq Q(\boldsymbol{\theta}^n|\boldsymbol{\theta}^n)$ and (26), the surrogate optimization guarantees convergence to a local optimum and a monotonic increase in the original likelihood L,

$$L\left(\boldsymbol{y}|\boldsymbol{K}\boldsymbol{\beta}(\boldsymbol{\theta}^{n+1}|\boldsymbol{z})\right) \geq L\left(\boldsymbol{y}|\boldsymbol{K}\boldsymbol{\beta}(\boldsymbol{\theta}^n|\boldsymbol{z})\right). \quad (29)$$

C. Summary of the Algorithm and Implementations

A pseudo-code of the proposed algorithm is provided in Algorithm 1. Each iteration of the algorithm consists of two separate steps:

1) Image reconstruction: Obtain an intermediate kernel coefficient image update $\hat{\boldsymbol{\alpha}}^{n+1}$ from the projection data $\boldsymbol{y}$ using KEM in (25);
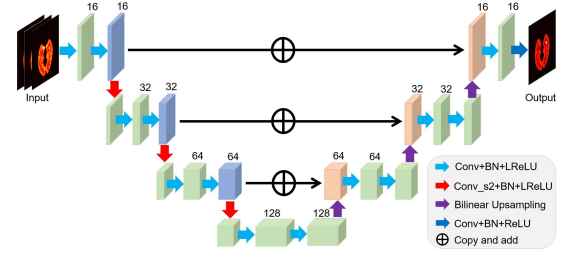

Fig. 3. Illustration of the modified residual U-net $\beta(\theta|z)$ used in this work.

2) Neural-network learning: Find a CNN approximation of the intermediate image $\hat{\boldsymbol{\alpha}}^{n+1}$ using the image-domain maximum-likelihood optimization in (28).

We call this algorithm *Neural KEM* to reflect the fact that neural-network learning is used following the KEM update. Compared to ADMM, the Neural KEM algorithm does not need to tune a hyperparameter and is easier to use.

The proposed algorithm is applicable to different neural network architectures that are suitable for image representation. In our work, a popular residual U-net (e.g., used in [18]) is used for neural network learning and is illustrated in Fig. 3. The network is available in both 2D and 3D versions for learning 2D and 3D images, respectively. It consists of the following operations: 1) 3×3 ($\times 3$) 2D (3D) convolutional layer, 2) 2D (3D) batch normalization (BN) layer, 3) leaky rectified linear unit (LReLU) layer, 4) 3×3 ($\times 3$) convolutional layer with stride 2×2 ($\times 2$) for down-sampling, 5) 2×2 ($\times 2$) bilinear (trilinear) interpolation layer for up-sampling, 6) identity mapping layer that adds feature maps from left-side encoder path to the right-side decoder path. In addition, a ReLU layer is used before the output in order to satisfy the non-negative constraint on the kernel coefficient image. The total number of model parameters in the 3D U-net is about 1.3 million.

Algorithm 1 Neural KEM for PET Reconstruction

1: Input parameters: Maximum iteration number `MaxIt`, initial $\boldsymbol{\theta}^1$.

2: **for** $n = 1$ to `MaxIt` **do**

3:     Obtain an intermediate coefficient image update:
$$\hat{\boldsymbol{\alpha}}^{n+1} = \frac{\boldsymbol{\alpha}^n}{\boldsymbol{w}} \cdot \left(\boldsymbol{K}^T \boldsymbol{P}^T \frac{\boldsymbol{y}}{\boldsymbol{PK}\boldsymbol{\alpha}^n + \boldsymbol{r}} \right),$$
    with $\boldsymbol{\alpha}^n = \boldsymbol{\beta}(\boldsymbol{\theta}^n|\boldsymbol{z})$

4:     Perform a neural network learning by maximizing:
$$Q(\boldsymbol{\theta}|\boldsymbol{\theta}^n) \triangleq \sum_j w_j \left(\hat{\alpha}_j^{n+1} \log \beta_j(\boldsymbol{\theta}|\boldsymbol{z}) - \beta_j(\boldsymbol{\theta}|\boldsymbol{z}) \right)$$

5: **end for**

6: **return** $\hat{x} = \boldsymbol{K}\boldsymbol{\beta}(\hat{\boldsymbol{\theta}}|\boldsymbol{z})$

IV. VALIDATION USING 2D COMPUTER SIMULATION

A. Simulation Setup

We conducted a two-dimensional (2D) computer simulation study to validate the proposed method in dynamic PET image reconstruction. Dynamic scans were simulated for a GE DST whole-body PET scanner using a Zubal head phantom shown in Fig. 4a. The phantom is composed of gray matter, white
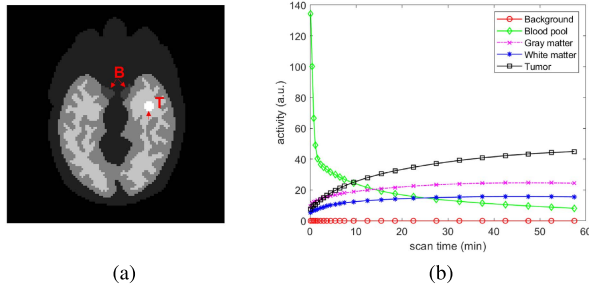
(a)   (b)

Fig. 4. Digital phantom and time activity curves used in the simulation studies. (a) Zubal brain phantom; 'B' represents the blood ROI and 'T' is the tumor ROI. (b) Regional time activity curves.

matter, blood pools (18mm in long axis) and a tumor (15 mm in diameter). The detector system consists of 280 detector blocks, arranged as four rings of 70 blocks each. The width of the block is 38.35mm. The scanner consists of 10,080 BGO crystals. A one-hour dynamic scan was divided into 24 time frames: 4×20s, 4×40s, 4×60s, 4×180s, and 8×300s. The pixel size is 3×3 mm^2 and the image size is 111×111. The time activity curve of ^{18}F-FDG in each region is shown in Fig. 4b. Dynamic activity images were first forward projected to generate noise-free sinograms. Poisson noise was then introduced. Scatters were simulated using the SimSET package [39] using the cylindrical scanner model with no block and gap effects included. We also included 20% uniform random events. Attenuation map, mean scatters and randoms were used in all reconstruction methods to obtain quantitative images. The expected total number of events over 60 min was 8 million. Ten noisy realizations were simulated and each was reconstructed independently for comparison.

B. Reconstruction Methods

We compared the proposed neural KEM with four different reconstruction methods: (1) standard ML-EM, (2) KEM [2], (3) DIP reconstruction by ADMM [18], and (4) DIP reconstruction by the optimization transfer (OT) algorithm, which is equivalent to the neural KEM with $\boldsymbol{K} = \boldsymbol{I}$.

In the kernel-based methods (regular KEM and neural KEM), three 20-minute composite frames were used to generate the image prior data $\boldsymbol{z}$ as used in [2]. The radial Gaussian kernel function $\kappa(\boldsymbol{f}_j, \boldsymbol{f}_l) = \exp(-||\boldsymbol{f}_j - \boldsymbol{f}_l||^2/2\sigma^2)$ was used. Pixel intensity values extracted from the composite images were used to form the feature vector $\boldsymbol{f}$ for generating the kernel matrix $\boldsymbol{K}$ using $\sigma = 1$ and kNN with k=48 which were the same as used in [2].

For the DIP reconstruction by ADMM [18], within each outer iteration, 4 iterations were used for solving (14) and 50 iterations were used for solving (15). These settings were empirically optimized for obtaining stable results according to image mean squared error (MSE; defined in next subsection) in our experiments. The effect of the ADMM hyper-parameter ρ was also investigated and reported. $\rho = 0.05$ was chosen for nearly optimal image MSE.

The input of CNN in both the DIP methods and neural KEM was set to the composite image prior $\boldsymbol{z}$. A ML-EM image was also tested but resulted in worse results. For implementation,

the tomographic reconstruction step was implemented in MATLAB and the neural-network learning step was implemented in PyTorch, both on a PC with an Intel i9-9920X CPU with 64GB RAM and a NVIDIA GeForce RTX 2080Ti GPU. The Adam algorithm was used with a learning rate 10^{-3} for neural network learning. All reconstructions were run for 60 iterations with a uniform initial image. The subiteration number for the neural network learning step was empirically optimized to be 150 in both the DIP by OT and neural KEM method for image MSE. The effect of this subiteration number was also investigated.

C. Evaluation Metrics

Different methods were first compared using image MSE defined by

$$\text{MSE}(\hat{\boldsymbol{x}}_m) = 10 \log_{10} \left(||\hat{\boldsymbol{x}}_m - \boldsymbol{x}_m^{\text{true}}||^2 / ||\boldsymbol{x}_m^{\text{true}}||^2 \right) (\text{dB}), \quad (30)$$

where $\hat{\boldsymbol{x}}_m$ is an image estimate of frame m obtained with one of the reconstruction methods and $\boldsymbol{x}_m^{\text{true}}$ denotes the ground truth image. The ensemble bias and standard deviation (SD) of the mean intensity in regions of interest (ROIs) were also calculated to evaluate ROI quantification,

$$\text{Bias} = \frac{1}{c^{\text{true}}} \left| \overline{c} - c^{\text{true}} \right|, \quad \text{SD} = \frac{1}{c^{\text{true}}} \sqrt{\frac{1}{N_r - 1} \sum_{i=1}^{N_r} |c_i - \overline{c}|}, \quad (31)$$

where c^{true} is the noise-free intensity and $\overline{c} = \frac{1}{N_r} \sum_{i=1}^{N_r} c_i$ denotes the mean of N_r realizations. c_i is the mean ROI uptake in the ith realization and $N_r = 10$ in this study.

D. Comparison for Reconstructed Image Quality

Fig. 5 shows the true activity images and reconstructed images at iteration 60 by five different reconstruction methods for frame 2 (early 20-s frame, low count level), frame 12 (middle 1-min frame, moderate count level) and frame 24 (late 3-min frame, relatively high count level), respectively. The results of MSE in dB are included. As expected, kernel-based methods (regular KEM and neural KEM) achieved a better image quality with lower MSE as compared to the methods without kernel ((b), (d) and (e)). The DIP by ADMM [18] and DIP by OT both suppressed noise well but also led to over-smoothness. The proposed neural KEM was less noisy than the regular KEM due to the added level of regularization from the deep coefficient prior on $\boldsymbol{\alpha}$ and demonstrated better detail preservation than the DIP methods due to the additional structural information embedded in the kernel matrix $\boldsymbol{K}$.

Fig. 6(a) and Fig. 6(b) further show image MSE as a function of iteration number for the two different frames (frame 2 and frame 12). For the DIP reconstruction, the ADMM algorithm demonstrated a relatively faster convergence rate in early iterations than the KEM and DIP by OT. This is mainly because four sub-iterations were used for the tomographic reconstruction step in the ADMM algorithm while one sub-iteration was used for other algorithms. The DIP by OT was either comparable to (Fig. 6(b)) or better than (Fig. 6(a)) the DIP by ADMM. Note that here the DIP
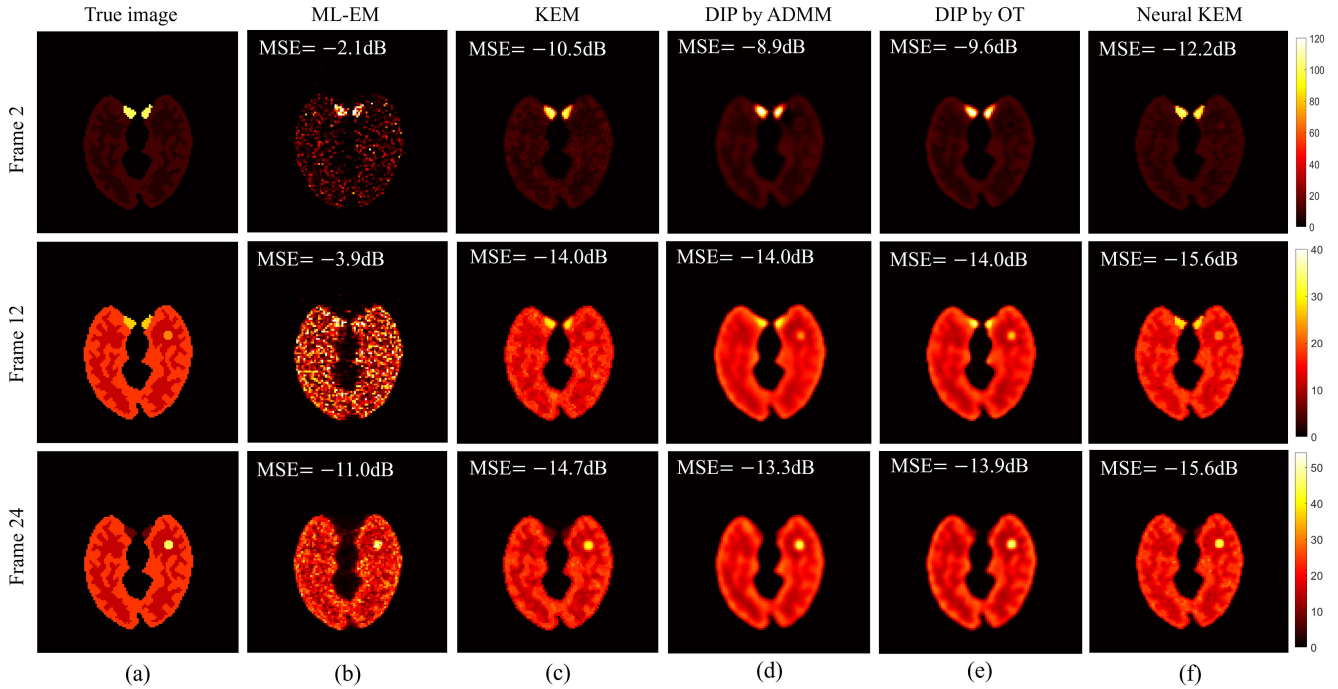
Fig. 5. True activity images and reconstructed images by different reconstruction methods for frame 2 (top row), frame 12 (middle row) and frame 24 (bottom row). (a) True images, (b) ML-EM, (c) KEM [2], (d) DIP by ADMM [18], (e) DIP by OT, (f) Proposed neural KEM.
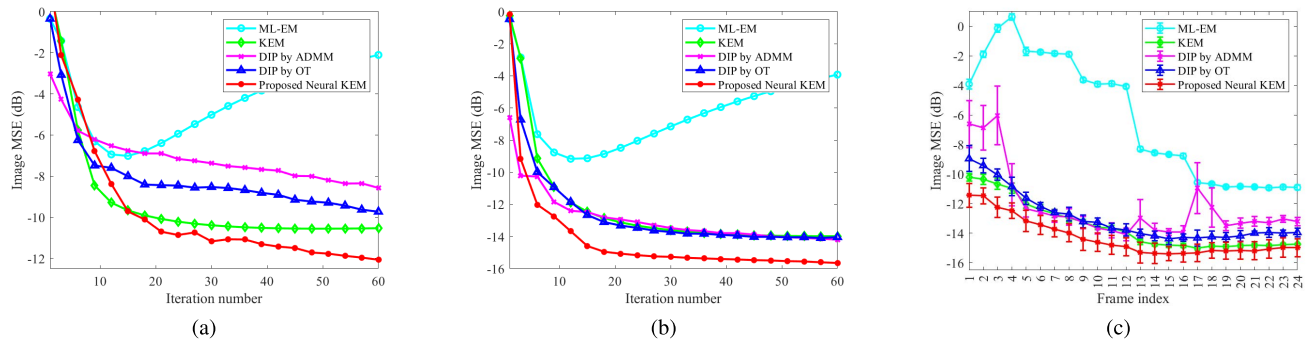

Fig. 6. Plots of image MSE for different reconstruction methods. (a-b) image MSE as a function of iteration number for frame 2 (a) and frame 12 (b); (c) image MSE of all time frames. The error bars in (c) were obtained from 10 realizations.

by ADMM and DIP by OT were not always close to each other. This can be explained by that the neural network model is nonlinear and different algorithms are not guaranteed to provide the same solution.

Fig. 6(c) shows the plots of image MSE for all time frames reconstructed by different methods with 60 iterations. Error bars were calculated over 10 noisy realizations. The DIP by ADMM showed a slightly unstable behavior across different frames. This is likely because a single value of the hyperparameter ρ has varying efficacy for different time frames. The regular KEM was either equivalent to or slightly better than the DIP by OT. The proposed neural KEM further improved all the frames as compared to the regular KEM.

E. Comparison for ROI Quantification

Fig. 7 shows the trade-off between the bias and SD of different methods for ROI quantification in the blood and tumor regions. The two ROIs are the same as the anatomical regions marked in Fig. 4a. The curves were obtained by

varying the iteration number from 10 to 60 iterations with an interval of 10 iterations. As expected, the bias decreases as the iteration number increases in all the methods. For the blood ROI quantification, the DIP by OT was better than the DIP by ADMM because of the improved convergence and stability by OT in these three cases. For the tumor ROI quantification, the DIP by OT was either better than the DIP by ADMM in the low and medium cases (frame 2 and frame 12) or comparable in the high-count case (frame 24).

The proposed neural KEM achieved the lowest SD and bias simultaneously for frame 2 and frame 12 after 20 iterations, demonstrating its advantage for low- and medium-count frames. In the high-count case (frame 24), the traditional ML-EM achieved the lowest bias for small targets (i.e., the tumor and blood ROIs) due to a good recovery of the contrast at a high iteration number. The neural KEM was with a higher bias due to oversmoothness in this high-count case but it was still better than the regular KEM and DIP methods.

(a) Blood ROI quantification
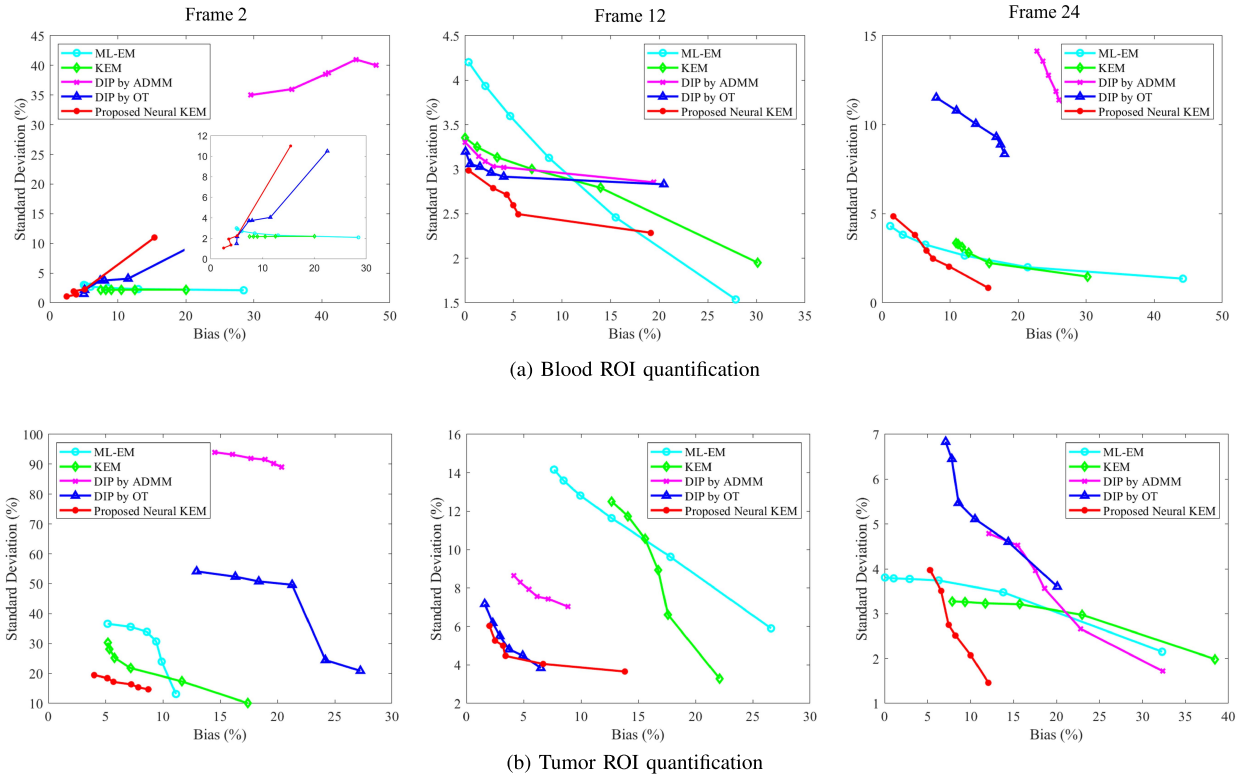

(b) Tumor ROI quantification

Fig. 7. Plots of bias-SD trade-off for ROI quantification in frame 2, 12 and 24 by varying the iteration number from 10 to 60 (i.e., from rightmost to leftmost on each curve). (a) Blood ROI quantification, (b) tumor ROI quantification. A zoom-in is included for frame 2 in (a).

F. Effect of Method Parameters

To demonstrate the challenge for choosing a proper ADMM parameter ρ in the DIP method, Fig. 8(a) shows the iteration-based MSE result for a range of ρ values for frame 12. An inappropriate ρ resulted in instability and oscillations across iterations. $\rho = 0.05$ was a good choice for this frame but resulted in poor MSE for some of other frames, as shown in Fig. 6(c). The OT algorithm avoids this difficult-to-tune parameter.

Similar to other methods that use neural network-based deep prior, the proposed neural KEM involves a sub-iteration number that needs to be determined for the neural-network learning step. Fig. 8(b) shows the effect of this sub-iteration number on the final image MSE for frame 2 (low-count), frame 12 (medium-count) and frame 24 (high-count). The result suggests a reasonable choice is 150 iterations, which also worked well for other frames. We also found that the CNN training is stable when the learning rate in the Adam optimizer ranges from 10^{-4} to 10^{-2}.

V. VALIDATION USING 3D COMPUTER SIMULATION

A. Simulation Setup

We also performed a fully-3D computer simulation study using an XCAT heart phantom for a GE Discovery 690 PET scanner. This scanner has 13,824 LYSO crystals, arranged in 24 ring detectors. The detection unit is composed of blocks consisting of 9×6 crystals, each containing a total of 64 blocks per ring. The sinogram size is $381 \times 553 \times 288$ and the image size is $137 \times 137 \times 47$. The voxel size was $4.0 \times 4.0 \times 3.3$ mm^3.
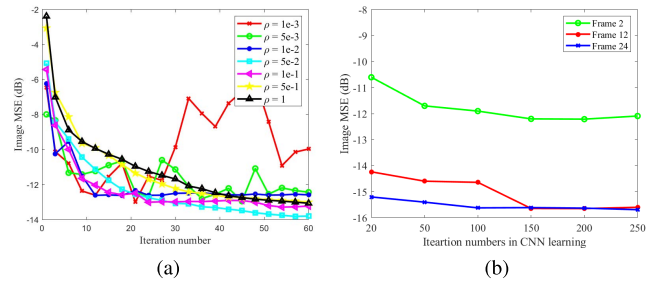

Fig. 8. Effect of (a) ADMM hyper-parameter ρ of the DIP method and (b) CNN learning subiterations of the neural KEM on the image MSE.

A one-hour dynamic ^{18}F-FDG scan was simulated using 49 time frames: 30×10s, 10×60s, and 9×300s. Here the framing scheme is adapted to capture the fast kinetics in the heart. The TACs of different regions were extracted from a real patient FDG PET scan to generate noise-free dynamic activity images. The images were then forward projected to generate noise-free sinograms. No time-of-flight information was simulated. To reduce time, SimSET-based simulation was not used for scatter simulation here. Instead, scattered and random events were simulated using a 40% uniform sinogram. Poisson noise was then generated with 1.25 billion expected events over 1 hour. Ten noisy realizations were simulated.

B. Reconstruction Methods

For the kernel methods, the kernel matrix $\boldsymbol{K}$ was built using four composite frames (one 5-min frame, one 15-min frame and two 20-min frames). Here compared to the 3 composite
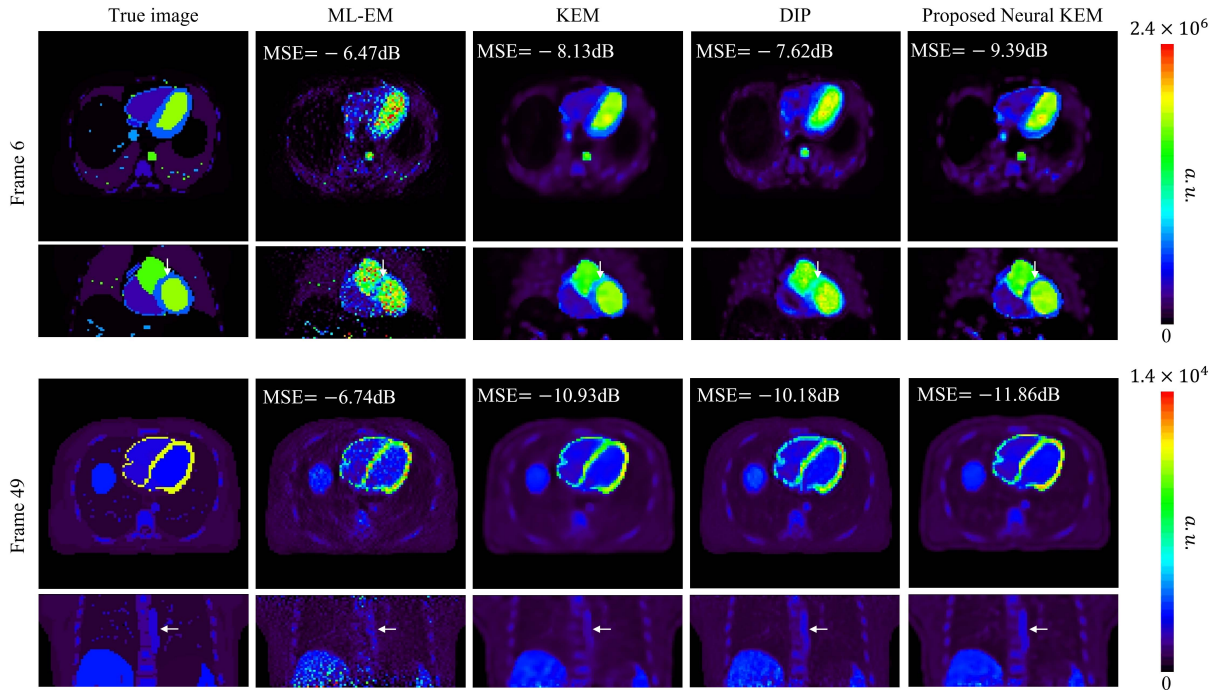
Fig. 9. True activity images and reconstructed images by different methods in a 3D computer simulation study. A low-count case (frame 6, top row) and a high-count case (frame 49, bottom row) are shown. Each image is shown in transverse and coronal views.
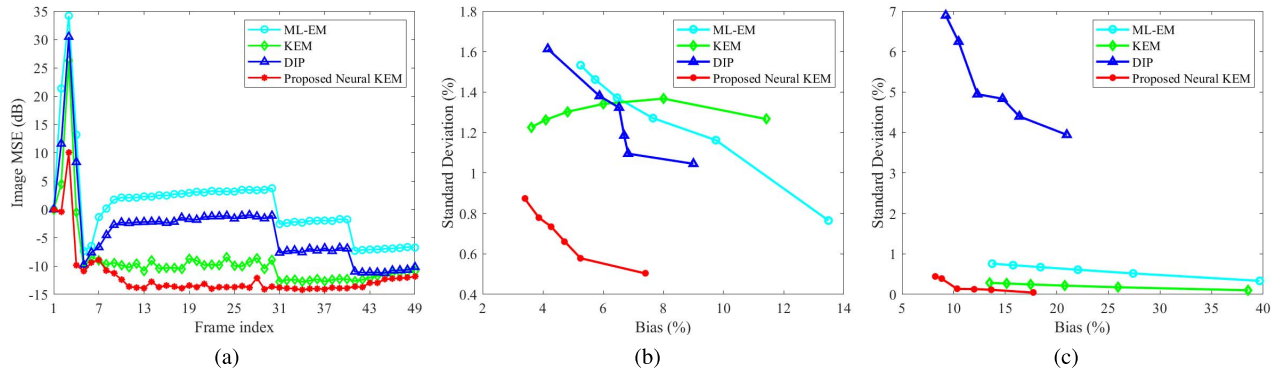

Fig. 10. Quantitative results from the 3D simulation study. (a) Plots of image MSE across all frames; (b-c) Plots of bias-SD trade-off for ROI quantification by varying the iteration number from 10 to 60 (i.e., from rightmost to leftmost on each curve). (b) Aorta ROI in a low-count case (frame 6), (c) myocardium ROI in a high-count case (frame 49).

frames used for the 2D simulation study, the first 20-min was divided into two shorter composite frames to better capture the early dynamic information for reconstructing the data of higher temporal resolution. The kNN search was performed in a $9 \times 9 \times 9$ local region with $k = 50$ nearest neighbors to reduce the computation time. The 3D version of U-net was used in the neural network-based methods with the 4 composite images as the network input. Similar to the 2D simulation study, 150 iterations were used for the neural-network learning step within each outer iteration. The proposed neural KEM was compared with ML-EM, regular KEM and DIP by OT using the image MSE and ROI bias-SD metrics. All the methods were run for 60 iterations starting from a uniform initial image.

C. Evaluation Results

Fig. 9 shows the true 3D activity images and reconstructed images at iteration 60 by different reconstruction methods for

frame 6 (early 10-s frame, low-count level) and frame 49 (late 5-min frame, high-count level). The proposed neural KEM achieved the lowest image MSE for both frames.

Fig. 10(a) shows the plots of image MSE for all frames reconstructed by different methods with 60 iterations. The two kernel-based methods (regular KEM and neural KEM) demonstrated a substantial improvement as compared to the ML-EM and DIP methods. The neural KEM was further better than the regular KEM particularly for those low-count frames.

Fig. 10(b) and fig. 10(c) show the trade-off between the bias and SD of different methods for ROI quantification in an aorta ROI of 3077 voxels and a myocardium ROI of 3744 voxels, respectively. The result of this 3D simulation study here is consistent with that of the 2D simulation study. The DIP method demonstrated a poor ROI bias-SD performance for the myocardial ROI quantification even though the associated image MSE was better than the ML-EM. In comparison, the

(a) Late frame at $t = 3300 - 3600s$

(b) Early frame at $t = 18 - 20s$

(c) Early frame at $t = 36 - 38s$
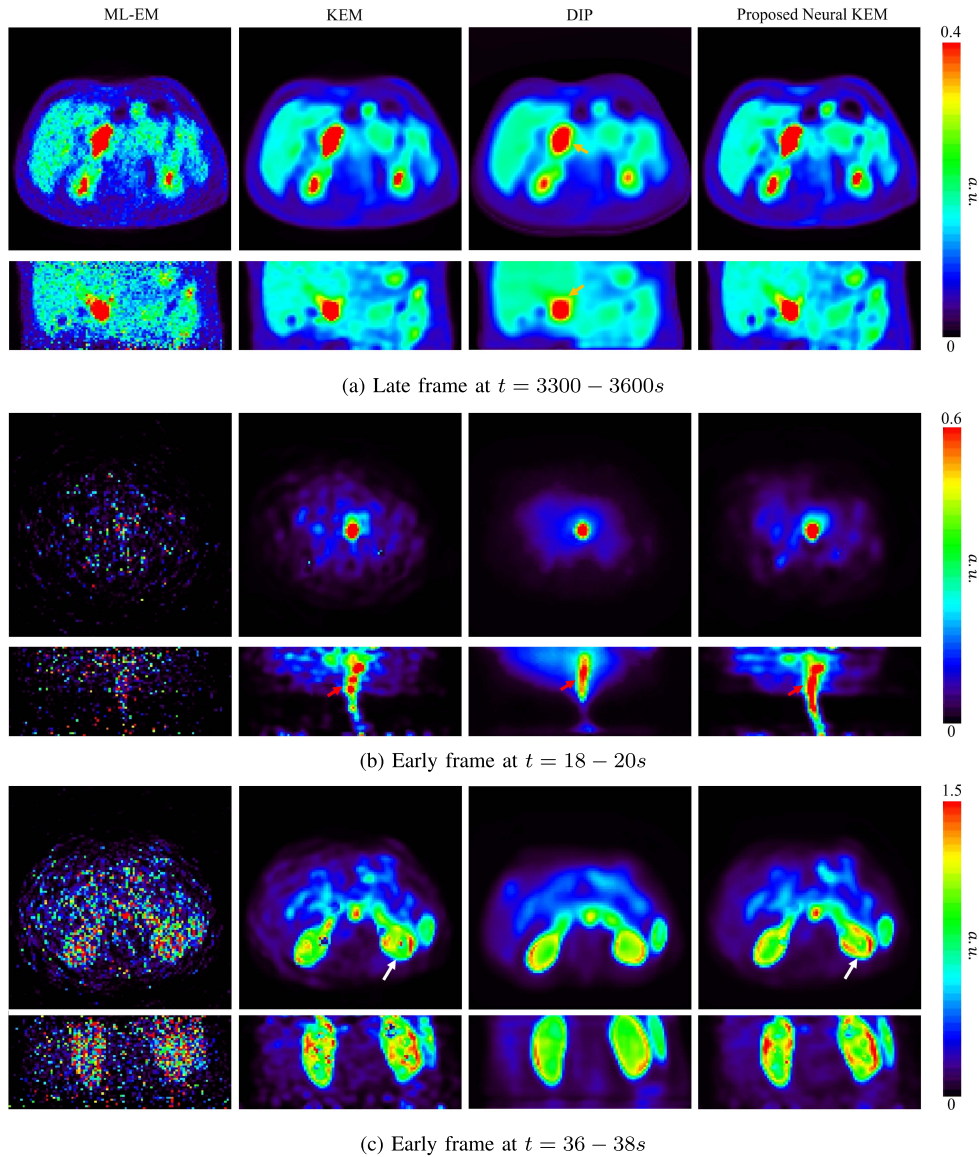
Fig. 11. 3D image reconstructions of (a) a late frame at $t = 3300 - 3600s$ and two early HTR frames at (b) $t = 18 - 20s$ and (c) $t = 36 - 38s$ by different methods. Here different frames are displayed for visualizing different targets of interest, i.e. tumor in (a), aorta in (b), and kidney in (c). Each reconstruction is shown in the transverse and coronal views.

proposed neural KEM achieved the best trade-off among the different reconstruction methods.

VI. APPLICATION TO REAL PATIENT DATA

A. Patient Data Acquisition

We have further applied the neural KEM to dynamic PET imaging for a real patient dataset. A cancer patient scan was performed on the GE Discovery 690 PET/CT scanner at the UC Davis Medical Center. The PET scan started right at the injection of 10 mCi ^{18}F-FDG and lasted 60 minutes. As our simulation studies have indicated that the neural KEM mainly benefits low-count frames, here we focus on high-temporal resolution dynamic imaging. The one-hour data are divided into 97 time frames following the schedule $60 \times 2s$, $18 \times 10s$, $10 \times 60s$, and $9 \times 300s$. A CT scan was acquired for attenuation correction. The projection data size was $381 \times 553 \times 288$ and the image size was $192 \times 192 \times 47$. All data corrections,

including normalization, attenuation correction, scatter correction and random correction, were extracted using the vendor software and included in the reconstruction process. Similar to the 3D simulation study, four composite frames (one 5-min frame, one 15-min frame and two 20-min frames)) were used to build the kernel matrix for this high-temporal resolution reconstruction. Other algorithm settings were also the same as used for the 3D simulation study described in Section V.B.

For ROI analyses, three ROIs were manually drawn on the corresponding CT images in the descending aorta, the tumor, and the kidney cortex regions. The volume of the blood ROI, tumor ROI and kidney ROI is 9 cm^3, 12 cm^3, and 15 cm^3, respectively.

B. Results of Reconstructed PET Images

Fig. 11 shows the comparison of different reconstructions for a late 5-minute frame ($t = 3300 - 3600s$) and two early
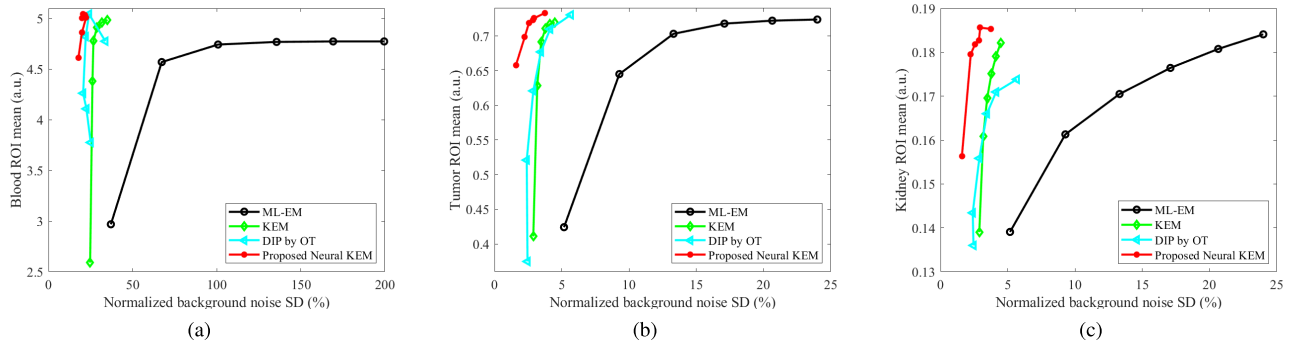
Fig. 12.   Plots of ROI mean of (a) the blood region, (b) tumor, and (c) kidney cortex versus liver background noise by varying iteration number from 10 to 60.

2-s frames ($t = 18 - 20s$, $t = 36 - 38s$). Each reconstruction is shown in the transverse and coronal views.

For the 5-minute frame which has a relatively high count level, the ML-EM reconstruction had good contrast in the tumor region though still contained high noise in the normal liver parenchyma. The DIP by OT caused over-smoothness and demonstrated a distortion in the tumor as pointed by the arrow in Fig. 11(a). With a preserved tumor shape similar to that of the ML-EM, both the regular KEM and neural KEM suppressed noise well and provided similar results in this high-count reconstruction, though the latter may have a slightly higher risk of oversmoothing small targets or sharp edges for higher count data due to the additional regularization from the use of deep coefficient prior.

For the 2-s frames, the ML-EM reconstructions were extremely noisy. The DIP by OT significantly reduced noise but tended to over-smooth the images. In Fig. 11(b), KEM resulted in a discontinuous aorta while the proposed neural KEM showed a more natural shape. In Fig. 11(c), the neural KEM also showed a more continuous renal cortex, which may benefit parametric imaging as will be presented in the next subsection. To sum up, the proposed neural KEM not only suppressed the noise in the background regions but also preserved structural contrast and details, though it cannot exclude a potential risk of over-regularization similar to any other methods that includes a regularization. Here the neural KEM and DIP by OT showed different anatomical structures in these two low-count frames. While there is no ground truth and the ML-EM was too noisy to provide a reference, the result from the high-count frame shown in Fig. 11(a) may imply the result by the neural KEM is more likely to be close to the truth of the low-count frames.

Fig. 12 further shows a quantitative comparison of different methods for ROI quantification in the blood ROI in frame 15 (where the uptake in the blood reaches its maximum), and in the tumor ROI and kidney ROI in the last frame (55-60 min). Here the ROI mean is plotted versus normalized background noise SD by varying the iteration number from 10 to 60. The proposed neural KEM achieved the best trade-off among all the other methods.

C. Demonstration for Parametric Imaging

Parametric imaging was also performed for the dynamic images of the same subject. Because different reconstruction

methods mainly make a difference for early-time frames which have a low count level (Fig. 11), here we focused on parametric imaging of early-dynamic data. A two-tissue compartment model with voxel-wise time delay estimation [41] was used to generate parametric images from the early-dynamic data. For each reconstruction method, the blood input function was derived from the descending aorta ROI.

Fig. 13 shows the parametric images of fractional blood volume v_b and FDG delivery rate K_1 generated from the early 120s data. The CT images are also shown for reference of anatomy. The ML-EM result suffered from heavy noise. The KEM result demonstrated a significant improvement but still suffered from noise-induced artifacts. The proposed neural KEM showed more complete and regular kidney cortex structures that seem consistent with the kidney anatomy and function [42]. The K_1 by DIP was much lower in the kidney cortex region, which can be explained by the underestimation of the renal uptake in the DIP-reconstructed activity image as shown in Fig. 11(c) and Fig. 12(c).

VII. Discussion

This work proposed an implicit regularization for improving the kernel method using deep coefficient prior and developed a neural KEM algorithm for neural-network based tomographic reconstruction. Because the loss function (24) used for the CNN learning is derived from the optimization transfer theory, the proposed neural KEM is thus guaranteed to monotonically increase the data likelihood. Compared to the ADMM used in most DIP reconstructions [18], [19], [21], [23], the optimization transfer algorithm does not introduce an additional hyper-parameter. The results shown in Fig. 6 and Fig. 7 indicate a more stable performance of the optimization transfer algorithm than the ADMM.

Our studies showed mixed results for comparing DIP with the standard KEM, while DIP was reported superior over KEM in [18] for MRI-guided PET image reconstruction. This is likely due to the use of different sources of image prior (MRI vs composite images of dynamic PET) and the fact that dynamic PET consists of frames of a range of count levels. The KEM or DIP alone demonstrated an instability (too noisy or too smooth) for dynamic PET image reconstruction. By combining them together, the proposed neural KEM achieved a much better performance than each individual method.
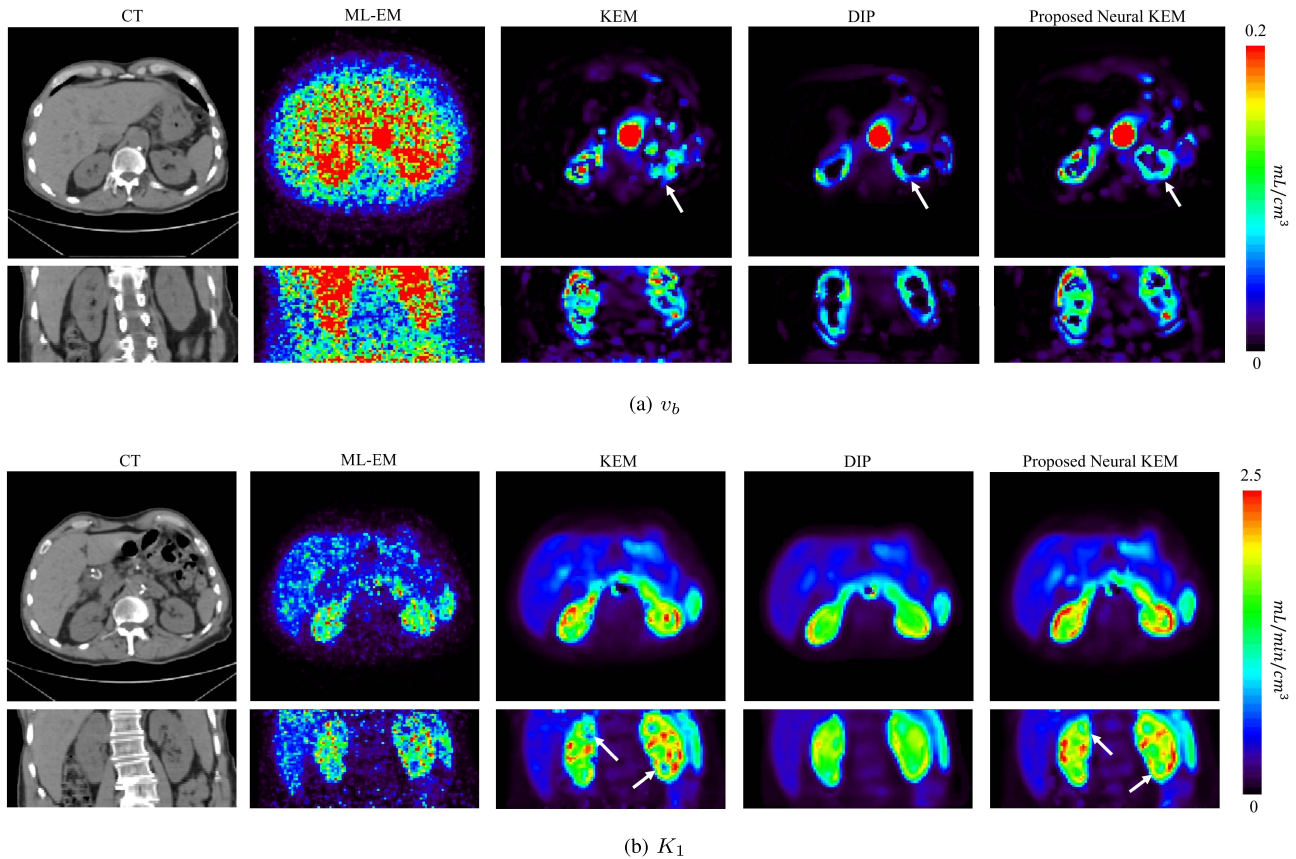
(a) v_b

(b) K_1

Fig. 13. Parametric images of (a) v_b and (b) K_1 generated from the early-dynamic images reconstructed using ML-EM, KEM, DIP and the proposed neural KEM. Each image is shown in transverse and coronal views.

The neural KEM in this work focused on frame-by-frame image reconstruction in the spatial domain but can be potentially extended to more general cases. For example, a spatiotemporal kernel method [3] allows both spatial and temporal correlations to be encoded in the kernel matrix. The proposed neural KEM algorithm may be combined with the spatiotemporal kernel method to further improve the dynamic image reconstruction of high-temporal resolution data. In addition, the construction of a spatial kernel itself can also be modified by using a different kernel function, e.g., using a pre-defined wavelet representation [40] or a neural-network representation [43]. The kernel construction can be further trained using deep learning as demonstrated in our recent work [44]. It is worth noting that all these methods are aimed at improving K and are therefore complementary to the proposed neural KEM which improves α in (5). Our preliminary results from computer simulation (not shown) have suggested that the observed benefit of using the deep coefficient prior on α in this paper can be also transferred into those methods that use a modified or trained kernel. A detailed study will be reported in our future work.

Compared to the standard KEM, the neural KEM introduces a nonlinear step (i.e., the neural network learning step), which brings the image quality improvement but adds extra computational cost. For example, the standard KEM took 30 minutes while the neural KEM took 70 minutes for reconstructing one frame in the real data study. A potential way to reduce the computational burden is to accelerate the speed of the neural network learning step either by improving the learning algorithm or by using pretraining for a better initial.

In this study, we demonstrated the performance of the proposed algorithm on conventional PET scanners. The recent advent of total-body PET scanners (e.g., [45], [46], [47], [48], [49]) has made it even more feasible to pursue low-dose dynamic imaging and high-temporal resolution dynamic imaging, especially for the entire body simultaneously. Our future work will also implement and evaluate the proposed neural KEM on total-body PET for parametric imaging.

VIII. CONCLUSION

In this paper, we have developed a neural KEM algorithm that combines the kernel method with deep coefficient prior. The algorithm is enabled by optimization transfer, leading to an easy-to-implement modularized implementation. Computer simulations and real patient data have demonstrated the improvement of the neural KEM over conventional KEM and DIP methods for dynamic PET imaging.

ACKNOWLEDGMENT

REFERENCES

[1] J. Qi and R. M. Leahy, "Iterative reconstruction techniques in emission computed tomography," *Phys. Med. Biol.*, vol. 51, no. 15, pp. R541–R578, Aug. 2006.

[2] G. B. Wang and J. Qi, "PET image reconstruction using kernel method," *IEEE Trans. Med. Imag.*, vol. 34, no. 1, pp. 61–71, Jan. 2015.

[3] G. B. Wang, "High temporal-resolution dynamic PET image reconstruction using a new spatiotemporal kernel method," *IEEE Trans. Med. Imag.*, vol. 38, no. 3, pp. 664–674, Mar. 2019.

[4] W. Hutchcroft, G. B. Wang, K. T. Chen, C. Catana, and J. Qi, "Anatomically-aided PET reconstruction using the kernel method," *Phys. Med. Biol.*, vol. 61, no. 18, pp. 6668–6683, Sep. 2016.

[5] P. Novosad and A. J. Reader, "MR-guided dynamic PET reconstruction with the kernel method and spectral temporal basis functions," *Phys. Med. Biol.*, vol. 61, no. 12, pp. 4624–4645, 2016.

[6] J. Bland et al., "MR-guided kernel EM reconstruction for reduced dose PET imaging," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 2, no. 3, pp. 235–243, May 2018.

[7] K. Gong, Z. Cheng-Liao, G. B. Wang, K. T. Chen, C. Catana, and J. Qi, "Direct Patlak reconstruction from dynamic PET data using the kernel method with MRI information based on structural similarity," *IEEE Trans. Med. Imag.*, vol. 37, no. 4, pp. 955–965, Apr. 2018.

[8] D. Deidda et al., "Hybrid PET-MR list-mode kernelized expectation maximization reconstruction," *Inverse. Probls.*, vol. 35, no. 4, 2019, Art. no. 044001.

[9] D. Deidda et al., "Effect of PET-MR inconsistency in the kernel image reconstruction method," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 4, pp. 400–409, Jul. 2019.

[10] X. Zhang et al., "Total-body dynamic reconstruction and parametric imaging on the uEXPLORER," *J. Nucl. Med.*, vol. 61, no. 2, pp. 285–291, 2020.

[11] D. L. Snyder and M. I. Miller, "The use of sieves to stabilize images produced with the EM algorithm for emission tomography," *IEEE Trans. Nucl. Sci.*, vol. NS-32, no. 5, pp. 3864–3872, Oct. 1985.

[12] J. E. Bowsher, V. E. Johnson, T. G. Turkington, R. J. Jaszczak, C. E. Floyd, and R. E. Coleman, "Bayesian reconstruction and use of anatomical a priori information for emission tomography," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 673–686, Oct. 1996.

[13] C. Chan, R. Fulton, D. D. Feng, and S. Meikle, "Regularized image reconstruction with an anatomically adaptive prior for positron emission tomography," *Phys. Med. Biol.*, vol. 54, no. 24, pp. 400–7379, 2009.

[14] J. Cheng-Liao and J. Qi, "PET image reconstruction with anatomical edge guided level set prior," *Phys. Med. Biol.*, vol. 56, no. 21, p. 6899, 2011.

[15] K. Vunckx et al., "Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 599–612, Mar. 2012.

[16] K. Kim et al., "Penalized PET reconstruction using deep learning prior and local linear fitting," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1478–1487, Jun. 2018.

[17] N. Xie et al., "Penalized-likelihood PET image reconstruction using 3D structural convolutional sparse coding," *IEEE Trans. Bio-Med. Eng.*, vol. 69, no. 1, pp. 4–14, Jan. 2022.

[18] 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.

[19] 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.

[20] K. Gong, C. Catana, J. Qi, and Q. Li, "Direct reconstruction of linear parametric images from dynamic PET using nonlocal deep image prior," *IEEE Trans. Med. Imag.*, vol. 41, no. 3, pp. 680–689, Mar. 2022.

[21] Z. Xie et al., "Generative adversarial network based regularized image reconstruction for PET," *Phys. Med. Biol.*, vol. 65, no. 12, 2020, Art. no. 125016.

[22] 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.

[23] Z. Xie, T. Li, X. Zhang, W. Qi, E. Asma, and J. Qi, "Anatomically aided PET image reconstruction using deep neural networks," *Med. Phys.*, vol. 48, no. 9, pp. 5244–5258, Sep. 2021.

[24] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," *Int. J. Comput. Vis.*, vol. 128, no. 7, pp. 1867–1888, Mar. 2020.

[25] G. B. Wang and J. Qi, "Iterative nonlinear least squares algorithms for direct reconstruction of parametric images from dynamic PET," in *Proc. IEEE ISBI*, May 2008, pp. 1031–1034.

[26] G. B. Wang and J. Qi, "An optimization transfer algorithm for nonlinear parametric image reconstruction from dynamic PET data," *IEEE Trans. Med. Imag.*, vol. 31, no. 10, pp. 1977–1986, Oct. 2012.

[27] J. A. Fessler, "Optimization transfer approach to joint registration/reconstruction for motion-compensated image reconstruction," in *Proc. IEEE ISBI*, Apr. 2010, pp. 596–599.

[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] A. J. Reader, G. Corda, A. Mehranian, C. Costa-Luis, S. Ellis, and J. A. Schnabel, "Deep learning for PET image reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 5, no. 1, pp. 1–24, Jan. 2021.

[30] K. Gong, E. Berg, S. R. Cherry, and J. Qi, "Machine learning in PET: From photon detection to quantitative image reconstruction," *Proc. IEEE*, vol. 108, no. 1, pp. 51–68, Jan. 2020.

[31] T. Wang et al., "Machine learning in quantitative PET: A review of attenuation correction and low-count image reconstruction methods," *Phys. Med.*, vol. 76, pp. 294–306, Aug. 2020.

[32] G. Wang, J. C. Ye, and B. De Man, "Deep learning for tomographic image reconstruction," *Nature Mach. Intell.*, vol. 2, no. 12, pp. 737–748, Dec. 2020.

[33] G. Wang, M. Jacob, X. Mou, Y. Shi, and Y. C. Eldar, "Deep tomographic image reconstruction: Yesterday, today, and tomorrow—Editorial for the 2nd special issue 'machine learning for image reconstruction,'" *IEEE Trans. Med. Imag.*, vol. 40, no. 11, pp. 2956–2964, Nov. 2021.

[34] I. Häggström, 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.

[35] H. Lim, I. Y. Chun, Y. K. Dewaraja, and J. A. Fessler, "Improved low-count quantitative PET reconstruction with an iterative neural network," *IEEE Trans. Med. Imag.*, vol. 39, no. 11, pp. 3512–3522, Nov. 2020.

[36] 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, Jan. 2021.

[37] L. A. Shepp and Y. Vardi, "Maximum likelihood reconstruction for emission tomography," *IEEE Trans. Med. Imag.*, vol. MI-1, no. 2, pp. 113–122, Oct. 1982.

[38] J. H. Friedman, J. L. Bentley, and R. A. Finkel, "An algorithm for finding best matches in logarithmic expected time," *ACM Trans. Math. Softw.*, vol. 3, no. 3, pp. 209–226, 1997.

[39] R. L. Harrison et al., "Preliminary experience with the photon history generator module of a public-domain simulation system for emission tomography," in *Proc. Conf. Rec. IEEE Nucl. Sci. Symp. Med. Imag. Conf.*, Nov. 1993, pp. 1154–11581.

[40] Z. Ashouri, G. Wang, R. M. Dansereau, and R. A. Dekemp, "Evaluation of wavelet kernel-based PET image reconstruction," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 6, no. 5, pp. 564–573, May 2022.

[41] G. B. Wang et al., "Total-body PET multiparametric imaging of cancer using a voxelwise strategy of compartmental modeling," *J. Nucl. Med.*, vol. 63, no. 8, pp. 1274–1281, 2022.

[42] C. Ruiz-Bedoya et al., "11C-PABA as a PET radiotracer for functional renal imaging: Preclinical and first-in-human study," *J. Nucl. Med.*, vol. 61, no. 11, pp. 1665–1671, 2020.

[43] S. Li and G. B. Wang, "Modified kernel MLAA using autoencoder for PET-enabled dual-energy CT," *Phil. Trans. Roy. Soc. A, Math., Phys. Eng. Sci.*, vol. 379, no. 2204, Aug. 2021, Art. no. 20200204.

[44] S. Li and G. B. Wang, "Deep kernel representation for image reconstruction in PET," *IEEE Trans. Med. Imag.*, vol. 41, no. 11, pp. 3029–3038, Nov. 2022.

[45] S. Cherry, R. D. Badawi, J. S. Karp, W. W. Moses, P. Price, and T. Jones, "Total-body imaging: Transforming the role of positron emission tomography," *Sci. Transl. Med.*, vol. 9, no. 381, Mar. 2017, Art. no. eaaf6169.

[46] R. Badawi et al., "First human imaging studies with the EXPLORER total-body PET scanner," *J. Nucl. Med.*, vol. 60, no. 3, pp. 299–303, 2019.

[47] B. Spencer et al., "Performance evaluation of the uEXPLORER total-body PET/CT scanner based on NEMA NU 2–2018 with additional tests to characterize PET scanners with a long axial field of view," *J. Nucl. Med.*, vol. 62, no. 6, pp. 861–870, 2021.

[48] J. Karp et al., "Imaging performance of the PennPET Explorer scanner," *J. Nucl. Med.*, vol. 59, p. 222, May 2018.

[49] A. Pantel et al., "PennPET Explorer: Human imaging on a whole-body imager," *J. Nucl. Med.*, vol. 61, no. 1, pp. 144–151, 2020.