

SPECT Reconstruction With a Trained Regularizer Using CT-Side Information: Application to ^{177}Lu SPECT Imaging

Hongki Lim , Member, IEEE, Yuni K. Dewaraja , Member, IEEE, and Jeffrey A. Fessler , Fellow, IEEE

Abstract—Improving low-count SPECT can shorten scans and support pre-therapy theranostic imaging for dosimetry-based treatment planning, especially with radionuclides like ^{177}Lu known for low photon yields. Conventional methods often underperform in low-count settings, highlighting the need for trained regularization in model-based image reconstruction. This article introduces a trained regularizer for SPECT reconstruction that leverages segmentation based on CT imaging. The regularizer incorporates CT-side information via a segmentation mask from a pre-trained network (nnUNet). In this proof-of-concept study, we used patient studies with ^{177}Lu DOTATATE to train and tested with phantom and patient datasets, simulating pre-therapy imaging conditions. Our results show that the proposed method outperforms both standard unregularized EM algorithms and conventional regularization with CT-side information. Specifically, our method achieved marked improvements in activity quantification, noise reduction, and root mean square error. The enhanced low-count SPECT approach has promising implications for theranostic imaging, post-therapy imaging, whole body SPECT, and reducing SPECT acquisition times.

Index Terms—Anatomical information, emission tomography, low-count quantitative SPECT, segmentation.

I. INTRODUCTION

SINGLE Photon Emission Computed Tomography (SPECT) is a nuclear imaging method that captures three-dimensional images using gamma-ray emitting radiotracers.

Manuscript received 28 February 2023; revised 15 August 2023; accepted 20 September 2023. Date of publication 25 September 2023; date of current version 4 October 2023. This work was supported in part by the National Cancer Institute under Grant R01CA240706, in part by the National Research Foundation of Korea funded by the Korea Government (MSIT) under Grant 2022R1F1A1069055, in part by the MSIT (Ministry of Science and ICT), Korea, through the Innovative Human Resource Development for Local Intellectualization support program under Grant IITP-2023-RS-2023-00259678 supervised by the IITP (Institute for Information & communications Technology Planning & Evaluation), in part by by Korea Institute of Energy Technology Evaluation and Planning (KETEP) funded by the Korea Government (MOTIE), Graduate School of Digital-based Sustainable Energy Process Innovation Convergence, under Grant RS-2023-00243974, and in part by Inha University Research Grant. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Gregory T. Buzzard. (*Corresponding author: Hongki Lim.*)

Hongki Lim is with the Department of Electronic Engineering, Inha University, Incheon 22212, South Korea (e-mail: hklim@inha.ac.kr).

Yuni K. Dewaraja is with the Department of Radiology, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: yuni@med.umich.edu).

Jeffrey A. Fessler is with the Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: fessler@umich.edu).

Digital Object Identifier 10.1109/TCI.2023.3318993

This technique visualizes functional processes within the body, aiding in diagnosis and treatment. Improving low-count SPECT to reduce noise and scan time, or enabling whole-body imaging is particularly relevant when imaging radionuclides with low photon intensities such as ^{177}Lu (208 keV gamma: 10%) [1] and ^{90}Y ($\leq 5\%$ bremsstrahlung yield) [2]. In dosimetry, low count-rates are often encountered when imaging at later time points to determine pharmacokinetics following a therapeutic administration. Furthermore, ultra-low count rates are encountered when imaging pre-therapy for dosimetry based treatment planning. However, quantifying images at low count-rates is difficult due to the high level of noise, particularly for voxel-level dose estimation.

To improve reconstruction for low-count SPECT, regularized model-based image reconstruction (MBIR) has been widely used. “Learned” regularizers in MBIR have significantly improved the quality and quantification of medical image reconstruction, as seen in applications of CT [3], [4], [5] and MRI [6], [7], [8], [9], compared to “mathematically designed” regularizers. Several studies have applied these techniques to emission tomography [10], [11], [12], [13]. Although there has been much research on incorporating anatomical information into emission tomography [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], most existing “trained” regularizers for emission tomography have yet to explore the utilization of anatomical information from dual-modality systems such as SPECT-CT or PET-MR. [40] is a learning-based method that utilizes side information from MR images to assist in PET image deblurring and super-resolution. However, this approach is categorized as an image-domain post-processing technique, while the proposed method addresses the problem of measurement to image.

Incorporating anatomical information into emission tomography can be beneficial because anatomical images have higher resolution than emission images and provide more accurate boundary information. Many existing methods that use anatomical information in emission tomography rely on regularizers that promote spatial smoothness within organ boundaries while restricting smoothing across boundaries. This is because the distribution of PET or SPECT tracer is likely to be smooth within organ boundaries. Many learning-based methods for medical image reconstruction are conceptualized as ‘unrolled’ versions

of iterative optimization methods [41]. Here, the traditional alternating minimization of iterative algorithms is laid out in a sequence, with certain sub-problems tackled by neural networks, characterizing a data-driven approach. Building on this foundation, our article reexamines the ‘using anatomical information’ approach to SPECT reconstruction.

We propose a trained version of [15], [31] that used a modified quadratic regularizer with finite differences between neighboring voxel values. The modification was the incorporation of CT-side information as a regularization weight, which was set to zero at the boundary to prevent penalizing the differences in voxel values at the boundary location. Instead of the conventional finite differencing matrix, we use a trained convolutional operators for a more data-driven regularization. Our treatment of CT-side information also diverges from previous methods. Instead of using the CT segmentation mask for boundary indication, our method adopts a data-driven approach, training a function to transform the CT segmentation outcomes for SPECT regularization. Moreover, our segmentation process is fully automated and driven by a cutting-edge segmentation network that processes CT images. This is in contrast to the manually drawn segmentations employed in [15], [31].

Our proposed method can be easily extended to an end-to-end framework by jointly training the parameters in the reconstruction and analysis tasks. [42] proposed an end-to-end approach for jointly performing reconstruction and analysis tasks, and our proposed method differs in that the output of the analysis task is fed back into the reconstruction process, making it a feedback system.

This article is a proof-of-concept study where we aimed to enhance the performance of SPECT/CT imaging, particularly in the context of ^{177}Lu DOTATATE therapy, a widely-used targeted radionuclide therapy for neuroendocrine tumors. Although ^{177}Lu is increasingly adopted in these therapies due to its dual benefits—delivering therapeutic effects and facilitating SPECT imaging—capturing accurate images remains a challenge, especially under low-count conditions commonly encountered in pre-therapy diagnostic imaging and in post-therapy imaging at late time points. Although there exists potential for the method to be generalized across other radionuclides, our research concentrates on ^{177}Lu to maintain focus and to provide a comprehensive examination within this context. Explorations with radionuclides beyond ^{177}Lu , such as ^{90}Y , and their associated nuances are considered pivotal directions for our subsequent research.

Part of this work was presented at the SPIE Medical Imaging 2020 [43]. We significantly extended this work with more sophisticated methods, including how we incorporate analysis result into the reconstruction process. Our current method provides a more flexible, data-driven approach, allowing the segmentation mask to be used for regularization in a more adaptive manner. Moreover, we applied the proposed method to measured ^{177}Lu SPECT data and added detailed analysis of the proposed method, including a comparison between different variants of the method.

The rest of this article is organized as follows. Section II describes the proposed method, a trained regularizer using CT-side

information for SPECT MBIR. Section II also provides details on how the ^{177}Lu SPECT simulation and real measurement datasets are created and what evaluation metrics are used to measure performance. Section III presents the results of different reconstruction methods on simulation and measurement data. Section IV discusses how the proposed method utilizes CT-side information and compares its variants. Section V concludes with future works.

II. METHODS

This section presents the background of the problem and reviews other work related to the proposed method, such as conventional regularized MBIR methods using CT-side information. Then, a detailed derivation is provided that inspires the final form of the proposed method and the method for obtaining CT-side information is discussed. Additionally, details on the simulation and measurement data and the evaluation metrics used to assess the efficacy of each reconstruction algorithm are provided.

A. Background

Regularized MBIR has been used widely to improve reconstruction quality of low-count SPECT:

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x} \geq 0} f(\mathbf{x}) + \beta R(\mathbf{x})$$

$$f(\mathbf{x}) := \mathbf{1}^T (\mathbf{A}\mathbf{x} + \bar{\mathbf{r}}) - \mathbf{y}^T \log(\mathbf{A}\mathbf{x} + \bar{\mathbf{r}}), \quad (1)$$

where $\mathbf{x} \in \mathbb{R}^{n_p}$ is the unknown SPECT image, $f(\mathbf{x})$ is Poisson negative log-likelihood for measurement $\mathbf{y} \in \mathbb{R}^{n_d}$ and estimated measurement means $\bar{\mathbf{y}}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \bar{\mathbf{r}}$, the matrix $\mathbf{A} \in \mathbb{R}^{n_d \times n_p}$ denotes the system model, $\bar{\mathbf{r}} \in \mathbb{R}^{n_d}$ denotes the mean background events such as the additive scatter contribution, $R(\mathbf{x})$ is a regularization term to control the noise, and β is the regularization parameter. Most existing mathematically designed SPECT regularizers penalize differences between neighboring pixels [44], [45]. A typical form of such regularizer includes $R(\mathbf{x}) = \frac{\beta}{2} \sum_{k=1}^K \|\check{\mathbf{c}}_k * \mathbf{x}\|_2^2$, where often $K = 3$ because $\check{\mathbf{c}}_k$ is a $[1, -1]$ finite difference filter along one of x, y , or z directions. Therefore, $R(\mathbf{x})$ penalizes the differences between adjacent voxels without taking into account anatomical boundaries. A conventional non-trained quadratic regularizer using CT-side information [15], [31] modifies the quadratic regularizer as follows:

$$R(\mathbf{x}; \check{\mathbf{w}}_k) = \frac{\beta}{2} \sum_{k=1}^K \|\check{\mathbf{c}}_k * \mathbf{x}\|_{\check{\mathbf{w}}_k}^2, \quad (2)$$

where $\check{\mathbf{W}}_k = \text{diag}\{\check{\mathbf{w}}_k\} \in \mathbb{R}^{n_p \times n_p}$ and $\check{\mathbf{w}}_k \in \mathbb{R}^{n_p}$ is a boundary indicator image with 0 values at anatomical boundaries. Because standard regularized SPECT reconstruction methods, such as EM-based algorithms, update the image at each iteration with $\nabla f(\mathbf{x})$ and $\nabla R(\mathbf{x}) = \beta \sum_{k=1}^K \check{\mathbf{c}}_k * (\check{\mathbf{W}}_k (\check{\mathbf{c}}_k * \mathbf{x}))$ where $\check{\mathbf{c}}_k$ is flipped convolutional kernel of $\check{\mathbf{c}}_k$, $\nabla R(\mathbf{x})$ is zero-valued where $\check{\mathbf{w}}_k$ is zero-valued, indicating the boundary region. Thus,

incorporating the CT-side information $\tilde{\mathbf{w}}_k$ prevents the regularizer from penalizing differences between adjacent voxels only where the anatomical boundaries are present.

B. A Trained Regularization Method Using CT-Side Information

Using “hand-crafted” filters, as in (2), is unlikely to be optimal, especially when the filters can be “trained” through data. Recently, applying trained regularizers to MBIR has significantly improved image quality and quantification in medical image reconstruction [46], [47]. However, most existing trained regularizers for emission tomography do not fully exploit the information from dual-modality systems such as SPECT-CT or PET-MR. In the present work, instead of using single-modality based regularizer $R(\mathbf{x})$, we propose to use multi-modality based regularizer $R(\mathbf{x}; \mathbf{s})$ that exploits the anatomical side information provided by segmentation “mask” variable \mathbf{s} . A segmentation network generates \mathbf{s} , with the input being a CT image acquired (nearly) simultaneously with the SPECT measurement. The regularization term is composed of trained convolutional operations and a 1-norm to promote sparsity:

$$R(\mathbf{x}; \mathbf{s}) = \min_{\mathbf{z}} \sum_{k=1}^K \frac{1}{2} \|\mathbf{c}_k * \mathbf{x} - \mathbf{z}_k\|_{\mathbf{W}_k}^2 + \alpha_k \|\mathbf{z}_k\|_1, \quad (3)$$

where $\mathbf{W}_k = \text{diag}\{\mathbf{w}_k\}$, $\mathbf{w}_k = g_k(\mathbf{c}_k * \mathbf{s})$, $\{\mathbf{c}_k \in \mathbb{R}^{n_c \times n_c \times n_c} : k = 1, \dots, K\}$ is a set of 3D convolution filters, $*$ denotes convolution operation, $\{\mathbf{z}_k \in \mathbb{R}^{n_p} : k = 1, \dots, K\}$ is a set of sparse codes, and K is the number of filters. As illustrated in Fig. 3, $g_k(\cdot)$ is a function for filtered segmentation mask and we define it as follows:

$$g_k(\nu) = \frac{2e^{-\kappa_k \nu^2}}{e^{-\kappa_k \nu^2} + 1}, \quad (4)$$

where $\kappa_k \in \mathbb{R}$ is a trainable parameter.

The way we obtain \mathbf{w}_k is inspired by $\tilde{\mathbf{w}}_k$ in (2) because $\tilde{\mathbf{w}}_k$ can be obtained by $\tilde{\mathbf{w}}_k = \check{g}(\tilde{\mathbf{c}}_k * \mathbf{s})$. Here, $\check{g}(\cdot)$ is equivalent to $g_k(\cdot)$ in (4), except there is no trainable parameter κ_k ($\kappa_k = 1$).

A traditional optimization approach for solving (1) with (3) uses a block coordinate descent algorithm that alternatively updates $\{\mathbf{z}_k\}$ and \mathbf{x} :

$$\mathbf{z}_k^{(n+1)} = \arg \min_{\mathbf{z}_k} \frac{1}{2} \|\mathbf{c}_k * \mathbf{x}^{(n)} - \mathbf{z}_k\|_{\mathbf{W}_k}^2 + \alpha_k \|\mathbf{z}_k\|_1 \quad (5)$$

$$= \mathcal{T}(\mathbf{c}_k * \mathbf{x}^{(n)}, \alpha_k \odot \mathbf{w}_k) \quad (6)$$

$$\mathbf{x}^{(n+1)} = \arg \min_{\mathbf{x}} f(\mathbf{x}) + \frac{\beta}{2} \sum_{k=1}^K \|\mathbf{c}_k * \mathbf{x} - \mathbf{z}_k^{(n+1)}\|_{\mathbf{W}_k}^2, \quad (7)$$

where \odot is element-wise division and $\mathcal{T}(\cdot, \cdot)$ is the element-wise soft thresholding operator:

$$\mathcal{T}(t, \mathbf{q})_j := \text{sign}(t_j) \max(|t_j| - q_j, 0). \quad (8)$$

When trained filters $\{\mathbf{c}_k\}$ satisfy $\sum_{k=1}^K \tilde{\mathbf{c}}_k * (\mathbf{w}_k \odot (\mathbf{c}_k * \mathbf{x})) = \mathbf{x}$ where $\tilde{\mathbf{c}}_k$ is a reversed version of \mathbf{c}_k and \odot denotes element-wise multiplication, the updates in (6) and (7) can be

Algorithm 1: EM-Based SPECT Reconstruction Algorithm Using Trained Regularizer with CT-Side Information.

Require:

$$\{\mathbf{c}_k^{(n)}, \mathbf{d}_k^{(n)}, \alpha_k^{(n)} : n = 1, \dots, T\}, \mathbf{y}, \bar{\mathbf{r}}, \mathbf{A}, c$$

Initialize:

$\mathbf{x}^{(0)}$ using EM algorithm

Calculate $a_j = \sum_{i=1}^{n_d} a_{ij}$

for $n = 0, \dots, T - 1$ **do**

$$\mathbf{u}^{(n+1)} = \sum_{k=1}^K \mathbf{d}_k^{(n+1)} * (\mathbf{W}_k (\mathcal{T}(\mathbf{c}_k^{(n+1)} * \mathbf{x}^{(n)}, \alpha_k^{(n+1)} \odot \mathbf{w}_k)))$$

$$\beta^{(n+1)} = \frac{\|a_j - \sum_{i=1}^{n_d} a_{ij} \frac{y_i}{\bar{y}_i(\mathbf{x}^{(n)})}\|_2}{\|\mathbf{x}^{(n)} - \mathbf{u}^{(n+1)}\|_2} \cdot c$$

for $n' = 0, \dots, T' - 1$ **do**

$$\delta = \frac{1}{2}(a_j - \beta^{(n+1)} u_j^{(n+1)})$$

$$\gamma = x_j^{(n')} \left(\sum_{i=1}^{n_d} a_{ij} \frac{y_i}{\bar{y}_i(\mathbf{x}^{(n')})} \right)$$

$$x_j^{(n'+1)} = \begin{cases} \frac{\sqrt{\delta^2 + \beta^{(n+1)} \gamma} - \delta}{\beta^{(n+1)}}, & \delta < 0 \\ \frac{\gamma}{\sqrt{\delta^2 + \beta^{(n+1)} \gamma} + \delta}, & \delta \geq 0 \end{cases}$$

end for

$$\mathbf{x}^{(n+1)} = \mathbf{x}^{(T')}$$

end for

rewritten as the following equivalent variable updates:

$$\mathbf{u}^{(n+1)} = \sum_{k=1}^K \tilde{\mathbf{c}}_k * \left(\mathbf{W}_k \left(\mathcal{T}(\mathbf{c}_k * \mathbf{x}^{(n)}, \alpha_k \odot \mathbf{w}_k) \right) \right)$$

$$\mathbf{x}^{(n+1)} = \arg \min_{\mathbf{x}} f(\mathbf{x}) + \frac{\beta}{2} \|\mathbf{x} - \mathbf{u}^{(n+1)}\|_2^2, \quad (9)$$

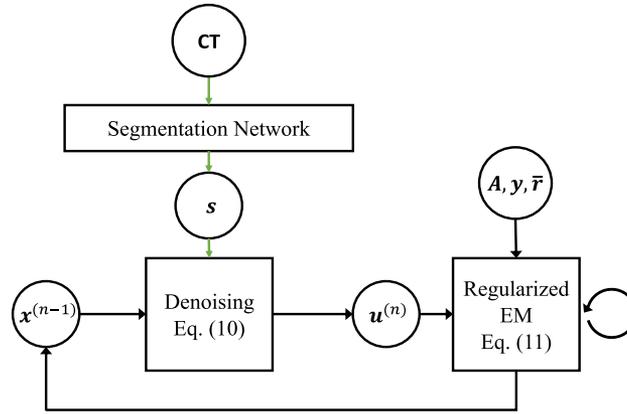
For the trained approach considered in this work, we use separate decoding filters $\{\mathbf{d}_k\}$ instead of using $\{\tilde{\mathbf{c}}_k\}$ to have more trainable parameters. Moreover, we use iteration-dependent filters and thresholding values so that the learned method can handle different levels (or types) of artifact at each iteration. Inspired by (9), hereafter, we define the variable updates for SPECT reconstruction as follows:

$$\begin{aligned} \mathbf{u}^{(n+1)} &= \sum_{k=1}^K \mathbf{d}_k^{(n+1)} * \left(\mathbf{W}_k \left(\mathcal{T}(\mathbf{c}_k^{(n+1)} * \mathbf{x}^{(n)}, \alpha_k^{(n+1)} \odot \mathbf{w}_k^{(n+1)}) \right) \right) \end{aligned} \quad (10)$$

$$\mathbf{x}^{(n+1)} = \arg \min_{\mathbf{x}} f(\mathbf{x}) + \frac{\beta}{2} \|\mathbf{x} - \mathbf{u}^{(n+1)}\|_2^2, \quad (11)$$

where separate encoding and decoding filters $\{\mathbf{c}_k^{(n)}\}$ and $\{\mathbf{d}_k^{(n)}\}$ are learned for the n th iteration during training. Our proposed framework alternatively updates the reconstructed image \mathbf{x} and the denoised image \mathbf{u} using convolution filters and soft-thresholding values that are trained at each iteration. To solve the \mathbf{x} -update in (11), we iteratively find a minimizer of the EM-surrogate of $f(\mathbf{x})$ [48]. We also apply an adaptive regularization parameter scheme proposed in [12] to automatically choose β

(a) High-level block diagram of our proposed method.



(b) Detailed block diagram of Denoising block in Fig. 1(a).

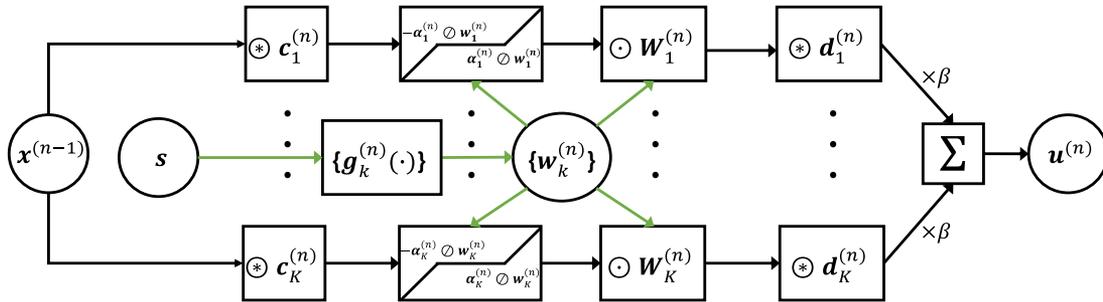


Fig. 1. (a) High-level overview of the proposed method: Our method operates iteratively, alternating between denoising and a regularized EM update. During the denoising step, we leverage the CT segmentation results. (b) Detailed block diagram of the denoiser, as referenced in (10). The output from the regularized EM ($\mathbf{x}^{(n-1)}$) undergoes convolution with a set of K filters ($\mathbf{c}_k^{(n)}$), is then soft-thresholded, and subsequently multiplied by the transformed segmentation mask ($\mathbf{w}_k^{(n)}$). This output is further convolved with another set of K filters ($\mathbf{d}_k^{(n)}$). The use of CT-side information in SPECT reconstruction is emphasized by green arrows.

regardless of count-level:

$$\beta^{(n+1)} = c \cdot \frac{\left\| a_j - \sum_{i=1}^{n_d} a_{ij} \frac{y_i}{\bar{y}_i(\mathbf{x}^{(n)})} \right\|_2}{\left\| \mathbf{x}^{(n)} - \mathbf{u}^{(n+1)} \right\|_2}, \quad (12)$$

where c is a user-defined parameter. Fig. 1 illustrates the corresponding architecture and Algorithm 1 provides a detailed pseudocode of the proposed method. In Algorithm 1, T is the total number of iterations where each iteration performs (10) and (11) and T' is the number of iterations for solving the \mathbf{x} -update (11).

In the training phase, we train the set of filters $\{\mathbf{c}_k\}$, $\{\mathbf{d}_k\}$, soft-thresholding values $\{\alpha_k\}$, and the parameters $\{\kappa_k\}$ for the function in (4) to map the estimated image from the previous iteration into a high quality image (e.g., EM reconstruction with high-count measurement) at each iteration:

$$\arg \min_{\{\mathbf{c}_k\}, \{\mathbf{d}_k\}, \{\alpha_k\}, \{\kappa_k\}} \sum_{l=1}^L \left\| \mathbf{x}_{\text{true}, l} - \sum_{k=1}^K \mathbf{d}_k * \mathbf{W}_k \left(\mathcal{T} \left(\mathbf{c}_k * \mathbf{x}_l^{(n)}, \alpha_k \odot \mathbf{w}_k \right) \right) \right\|_2^2$$

where L is the total number of training samples, $\{\mathbf{x}_{\text{true}, l} : l = 1, \dots, L\}$ is a set of high quality images and $\{\mathbf{x}_l^{(n)} : l = 1, \dots, L\}$ is a set of images estimated in the n th iteration. While end-to-end training is generally preferred, we opted for iteration-by-iteration training due to technical challenges in integrating the system model \mathbf{A} implementation within our deep learning-based framework.

C. Application to ^{177}Lu SPECT Imaging

To demonstrate performance, we applied the proposed method to patient SPECT/CT images acquired after ^{177}Lu DOTATATE, a targeted radionuclide therapy (TRT) for treatment of neuroendocrine tumors (NETs). This therapy is administered to patients in four cycles with a fixed activity, typically 7.4 GBq/cycle. ^{177}Lu is delivered preferentially to tumor cells by DOTATATE, which binds to the somatostatin receptors expressed by most well-differentiated NETs. ^{177}Lu , which is increasingly used in TRT, emits both beta particles that deliver the therapeutic effect and gamma rays suitable for SPECT imaging of the activity distribution for visual assessment and dosimetry. Multiple studies report on ^{177}Lu SPECT/CT imaging after ^{177}Lu DOTATATE therapy cycles for the purposes of performing dosimetry to adjust subsequent cycles on an individual

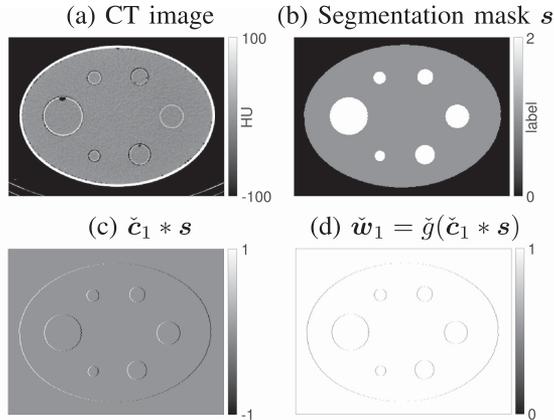


Fig. 2. Illustration of how CT-side information is incorporated into the conventional SPECT reconstruction method [15]. \hat{c}_1 is a [1,-1] finite difference filter that is convolved along the x -axis. \hat{w}_1 is zero-valued in the boundary region.

basis [49], [50]. However, currently pre-therapy ^{177}Lu imaging is rarely performed relying instead on ^{68}Ga DOTATATE PET/CT imaging to select patients for therapy. This is because of the superior spatial resolution of PET and the challenges of ^{177}Lu imaging under the very low count conditions encountered in (diagnostic) pre-therapy imaging. ^{68}Ga PET imaging however, is not well-suited for dosimetry because of the short half-life of ^{68}Ga and there are distinct advantages of using the same compound for both pre- and post therapy imaging. Pre-therapy ^{177}Lu imaging with tracer quantities of activity (~ 210 Mbq) has been reported for both ^{177}Lu DOTATATE [51] and ^{177}Lu PSMA [52] therapies. Furthermore, very low count conditions are encountered even after therapy administration when imaging at late time points to determine pharmacokinetics for dosimetry.

D. Neural Network for CT Segmentation

For the image segmentation network, we used nnUNet [53] based on the recommendations from [54]. Ref. [54] comprehensively compared recent state-of-art medical image segmentation methods and reported that a vanilla version of nnUNet performed reasonably well compared to other heavier models. We used the pre-trained version of nnUNet provided by the article's authors. The authors provided many pre-trained models, and we used the 3D full resolution mode of the model trained for the Task 3 of Medical Segmentation Decathlon [55] that contains different types of liver tumors, including hepatocellular carcinoma, cholangiocarcinoma, and metastatic tumors [56]. ^{177}Lu DOTATATE is for the treatment of neuroendocrine tumors, mostly located in the liver (83% of ^{177}Lu DOTATATE patients have liver metastases [57]), and the healthy liver also has relatively high uptake. Fig. 4(a) and (b) shows an example of a ^{177}Lu patient CT image and the corresponding segmentation prediction by nnUNet.

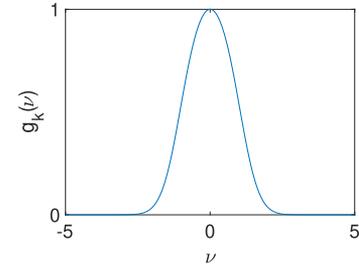


Fig. 3. Function $g_k(\cdot)$ for $\kappa_k = 1$. κ_k is a trainable parameter that changes the width of the function.

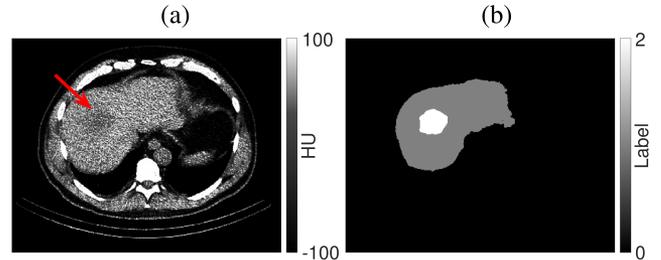


Fig. 4. (a) One CT image of a ^{177}Lu patient acquired on a Symbia SPECT-CT; (b) Corresponding segmentation result generated by nnUNet.

TABLE I
SEVERAL KEY SETTINGS FOR SIMULATION DATA

Activity concentration ratio (hot:warm)	$4 \pm \mathcal{N}(0, 1)$
Primary counts	$800\text{K} \pm \text{round}(\mathcal{N}(0, 10^{10}))$
Scatter fraction (%)	$54.5 \pm \mathcal{N}(0, 10^2)$

Each data sample has different values for key settings.

E. Dataset

To validate the efficacy of the proposed method, we applied the proposed reconstruction method to both simulation and clinical data.

1) *Simulation Data:* We used LiTS (Liver Tumor Segmentation) dataset [56] for the simulation study. Among 130 training samples in 3D, we used 30 samples for training denoiser in (10), and 10 samples for testing the performance of different algorithms. We used the label images of LiTS dataset to generate the true SPECT image by changing the tumor value to $4 \pm \mathcal{N}(0, 1)$ to set the tumor-to-liver ^{177}Lu uptake ratio, because 4:1 is a typical value in patients studies. We downsampled the CT-sized label image to SPECT-sized image with a voxel size $4.8 \times 4.8 \times 4.8$ (mm³). The number of slices is different between each 3D sample in the LiTS dataset. Fig. 5(a) and (b) show the CT image provided by the dataset and the segmentation mask (s in (3)) predicted by nnUNet. Fig. 5(c) is the true SPECT image that we generated based on label. We simulated the extremely low-count realizations with realistic total primary counts and scatter fraction based on numbers from patient imaging (summarized in Table I) using the system matrix generated by Michigan Image Reconstruction Toolbox.¹

¹[Online]. Available: <https://web.eecs.umich.edu/fessler/code/>

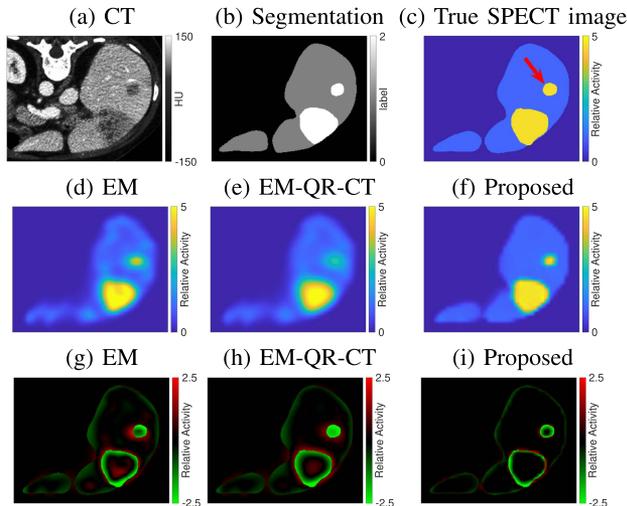


Fig. 5. Simulation result: (a) CT slice corresponding to the true SPECT image in (c). (b) Segmentation estimated by nnUNet. (d)–(f) Reconstructed images. (g)–(i) Difference maps compared to GT.

TABLE II
PROJECTION PRIMARY COUNTS (COUNT RATE \times DURATION) RANGE FOR CLINICAL DATA, RESAMPLED TO SIMULATE A 1-MINUTE EQUIVALENT SCAN

	Day0	Day5
Patient	492K - 1.1M	150K - 250K
Sphere Phantom	1.3M	

The data includes studies from 6 patients across multiple acquisition time points, as well as a phantom study.

2) *Clinical Patient Data and Physical Phantom Measurement Data*: To train the denoising network with the given segmentation mask, we used four ^{177}Lu patient studies with multiple acquisitions. The patient images were acquired at 1-5 days after 7.4 GBq ^{177}Lu DOTATATE therapy. To generate low-count realizations in the training dataset (4 patients, 2 cycles, total number of samples $L=25$), we resampled [58] the post-therapy 25-minute measurement data using Poisson resampling with a rate of 4% to generate 1-minute equivalent scan, as ultra-low count rates are to be expected in pre-therapy imaging application where the administered activity will be only a fraction of that used for therapy. Table II summarizes the primary counts. To test and compare the different reconstruction methods, we used a ^{177}Lu phantom measurement and two patient studies that were not used for training. As in the training data, patient images were acquired at multiple time points, and the total number of patient projection data we acquired and resampled for testing was 8. The phantom has six hot spheres (2,4,8,16,30 and 113 mL) in a warm background and its CT image is shown in Fig. 7(a). Including this one additional phantom measurement, the total number of testing samples is 9. For the phantom study, we used the segmentation mask manually drawn on the CT because the geometry of phantom is far different from the human abdomen, therefore the trained network does not work for the phantom. We did not quantitatively evaluate the output of nnUNet for segmentation since the ground truth is unknown in patients. Our aim is to improve the reconstruction, therefore we checked for any obvious failed cases through visual inspection.

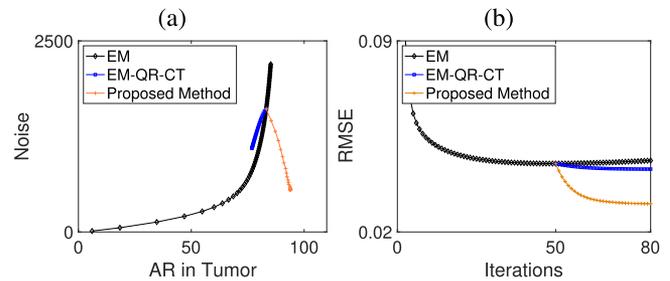


Fig. 6. Simulation result: (a) Noise vs activity recovery in tumor. (b) RMSE vs iteration. Averaged 10 test samples. The true image was the GT.

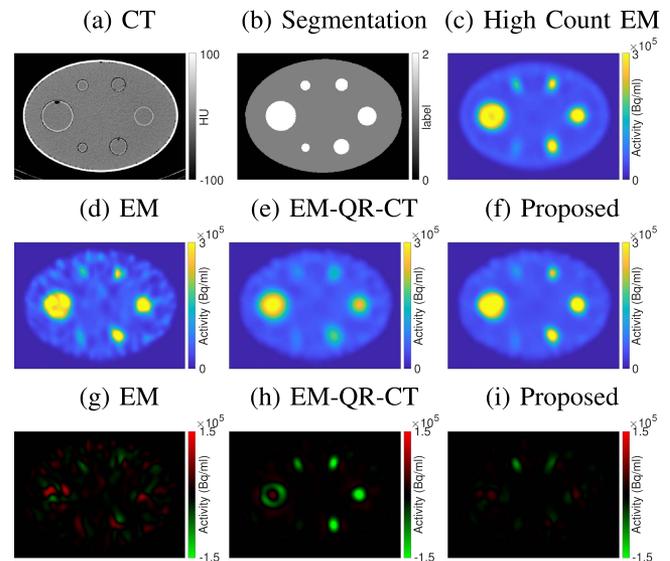


Fig. 7. Physical phantom result: (a) CT slice corresponding to the high-count EM image in (c). (b) Manually drawn segmentation mask. (d)–(f) Reconstructed images. (g)–(i) Difference maps compared to high count EM image.

F. Training SPECT Denoiser

We trained K sets of 3D convolutional filters ($\{\mathbf{c}_k^{(n)}\}, \{\mathbf{d}_k^{(n)}\} \in \mathbb{R}^{3 \times 3 \times 3}$), thresholding values ($\alpha_k^{(n)}$), and parameter for ($\kappa_k^{(n)}$) for each iteration using the PyTorch (version 1.2) [59] deep-learning library. We trained a 10 outer-iteration networks where each outer-iteration has $K = 49$ sets of thresholding values and convolutional encoding/decoding filters for simulation data, and $K = 100$ sets for real data. Each outer-iteration used 3 inner-iterations for the x -update. We used the Adam optimization method to train the network. The denoising network was trained for 150 epochs in the first outer iteration and 50 epochs in the rest of the iterations, with the parameters initialized using the values from the previous iteration. We applied a learning rate decay method (learning rate = learning rate \times 0.9 at every 20 epoch).

G. Evaluation Metrics

We evaluated SPECT reconstruction performance with activity recovery (AR) of tumor or hot region, image-ensemble-noise in warm background region (healthy liver) across realizations,

and root mean square error (RMSE):

$$\text{AR}(\%) = \frac{\hat{C}_{\text{tumor}}}{\bar{C}_{\text{tumor}}} \times 100$$

$$\text{Noise}_{\text{BKG}}(\%) = \frac{\sqrt{\frac{1}{J_{\text{BKG}}} \sum_{j \in \text{BKG}} \left(\frac{1}{M-1} \sum_{m=1}^M (\hat{\mathbf{a}}_m[j] - \frac{1}{M} \sum_{m'=1}^M \hat{\mathbf{a}}_{m'}[j])^2 \right)}}{\frac{1}{J_{\text{BKG}}} \sum_{j \in \text{BKG}} \frac{1}{M} \sum_{m=1}^M \hat{\mathbf{a}}_m[j]} \times 100,$$

where \bar{C}_{VOI} is true mean counts in the volume of interest (VOI), \hat{C}_{VOI} is estimated mean counts in the VOI, M is total number of realizations, and J_{BKG} is the total number of voxels in background liver region. In the simulation data evaluation, we considered the image used to simulate projections as the true image. In the clinical data evaluation, we considered the high-count EM image obtained after 50 iterations as the true image, as there is no ground truth (GT) available for the patient data.

III. RESULTS

We compared standard EM, the regularized EM method with CT-side information (EM-QR-CT) which uses the quadratic regularization $R(\mathbf{x}; \tilde{\mathbf{w}}_k)$ in (2), and the proposed method. For the regularized methods, including the proposed method, we used 50 iterations of EM algorithm to obtain the initial image $\mathbf{x}^{(0)}$. We used 30 iterations for the conventional regularizers and 10 (3) outer (inner) iterations for the proposed method.

A. Results on Simulation Data

Fig. 5(d)–(f) visually compares the reconstructed images using different methods. Fig. 5(a) shows the CT slice where the reconstructed images are displayed. Fig. 5(b) is the segmentation mask predicted by nnUNet, and both EM-QR-CT and the proposed method use this mask in their reconstruction process. Fig. 5(g)–(i) display the difference maps between the reconstructed images and the true image in Fig. 5(c). In this comparison, the regularization parameter c in (12) is set to 2^{-8} in EM-QR-CT and 2^{-6} in the proposed method. The regularization parameter values are chosen based on the results shown in Fig. 12. Because we simulated very low-count setting, the EM reconstructed image is noisy in the uniform regions. EM-QR-CT reduces the noise, but it also removes fine details in the image, as can be seen in the small tumor region indicated by the red arrow in Fig. 5(c). The proposed method reduces noise and preserves fine details simultaneously. The reconstructed image using the proposed method is closest to the true image.

The numerical evaluation results in Fig. 6(a) align with the visual comparison. While EM algorithm improves the recovery in the hot (tumor) region, it also increases the noise as it iterates. EM-QR-CT reduces noise, but it worsens the recovery in the hot region. In contrast, the proposed method enhances both noise reduction and recovery in the tumor region simultaneously. Compared to EM-QR-CT and EM, the proposed method improved RMSE by 29.5% and 34.3%, respectively where the iteration numbers were 80 for all methods.

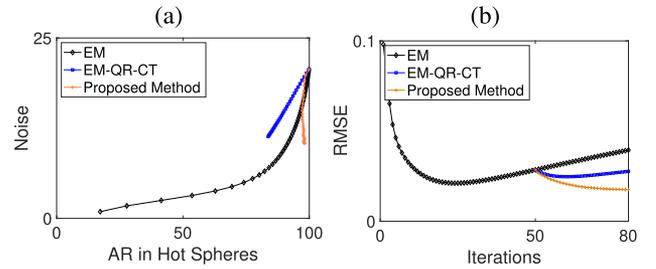


Fig. 8. Physical phantom result: (a) Noise vs Activity recovery in hot spheres (6 spheres averaged). (b) RMSE vs iteration. The high count EM image was the GT.

B. Results on Measured Data

For the phantom study, Fig. 7(c) shows the reconstructed image obtained after 50 iterations of EM algorithm, using high-count projection data. For this study, a segmentation mask (Fig. 7(b)) was manually drawn on the CT (Fig. 7(a)) because the geometry of phantom is different from the human abdomen, and thus, the trained nnUNet cannot be applied. The manually drawn mask was used for the EM-QR-CT and the proposed method. Fig. 7(d)–(f) show the reconstructed images using different methods. Similar to the visual comparison in the simulation study, EM exhibits a noisy texture in the uniform background, and EM-QR-CT improves the noise while overly smoothing the small hot spheres. The visual quality of the proposed method is closest to that of high-count EM. The error maps in Fig. 7(g)–(i) shows that the underestimation and overestimation are scattered across the entire regions in the EM image, while underestimation in the hot regions is evident in the EM-QR-CT image. Compared to the other methods, the proposed method shows the least errors, and there is no obvious pattern in the error map. The EM image may appear to depict the smallest sphere with greater clarity. However, it is important to note that this perceived clarity is a result of the overestimation of the AR by the EM algorithm. For the smallest sphere, the EM algorithm overestimates the AR by roughly 35%, whereas our proposed method underestimates it by approximately 15%. Even though both methods exhibit biases, the bias in our proposed method is substantially lower than that observed with the EM algorithm.

Fig. 8 shows the numerical evaluation results for the physical phantom study. The trend of plots is similar to that of the simulation results. In Fig. 8(a), EM increases the recovery while also increasing the noise and EM-QR-CT improves the noise while decreasing the recovery. The proposed method improves the noise while maintaining the recovery in the hot spheres. In Fig. 8(b), the proposed method achieved a lower RMSE compared to the other methods.

Fig. 9(d)–(f) show the reconstructed images for patient data. The reconstructed images are visualized with a coronal view. The image qualities of each algorithm is similar to the results of the previous simulation and physical phantom studies. EM-QR-CT generates a relatively less noisy image in the uniform region compared to EM, but the region near the tumor is overestimated. This is because the hot region is highly overestimated in the low-count EM and EM-QR-CT spreads the counts into adjacent

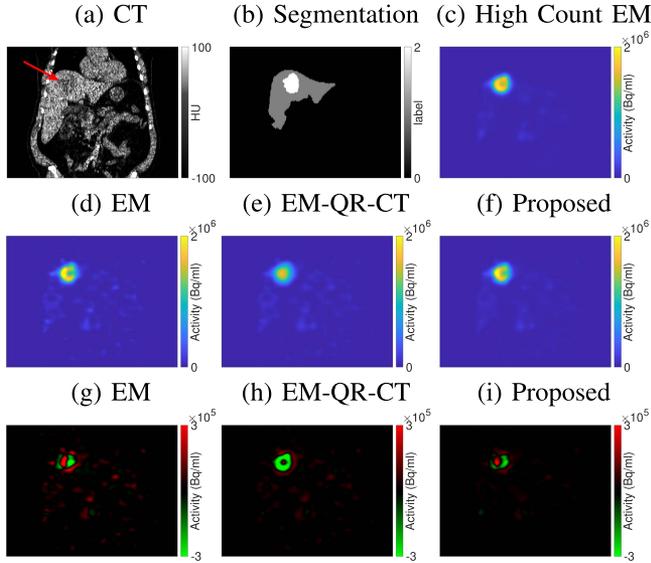


Fig. 9. Patient result: (a) CT slice corresponding to the high-count EM image in (c). (b) Segmentation estimated by nnUNet. (d)–(f) Reconstructed images. (g)–(i) Difference maps compared to high count EM.

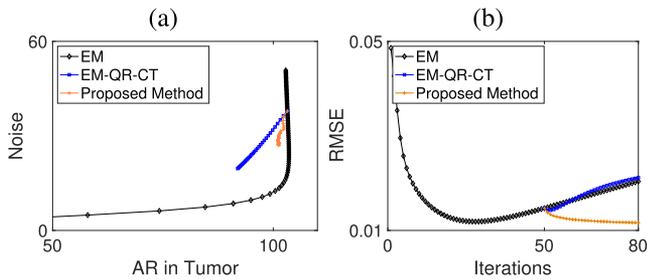


Fig. 10. Patient result: (a) Noise vs activity recovery in tumor. (b) RMSE vs iteration. Averaged 8 test samples. Patient high count EM images were used as ground truth.

voxels. As shown in Fig. 9(i), the proposed method does not spread the overestimated counts.

Fig. 10 shows the averaged numerical evaluation results for 8 patient samples. As shown in Fig. 10(a), the proposed method improves the positive bias in tumor region while reducing the noise. In Fig. 10(b), the proposed method achieved a lower RMSE by 40.4% compared to EM-QR-CT and by 42.9% compared to EM (iteration number = 80 for all methods).

IV. DISCUSSION

This work demonstrated the effectiveness of a sparsity-based trained regularizer using anatomical information, both qualitatively and quantitatively, in ^{177}Lu SPECT-CT imaging under very low-count conditions, typical for diagnostic imaging, and compared it to traditional non-trained regularizer using anatomical information. The proposed approach uses a learned denoising neural network that leverages information from a dual modality system such as SPECT-CT. Specifically, the proposed

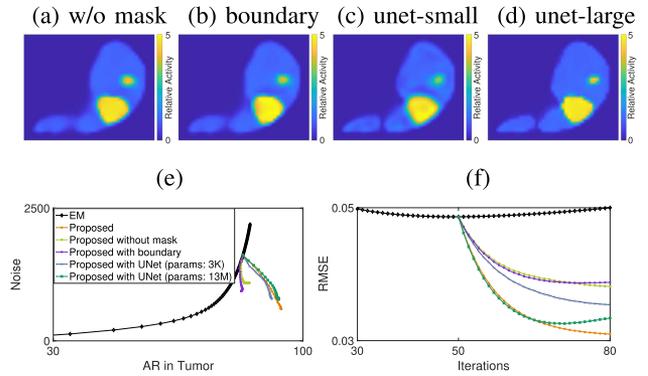


Fig. 11. Performance evaluation of variants of the proposed method.

framework employs a segmentation mask generated by a segmentation neural network as a weight for the trained regularizer. The proposed approach improved the recovery accuracy in the volume of interest for ^{177}Lu SPECT-CT simulation data compared to other non-trained regularizers. The improvements were also observed for ^{177}Lu SPECT-CT phantom measurements and patient measurement data.

Fig. 11 compares different variants of the proposed methods. The “Proposed without mask” method replaces $R(\mathbf{x}; \mathbf{s})$ in (3) with $R(\mathbf{x})$ [12] as follows:

$$R(\mathbf{x}) = \min_{\mathbf{z}} \sum_{k=1}^K \frac{1}{2} \|\mathbf{c}_k * \mathbf{x} - \mathbf{z}_k\|_2^2 + \alpha_k \|\mathbf{z}_k\|_1.$$

All other parts of the algorithm and training settings remain unchanged. The “Proposed with boundary” method [43] replaces \mathbf{w}_k with the boundary image $\tilde{\mathbf{w}}$. $\tilde{\mathbf{w}}$ is calculated as follows:

$$\tilde{\mathbf{w}} = (\tilde{\mathbf{w}}_1^{\circ 2} + \tilde{\mathbf{w}}_2^{\circ 2} + \tilde{\mathbf{w}}_3^{\circ 2})^{\frac{1}{2}},$$

where \circ is Hadamard (element-wise) power, $\tilde{\mathbf{w}}_k$ is obtained with finite differencing matrix along one of $x, y,$ or z directions. The “Proposed with UNet” replaces the denoiser in (10) with the UNet [60]. The input to UNet is a concatenation of $\mathbf{x}^{(n)}$ and \mathbf{s} in the channel dimension. We evaluated two different sizes of UNet. One model has 3254 parameters, which is similar in size to the proposed method ($2744, 3 \times 3 \times 3 \times 49 \times 2 + 49 \times 2$). The other model has 13 million parameters. The evaluation results show that methods incorporating the full mask outperform the method without mask and the method using boundary. Among the methods using the full mask, the original proposed method performs better than the UNet-based method, especially when the number of parameters is similar. Even when UNet has many more parameters, the proposed method improved all the evaluation metrics further than UNet-based method.

Fig. 12 illustrates the impact of the regularization parameter c on the reconstruction performance. For the EM-QR-CT method, no value of the regularization parameter improved both noise and recovery accuracy. A value of $c = 2^{-8}$ appeared to be optimal because it reduces noise without sacrificing accuracy much, resulting in the lowest RMSE. In the proposed method, all values larger than $c = 2^{-10}$ improve both noise and accuracy. However, the RMSE value slightly increases in later iterations for larger

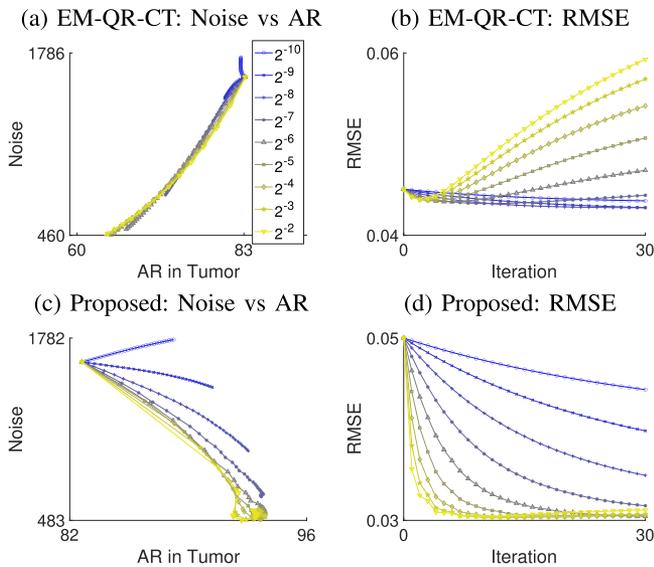


Fig. 12. Impact of parameter c on the performance in regularized methods.

values of c . This is due to the regularizer having a tendency to overfit to the training data, and too large of a c value forcing the neglect of the data fidelity term. Finding the appropriate value of c could enhance the generalization capability.

Using anatomical information would be more beneficial when the anatomical and emission images align well, as in the simulation study. Because nnUNet is trained on the LiTS dataset (Task 3 dataset of the Medical Segmentation Decathlon slightly modified LiTS dataset), the segmentation performance in the simulation study is better than in clinical data. Additionally, the way we generated the simulation dataset assumes that the anatomical liver and tumor regions match with the ^{177}Lu uptakes. However, the side information can be mismatched due to misregistration or differences between information captured by the functional (SPECT) and the anatomical (CT) imaging modalities. For example, anatomical image might show no anomalous voxels where functional images have high uptake, and vice versa. Moreover, in worst case, the segmentation method might fail to predict the volume of interest correctly especially when performed on CT of SPECT/CT that are not of diagnostic quality. In our clinical study, there were several obvious failed cases in the training and testing data (2 cases in training, 1 case in testing). Fig. 13 displays the CT image of failed case in test data, corresponding segmentation result, and the evaluation metrics for a failed sample (Patient6 Cycle1 Day1) and another sample (Patient6 Cycle1 Day2, a scan of the same patient at a different time point) whose segmentation result is correct. Despite the incorrect segmentation result for the Patient6 Cycle1 Day1 case, the proposed method improved the RMSE (11.12%) compared to the initial image $x^{(0)}$ after 50 iterations of the EM algorithm, though the improvement was lower than for the Patient6 Cycle1 Day2 case (24.65%). One possibility to enhance the performance of the segmentator is to use the diagnostic quality baseline CT images instead of the CT of SPECT/CT. However, this introduces additional challenges with misregistration and changes in the anatomy due to the fact

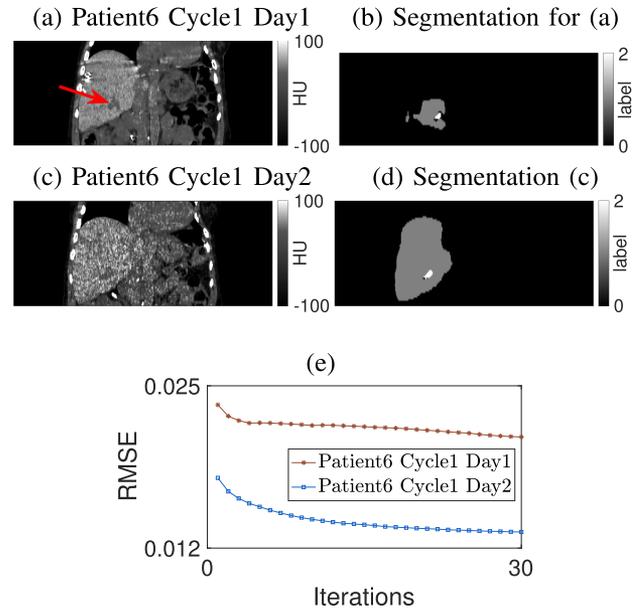


Fig. 13. (a)–(b) Patient6 Cycle1 Day1: CT and corresponding segmentation. (c)–(d) Patient6 Cycle1 Day2: CT and corresponding segmentation. (e) Patient6 Cycle1 Day1 and Patient6 Cycle1 Day2 are the scans of identical patient at different time points, however, greater RMSE improvement is made when the segmentation result is correct. The discrepancy in image quality between (a) and (c) is due to the different values of mAs used to reduce the CT exposure. Specifically, (a) was captured using a higher mAs, whereas (c) was taken with a lower mAs. In two other instances from the training dataset, nnUNet’s segmentation underperformed on scans acquired with a reduced mAs setting, in contrast to the failed case in (a).

that the baseline CT is typically performed a few months before the SPECT.

To understand how the trained denoising network uses the segmentation mask, Fig. 14 visualizes the weights based on the segmentation mask for the sphere phantom and the denoising process. The columns of Fig. 14 represent one of the denoising steps in the denoiser of the proposed method. This work used 100 sets of filters and thresholding values, each trained to produce a denoised image and using the segmentation mask as prior information to locate the region of interest.

V. CONCLUSION

Exploiting the available information from a dual-modality system can benefit a trained regularizer. The proposed method employs the anatomical (CT) segmentation mask generated by segmentation neural network in a fully automatic manner. The CT segmentation informs the SPECT denoising network where the region of interest is likely located. For ^{177}Lu SPECT/CT under low count conditions, the proposed method achieved qualitative and quantitative improvements by incorporating the anatomical image information, compared to the non-trained quadratic regularizers using the CT-side information. In particular, the proposed method reduces noise without compromising recovery accuracy.

Although the data in our study originated from just four patients for training (two patients for testing) our patient SPECT/CT data sample size was 25 for training (eight for

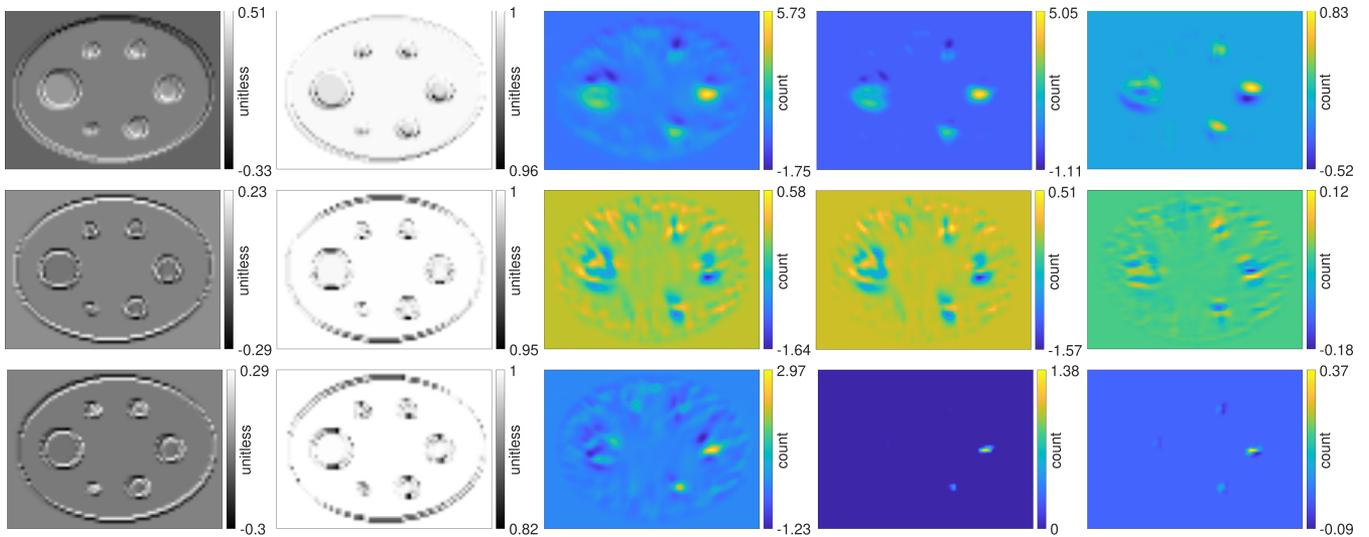


Fig. 14. Illustration of how CT-side information is incorporated into the proposed method. Each column corresponds to one of the steps in the denoiser (first column: $c_k^{(1)} * s$, second column: $w_k^{(1)} = g_k^{(1)}(c_k^{(1)} * s)$, third column: $c_k^{(1)} * x^{(0)}$, fourth column: $W_k^{(1)} \mathcal{T}(c_k^{(1)} * x^{(0)}, \alpha_k^{(1)} \circ w_k^{(1)})$, fifth column: $d_k^{(1)} * W_k^{(1)} \mathcal{T}(c_k^{(1)} * x^{(0)}, \alpha_k^{(1)} \circ w_k^{(1)})$). Each row corresponds to different k values.

testing) as we used data from multiple time points and multiple cycles corresponding to different count-rates. Moreover, owing to our iteration-by-iteration training, the receptive field of our denoiser during each iteration is confined to $5 \times 5 \times 5$. This approach closely aligns with a patch-based method. Given this perspective, even a single 3D image volume can be divided into a multitude of distinct $5 \times 5 \times 5$ patches. Given training volumes from four patients, our model is exposed to a considerable variety of patches. Nevertheless, stepping away from the patch-centric view, training on data from merely four patients might not capture the broad spectrum of data distributions, which could potentially limit the model's generalization capability. Lu-177 DOTATATE SPECT/CT data is not widely available because imaging is not currently required as part of the clinical protocol for administering this therapy. However, due to the recent interest in personalized treatment there is an increase in clinics performing imaging, hence more accessibility to data for training/testing is expected in the near future.

Moreover, further investigation of the potential challenges, such as the impact of additional challenges with misregistration between modalities, is necessary. Future work also includes training and testing on datasets that exhibit large discrepancies or misregistration between modalities due to motion, as well as training neural networks for both segmentation and denoising by combining loss functions with weighted combinations.

REFERENCES

- [1] M. Ljungberg, A. Celler, M. W. Konijnenberg, K. F. Eckerman, Y. K. Dewaraja, and K. Sjögreen-Gleisner, "MIRD pamphlet no. 26: Joint EANM/MIRD guidelines for quantitative ^{177}Lu SPECT applied for dosimetry of radiopharmaceutical therapy," *J. Nucl. Med.*, vol. 57, no. 1, pp. 151–162, 2016.
- [2] H. Lim, J. A. Fessler, S. J. Wilderman, A. F. Brooks, and Y. K. Dewaraja, "Y-90 SPECT ML image reconstruction with a new model for tissue-dependent bremsstrahlung production using CT information: A proof-of-concept study," *Phys. Med. Biol.*, vol. 63, no. 11, 2018, Art. no. 115001.
- [3] H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann, and M. Unser, "CNN-based projected gradient descent for consistent CT image reconstruction," *IEEE Trans. Med. Imag.*, vol. 37, no. 6, pp. 1440–1453, Jun. 2018.
- [4] J. C. Ye, Y. Han, and E. Cha, "Deep convolutional framelets: A general deep learning framework for inverse problems," *SIAM J. Imag. Sci.*, vol. 11, no. 2, pp. 991–1048, 2018.
- [5] Y. Zhang, D. Zeng, Z. Bian, H. Lu, and J. Ma, "Weighted tensor low-rankness and learnable analysis sparse representation model for texture preserving low-dose CT reconstruction," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 321–336, 2021.
- [6] K. Hammernik et al., "Learning a variational network for reconstruction of accelerated MRI data," *Magn. Reson. Med.*, vol. 79, no. 6, pp. 3055–3071, 2018.
- [7] H. K. Aggarwal, M. P. Mani, and M. Jacob, "MoDL: Model-based deep learning architecture for inverse problems," *IEEE Trans. Med. Imag.*, vol. 38, no. 2, pp. 394–405, Feb. 2019.
- [8] Y. Liu, Q. Liu, M. Zhang, Q. Yang, S. Wang, and D. Liang, "IFR-Net: Iterative feature refinement network for compressed sensing MRI," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 434–446, 2020.
- [9] Z. Ke et al., "Deep manifold learning for dynamic MR imaging," *IEEE Trans. Comput. Imag.*, vol. 7, pp. 1314–1327, 2021.
- [10] 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.
- [11] 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.
- [12] 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.
- [13] H. Lim, I. Y. Chun, J. A. Fessler, and Y. K. Dewaraja, "Improved low count quantitative SPECT reconstruction with a trained deep learning based regularizer," *J. Nucl. Med.*, vol. 60, 2019, Art. no. 42.
- [14] R. Leahy and X. Yan, "Incorporation of anatomical MR data for improved functional imaging with PET," in *Proc. Biennial Int. Conf. Inf. Process. Med. Imag.*, 1991, pp. 105–120.
- [15] J. A. Fessler, N. H. Clinthorne, and W. L. Rogers, "Regularized emission image reconstruction using imperfect side information," *IEEE Trans. Nucl. Sci.*, vol. 39, no. 5, pp. 1464–1471, Oct. 1992.
- [16] X. Ouyang, W. H. Wong, V. E. Johnson, X. Hu, and C.-T. Chen, "Incorporation of correlated structural images in PET image reconstruction," *IEEE Trans. Med. Imag.*, vol. 13, no. 4, pp. 627–640, Dec. 1994.
- [17] J. E. Bowsher, V. E. Johnson, T. G. Turkington, R. J. Jaszczyk, C. 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.

- [18] B. A. Ardekani, M. Braun, B. F. Hutton, I. Kanno, and H. Iida, "Minimum cross-entropy reconstruction of PET images using prior anatomical information," *Phys. Med. Biol.*, vol. 41, no. 11, 1996, Art. no. 2497.
- [19] S. Sastry and R. E. Carson, "Multimodality Bayesian algorithm for image reconstruction in positron emission tomography: A tissue composition model," *IEEE Trans. Med. Imag.*, vol. 16, no. 6, pp. 750–761, Dec. 1997.
- [20] B. Lipinski, H. Herzog, E. R. Kops, W. Oberschelp, and H. Muller-Gartner, "Expectation maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 129–136, Apr. 1997.
- [21] A. O. Hero, R. Piramuthu, J. A. Fessler, and S. R. Titus, "Minimax emission computed tomography using high-resolution anatomical side information and B-spline models," *IEEE Trans. Inf. Theory*, vol. 45, no. 3, pp. 920–938, Apr. 1999.
- [22] A. Rangarajan, T. Hsiao, and G. Gindi, "A Bayesian joint mixture framework for the integration of anatomical information in functional image reconstruction," *J. Math. Imag. Vis.*, vol. 12, no. 3, pp. 199–217, 2000.
- [23] 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, 2001, Art. no. 1.
- [24] J. E. Bowsher et al., "Utilizing MRI information to estimate F18-FDG distributions in rat flank tumors," in *Proc. IEEE Symp. Conf. Rec. Nucl. Sci.*, 2004, pp. 2488–2492.
- [25] K. Baete, J. Nuyts, W. V. Paesschen, P. Suetens, and P. Dupont, "Anatomical-based FDG-PET reconstruction for the detection of hypometabolic regions in epilepsy," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 510–519, Apr. 2004.
- [26] A. M. Alessio and P. E. Kinahan, "Improved quantitation for PET/CT image reconstruction with system modeling and anatomical priors," *Med. Phys.*, vol. 33, no. 11, pp. 4095–4103, 2006.
- [27] F. Bataille, C. Comtat, S. Jan, F. Sureau, and R. Trebossen, "Brain PET partial-volume compensation using blurred anatomical labels," *IEEE Trans. Nucl. Sci.*, vol. 54, no. 5, pp. 1606–1615, Oct. 2007.
- [28] J. Tang and A. Rahmim, "Bayesian PET image reconstruction incorporating anato-functional joint entropy," *Phys. Med. Biol.*, vol. 54, no. 23, 2009, Art. no. 7063.
- [29] L. Cao and J. Peter, "Bayesian reconstruction strategy of fluorescence-mediated tomography using an integrated SPECT-CT-OT system," *Phys. Med. Biol.*, vol. 55, no. 9, 2010, Art. no. 2693.
- [30] S. Somayajula, C. Panagiotou, A. Rangarajan, Q. Li, S. R. Arridge, and R. M. Leahy, "PET image reconstruction using information theoretic anatomical priors," *IEEE Trans. Med. Imag.*, vol. 30, no. 3, pp. 537–549, Mar. 2011.
- [31] Y. K. Dewaraja, K. F. Koral, and J. A. Fessler, "Regularized reconstruction in quantitative SPECT using CT side information from hybrid imaging," *Phys. Med. Biol.*, vol. 55, no. 9, 2010, Art. no. 2523.
- [32] J. Cheng-Liao and J. Qi, "PET image reconstruction with anatomical edge guided level set prior," *Phys. Med. Biol.*, vol. 56, no. 21, 2011, Art. no. 6899.
- [33] M. Shidahara et al., "Wavelet-based resolution recovery using an anatomical prior provides quantitative recovery for human population phantom PET [¹¹C] raclopride data," *Phys. Med. Biol.*, vol. 57, no. 10, 2012, Art. no. 3107.
- [34] V.-G. Nguyen and S.-J. Lee, "Incorporating anatomical side information into PET reconstruction using nonlocal regularization," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 3961–3973, Oct. 2013.
- [35] R. Loeb, N. Navab, and S. I. Ziegler, "Direct parametric reconstruction using anatomical regularization for simultaneous PET/MRI data," *IEEE Trans. Med. Imag.*, vol. 34, no. 11, pp. 2233–2247, Nov. 2015.
- [36] S. Chen, H. Liu, P. Shi, and Y. Chen, "Sparse representation and dictionary learning penalized image reconstruction for positron emission tomography," *Phys. Med. Biol.*, vol. 60, no. 2, 2015, Art. no. 807.
- [37] W. Hutchcroft, G. Wang, K. T. Chen, C. Catana, and J. Qi, "Anatomically-aided PET reconstruction using the kernel method," *Phys. Med. Biol.*, vol. 61, no. 18, 2016, Art. no. 6668.
- [38] F. Knoll, M. Holler, T. Koesters, R. Otazo, K. Bredies, and D. K. Sodickson, "Joint MR-PET reconstruction using a multi-channel image regularizer," *IEEE Trans. Med. Imag.*, vol. 36, no. 1, pp. 1–16, Jan. 2017.
- [39] K. Gong, J. Cheng-Liao, G. 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.
- [40] T.-A. Song, S. R. Chowdhury, F. Yang, and J. Dutta, "Super-resolution PET imaging using convolutional neural networks," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 518–528, 2020.
- [41] J. A. Fessler, "Optimization methods for magnetic resonance image reconstruction: Key models and optimization algorithms," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 33–40, Jan. 2020.
- [42] J. Adler, S. Lunz, O. Verdier, C.-B. Schönlieb, and O. Öktem, "Task adapted reconstruction for inverse problems," *Inverse Problems*, vol. 38, no. 7, 2022, Art. no. 075006.
- [43] H. Lim, Y. K. Dewaraja, and J. A. Fessler, "Joint low-count PET/CT segmentation and reconstruction with paired variational neural networks," *Proc. SPIE*, vol. 11312, pp. 199–205, 2020.
- [44] P. A. Wolf, J. S. Jørgensen, T. G. Schmidt, and E. Y. Sidky, "Few-view single photon emission computed tomography (SPECT) reconstruction based on a blurred piecewise constant object model," *Phys. Med. Biol.*, vol. 58, no. 16, 2013, Art. no. 5629.
- [45] S. Y. Chun, Y. K. Dewaraja, and J. A. Fessler, "Alternating direction method of multiplier for tomography with nonlocal regularizers," *IEEE Trans. Med. Imag.*, vol. 33, no. 10, pp. 1960–1968, Oct. 2014.
- [46] M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: A review," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 85–95, Nov. 2017.
- [47] D. Liang, J. Cheng, Z. Ke, and L. Ying, "Deep magnetic resonance image reconstruction: Inverse problems meet neural networks," *IEEE Signal Process. Mag.*, vol. 37, no. 1, pp. 141–151, Jan. 2020.
- [48] A. R. D. Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Trans. Med. Imag.*, vol. 14, no. 1, pp. 132–137, Mar. 1995.
- [49] M. Sandström, U. Garske-Román, S. Johansson, D. Granberg, A. Sundin, and N. Freedman, "Kidney dosimetry during ¹⁷⁷Lu-DOTATATE therapy in patients with neuroendocrine tumors: Aspects on calculation and tolerance," *Acta Oncologica*, vol. 57, no. 4, pp. 516–521, 2018.
- [50] A. Sundlöv et al., "Individualised ¹⁷⁷Lu-DOTATATE treatment of neuroendocrine tumours based on kidney dosimetry," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 44, pp. 1480–1489, 2017.
- [51] K. Kairemo and A. Kangasmäki, "4D SPECT/CT acquisition for 3D dose calculation and dose planning in ¹⁷⁷Lu-peptide receptor radionuclide therapy: Applications for clinical routine," in *Proc. Theranostics, Gallium-68, Other Radionuclides*, 2013, pp. 537–550.
- [52] L. Kabasakal et al., "Pre-therapeutic dosimetry of normal organs and tissues of ¹⁷⁷Lu-PSMA-617 prostate-specific membrane antigen (PSMA) inhibitor in patients with castration-resistant prostate cancer," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 42, pp. 1976–1983, 2015.
- [53] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, 2021.
- [54] D. Gut, Z. Tabor, M. Szymkowski, M. Rozynek, I. Kucybała, and W. Wojciechowski, "Benchmarking of deep architectures for segmentation of medical images," *IEEE Trans. Med. Imag.*, vol. 41, no. 11, pp. 3231–3241, Nov. 2022.
- [55] M. Antonelli et al., "The medical segmentation decathlon," *Nature Commun.*, vol. 13, no. 1, 2022, Art. no. 4128.
- [56] P. Bilic et al., "The liver tumor segmentation benchmark (LITS)," *Med. Image Anal.*, vol. 84, 2023, Art. no. 102680.
- [57] J. Strosberg et al., "Phase 3 trial of ¹⁷⁷Lu-Dotatate for midgut neuroendocrine tumors," *New England J. Med.*, vol. 376, no. 2, pp. 125–135, 2017.
- [58] D. White and R. S. Lawson, "A poisson resampling method for simulating reduced counts in nuclear medicine images," *Phys. Med. Biol.*, vol. 60, no. 9, 2015, Art. no. N167.
- [59] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Annu. Conf. Neural Inf. Process. Syst.*, 2019.
- [60] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, 2015, pp. 234–241.