

Improving PET Image Quality Using Deep Image Prior

August 2024

Fumio Hashimoto

Graduate School of

Science and Engineering

CHIBA UNIVERSITY

(千葉大学審査学位論文)

Improving PET Image Quality Using Deep Image Prior

August 2024

Fumio Hashimoto

Graduate School of

Science and Engineering

CHIBA UNIVERSITY

Table of Contents

Table of Contents	1
List of Figures	3
List of Tables	9
Abstract	10
1 Introduction	12
2 Basic principles	15
2.1 Positron emission tomography	15
2.1.1 Basics of positron emission tomography.....	15
2.1.2 Image reconstruction	17
2.2 Deep learning.....	22
2.2.1 Convolutional neural networks.....	22
2.2.2 Deep image prior	24
2.2.3 Deep learning for PET imaging.....	25
3 Conditional DIP for reconstructed PET images	28
3.1 Introduction	28
3.2 Methods	29
3.2.1 Conditional deep image prior	29
3.2.2 Network structures and optimization.....	34
3.3 Experimental setup	38
3.3.1 Computer simulation	38
3.3.2 Real preclinical PET data	42
3.3.3 Comparison algorithms	44
3.4 Results	46
3.4.1 Computer simulation	46
3.4.2 Real preclinical PET data	51
3.5 Discussion.....	57
4 DIPrecon: PET image reconstruction using conditional DIP.....	61
4.1 Introduction	61
4.2 Methods	62

4.2.1	Proposed end-to-end DIPrecon algorithm	62
4.2.2	Implementation	64
4.3	Experimental setup	66
4.3.1	Computer simulation	66
4.3.2	Real preclinical PET data	68
4.3.3	Comparison algorithms	69
4.4	Results	70
4.4.1	Computer simulation	70
4.4.2	Real preclinical PET data	74
4.5	Discussion.....	76
5	Expanding DIPrecon to 3D PET	79
5.1	Introduction	79
5.2	Methods	80
5.2.1	Block iterative algorithm for fully 3D PET data	80
5.2.2	Penalization for DIP optimization	82
5.2.3	Implementation.....	83
5.3	Experimental setup	85
5.3.1	Computer simulation	85
5.3.2	Real preclinical PET data	86
5.3.3	Comparison algorithms	87
5.4	Results	88
5.4.1	Computer simulation	88
5.4.2	Real preclinical PET data	94
5.5	Discussion.....	96
6	Conclusion	101
	Reference	103
	Publication.....	115
	Journal papers	115
	International conferences.....	119
	Acknowledgement.....	123

List of Figures

Figure 2.1: Schematic illustration of the principle of the PET system.....	16
Figure 2.2: Chemical structure of ^{18}F -FDG.....	16
Figure 2.3: Example of the FBP and some iterative PET image reconstruction algorithms, which were applied to the same normal and low-dose simulation dataset. The figure is reprinted from the work of Hashimoto et al. [22] (CC BY 4.0).....	21
Figure 2.4: Schematic illustration of the U-net architecture. The number of feature maps is denoted at the top of each box. The pixel size is indicated on the left side of each box. The arrows denote the different operations. The figure is reprinted from the work of Hashimoto et al. [79] (CC BY 4.0)	23
Figure 2.5: Classification of deep learning approaches for PET imaging in three categories; post-processing (denoising), direct reconstruction, and iterative reconstruction methods integrated with neural networks (NNs). The figure is reprinted from the work of Hashimoto et al. [22] (CC BY 4.0)	27
Figure 3.1: Procedure of the proposed conditional DIP framework for dynamic PET imaging. The arrows denote the different operations. The figure is reprinted from the work of Hashimoto et al. [79] (CC BY 4.0).....	30
Figure 3.2: Example of the original DIP and conditional DIP frameworks. The arrows denote the different operations. The figure is reprinted with a modification from the work of Hashimoto et al. [79] (CC BY 4.0)	31
Figure 3.3: Schematic illustration of the optimization trajectories of the conditional DIP framework for dynamic PET imaging. The gray region shows a manifold of degraded noisy dynamic PET images from the ground truth image. The trajectories of traditional optimization-based denoising, original and conditional DIP methods are shown in blue, green, and red lines, respectively. The arrows denote the different operations. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPeM)	33
Figure 3.4: Overview of the proposed 4DDIP framework. The proposed network consists of two modules: a feature extractor, and a reconstruction branch. In this	

framework, the common features of all dynamic PET images can be shared by one CNN architecture by end-to-end learning, and can improve denoising performance compared to 3D processing. The arrows denote the different operations. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPEM)	37
Figure 3.5: The regional time activity curves and plasma input function in the computer simulation.	40
Figure 3.6: Influence of the loss weights application on the loss. (a) The learning curves without applying the loss weights, and (b) with applying the loss weights. The loss weights reduced the variance in the loss curves, thereby contributing to the stabilization of network training. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPEM).....	47
Figure 3.7: Simulation results of (a) the dynamic PET image at the time frame 6, (b) time frame 26, and (c) Patlak influx constant image. The columns correspond to the ground truth, reconstructed images obtained by the OSEM algorithm, Gaussian filtering, guided filtering, and conditional DIP with the 3D U-Net and 4D branch CNN architectures (left to right). The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPEM)	48
Figure 3.8: Tradeoffs between the mean squared bias and variance for the simulation data at the tumor ROIs in (a) the time frame 26 and (b) Patlak influx constant image. The markers corresponding to 2.0, 2.5, 3.0, 3.5, and 4.0 mm FWHM for the Gaussian filtering, 0.1, 0.5, 0.75, 1.0 and 1.5 ϵ for the guided filtering, 20 000, 15 000, 10 000, 5000 and 500 epochs for the conditional DIP with 3D U-Net architecture, and 20 000, 10 000, 5000, 2500 and 1000 epochs for the conditional DIP with 4D branch CNN architecture. The images in Figure 3.7 are labeled by filled markers. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPEM).....	49
Figure 3.9: PSNRs and SSIMs for the simulation data.	50
Figure 3.10: Preclinical results of (a) the dynamic PET images at time frame 6, (b) the time frame 26, and (c) the Patlak influx constant images in ^{18}F -FDG. The columns correspond to the images processed by 3D DRAMA, Gaussian filtering, guided filtering, and conditional DIP with 3D U-Net and 4D branch CNN architectures (left to right). The figure is reprinted from the work of	

Hashimoto et al. [80] (© 2021 IPPEM).....	53
Figure 3.11: Tradeoffs between the putamen uptake and background standard deviation for preclinical ^{18}F -FDG data in (a) the time frame 26, and (b) Patlak influx constant images. The markers corresponding to 1.0, 2.0, 3.0, 4.0 and 5.0 mm FWHM for the Gaussian filtering, 0.1, 0.5, 0.75, 1.0 and 1.5 ϵ for the guided filtering, 1250, 1000, 750, 500 and 250 epochs for the conditional DIP with 3D U-Net architecture, and 7000, 5000, 4000, 3000 and 2000 epochs for the conditional DIP with 4D branch CNN architecture (upper right to lower left). The images in Figure 3.10 are labeled by filled markers. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPPEM).....	54
Figure 3.12: Preclinical results of (a) the dynamic PET images at the time frame 6, (b) time frame 26, and (c) non-displaceable binding potential images in ^{11}C -raclopride. The columns correspond to the images processed by the 3D DRAMA, Gaussian filtering, guided filtering, and conditional DIP with 3D U-Net and 4D branch CNN architectures (left to right). The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPPEM).....	55
Figure 3.13: Mean BP_{ND} values at the putamen ROI in the ^{11}C -raclopride experiment.	56
Figure 3.14: Regional TACs at (a) the putamen and (b) cerebellar cortex ROIs. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPPEM)..	56
Figure 4.1: Implementation overview of the proposed end-to-end PET image reconstruction algorithm using the conditional DIP framework. The voxel size is provided on the left side of the box. First, a reconstructed PET image was calculated using the conditional DIP framework. Then, the loss calculation was performed with the measured data and the estimated projection data from the DIP output. The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE)	65
Figure 4.2: Three orthogonal slices of the reconstructed images and their magnified images of the red squared regions by different reconstruction algorithms in the computer simulation. The columns correspond to the MR images, ground truth, reconstructed images obtained using FBP, and MLEM, hybrid DIPrecon, and the proposed end-to-end DIPrecon with random noise and MRI input (left-to-right). The figure is reprinted from the work of Hashimoto et al. [113] (© 2022	

IEEE) 71

Figure 4.3: Quantitative results of the PSNR (top) and SSIM (bottom) for the computer simulation with different reconstruction algorithms: the MLEM, hybrid DIPrecon (Hybrid), end-to-end DIPrecon with random noise input (End-to-end w/ noise) and with MRI input (End-to-end w/ MRI). The line within the box represents the median value, and the upper and lower lines of the box represent the 75th and the 25th percentiles, respectively. The upper and lower whiskers represent the maximum and minimum values, respectively. The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE)..... 72

Figure 4.4: Tradeoffs between the CRC and STD at the gray matter region for the computer simulation with different reconstruction algorithms: the MLEM, hybrid DIPrecon (Hybrid), end-to-end DIPrecon with random noise input (End-to-end w/ noise) and with MRI input (End-to-end w/ MRI). Markers are plotted every ten iterations from ten to 100 in the MLEM, every iteration from one to ten in the hybrid DIPrecon, every three epochs from 18 to 45 in the end-to-end DIPrecon with the random noise input, and every epoch from 33 to 42 in the end-to-end DIPrecon with the MRI input. The reconstructed images in Figure 4.2 are labeled by filled makers. The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE) 73

Figure 4.5: Three orthogonal slices of the reconstructed images by different reconstruction algorithms in the computer simulation. The columns correspond to the MR images, reconstructed images obtained using FBP, MLEM, hybrid DIPrecon, and the end-to-end DIPrecon with MRI input (left-to-right). The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE) . 74

Figure 4.6: Tradeoffs between the striatum uptake and STD for real preclinical data with different reconstruction algorithms: the MLEM, hybrid DIPrecon (Hybrid), and end-to-end DIPrecon with MRI input (End-to-end w/ MRI). Markers are plotted every ten iterations from ten to 100 in MLEM, every iteration from one to ten in the hybrid DIPrecon, every three epochs from 9 to 18, and every epoch from 21 to 30 in the end-to-end DIPrecon with MRI input. The corresponding reconstructed images in Figure 4.5 are labeled by filled markers. The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE) 75

Figure 5.1: Implementation overview of the proposed fully 3D end-to-end PET image reconstruction using the conditional DIP framework. First, a PET image

was calculated using the conditional DIP framework. Then, loss calculation was performed with each measured block sinogram and the estimated block sinogram from the DIP output. Finally, the network parameters were sequentially updated in a mini-batch optimization manner. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM) 84

Figure 5.2: Reconstructed results of the human brain ^{18}F -FDG computer simulation with different regularization parameters β with two subsets. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM) 89

Figure 5.3: Impact of the settings of the regularization parameter β setting with two subsets, in the PSNR for the proposed fully 3D end-to-end DIPrecon algorithm. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM) 90

Figure 5.4: Transaxial and coronal slices of the reconstructed images and their magnified images of the red squared regions by different reconstruction algorithms in the computer simulation. The columns correspond to the MR images, ground truth, reconstructed images obtained using the MLEM, MAPEM, hybrid DIPrecon, and the proposed end-to-end DIPrecon without RDP (Proposed w/o RDP) and with RDP (Proposed w/ RDP) (left-to-right). The PSNR values for each algorithm are provided below the name of the algorithm. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM) 90

Figure 5.5: Mean tradeoffs between the tumor CRC and STD for the human brain ^{18}F -FDG computer simulation. Markers are plotted every ten iterations from one to 100 in the MLEM and MAPEM algorithms, every ten iterations from one to 200 in the hybrid DIPrecon algorithm, every two epochs from 1 to 20 in the proposed algorithm without RDP (Proposed w/o RDP), and every three epochs from 1 to 50 in the proposed algorithm with RDP (Proposed w/ RDP). The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM) 91

Figure 5.6: Impact of the RDP term on the PSNR for the proposed fully 3D end-to-end DIPrecon algorithm. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM) 91

Figure 5.7: Mean and standard deviation PET images of ten independent and

identically distributed samples from the same digital brain phantom for different reconstruction algorithms. From left to right, the MAPEM with the RDP, hybrid DIPrecon, proposed end-to-end DIPrecon without RDP (Proposed w/o RDP), and proposed end-to-end DIPrecon with RDP (Proposed w/RDP). The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPEM) 92

Figure 5.8: Quantitative results of ten independent and identically distributed samples from the same digital brain phantom for different reconstruction algorithms. From left to right, the MAPEM with the RDP, hybrid DIPrecon, proposed end-to-end DIPrecon without RDP (Proposed w/o RDP), and proposed end-to-end DIPrecon with RDP (Proposed w/RDP). The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPEM) 92

Figure 5.9: Reconstructed results of the human brain computer simulation with ^{18}F -FDG contrast for different number of subsets. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPEM) 93

Figure 5.10: Impact of the number of subsets in terms of the PSNR for the proposed fully 3D end-to-end DIPrecon algorithm. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPEM)..... 93

Figure 5.11: Reconstruction results of the reconstructed images and their magnified images of the red squared regions by different reconstruction algorithms for the real preclinical low-dose ^{18}F -FDG PET data. The columns correspond to the MR images, reconstructed images obtained using the MLEM with full count data and low-dose data, MAPEM, hybrid DIPrecon, and the proposed end-to-end DIPrecon with RDP (Proposed w/ RDP) (left-to-right). The PSNR values for each algorithm are provided below the name of the algorithm. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPEM) 94

Figure 5.12: Line profiles for the real preclinical low-dose ^{18}F -FDG PET data by different reconstruction algorithms. The profile line is represented in the upper left. Triangular arrows indicate the putamen areas. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPEM) 95

List of Tables

Table 3.1: The kinetic parameters used in this computer simulation.	40
Table 3.2: Coincidence events of each time frame in the experiments.	41

Abstract

Although positron emission tomography (PET) serves as a powerful tool for in vivo diagnosis, the impact of the noise on reconstructed images presents a substantial challenge due to the fewer counts present in acquired data. Such noise potentially degrades the accuracy of quantitative analysis and the detectability of lesions. Recently, several deep learning-based post-processing and reconstruction have been developed to reduce the noise in reconstructed images. However, deep learning may not provide expected performance when a domain gap occurs between training and test data. To solve the challenge, deep image prior (DIP) has attracted attention for denoising task. DIP is a kind of unsupervised method that solves inverse problems by using the inductive bias of the network structure as a regularization without prior training datasets. This thesis aims at improving PET image quality using the DIP, which does not depend on the domain of training data such as types of PET scanners and tracers. The first part of the thesis introduces a PET image denoising method using a conditional DIP and proposes a novel network architecture for 4D dynamic PET denoising. The second part of the thesis proposes the iterative PET image reconstruction incorporating the conditional DIP in an end-to-end manner. Then, the proposed reconstruction algorithm expands 2D PET data to practical, fully 3D PET data. The main advantages of the proposed denoising method are its fast computation time and easy implementation, which can potentially be helpful in clinical situations. On the other hand, the proposed reconstruction method, which performs optimization while measuring consistency with measured data, can reduce the loss of detailed information such as small structures due to over-smoothing.

Positron emission tomography (PET) は生体内の薬物動態を非侵襲に観察できる強力な手法であるが、PET は測定できるカウントが少なく、再構成画像はノイズの影響を受け、解析精度や病変検出能を低下させる恐れがある。近年、再構成画像のノイズを低減するために深層学習を利用した後処理や画像再構成が多く開発されている。ただし、深層学習は学習データとテストデータでドメインギャップが生じた場合、期待した性能が得られないことがある。この問題を解決する方法として、deep image prior (DIP) を利用したノイズ低減法が注目されている。DIP は教師なし学習の一種であり、学習データセットを利用せず、ネットワーク構造に起因する帰納バイアスを正則化として利用し逆問題を解く手法である。そこで本論文では、DIP を用いることで学習データのドメイン、すなわち PET 装置や薬剤の種類に依存しない画質改善手法を開発することを目的とする。最初に、条件付き DIP を用いた PET 画像ノイズ除去法を提案し、それを 4 次元 PET データへ拡張するための新規ネットワーク構造を考案した。次に、条件付き DIP を組み込んだ end-to-end 型の反復 PET 画像再構成を提案し、それを 2 次元 PET から 3 次元 PET データへ拡張する実用的な手法を開発した。提案ノイズ除去法は、計算が高速かつ実装が簡単な点が最大の特徴であり、実際の臨床現場での普及が期待できる。一方、計測データとの整合性を測りながら最適化を行う提案再構成法は、過度な平滑化によって微細構造などの情報が消失するのを防ぐ効果が期待できる。

1 Introduction

Positron emission tomography (PET) is a molecular imaging technique to visualize and quantify the distribution of radioactive tracers labeled with positron-emitting radioisotopes (RIs), such as ^{11}C , ^{13}N , ^{15}O , and ^{18}F , injected into living human bodies [1]. It enables the observation of various biochemical processes in vivo from blood flow, glucose metabolism to neural receptor activity [2]. Thus, it is utilized not only in the diagnosis of cancer [3,4], and neurodegenerative diseases such as Alzheimer's disease [5,6], but also in fundamental neuroscience research, particularly in study of higher brain functions [7].

While PET serves as a powerful tool for definitive in vivo diagnosis, the impact of the noise on reconstructed images presents a substantial challenge relative to other tomographic modalities, such as X-ray computed tomography (CT) [8]. This is attributable to the fewer counts present in the acquired data. Such noise potentially degrades the accuracy of quantitative analysis and the detectability of the lesions, which may cause missed lesions [9-11].

A straightforward strategy for keeping PET image quality is to increase the amount of radioactive tracer injected into the living human body. However, it is sometimes difficult to apply this way due to the increased radiation exposure, which may potentially increase lifetime cancer risk [12] and the limitations in high count-rate capabilities of PET systems [13]. An alternative solution is to extend the PET scanning time; this could lead to psychological discomfort for patients, and some may find it challenging to stay motionless for extended periods during the PET examination [14]. Therefore, there is a need for noise reduction strategies that neither increase the radiation exposure nor extend

the scanning time. It would not be an overstatement to say that the evolution of PET imaging has been a continuous struggle to overcome image noise.

From hardware perspective, improving the quality of PET images primarily involves developing high-sensitivity PET scanners. This improvements can be achieved by using radiation detectors of greater sensitivity [15,16] and optimization of scanner geometry to detect larger solid angles [17-20]. Alongside hardware advancements, there are significant efforts to improve PET image quality through advanced image denoising and reconstruction techniques [21-23]. This thesis focuses on PET image denoising and reconstruction techniques for improving PET image quality without PET instrumentation modifications.

Techniques for improving PET image quality through denoising and reconstruction have traditionally been realized by handcrafted filters or artificial prior information. These have included algorithms such as the Gaussian filter, guided filter [24], nonlocal means filter [25,26], block-matching filter [27] for post-denoising. For iterative reconstruction, priors such as the Gibbs prior [28] and patients' anatomical information [29,30] have been utilized. The advent of deep learning has revolutionized PET image denoising and reconstruction by introducing data-driven approaches that learn optimal denoising and reconstruction strategies directly from the huge datasets [31,32]. However, to apply deep learning to medical imaging, including PET, it is necessary to overcome the major hurdle of acquiring a large number of high-quality training datasets. In addition, negative effects on PET image quality should be considered for PET imaging if there is a domain gap between training and testing datasets resulting from different PET scanners or other radioactive tracers.

The work presented in the thesis aims to improve PET image quality in the

frameworks of the PET image denoising and reconstruction, using a deep image prior (DIP) [33,34]. The DIP realizes an unsupervised deep learning method for solving inverse problems such as denoising, which works a neural network structure as an intrinsic regularizer and does not require the preparation of a prior training dataset. Therefore, it is expected to solve the domain gap problem mentioned above. In this thesis, Chapter 2 describes the basic principles of the PET imaging and deep learning technology, especially image reconstruction and the DIP. Chapter 3 presents about the PET image denoising method using the conditional DIP and the application to dynamic PET imaging. Chapter 4 discusses the 2D PET image reconstruction using the conditional DIP, and Chapter 5 explores its extension to fully 3D PET data.

2 Basic principles

2.1 Positron emission tomography

2.1.1 Basics of positron emission tomography

A distinctive advantage of PET over other tomographic scanners such as X-ray CT and magnetic resonance imaging (MRI) is its superior molecular sensitivity, which enables the precise quantification of biological functions in the living body [1].

Figure 2.1 shows a schematic illustration of the principle of the PET system. Within the sequence of PET scans, the process begins with the generation of positron-emitting RIs, typically synthesized in a cyclotron. After their production, these RIs are chemically incorporated into radiopharmaceutical compounds, called PET tracers. These PET tracers are then administered to the patient body, with the choice of PET tracers allowing for targeted investigation of biological functions with specific organs, including blood flow, metabolism, and receptor activity. Figure 2.2 shows the PET tracer ^{18}F -fluoro-2-deoxy-D-glucos (^{18}F -FDG), which is commonly utilized for cancer diagnosis as well as cardiac and brain disorders [35-37]. Positrons emitted from the RI source interact with electrons within the body, leading to annihilation events. Each annihilation produces a pair of 511 keV gamma rays, which are emitted in opposite direction. These gamma rays are simultaneously detected by the radiation detectors of the PET scanner in what are called coincidence events. The data from these coincidence events are then transferred to a computer. Ultimately, PET images are computed through an image reconstruction process.

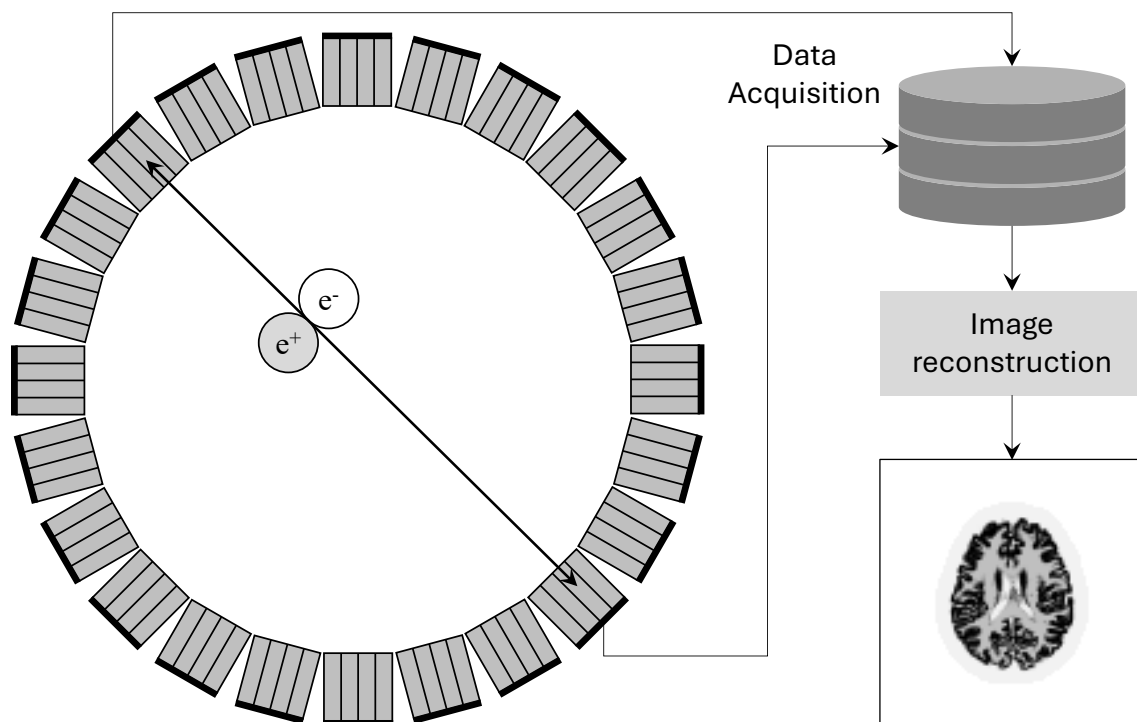


Figure 2.1: Schematic illustration of the principle of the PET system.

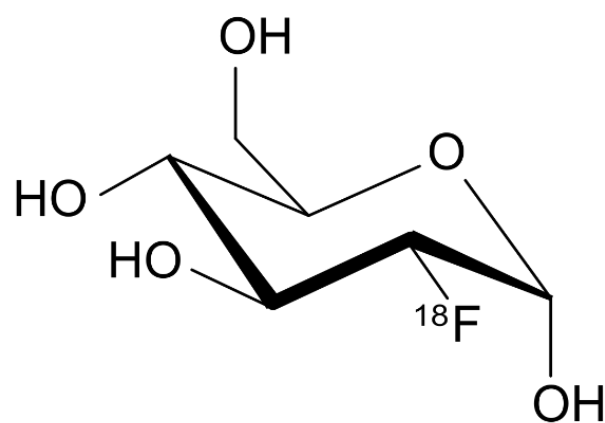


Figure 2.2: Chemical structure of ^{18}F -FDG.

2.1.2 Image reconstruction

In the early stages of PET image reconstruction, analytical reconstruction techniques, notably filtered backprojection (FBP), were predominant [38-40]. Assuming that the two-dimensional (2D) distribution of radioactive tracers within the body can be represented by a continuous function $X(u, v)$, the measurement data $Y(r, \phi)$, which is also known as a sinogram, can be expressed by the following equation [41]:

$$Y(r, \phi) = \int_{-\infty}^{\infty} X(r \cos \phi - s \sin \phi, r \sin \phi + s \cos \phi) ds. \quad (2.1)$$

The sinogram represents a series of integrals computed along the s-axis of the image $X(u, v)$ that has been rotated by an angle ϕ . This integral transformation process is commonly referred to as the Radon transform [42]. The foundational concept of FBP lies in the projection-slice theorem, which elucidates the direct relationship between the 2D Fourier transform of the image $X(u, v)$ and the 1D Fourier transform of the projection data $Y(r, \phi)$. FBP is calculated as the follows:

$$\begin{aligned} X(u, v) &= \int_0^\pi Y_{filtered}(r, \phi) \big|_{r=u \cos \phi + v \sin \phi} d\phi, \\ Y_{filtered}(r, \phi) &= \int_{-\infty}^{+\infty} G(\xi, \phi) |\xi| \exp(2\pi i \xi r) d\xi, \\ G(\xi, \phi) &= \int_{-\infty}^{+\infty} Y(r, \phi) \exp(-2\pi i \xi r) dr, \end{aligned} \quad (2.2)$$

where i and ξ represent the imaginary unit and variable of the frequency domain. $|\xi|$ is the high pass filter, which is known as a ramp filter. The ramp filter is delivered analytically, but its amplification of high-frequency components results in severe noise. Thus, various frequency cutoff filters, such as the Shepp-Logan and hamming filters [43], have been proposed to mitigate high-frequency noise, though this comes at the cost of diminished spatial resolution. Analytical methods like the FBP are lauded for their rapid

processing, linearity, and quantitative accuracy. Nonetheless, they are prone to noise interference, resulting in streak artifacts in images under low-count situations, as illustrated in Figure 2.3.

Developments in PET image reconstruction have led to the advent of iterative reconstruction techniques, among which the maximum likelihood expectation maximization (MLEM) algorithm [44-46] is a prominent example. The MLEM algorithm integrates statistical and physical models directly into the image reconstruction process. Specifically, the MLEM algorithm models the relationship between the image and sinogram through Poisson distribution and a system of linear equations [47].

$$\mathbf{y} = \text{Poisson}(\mathbf{Ax} + \bar{\mathbf{b}}), \quad (2.3)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_J)^T$ represents a vector of voxel values within the image, $\mathbf{y} = (y_1, y_2, \dots, y_I)^T$ represents a vector of the projection data value, and vector $\bar{\mathbf{b}} = (\bar{b}_1, \bar{b}_2, \dots, \bar{b}_I)^T$ represents an expected background components, such as scatter and random coincidence events, and $\mathbf{A} \in \mathbb{R}^{I \times J}$ denotes a system matrix, with individual elements, a_{ij} , which represents the probability that a pair of gamma rays originating from the j -th voxel are detected along the i -th line-of-response (LOR). The Poisson negative log-likelihood function of projection data $\mathbf{y}$ under image $\mathbf{x}$, which is formulated as follows,

$$L(\mathbf{y}|\mathbf{x}) = -\log P(\mathbf{y}|\mathbf{x}) = C - \sum_{i=1}^I \left\{ y_i \log \left(\sum_{j=1}^J a_{ij} x_j + \bar{b}_i \right) - \left(\sum_{j=1}^J a_{ij} x_j + \bar{b}_i \right) \right\}, \quad (2.4)$$

where $P(\mathbf{y}|\mathbf{x})$ is the probability and C is a constant value. The MLEM algorithm estimates an image by minimizing (2.4) using following iterative updates,

$$x_j^{(k+1)} = \frac{x_j^{(k)}}{\sum_{i=1}^I a_{ij}} \sum_{i=1}^I \frac{a_{ij} y_i}{\sum_{j'=1}^J a_{ij'} x_{j'}^{(k)} + \bar{b}_i}, \quad (2.5)$$

where k denotes the current number of iterations. The MLEM algorithm enhances PET image quality over the FBP algorithm by incorporating the statistical noise model for PET, as illustrated in Figure 2.3. Subsequent to the MLEM algorithm, block iterative algorithms, such as the ordered subset expectation maximization (OSEM) [48], were proposed as an accelerated algorithm, which divides the projection data into subsets and iteratively updates the image with each subset [49,50]. The iterative algorithms can be extended to 3D PET data and incorporated physical models by accurately modeling the system matrix [51-54].

The point spread function (PSF) reconstruction methods were developed to integrate the PSF into iterative reconstruction for dedicated and whole-body PET scanners [53]. For example, the PSF can be modeled in image space as follows,

$$\mathbf{x}^{n+1} = \frac{\mathbf{x}^n}{\mathbf{H}^T \mathbf{A}^T \mathbf{1}} \mathbf{H}^T \mathbf{A}^T \frac{\mathbf{y}}{\mathbf{A} \mathbf{H} \mathbf{x}^n \bar{\mathbf{b}}}, \quad (2.6)$$

where $\mathbf{H}$ represents a matrix of PSF kernel in the image space. The PSF reconstruction can reduce statistical noise and enhance image contrast and spatial resolution.

The MLEM algorithm has an unfavorable characteristic in which noise and edge artifacts tend to increase as the number of iterations increases [55]. Therefore, practical solutions often involve terminating the iterations early or applying post-filtering with the Gaussian filter to the reconstructed image. The maximum a posteriori (MAP) reconstruction is an alternative solution, which integrates image priors such as the smoothness of the image to achieve better noise and contrast characteristics than the MLEM algorithm [56-59]. The posterior probability of the PET image $\mathbf{x}$ given emission

data $\mathbf{y}$ can express through Bayes' theorem as follows,

$$P(\mathbf{x}|\mathbf{y}) = \frac{P(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})}, \quad (2.7)$$

where $P(\mathbf{x})$ represents the prior probability of the PET image, called the Gibbs distribution, as follows,

$$P(\mathbf{x}) = \frac{1}{Z} \exp(-\beta U(\mathbf{x})), \quad (2.8)$$

where Z represents a partition function, and U represents an energy function. β is a hyperparameter to adjust the influence of the prior distribution. The negative log-posterior likelihood can be defined as follows,

$$-\log P(\mathbf{y}|\mathbf{x}) - \log P(\mathbf{x}) = L(\mathbf{y}|\mathbf{x}) + \beta U(\mathbf{x}). \quad (2.9)$$

A typical energy function for the Gibbs distribution can be expressed as follows,

$$U(\mathbf{x}) = \sum_j \sum_{j' \in N_j} \omega_{jj'} V(x_j - x_{j'}), \quad (2.10)$$

where V represents a potential function. N_j represents a set of neighboring voxels for the j -th voxel. $\omega_{jj'}$ is a weight between neighboring voxels. For example, the quadratic prior is often used as the potential function as follows,

$$V(x_j - x_{j'}) = (x_j - x_{j'})^2. \quad (2.11)$$

The MAPEM algorithm using Green's one-step-late method [58] estimates the image by minimizing (2.9) using following iterative updates,

$$x_j^{(k+1)} = \frac{x_j^{(k)}}{\sum_{i=1}^I a_{ij} - \beta \left. \frac{\partial U(\mathbf{x})}{\partial x_j} \right|_{\mathbf{x}=\mathbf{x}^{(k)}}} \sum_{i=1}^I \frac{a_{ij} y_i}{\sum_{j'=1}^J a_{ij'} x_{j'}^{(k)} + \bar{b}_i}, \quad (2.12)$$

Figure 2.3 shows the reconstructed result of the use of the MAPEM with the quadratic prior in low-count situations. The MAPEM algorithm provides a smoother PET image

than the MLEM algorithm.

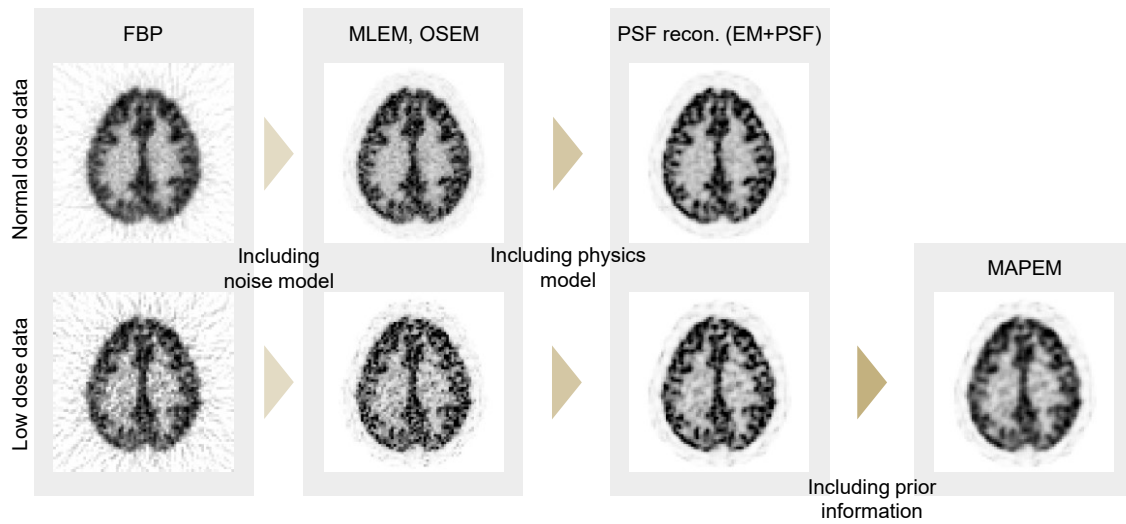


Figure 2.3: Example of the FBP and some iterative PET image reconstruction algorithms, which were applied to the same normal and low-dose simulation dataset. The figure is reprinted from the work of Hashimoto et al. [22] (CC BY 4.0)

2.2 Deep learning

Deep learning is a state-of-the-art subset of machine learning that operates through artificial neural networks with multiple layers, mimicking the human brain's structure and function [31,32]. In general, training deep learning models involves feeding them large amounts of training datasets and adjusting trainable parameters using an optimization algorithm such as stochastic gradient descent (SGD) [60]. This process minimizes a loss function, gradually improving the model's accuracy. Deep learning has significantly advanced fields such as computer vision [61] and natural language processing [62], achieving remarkable results in image recognition and generation tasks that were once challenging for traditional machine learning methods [63]. This thesis focuses on convolutional neural network (CNN) architectures, which are widely used in image processing tasks.

2.2.1 Convolutional neural networks

CNN is one of the specialized neural networks widely utilized in image processing [64,65]. The most critical component of CNNs is a convolution layer, which performs trainable convolution operations and then captures complex details in input data. These convolution kernels are optimized through training, and they become finely tuned to recognize various features from simple edges to complex shapes and textures. Another important component of CNN is nonlinear activation functions, such as the rectified linear unit (ReLU), which avoid the vanishing gradient problem in deep networks with many layers to enhance the ability to capture nonlinear relationships between input and output data. Following convolution layers, CNNs often use pooling layers that perform

downsampling to reduce the spatial resolutions of the feature maps. The pooling layers reduce the computational cost as well as control overfitting of the network. The CNN architectures are composed of combinations of these components, and deeper networks can generally learn more complex relationships.

Figure 2.4 shows the schematic illustration of the U-Net architecture [66,67], a CNN structure commonly used in image-to-image translation tasks for medical imaging [68], such as segmentation [69,70], denoising and image reconstruction [21,71,72]. The U-Net architecture consists of encoding and decoding paths. The encoder path typically increases the number of feature maps while downsampling the spatial size of the feature maps, and the decoder path conversely reduces the feature maps while upsampling the spatial size of the feature maps. Crucially, it integrates skip connections that directly concatenate or add feature maps from the encoder to the decoder at the same spatial resolutions, thereby preserving fine-grained details for accurate image-to-image translation.

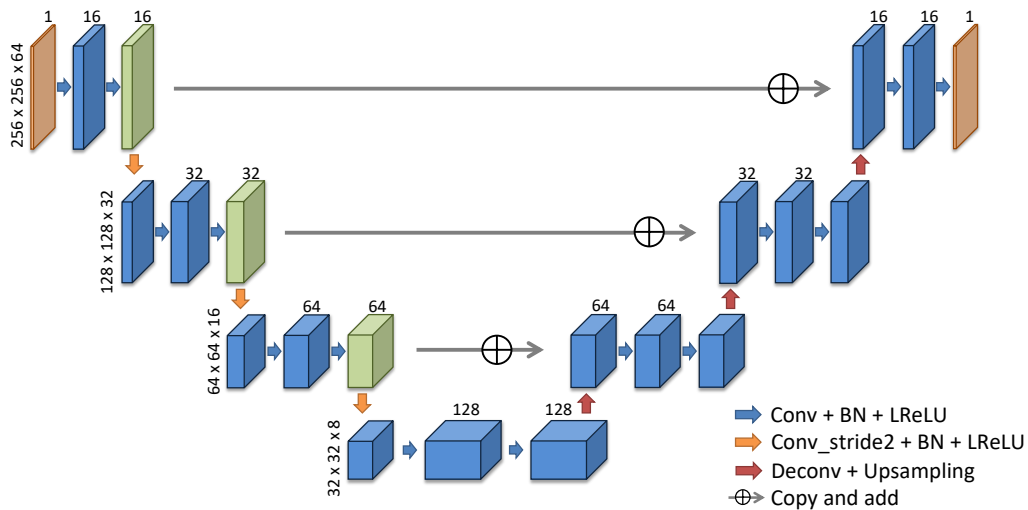


Figure 2.4: Schematic illustration of the U-net architecture. The number of feature maps is denoted at the top of each box. The pixel size is indicated on the left side of each box. The arrows denote the different operations. The figure is reprinted from the work of Hashimoto et al. [79] (CC BY 4.0)

2.2.2 Deep image prior

The DIP framework [33,34] is an innovative approach to image processing in which the structure of CNNs acts as intrinsic prior knowledge or regularizers, without preparing any training datasets. The DIP can solve various inverse problem tasks such as denoising, super-resolution, and inpainting. The training process of the DIP starts with the initialized network f with trainable parameters θ , taking random noise $\mathbf{z}$ as input and degraded image $\mathbf{x}_0$ as target, and then iteratively optimizes the network parameters as follows.

$$\begin{aligned}\theta^* &= \underset{\theta}{\operatorname{argmin}} E(\mathbf{x}_0; f(\theta|\mathbf{z})), \\ \mathbf{x}^* &= f(\theta^*|\mathbf{z}),\end{aligned}\tag{2.13}$$

where $E(\cdot)$ is a loss function, such as the mean squared error (MSE). This optimization intuitively converges to the original degraded image $\mathbf{x}_0$. However, the optimization approaches an undegraded image due to the inductive bias of the CNN structure. Thus, the DIP shows the potential to produce high-quality image restorations without preparing huge training datasets. This has great potential in situations where it is difficult to prepare large amounts of datasets, and expands the range of practical applications of deep learning in the medical imaging field.

2.2.3 Deep learning for PET imaging

The landscape of PET image reconstruction is currently experiencing a profound transformation, driven by integrating state-of-the-art deep learning algorithms alongside advances in computer vision [21,22]. Figure 2.5 delineates a classification of deep learning approaches for PET imaging, organized into three principal categories. (1) The first category is deep learning methods for PET image denoising. (2) The second category includes direct PET image reconstruction methods which are data-driven methods to learn a direct mapping from measured data to reconstructed images using huge training datasets of sinograms and corresponding PET images. (3) The third category is an iterative reconstruction which is a hybrid approach that uses iterative image reconstruction frameworks integrated with neural networks.

(1) Post processing (denoising) method [23]

The post processing (denoising) task is an inverse problem that restores a clean image from a noisy image that is mixed with image noise complicated by the reconstruction. The post processing using deep learning learns the nonlinear relationship between the clean PET images $\mathbf{x}$ and the noisy images $\hat{\mathbf{x}}$ as described in the following optimization problem,

$$\theta^* = \underset{\theta}{\operatorname{argmin}} E(f(\theta|\hat{\mathbf{x}}), \mathbf{x}), \quad (2.14)$$

The goal of the post processing methods using deep learning is to obtain the nonlinear mappings from low-dose to high-dose PET images in a data-driven manner. Compared to traditional post processing methods, the deep learning approaches demonstrate better denoising performance while maintaining spatial resolution and quantitative performance.

(2) Direct image reconstruction method [73,74]

The direct PET image reconstruction using deep learning learns direct mappings from the measured data $\mathbf{y}$ to the reconstructed PET images $\mathbf{x}$ through neural networks f in a data-driven manner, as follows,

$$\theta^* = \underset{\theta}{\operatorname{argmin}} E(f(\theta|\mathbf{y}), \mathbf{x}), \quad (2.15)$$

The direct PET image reconstruction is distinct from other methods because it aims to explore a way to reconstruct PET images using only training datasets, without relying on any physical models such as a forward or back-projection operations.

(3) Iterative reconstruction method [21,22]

The iterative PET image reconstruction using deep learning is a hybrid framework that integrates existing iterative reconstructions such as the MLEM and MAPEM algorithms, which is based on a statistical and physical models, with deep learning. The iterative reconstruction using deep learning can be categorized into two main ways: a synthetic-based reconstruction using an equality constraint and an analysis-based reconstruction using image priors R , as follows,

$$\min_{\mathbf{x}, \mathbf{z}} L(\mathbf{y}|\mathbf{x}) \quad s. t. \quad \mathbf{x} = f(\theta|\mathbf{z}), \quad (2.16)$$

or

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} (L(\mathbf{y}|\mathbf{x}) + \beta R(\mathbf{x})) \quad (2.17)$$

The iterative PET image reconstruction using deep learning can improve PET image quality through neural networks as constraints or image priors while measuring the consistency between the measured data and the reconstructed image.

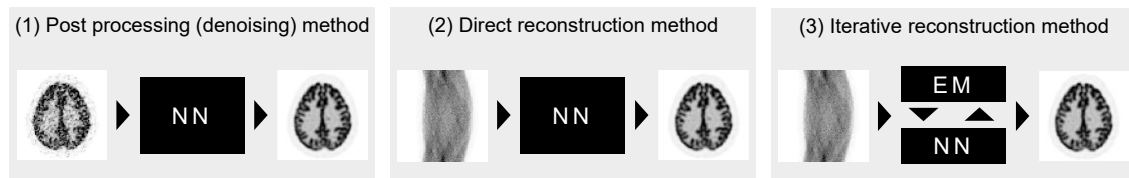


Figure 2.5: Classification of deep learning approaches for PET imaging in three categories; post-processing (denoising), direct reconstruction, and iterative reconstruction methods integrated with neural networks (NNs). The figure is reprinted from the work of Hashimoto et al. [22] (CC BY 4.0)

3 Conditional DIP for reconstructed PET images

3.1 Introduction

The DIP framework can address inverse problems, such as denoising, without the requirement for prior training datasets. This capability has the potential to address the fundamental challenges in the field of PET imaging, where the acquisition of large, tracer- and scanner-specific clinical datasets is often prohibitive.

In this chapter, we introduce a conditional DIP framework for PET image denoising. Furthermore, we expand the conditional DIP to 4D dynamic PET imaging and evaluate the denoising performance of PET imaging using computer simulation and real preclinical data scanned by an animal PET scanner, SHR-38000, Hamamatsu Photonics K.K.

3.2 Methods

3.2.1 Conditional deep image prior

In the image synthesis task of no-flash/flash images [75] using the DIP framework, as demonstrated by Ulyanov et al. [33,34], high-quality images can be generated using flash images as the network input instead of random noise. This suggests that utilizing prior information, such as the same patients' X-ray CT or MR images, related to the noisy PET image as neural network input, may potentially enhance PET image quality in the denoising task of the DIP. The PET image denoising task for a conditional DIP framework using prior information $\mathbf{g}$ is represented as follows [76-78],

$$\begin{aligned}\theta^* &= \underset{\theta}{\operatorname{argmin}} E(\mathbf{x}_0; f(\theta|\mathbf{g})), \\ \mathbf{x}^* &= f(\theta^*|\mathbf{g}),\end{aligned}\tag{3.1}$$

where f is a neural network with trainable parameters θ , $\mathbf{x}_0$ is noisy PET image, and $E(\cdot)$ is a loss function. In general, PET scans are often followed by X-ray CT or MRI scans using PET/CT and PET/MRI scanners to collect anatomical information for diagnostic purposes and attenuation correction. Therefore, these CT and MR images can be used as prior information $\mathbf{g}$. However, it is often difficult to install additional CT or MRI mechanisms into brain-dedicated PET and animal PET scanners due to design constraints, which may not acquire anatomical information.

In this chapter, we propose a conditional DIP framework that operates without such additional anatomical information for dynamic PET imaging [79,80]. In dynamic PET imaging, noisy and short-duration PET images are reconstructed in a time series to track the kinetics of PET tracers within the body. Therefore, the proposed method uses a static PET image, reconstructed by integrating emission data from dynamic PET data frames as

prior information, instead of anatomical information. Then, the conditional DIP framework denoises the dynamic PET images.

Figure 3.1 illustrates a example of the proposed conditional DIP framework for dynamic PET images. In this procedure, static PET images are input to the network, and followed by calculating the loss between the network output and the target dynamic PET images. Finally, the trainable parameters θ of the network are updated to denoise the dynamic PET images. The above process is repeated until the optimization is completed.

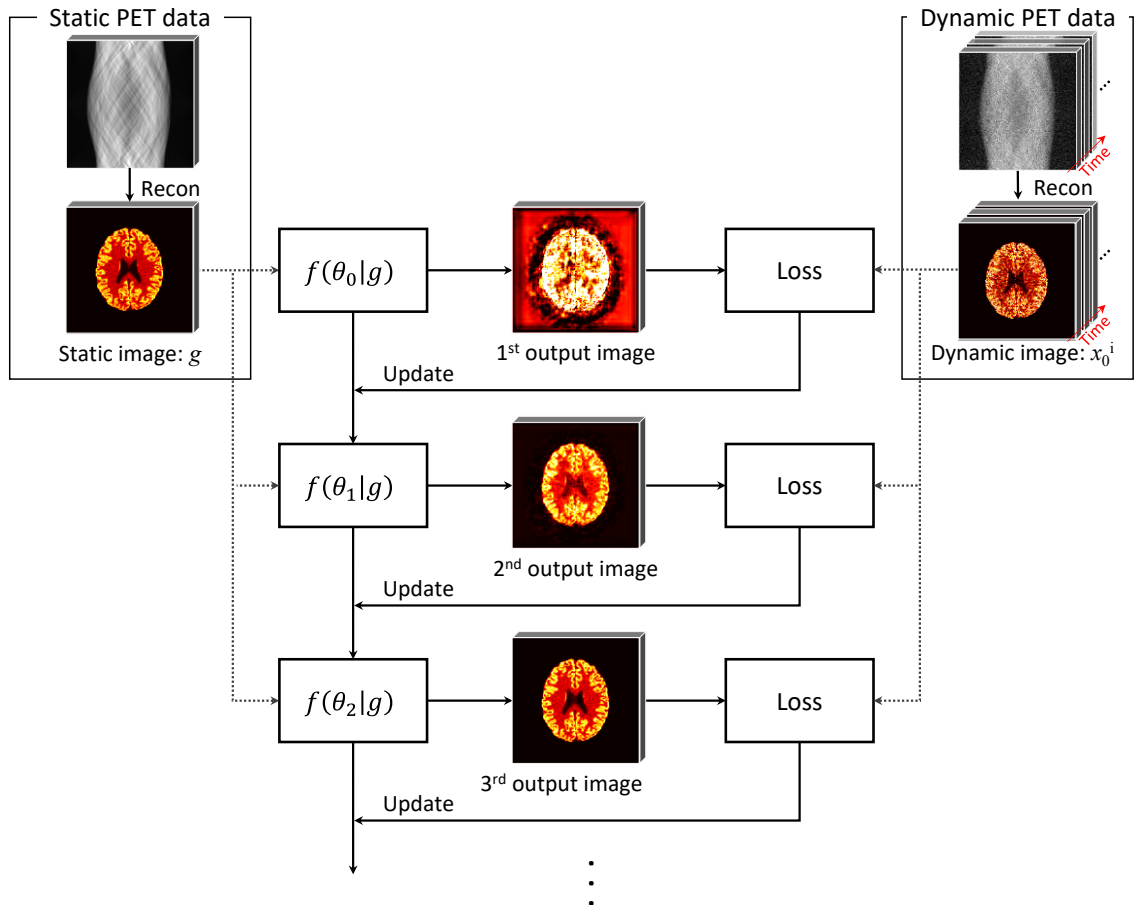


Figure 3.1: Procedure of the proposed conditional DIP framework for dynamic PET imaging. The arrows denote the different operations. The figure is reprinted from the work of Hashimoto et al. [79] (CC BY 4.0)

Figure 3.2 shows a simple example of the original DIP and conditional DIP frameworks for dynamic PET imaging. The experimental condition of this example was the same as in 3.3.1. The original DIP framework when random noise is utilized as the network input, obtained blurred images without keeping the edge information, such as cortex structures. In contrast, the conditional DIP framework when static image is utilized as the network input, not only obtained a comparable level of image smoothness but also kept edge details. This demonstration supports the effectiveness of the conditional DIP framework for dynamic PET imaging.

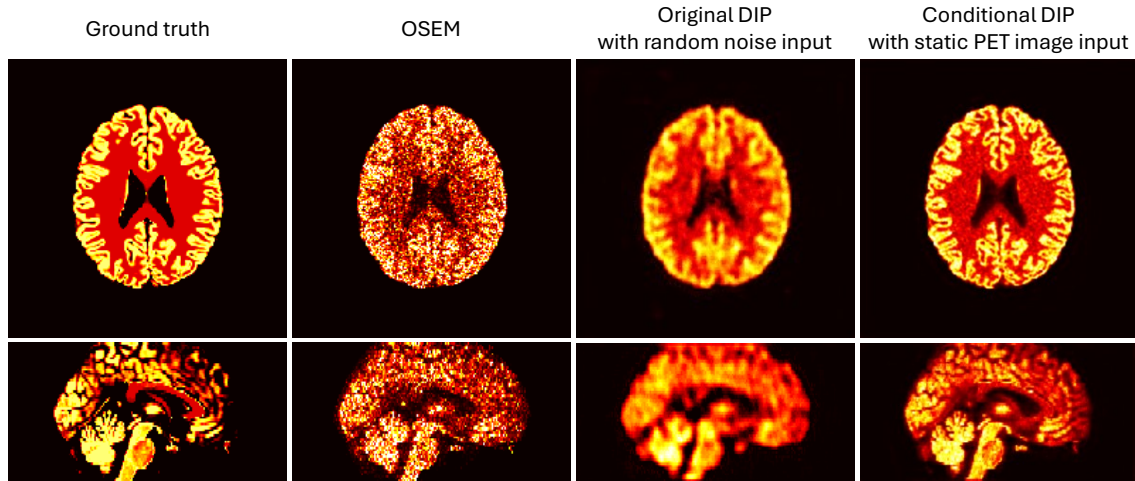


Figure 3.2: Example of the original DIP and conditional DIP frameworks. The arrows denote the different operations. The figure is reprinted with a modification from the work of Hashimoto et al. [79] (CC BY 4.0)

Figure 3.3 illustrates optimization trajectories of the conditional DIP framework for dynamic PET imaging. In the PET imaging system, dynamic PET images suffer from various artifacts due to scatter, attenuation, statistical noise, and other reconstruction errors, which degrade the reconstructed images, as illustrated in the gray manifold region.

Traditional optimization-based denoising methods without any prior information typically diverge significantly from the ground truth image, as represented by the blue trajectory. In contrast, the trajectory of the original DIP denoising tends to conform more closely to the ground truth image due to the inductive bias of the CNN structure¹, with early stopping of the optimization process yielding enhanced PET image quality, as indicated by the green trajectory. Using the static PET image as prior information provides an advantage whereby the initial point of the optimization is positioned nearer to the ground truth point compared to random noise input. Consequently, the optimization using the static PET image input is expected to improve PET image quality compared to the original DIP with random input, as shown by the red trajectory.

¹ Inductive biases in CNNs encompass local connectivity, shared weights, and spatial hierarchies. These biases enable CNNs to efficiently discern patterns rather than noise in images.

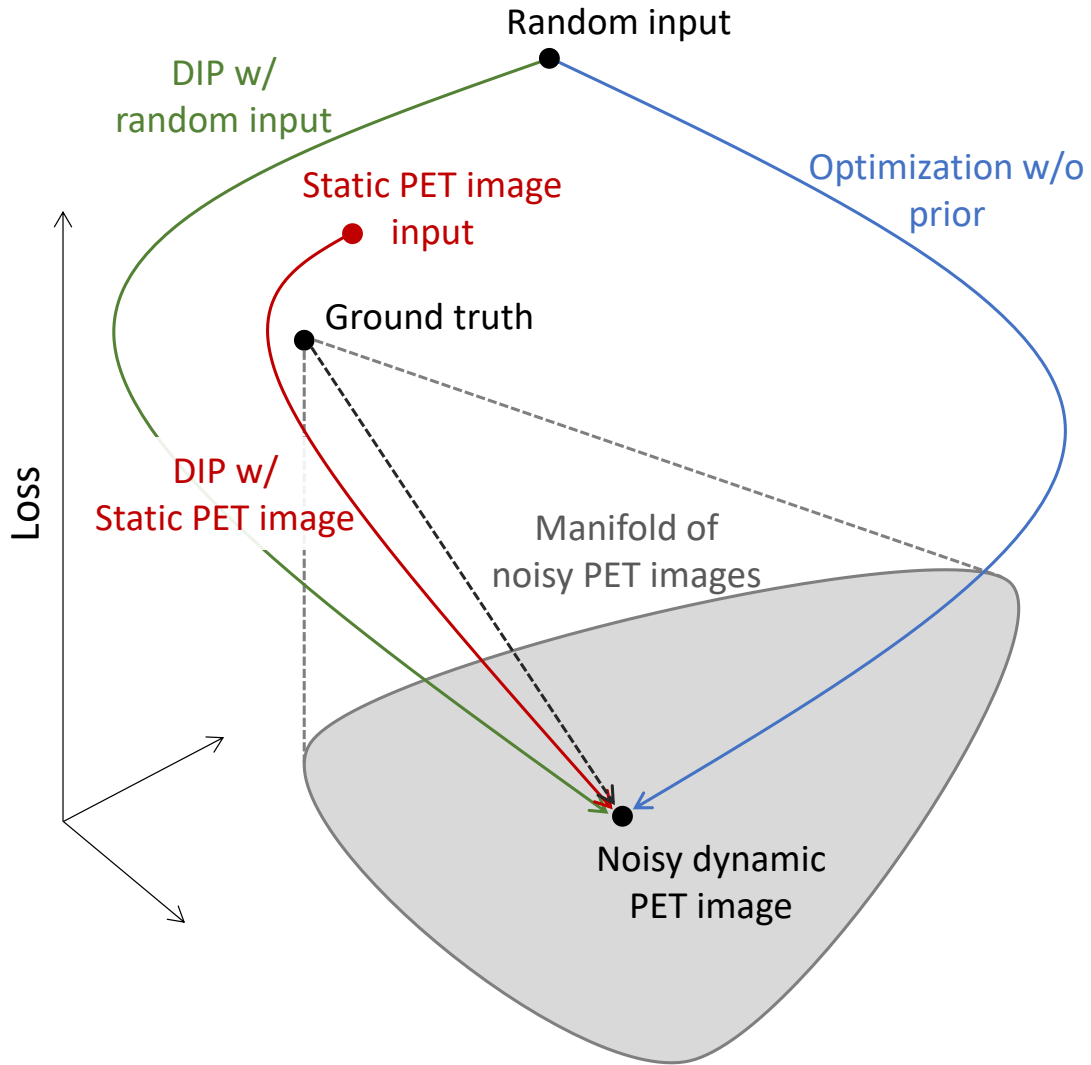


Figure 3.3: Schematic illustration of the optimization trajectories of the conditional DIP framework for dynamic PET imaging. The gray region shows a manifold of degraded noisy dynamic PET images from the ground truth image. The trajectories of traditional optimization-based denoising, original and conditional DIP methods are shown in blue, green, and red lines, respectively. The arrows denote the different operations. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPPEM)

3.2.2 Network structures and optimization

In this study, we employed two network structures: the 3D U-Net architecture [79] illustrated in Figure 2.4, and 4D branch CNN architecture [80] for dynamic PET image denoising.

The encoding path of the 3D U-Net architecture consists of sequential application of two sets of $3 \times 3 \times 3$ convolution layers in three dimensions. Each convolution layer is followed by batch normalization (BN) [81] and a Leaky Rectified Linear Unit (LReLU). Then, a downsampling operation is executed by a $3 \times 3 \times 3$ convolution layer with a stride of two, followed by application of the BN and LReLU. Concomitantly, with each downsampling operation, the feature channel is augmented doubled. The decoder path is constructed with a $3 \times 3 \times 3$ deconvolution layer and two $3 \times 3 \times 3$ convolution layers, followed by the BN and LReLU. In addition, feature maps from the encoder path are linked to the corresponding decoder path via skip connections. Finally, the output layer is activated by a linear function.

The 4D branch CNN architecture comprises two components: a feature extractor and multiple reconstruction branch modules. The feature extractor module is designed to share common spatial and temporal features from 4D dynamic PET images. Subsequently, each reconstruction branch module employs these common features output from the feature extractor module, to reconstruct each dynamic frame independently. The feature extractor module employs the same structure as the 3D U-Net in Figure 2.4. The reconstruction branch module is composed of a sequence of three $3 \times 3 \times 3$ convolution layers, with each followed by the LReLU, and one $3 \times 3 \times 3$ convolution layer through a sigmoid activation for output operations.

The 3D U-Net architecture independently performs the conditional DIP optimization for each frame by Equation (3.1). The 4D branch CNN architecture performs conditional DIP optimization for all dynamic frames in an end-to-end manner as follows,

$$\begin{aligned}\theta^* &= \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^T w_i E\left((\mathbf{x}_0)_i; f(\theta_{F,R_i}|\mathbf{g})\right), \\ \mathbf{x}^* &= f(\theta^*|\mathbf{g}).\end{aligned}\tag{3.2}$$

where $\mathbf{x}_0$ represents the 4D dynamic PET images at i -th time frame ($i = 1, 2, \dots, T$). θ_F and θ_R represent trainable parameters of the feature extractor and reconstruction branch modules.

A sequence of dynamic PET images exhibits significant variation in amplitude across individual frames, resulting in disparate magnitudes of the loss function for each time frame. Inspired by the weighted heterogeneous learning method proposed by Fukui et al. [82], we introduce loss weights w , aimed at equalizing the loss contributions from each frame to mitigate this discrepancy and improve stability during network training, as follows,

$$w_i = \frac{\min_k \frac{1}{N} \sum_{j=1}^N E_{k,j}}{\frac{1}{N} \sum_{j=1}^N E_{k,j}},\tag{3.3}$$

In Equation (3.3), the denominator represents the mean value of the training loss E across epochs up to N for each time frame k , while the numerator corresponds to the minimum value of the mean training loss. This framework serves to normalize the loss function relative to the loss magnitude at each time frame. In this study, the loss weights are pre-calculated over $N = 50$ epochs.

In this study, we used the MSE for the loss function. The SGD with a momentum of

0.9 and a learning rate of 0.01 were employed for the optimization. The trainable parameters of the networks were initialized using He initialization [83]. The experiments were conducted on a computer using the Ubuntu 16.04 with an NVIDIA Quadro RTX 8000 graphics processing unit (GPU) with 48 GB of memory. The network models were implemented using Keras (www.keras.io) with TensorFlow (www.tensorflow.org) as the backend.

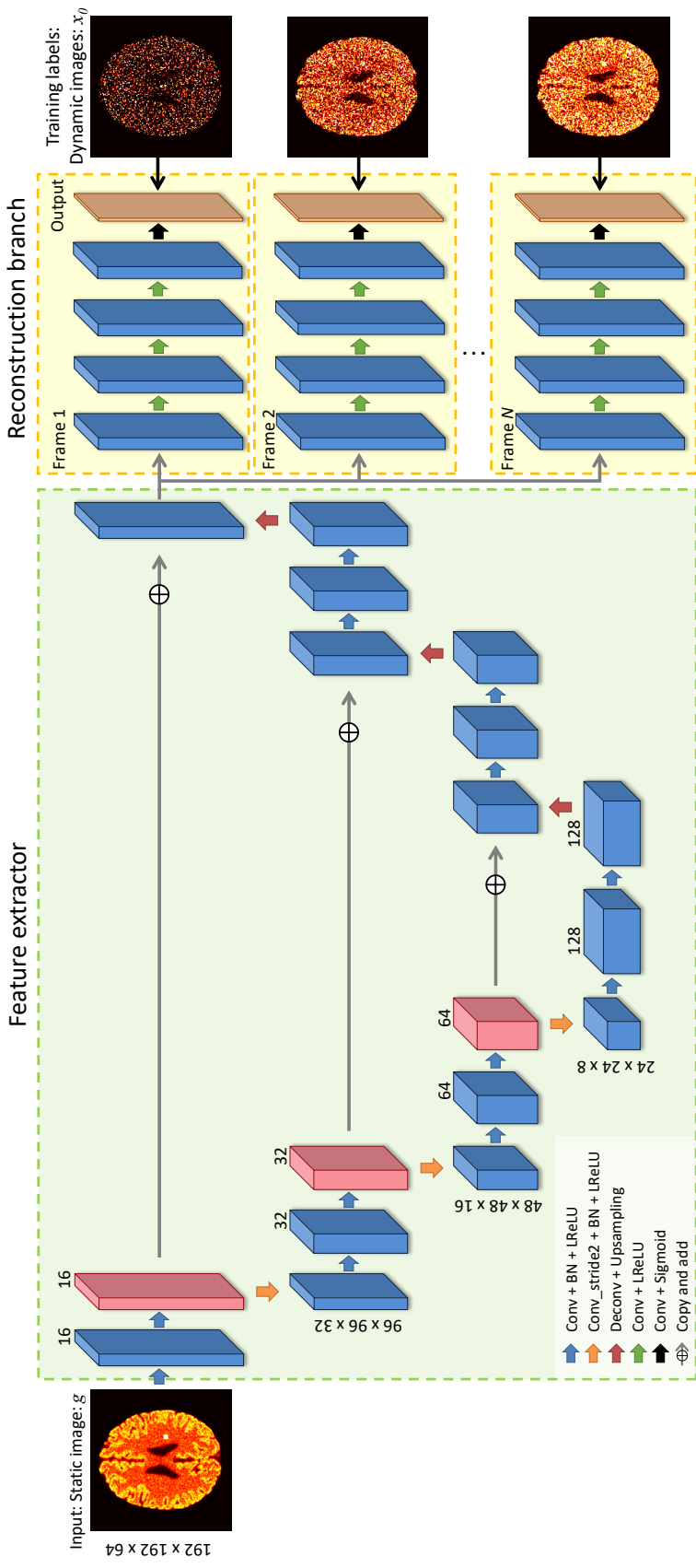


Figure 3.4: Overview of the proposed 4DDIP framework. The proposed network consists of two modules: a feature extractor, and a reconstruction branch. In this framework, the common features of all dynamic PET images can be shared by one CNN architecture by end-to-end learning, and can improve denoising performance compared to 3D processing. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPeM)

3.3 Experimental setup

This subsection describes the details of experimental setup of the computer simulation and real preclinical data to evaluate the performance of conditional DIP denoising for dynamic PET imaging.

3.3.1 Computer simulation

The experimental setup for the computer simulation was implemented with references to [79,80]. This study used a digital 3D brain phantom from the Brainweb database [84,85]. In this computer simulation, the time activity curves (TACs) of the glucose metabolism of ^{18}F -FDG in gray matter, white matter, and tumor were generated by Feng's method [86], based on a two-tissue compartment model [87]. Table 3.1 shows the kinetic parameters, K_1 , k_2 , k_3 and k_4 , and Figure 3.4 shows a plasma input function and regional TACs. The dynamic PET scan consisted of 30-time frames of 4×20 s, 4×40 s, 4×60 s, 4×180 s, and 14×300 s over 90 minutes, started just after the bolus injection of the PET tracer. The regional TACs were averaged in each time frame, and then the sinograms of the dynamic PET scan were calculated using a simple forward projection method. The sinogram and PET image sizes were $192 \text{ angles} \times 192 \text{ bins} \times 64 \text{ slices}$ and $192 \times 192 \times 64 \text{ voxels}$ with $1.0 \times 1.0 \times 2.0 \text{ mm/voxel}$. Poisson noise was added with approximately 10^9 coincidence events over 90 minutes, as listed in Table 3.2. In order to evaluate the simplified performance of the conditional DIP denoising, an ideal PET scanner without any attenuation and scatter was assumed. The image reconstruction was performed using the OSEM algorithm with six iterations and 16 subsets. We calculated parametric images using the Patlak plot graphical analysis [88] by Pmod 3.802

(PMOD Technologies Ltd, Zurich, Switzerland) [89]. The influx rate K_i was calculated from voxel-wise TACs, which corresponds the transfer rate of the irreversible tracer from blood vessels to brain tissue, as follows.

$$K_i = \frac{K_1 k_3}{k_2 + k_3}. \quad (3.4)$$

The start time of the linear phase after equilibration t^* was set to 20 minutes.

The mean squared bias and variance were computed as indicators of quantitative accuracy and noise property, as follows,

$$\text{Bias}^2 = \frac{\sum_{j \in \mathbf{R}} (x_j - (x_{\text{GT}})_j)^2}{\sum_{j \in \mathbf{R}} ((x_{\text{GT}})_j)^2}, \quad (3.5)$$

$$\text{Variance} = \frac{\sum_{j \in \mathbf{R}} (x_j - \bar{x})^2}{\sum_{j \in \mathbf{R}} ((x_{\text{GT}})_j)^2}, \quad (3.6)$$

where x and x_{GT} represent target and grand truth images, and $\mathbf{R}$ represents the target ROIs. In addition, the following indicators of the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) [90] were used for quantitative evaluation.

$$\text{PSNR} = 20 \log_{10} \left(\frac{\max(x_{\text{GT}})}{\sqrt{\frac{1}{N_{\mathbf{R}}} \sum_{j \in \mathbf{R}} (x_j - (x_{\text{GT}})_j)^2}} \right), \quad (3.7)$$

$$\text{SSIM} = \frac{1}{N_{\mathbf{R}}} \sum_{j \in \mathbf{R}} \frac{(2\mu_{jx}\mu_{jx_{\text{GT}}} + c_1)(2\sigma_{jxx_{\text{GT}}} + c_2)}{(\mu_{jx}^2 + \mu_{jx_{\text{GT}}}^2 + c_1)(\sigma_{jx}^2 + \sigma_{jx_{\text{GT}}}^2 + c_2)}, \quad (3.8)$$

where μ_{jx} , $\mu_{jx_{\text{GT}}}$ and σ_{jx} , $\sigma_{jx_{\text{GT}}}$ represent the mean values and standard deviations in the square region on the j -th voxel of x and x_{GT} , respectively. $\sigma_{jxx_{\text{GT}}}$ represents the covariance between x and x_{GT} . These indices are calculated within whole brain regions

R. c_1 and c_2 were defined as $(0.01L)^2$ and $(0.03L)^2$ with the dynamic range of the grand truth L .

Table 3.1: The kinetic parameters used in this computer simulation.

Regions	K1	k2	k3	k4
White matter	0.046	0.080	0.052	0.001
Gray matter	0.071	0.086	0.055	0.001
Tumor	20.082	0.055	0.085	0.001

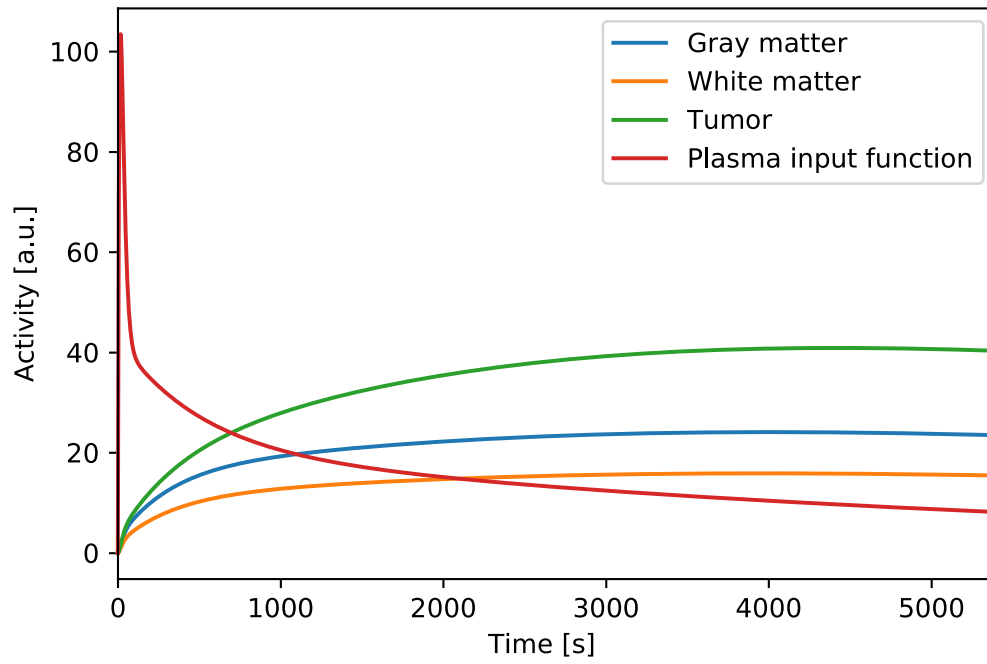


Figure 3.5: The regional time activity curves and plasma input function in the computer simulation.

Table 3.2: Coincidence events of each time frame in the experiments.

Time frame	Time duration	Computer simulation	Real preclinical data	
		^{18}F -FDG	^{18}F -FDG	^{11}C -raclopride
1	20	1.13E+05	2.24E+05	5.76E+05
2	20	4.85E+05	1.31E+06	3.50E+06
3	20	7.76E+05	2.11E+06	4.27E+06
4	20	9.71E+05	2.33E+06	4.40E+06
5	40	2.37E+06	5.10E+06	8.90E+06
6	40	2.84E+06	5.45E+06	8.97E+06
7	40	3.25E+06	5.71E+06	8.94E+06
8	40	3.62E+06	5.89E+06	8.79E+06
9	60	6.04E+06	9.06E+06	1.28E+07
10	60	6.68E+06	9.23E+06	1.22E+07
11	60	7.22E+06	9.37E+06	1.17E+07
12	60	7.70E+06	9.45E+06	1.11E+07
13	180	2.53E+07	2.87E+07	3.01E+07
14	180	2.78E+07	2.89E+07	2.57E+07
15	180	2.95E+07	2.89E+07	2.20E+07
16	180	3.08E+07	2.87E+07	1.88E+07
17	300	5.36E+07	4.73E+07	2.55E+07
18	300	5.57E+07	4.62E+07	1.97E+07
19	300	5.75E+07	4.50E+07	1.53E+07
20	300	5.90E+07	4.37E+07	1.19E+07
21	300	6.01E+07	4.22E+07	9.26E+06
22	300	6.11E+07	4.06E+07	7.26E+06
23	300	6.18E+07	3.91E+07	5.71E+06
24	300	6.22E+07	3.74E+07	4.50E+06
25	300	6.25E+07	3.57E+07	3.56E+06
26	300	6.26E+07	3.42E+07	2.83E+06
27	300	6.24E+07	3.27E+07	2.25E+06
28	300	6.22E+07	3.13E+07	1.80E+06
29	300	6.18E+07	3.00E+07	1.44E+06
30	300	6.12E+07	2.86E+07	1.15E+06

3.3.2 Real preclinical PET data

The real preclinical experiments were approved by the Animal Ethical Committee of the Central Research Laboratory, Hamamatsu Photonics K.K. (Approval number: HPK-2017-02).

The dynamic ^{18}F -FDG and ^{11}C -raclopride [91] PET scans of a conscious rhesus monkey brain were performed using an animal-dedicated PET scanner (SHR-38000, Hamamatsu Photonics K.K.) [92]. The axial and transaxial fields of view (FOV) of the PET scanner are 330 mm and 108 mm, respectively, and the spatial resolution at the center position is 2.3 mm. The head of rhesus monkey was fixed to a monkey chair using a head restraint on the monkey's skull. The head of monkey was set in the PET gantry in parallel with the orbitomeatal plane.

Initially, a 30-minute transmission scan was conducted with a ^{68}Ge – ^{68}Ga rod source, followed by a dynamic emission scan that spanned 90 minutes and was divided into 30 distinct time frames: 4×20 s, 4×40 s, 4×60 s, 4×180 s, and 14×300 s. The administered doses for ^{18}F -FDG and ^{11}C -raclopride were 194.7 MBq and 291.7 MBq, respectively. The arterial plasma input function for the ^{18}F -FDG PET scan was derived by calculating the ratio of the unmetabolized fraction to the total plasma activity. The image reconstruction was performed using a 3D dynamic row-action maximum-likelihood algorithm (DRAMA) algorithm [93] with two iterations and 60 subsets. The sinogram and PET image sizes were $360 \text{ angles} \times 360 \text{ bins} \times 103 \text{ slices}$ with span of three and a ring difference of 19, and $256 \times 256 \times 103 \text{ voxels}$ with $0.65 \times 0.65 \times 1.0167 \text{ mm/voxel}$. The PET images were cropped to $192 \times 192 \times 64 \text{ voxels}$ because of the GPU memory limitation. The detailed time frame and coincidence events in the experiments

are listed in Table 3.2.

For the ^{18}F -FDG data, the influx rate K_i was calculated using the Pmod software. The putamen and background white matter ROIs were manually set on the co-registered MR images, and mean putamen uptakes and background standard deviations were calculated for quantitative analyses.

For the ^{11}C -raclopride data, the non-displaceable binding potential BP_{ND} images were generated using the Logan plot graphical analysis [94] using the Pmod software, which calculate binding to dopamine D_2/D_3 receptors as follows,

$$\text{BP}_{\text{ND}} = \frac{k_3}{k_4}. \quad (3.9)$$

As a reference region, we used the cerebellar cortex region instead of an arterial blood input data. The start time of the linear phase after equilibration t^* was set to 20 minutes, and the efflux rate across the brain blood barrier k_2 was set to 0.15. The putamen and background white matter ROIs were manually set on the co-registered MR images.

3.3.3 Comparison algorithms

In the experiments, we compared the conditional DIP using 3D U-Net and 4D branch CNN architectures with following algorithms.

- Gaussian filtering: The Gaussian filtering is often used for post-denoising as a baseline method of the PET imaging. The 3D Gaussian filtering is calculated as follows,

$$(x_{GF})_i = \sum_{j \in \omega_i} \frac{\exp\left(-\frac{(i-j)^2}{2\sigma^2}\right)}{\sum_{k \in \omega_i} \exp\left(-\frac{(i-k)^2}{2\sigma^2}\right)} x_j, \quad (3.10)$$

where $\mathbf{x}$ is the noisy image. σ and ω represent the standard deviation of the Gaussian distribution and window.

- Guided filtering [24,95]: The guided filtering is widely used as a fast edge-preserving filter, which applies for edge-aware denoising as well as the image synthesis task of no-flash/flash image denoising. We applied 3D guided filtering which perform a linear transformation from a guidance image to the denoised output image. The 3D guided filtering is calculated as follows,

$$(x_{Guided})_i = \sum_j w_{guided}(I) x_j, \quad (3.11)$$

$$(w_{guided})_{ij}(I) = \frac{1}{|\omega|^2} \sum_{k \in \omega_i, k \in \omega_j} \left(1 + \frac{(I_i - \mu_k)(I_j - \mu_k)}{\sigma_k^2 + \varepsilon} \right), \quad (3.12)$$

where I is the guidance image. μ and σ^2 represent a mean and variance of the guidance image. $|\omega|$ is number of voxels in the window ω . ε represents a regularization parameter which penalizes a linear coefficient. In the

experiments, we used the same static PET image as the input of the conditional DIP as the guidance image for fair comparison.

3.4 Results

3.4.1 Computer simulation

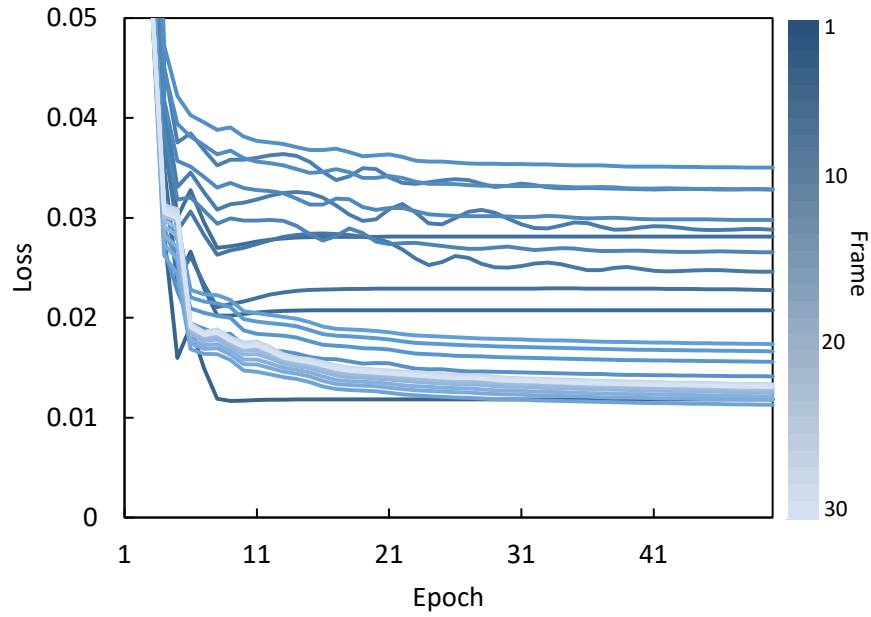
Figure 3.6 shows the loss curves for each time frame in the 4D branch CNN architecture. The differences in the loss curves increased, leading to unstable learning curves for the network when the loss weights were not applied for network training. Conversely, the differences in the loss curves decreased, and the learning curves of the network became stable when the loss weights were used for network training.

Figure 3.7 displays the transaxial and sagittal slices of the time frame 6, covering 120–160 s, time frame 26, covering 3900–4200 s, and the Patlak influx constant images, calculated by different denoising algorithms in the computer simulation. The qualitative assessment demonstrated that the resolution of the conditional DIP processed by the 4D branch CNN architecture is notably enhanced in both the early phase of time frame 6 and the Patlak influx constant images compared to other denoising algorithms.

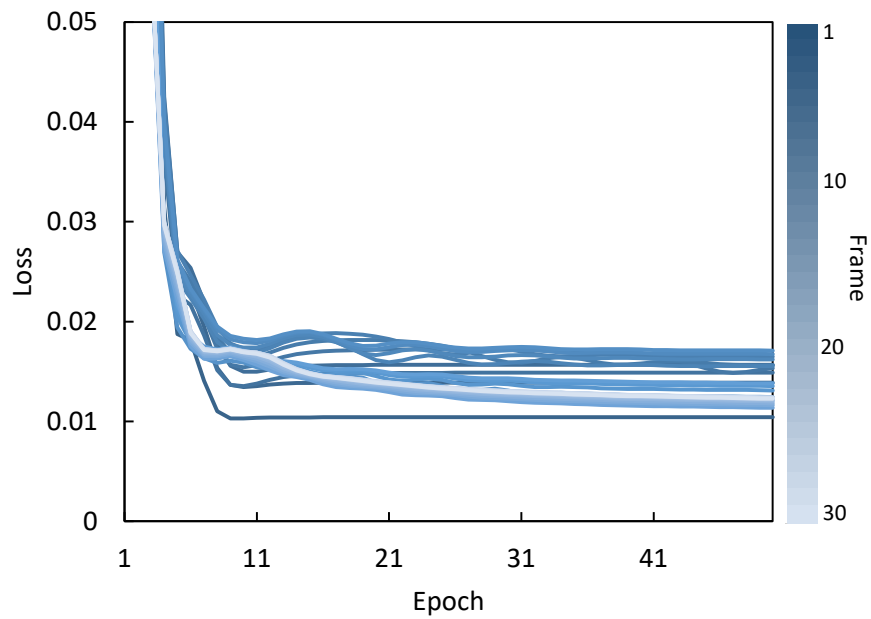
Figure 3.8 illustrates tradeoffs between the mean squared bias and variance for the tumor ROIs in the time frame 26 which is 65–70 min post-injection, as well as for the Patlak influx constant images. The tradeoff curves demonstrated that the conditional DIP processed by the 4D branch CNN architecture reduced bias and variance compared to different denoising algorithms.

Figure 3.9 shows the PSNR and SSIM across different denoising algorithms in the Patlak influx constant images. The PSNRs and SSIMs for both conditional DIP processed by the 3D U-Net and 4D branch CNN architectures surpass those achieved with the 3D Gaussian filtering and guided filtering. These results showed that the conditional DIP processed by the 4D branch CNN architecture outperforms other denoising algorithms in

terms of performance of noise suppression and quantitiveness.



(a) w/o loss weights



(b) w/ loss weights

Figure 3.6: Influence of the loss weights application on the loss. (a) The learning curves without applying the loss weights, and (b) with applying the loss weights. The loss weights reduced the variance in the loss curves, thereby contributing to the stabilization of network training. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPeM)

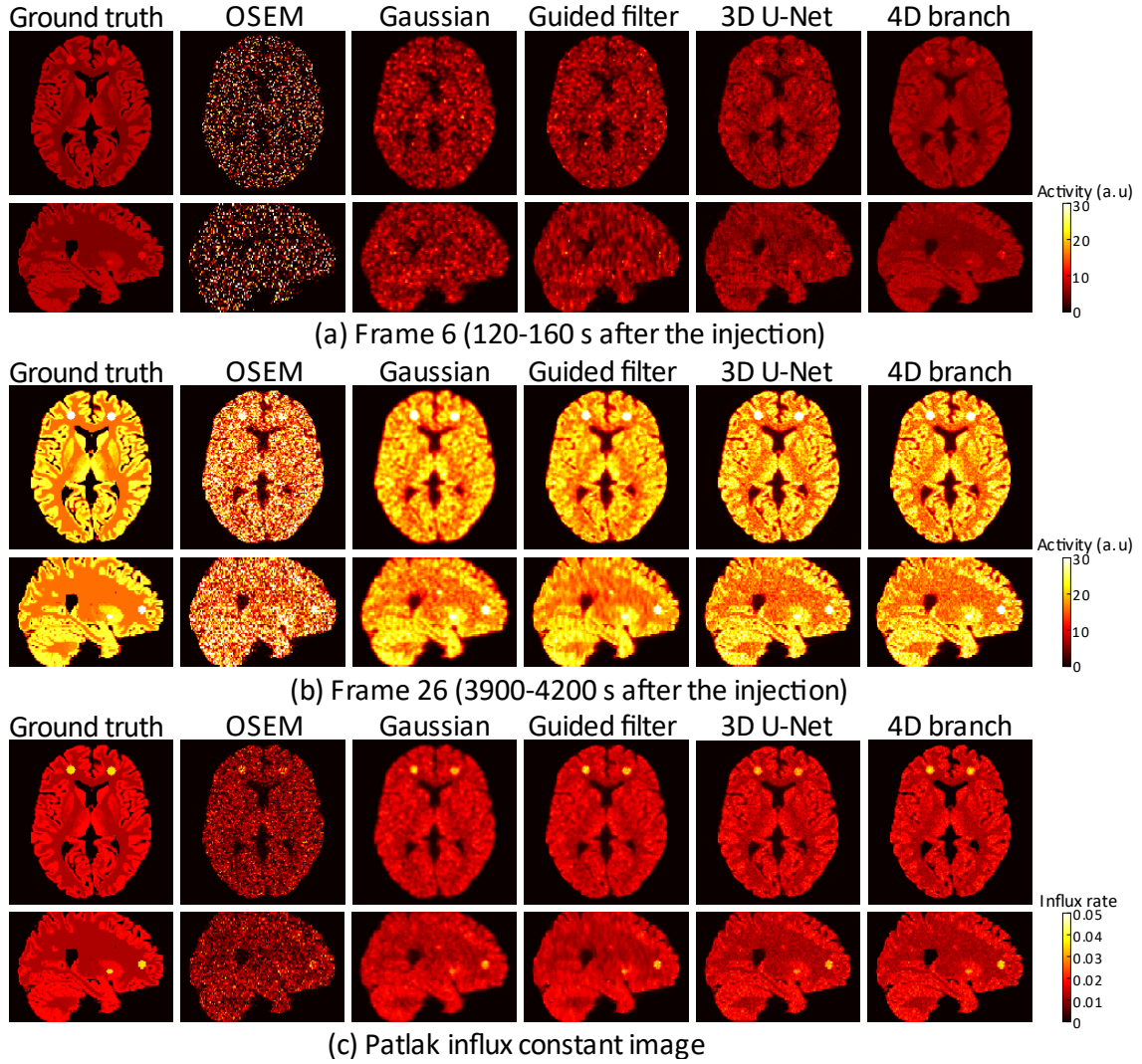


Figure 3.7: Simulation results of (a) the dynamic PET image at the time frame 6, (b) time frame 26, and (c) Patlak influx constant image. The columns correspond to the ground truth, reconstructed images obtained by the OSEM algorithm, Gaussian filtering, guided filtering, and conditional DIP with the 3D U-Net and 4D branch CNN architectures (left to right). The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPEM)

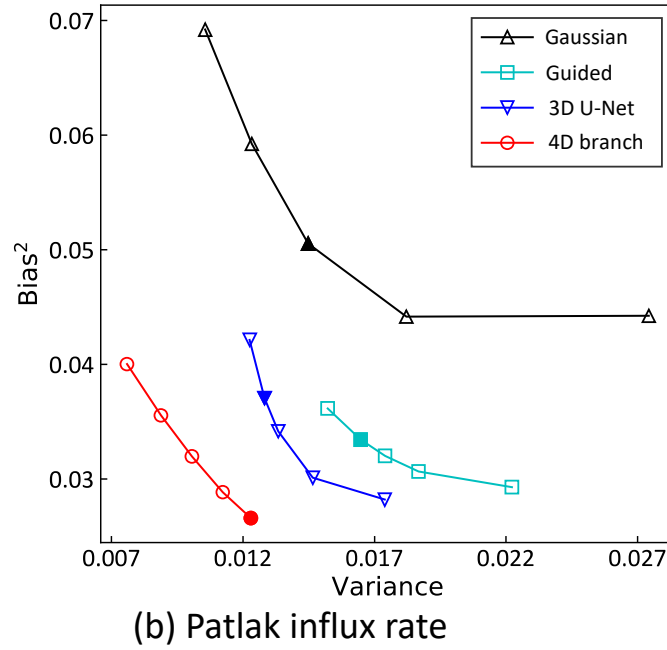
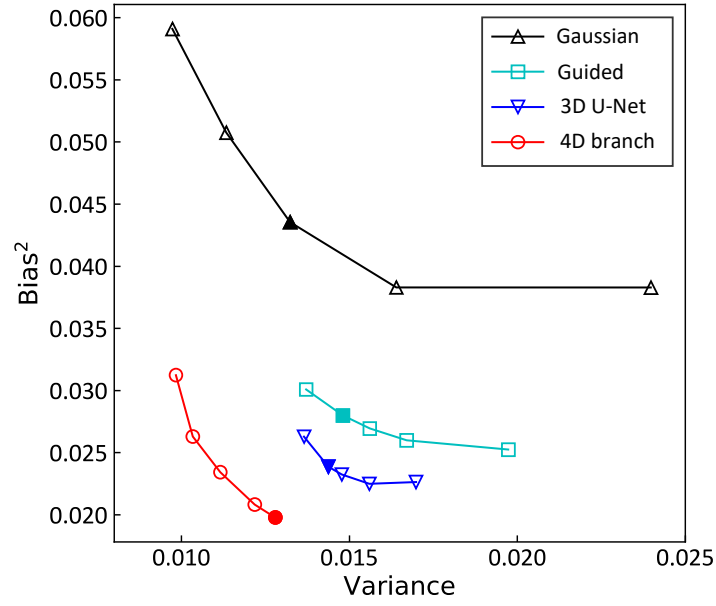


Figure 3.8: Tradeoffs between the mean squared bias and variance for the simulation data at the tumor ROIs in (a) the time frame 26 and (b) Patlak influx constant image. The markers corresponding to 2.0, 2.5, 3.0, 3.5, and 4.0 mm FWHM for the Gaussian filtering, 0.1, 0.5, 0.75, 1.0 and 1.5 ϵ for the guided filtering, 20 000, 15 000, 10 000, 5000 and 500 epochs for the conditional DIP with 3D U-Net architecture, and 20 000, 10 000, 5000, 2500 and 1000 epochs for the conditional DIP with 4D branch CNN architecture. The images in Figure 3.7 are labeled by filled markers. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPEM)

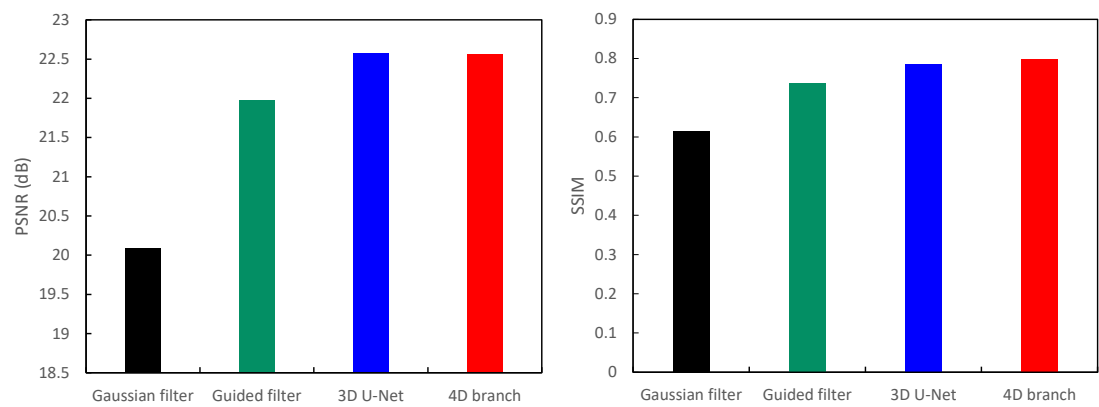


Figure 3.9: PSNRs and SSIMs for the simulation data.

3.4.2 Real preclinical PET data

Figure 3.10 displays the transaxial and sagittal slices of the time frame 6, covering 120–160 s, time frame 26, covering 3900–4200 s, and the Patlak influx constant images, calculated by different denoising algorithms in the real preclinical dynamic ^{18}F -FDG data. When the guided filtering was applied, which uses the same static PET image information for the guidance image as the conditional DIP, it resulted in the loss of structure details such as cortex region. Conversely, both the conditional DIP processed by the 3D U-Net and 4D branch CNN architectures successfully preserved the detailed structures. Notably, the proposed conditional DIP denoising with the 4D branch CNN managed to reduce statistical noise while preserving fine details.

Figure 3.11 illustrates tradeoffs between the average putamen uptake and the background standard deviation for the time frame 26, which is 65–70 min post-injection, as well as for the Patlak influx constant images. The conditional DIP denoising with the 4D branch CNN architecture demonstrated superior performance, exhibiting higher putamen uptake values and lower background standard deviation compared to different denoising algorithms.

Figure 3.12 displays the transaxial and sagittal slices of the time frame 6, covering 120–160 s, time frame 26, covering 3900–4200 s, and non-displaceable binding potential BP_{ND} images, calculated by different denoising algorithms in the real preclinical dynamic ^{11}C -raclopride data.

Figure 3.13 shows the mean BP_{ND} values at the putamen ROI. Through both visual assessment and quantitative analysis, the conditional DIP with the 4D branch CNN architecture is observed to enhance the resolution of dynamic PET images in the time

frame 26 and to yield a higher BP_{ND} in the non-displaceable binding potential images compared to different denoising algorithms.

Figure 3.14 illustrates the regional TACs for the putamen and cerebellar cortex regions treated by different denoising algorithms. The TACs denoised by the proposed conditional DIP with the 4D branch CNN architecture exhibit a smoother profile compared to those denoised by other algorithms. These findings indicate that the proposed conditional DIP with 4D branch CNN architecture is versatile and practical for PET tracers that exhibit not only a widespread distribution such as ^{18}F -FDG but also concentrated accumulation in brain regions such as ^{11}C -raclopride.

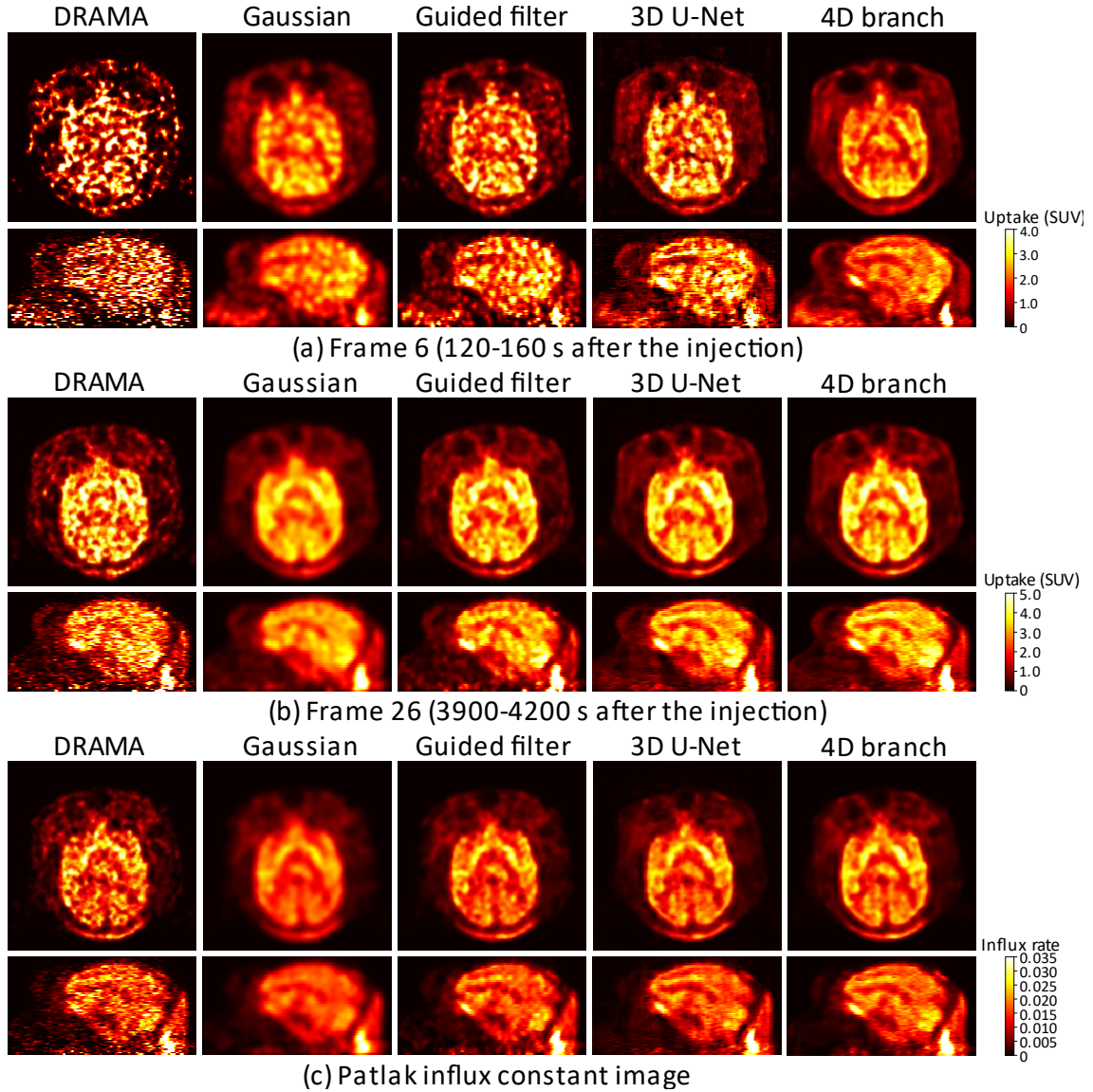


Figure 3.10: Preclinical results of (a) the dynamic PET images at time frame 6, (b) the time frame 26, and (c) the Patlak influx constant images in ^{18}F -FDG. The columns correspond to the images processed by 3D DRAMA, Gaussian filtering, guided filtering, and conditional DIP with 3D U-Net and 4D branch CNN architectures (left to right). The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPEM)

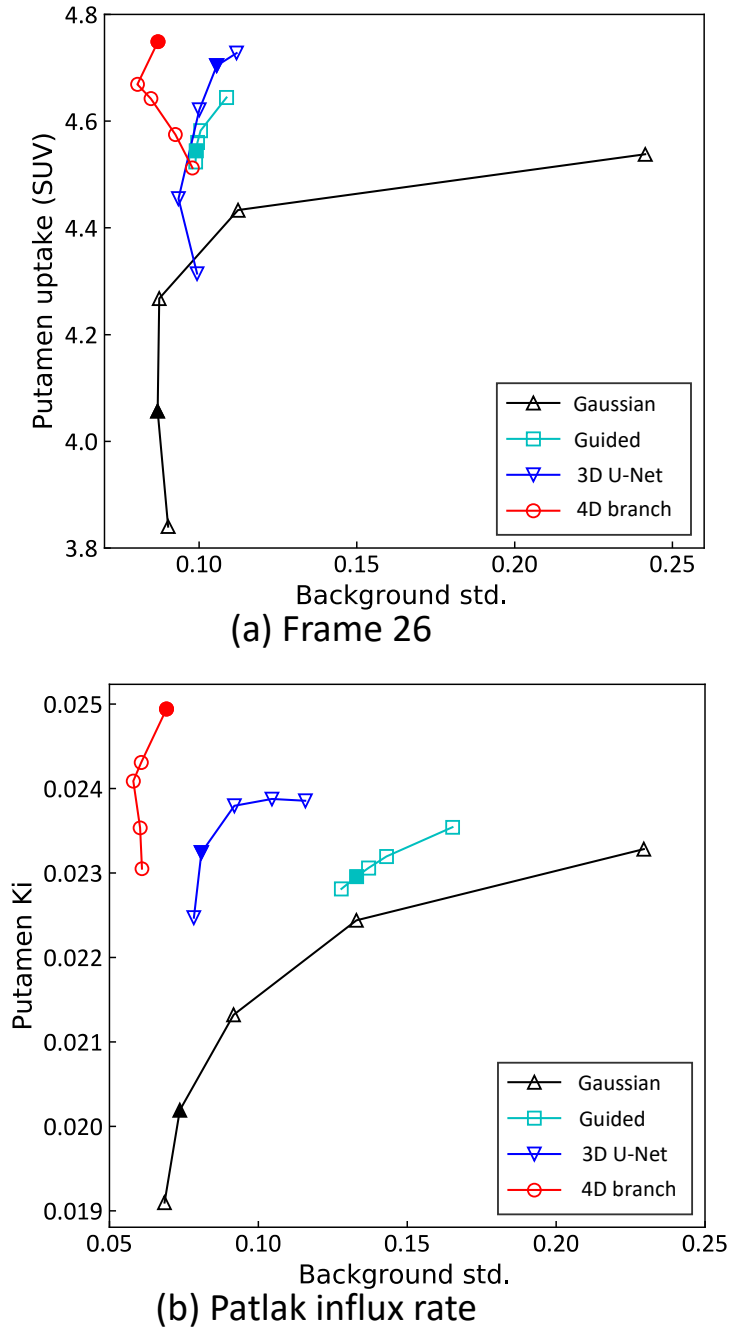


Figure 3.11: Tradeoffs between the putamen uptake and background standard deviation for preclinical ^{18}F -FDG data in (a) the time frame 26, and (b) Patlak influx constant images. The markers corresponding to 1.0, 2.0, 3.0, 4.0 and 5.0 mm FWHM for the Gaussian filtering, 0.1, 0.5, 0.75, 1.0 and 1.5 ϵ for the guided filtering, 1250, 1000, 750, 500 and 250 epochs for the conditional DIP with 3D U-Net architecture, and 7000, 5000, 4000, 3000 and 2000 epochs for the conditional DIP with 4D branch CNN architecture (upper right to lower left). The images in Figure 3.10 are labeled by filled markers. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPeM)

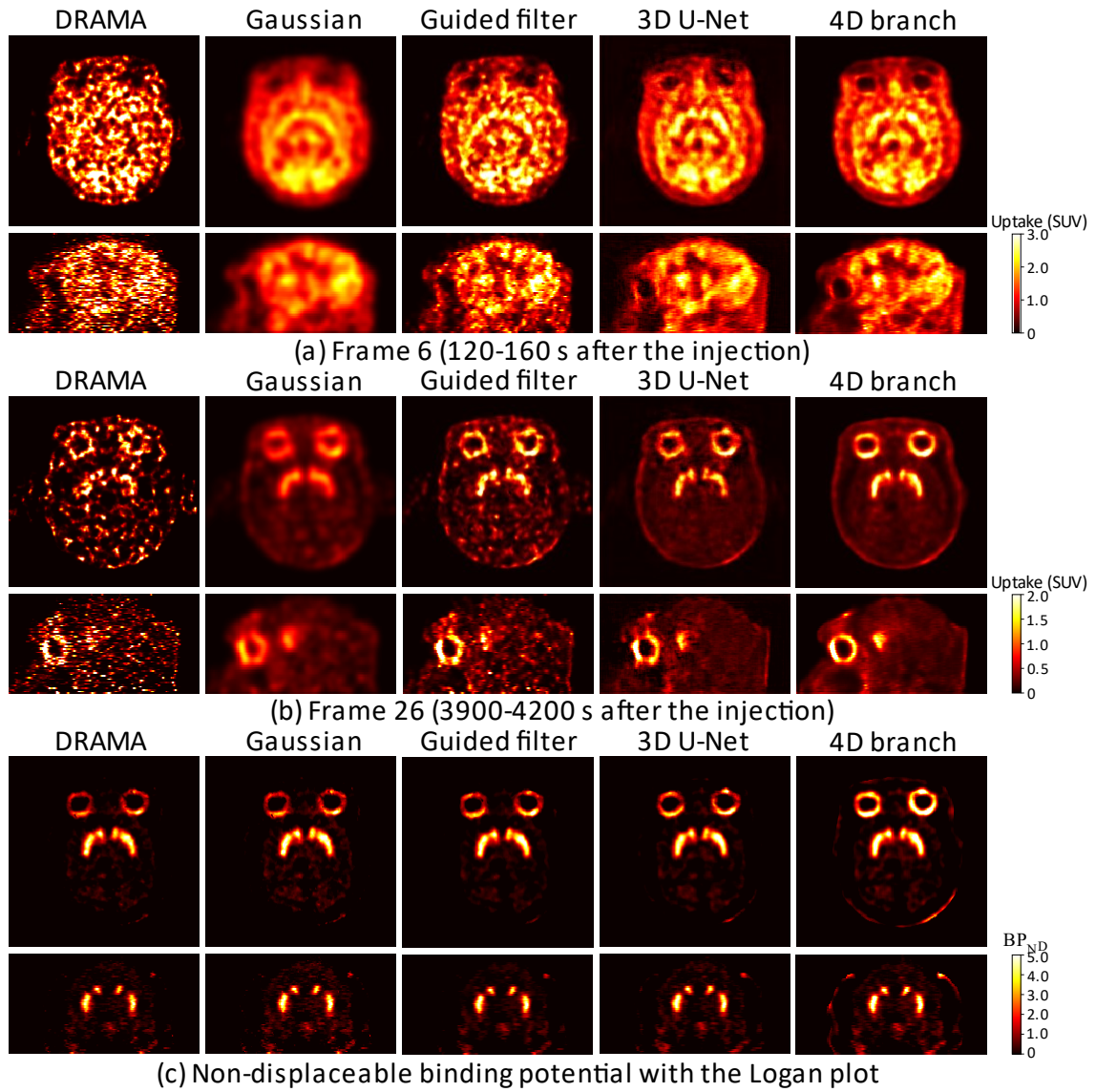


Figure 3.12: Preclinical results of (a) the dynamic PET images at the time frame 6, (b) time frame 26, and (c) non-displaceable binding potential images in ^{11}C -raclopride. The columns correspond to the images processed by the 3D DRAMA, Gaussian filtering, guided filtering, and conditional DIP with 3D U-Net and 4D branch CNN architectures (left to right). The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPEM)

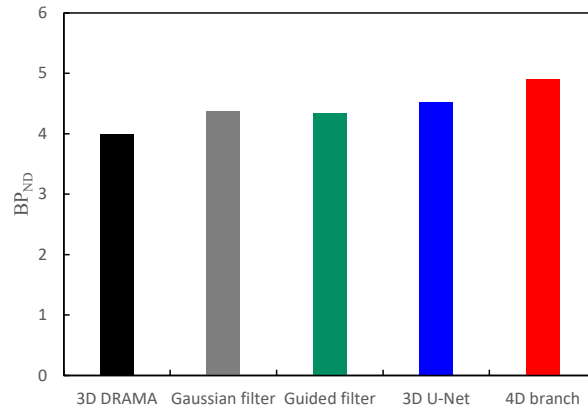
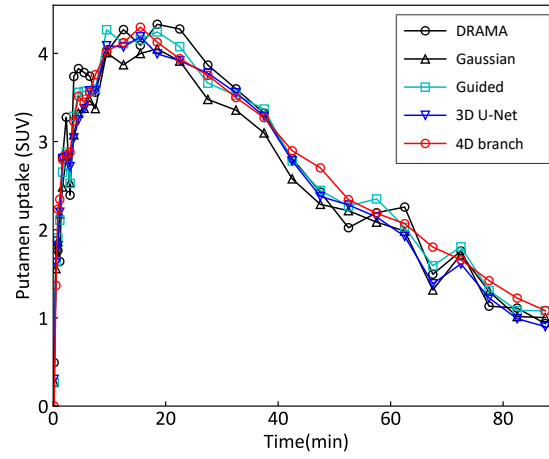
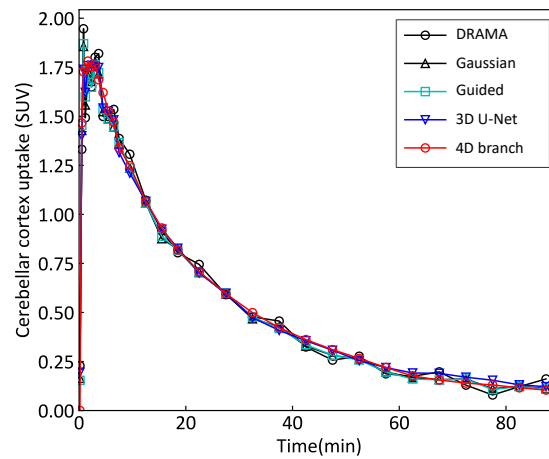


Figure 3.13: Mean BP_{ND} values at the putamen ROI in the ^{11}C -raclopride experiment.



(a) Putamen



(b) Cerebellar cortex

Figure 3.14: Regional TACs at (a) the putamen and (b) cerebellar cortex ROIs. The figure is reprinted from the work of Hashimoto et al. [80] (© 2021 IPPEM)

3.5 Discussion

Various supervised and self-supervised deep learning algorithms for PET image denoising have been developed [96-102]. For example, Gong et al. [96] proposed the supervised PET image denoising using the perceptual loss based on a pre-trained VGG network to enhance PET image quality [103]. Ote et al. [97] investigated the impact of four different network architectures on the supervised low-dose PET image denoising. However, the quantitative accuracy of the denoised PET images cannot be assured for unknown cases not covered in the training datasets, such as unfamiliar diseases, new PET tracers, and different human races. Furthermore, these methods require the preparation of extensive datasets. Our proposed conditional DIP framework can address these concerns as it is an unsupervised deep learning algorithm that relies solely on the subject's own data.

For the results of the computer simulation, the tradeoffs between the mean squared bias and variance for the tumor ROIs demonstrated that the proposed conditional DIP with the 4D branch CNN architecture can preserve the quantitative accuracy within the tumor region while minimizing image noise. This trend is particularly noticeable in earlier time frames, such as the time frame 6, compared to the guided filtering and conditional DIP with 3D U-Net architecture methods, which utilize the same static PET image information as the proposed conditional DIP denoising with the 4D branch CNN architecture. These are also evident in the quantitative results of the real preclinical PET data.

The enhanced performance of the proposed conditional DIP denoising with the 4D branch CNN architecture is attributed to its end-to-end framework, which includes the

feature extractor that facilitates the sharing of common information across different time frames. Common information extracted by the feature extractor are transferred to each reconstruction branch for reconstructing detailed information of each dynamic PET frame to improve the dynamic PET image denoising performance.

Previous research on conditional DIP for PET image denoising by Cui et al. [76] and Hashimoto et al. [79], has indicated that the conditional DIP outperforms other denoising algorithms such as the guided nonlocal means filtering [104], block-matching 4D filtering [105], and deep decoder denoising [106]. Given these findings, it seems reasonable that the proposed conditional DIP denoising with the 4D branch CNN architecture would have competitive or better results compared to these denoising algorithms.

In this chapter, we utilized the static PET image as the network input. The conditional DIP denoising is also available to use the X-ray CT or MRI images to enhance PET image denoising performance, in line with findings from previous research [76,78,107].

In the experiments of the real preclinical PET data, we used the ^{18}F -FDG, which showed no significant changes in the distribution of radioactivity between the early and later time frames, as well as a reversible-type PET tracer, ^{11}C -raclopride, that exhibits dynamic shifts in RIs, particularly in the striatum. The real preclinical data results indicated that the conditional DIP denoising with the 4D branch CNN architecture delivers greater resolution in the time frame 26 and a higher BP_{ND} values in the non-displaceable binding potential images compared to other denoising algorithms. The regional TAC for the putamen with the Gaussian filtering was relatively lower than that obtained with the conditional DIP denoising. Moreover, the regional TACs generated by the conditional DIP denoising with the 4D branch CNN architecture were smoother than those produced by other denoising algorithms. It is important to note that the regional

TACs of the cerebellar cortex were not underrepresented due to the comparatively large ROI. The results suggest that the conditional DIP denoising is capable of being effectively utilized for both nonreversible accumulate radioactive tracers, such as ^{18}F -FDG, and reversible radioactive tracers, such as ^{11}C -raclopride, with reduced bias. Consequently, we believe that the conditional DIP denoising framework is applicable to a wide array of PET tracers.

In small animal studies, an excessive dose of radioactive tracers might lead to an increased receptor occupancy by non-radioactive tracers in neuroreceptor imaging. The conditional DIP denoising framework has the potential to lower both the radiation dose and the dose of non-radioactive PET tracers, offering a solution to the issue of overdosage.

The original research by Ulyanov et al. [33,34] indicated that the performance of the DIP framework for tasks such as image denoising and inpainting was significantly influenced by the network structures. This is because the network structure has an intrinsic inductive bias. Our future work will focus on exploring an optimal network architecture that provides more robust regularization for PET image denoising.

A notable limitation of this study is the impact of involuntary patient movements during dynamic PET scanning. This real preclinical experiment can ignore the effect of movements because the monkey subjects were securely immobilized. However, for actual situation for human study, involuntary movements may result in blurred PET images due to positional errors across dynamic frames. According to the report by Zhao et al. [108], the conditional DIP tolerates certain degrees of misalignment between prior information and the target images. As an alternative, we can solve to use motion correction techniques [109-111] to mitigate above concern. Future research will explore the influence of patient movement on the image quality.

Our evaluation was confined to simulated and real preclinical brain PET data. To fully assess the effectiveness of the conditional DIP denoising framework, further evaluations incorporating other organs, human, and different radioactive tracer datasets are essential.

4 DIPrecon: PET image reconstruction using conditional DIP

4.1 Introduction

As explained in Chapter 3, the conditional DIP denoising framework has the potential to improve the image quality of PET images. The conditional DIP denoising is a post-processing method, which is characterized by fast calculations and is easy to implement in existing PET scanners. However, there is a risk of accidentally removing important image information, such as lesions, along with image noise.

In this chapter, we introduce an iterative PET image reconstruction incorporating the conditional DIP framework, called DIPrecon [112,113]. The DIPrecon incorporates the conditional DIP framework into image reconstruction to achieve powerful noise reduction while measuring consistency with measured emission data, fundamentally solving the above-mentioned challenges of post-processing. The advantage of the proposed reconstruction algorithm is that the end-to-end DIPrecon is formulated as a single optimization problem by incorporating the forward projection model into the loss function of the conditional DIP framework and can reduce the number of hyperparameters compared to a hybrid DIPrecon algorithm. We evaluate the proposed algorithm with conventional iterative reconstruction algorithms using Monte Carlo simulation data and real preclinical data.

4.2 Methods

4.2.1 Proposed end-to-end DIPrecon algorithm

For the proposed end-to-end reconstruction algorithm, the reconstructed PET image $\mathbf{x}$ is obtained using the conditional DIP denoising framework as follows,

$$\mathbf{x} = f(\theta|\mathbf{g}), \quad (4.1)$$

where f is a neural network with trainable parameters θ which are the representation of the reconstructed PET image $\mathbf{x}$. $\mathbf{g}$ represents prior information as the network input.

The PET image can be calculated by minimizing the constrained optimization problem using Equation (2.3) as follows,

$$\begin{aligned} \min E(\mathbf{Ax}, \mathbf{y}_0) \\ \text{s. t. } \mathbf{x} = f(\theta|\mathbf{g}), \end{aligned} \quad (4.2)$$

where E is the loss function such as the MSE or Poisson negative log likelihood function. $\mathbf{A}$ is a system matrix that denotes the probability of each voxel to each LOR, and $\mathbf{y}_0$ is the measured projection data. In this study, we substitute the above constraint into the objective function to compute reconstructed process in one step as follows,

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \|(\mathbf{A}f(\theta|\mathbf{g}) - \mathbf{y}_0) \odot \mathbf{m} \|, \quad (4.3)$$

$$\mathbf{x}^* = f(\theta^*|\mathbf{g}), \quad (4.4)$$

where $\odot$ represents the Hadamard product, and $\mathbf{m}$ is the binary mask of the detector gaps. The detector gaps are typically filled by an interpolation process in the sinogram space. In this study, inspired by the inpainting task of the DIP framework, the loss function was computed using only measured data. In addition, we used the MR images as the prior

information of the network input.

4.2.2 Implementation

Figure 4.1 illustrates an implementation overview of the proposed end-to-end DIPrecon algorithm. For this experiment, we utilized the same 3D U-Net architecture that was presented in Figure 2.4, with the only modification being the output convolution layer. The output layer consisted of a $1 \times 1 \times 1$ convolution layer with the ReLU activation function. We used a rotation-based method to implement the forward projection [114]. The proposed end-to-end DIPrecon algorithm performs the following steps. (1) The prior information of the MR image is input into the network. (2) The PET image is obtained from the network output. (3) The forward projection is calculated to obtain the estimated sinogram. (4) The loss is calculated with the measured sinogram and estimated sinogram. (5) The network parameters are updated through the back-propagation algorithm.

The limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) algorithm [115], a quasi-Newton method which uses an approximation of the inverse Hessian matrix (second-order gradient) of the loss function, was utilized in this implementation. The L-BFGS algorithm is known for its stable convergence and quicker performance compared to first-order gradient descent algorithms. Prior research on PET image reconstruction and denoising using the conditional DIP [76,78,112] have demonstrated that the L-BFGS algorithm offers improved computational time and convergence stability when compared to other first-order gradient descent methods, including the SGD, Nesterov’s accelerated gradient [116], and Adam [117].

We used the learning rate of 0.1. The experiments were conducted on a computer using the Ubuntu 18.04 with an NVIDIA Quadro RTX 8000 GPU with 48 GB of memory. The network models were implemented using PyTorch 1.7.1 (<https://pytorch.org/>).

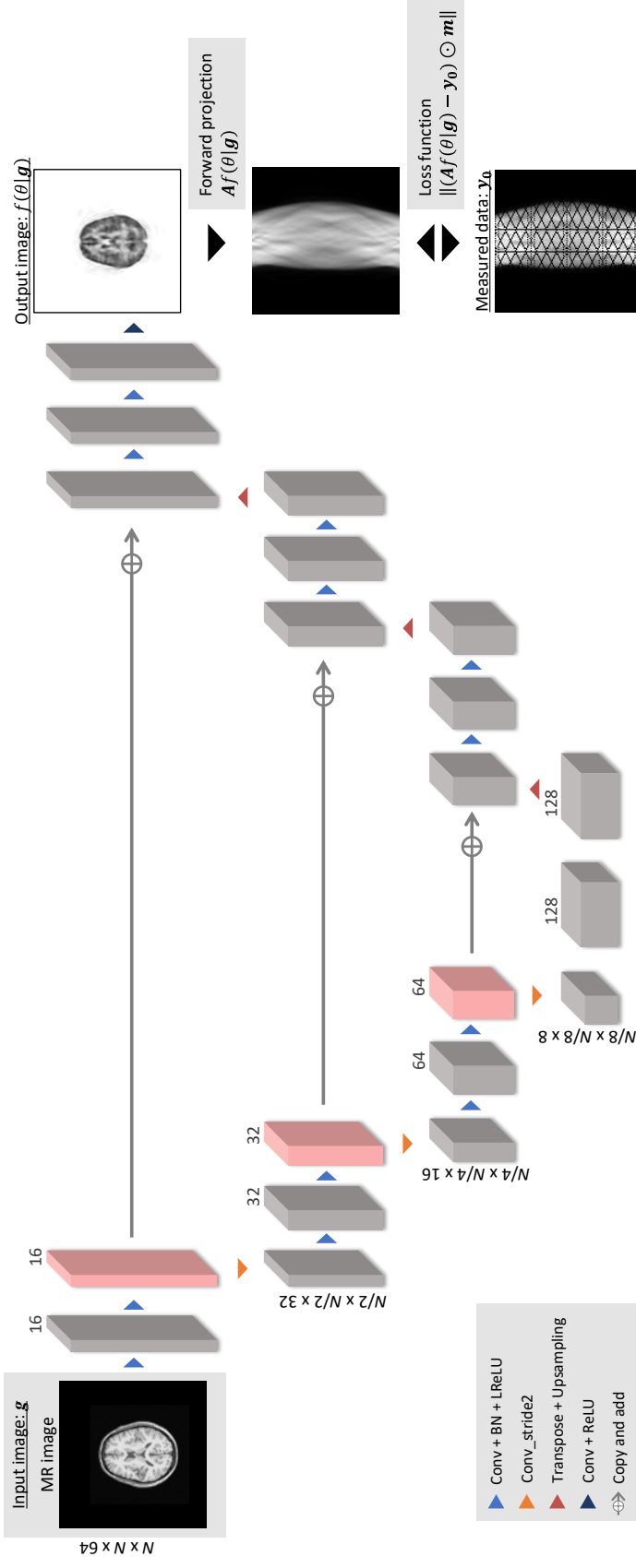


Figure 4.1: Implementation overview of the proposed end-to-end PET image reconstruction algorithm using the conditional DIP framework. The voxel size is provided on the left side of the box. First, a reconstructed PET image was calculated using the conditional DIP framework. Then, the loss calculation was performed with the measured data and the estimated projection data from the DIP output. The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE)

4.3 Experimental setup

4.3.1 Computer simulation

We utilized 20 brain phantoms from the BrainWeb database [84,85] and generated projection data through Monte Carlo simulation. The radioactivity contrast ratios for gray matter, white matter, and cerebrospinal fluid were set to 1:0.25:0.05, reflecting the typical distribution of the ^{18}F -FDG. The attenuation coefficients were assigned to 0.00958 mm^{-1} for soft tissue and 0.0151 mm^{-1} for bone, respectively.

For the experiments, we assumed a brain-dedicated PET scanner [111] as follows: It featured a detector ring with a diameter of 486.83 mm, composed of 28 detector units circumferentially and four units in the axial direction. Each detector unit housed a 16×16 array of cerium-doped lutetium–yttrium oxyorthosilicate (LYSO) crystals, with each crystal measuring $3.14 \times 3.14 \times 20 \text{ mm}$. The list-mode data of the 3D PET acquisition were converted into a 2D PET sinogram format using a single-slice rebinning method, with a maximum ring difference set to ± 15 . The scatter events were excluded from the list-mode data to simplify the simulation. The sinogram and PET image sizes were $128 \text{ angles} \times 128 \text{ bins} \times 64 \text{ slices}$, and $128 \times 128 \times 64 \text{ voxels}$ with $3.0 \times 3.0 \times 3.221 \text{ mm/voxel}$. The simulated sinogram had approximately 35.9 ± 1.59 million counts for each subject. Before the reconstruction, both component-based normalization and attenuation correction were performed on the sinogram space.

In this experiment, PSNR and SSIM were used for quantitative evaluation, as shown in Equations (3.7) and (3.8). Additionally, we computed the contrast recovery coefficient (CRC) and the background standard deviation (STD) as follows,

$$\text{CRC} = \left(\frac{\bar{a}}{\bar{b}} - 1 \right) / \left(\frac{\bar{a}_{\text{gt}}}{\bar{b}_{\text{gt}}} - 1 \right), \quad (4.5)$$

$$\text{STD} = \frac{1}{\bar{b}} \sqrt{\frac{1}{K_b} \sum_{k=1}^{K_b} (b_k - \bar{b})^2}, \quad (4.6)$$

where $\bar{a} = \frac{1}{K_a} \sum_{k=1}^{K_a} a_k$ and $\bar{b}$ represent the mean uptakes of the gray matter and white matter (background) over K_a and K_b ROIs. $\bar{a}_{\text{gt}}$ and $\bar{b}_{\text{gt}}$ represent the ground truth uptakes of the gray matter and background regions. We used 30 ROIs for gray matter and background, containing only 80% pixels of the gray matter were selected.

4.3.2 Real preclinical PET data

The real preclinical experiments were approved by the Animal Ethical Committee of the Central Research Laboratory, Hamamatsu Photonics K.K. (Approval number: HPK-2017-02).

We used the same real preclinical ^{18}F -FDG PET data as in subsection 3.3.2. The scatter correction was implemented through a convolution subtraction method, and attenuation correction was executed by reprojecting the transmission image into 3D sinogram space.

The acquired 3D PET emission data were converted into the 2D sinogram format utilizing the Fourier rebinning method. The sinogram and PET image sizes were $256 \text{ angles} \times 256 \text{ bins} \times 64 \text{ slices}$, and $256 \times 256 \times 64 \text{ voxels}$ with $0.65 \times 0.65 \times 1.0167 \text{ mm/voxel}$. A T1-weighted MR image was scanned on a different day and was manually registered to PET image by two radiological technologists.

4.3.3 Comparison algorithms

In the experiments, we compared the proposed end-to-end DIPrecon with the FBP using the Hanning filter, and MLEM with 100 iterations. These were then performed the Gaussian post-filtering of $\sigma = 1$ voxel. Furthermore, we performed the hybrid DIPrecon algorithm by Gong et al. [112] for comparison.

The hybrid DIPrecon algorithm defines the following constrained optimization problem,

$$\begin{aligned} & \max L(\mathbf{y}|\mathbf{x}) \\ \text{s.t. } & \mathbf{x} = f(\boldsymbol{\theta}|\mathbf{g})' \end{aligned} \quad (4.7)$$

where L is the Poisson log likelihood function. In the hybrid DIPrecon algorithm, Equation (4.7) is transformed into an unconstrained problem using the augmented Lagrangian format, and solved it by the alternating direction method of multipliers algorithm [118] in three steps as follows,

$$\mathbf{x}^{n+1} = \underset{\mathbf{x}}{\operatorname{argmax}} L(\mathbf{y}|\mathbf{x}) - \frac{\rho}{2} \|\mathbf{x} - f(\boldsymbol{\theta}^n|\mathbf{g}) + \boldsymbol{\mu}^n\|^2, \quad (4.8)$$

$$\boldsymbol{\theta}^{n+1} = \underset{\boldsymbol{\theta}}{\operatorname{argmin}} \|f(\boldsymbol{\theta}|\mathbf{g}) - (\mathbf{x}^{n+1} + \boldsymbol{\mu}^n)\|^2, \quad (4.9)$$

$$\boldsymbol{\mu}^{n+1} = \boldsymbol{\mu}^n + \mathbf{x}^{n+1} - f(\boldsymbol{\theta}^{n+1}|\mathbf{g}), \quad (4.10)$$

where $\boldsymbol{\mu}$ is a scaled dual variable and ρ is a positive constant. We used the same settings employed in the original paper by Gong et al. [112].

4.4 Results

4.4.1 Computer simulation

Figure 4.2 displays the three orthogonal slices of the reconstructed images and their magnified images by different reconstruction algorithms in the computer simulation. The proposed end-to-end DIPrecon algorithm using random noise input and MRI input obtained accurate PET images with preserving brain structures compared to both the FBP and MLEM algorithms. Notably, the proposed end-to-end DIPrecon algorithm with the MRI input generated more fine cortex structures than the other reconstruction algorithms. In addition, the proposed end-to-end DIPrecon algorithm provided finer edges compared to the hybrid DIPrecon algorithm.

Figure 4.3 illustrates the box plots of the PSNR and SSIM for the various reconstruction algorithms. The mean PSNR values for the FBP, MLEM, hybrid DIPrecon, and the proposed end-to-end DIPrecon with random noise and MRI input were 14.16, 14.18, 16.31, 15.40, and 16.31 dB, respectively, and the mean SSIM values were 0.588, 0.611, 0.786, 0.685, and 0.761, respectively. The PSNRs and SSIMs of the proposed end-to-end DIPrecon algorithm with MRI input were much higher than the FBP and MLEM algorithms. Note that the proposed end-to-end DIP recon algorithm also proved comparable to the hybrid DIPrecon algorithm.

Figure 4.4 illustrates the tradeoffs between the CRC and STD for the gray matter region by different reconstruction algorithms. The tradeoff curves demonstrated that the proposed end-to-end DIPrecon algorithm achieves the highest CRC and the hybrid DIPrecon algorithm provides better noise performance in STD. These quantitative results indicated that the proposed end-to-end DIPrecon algorithm provides better or compatible

performance compared with the FBP, MLEM, and hybrid DIPrecon algorithms.

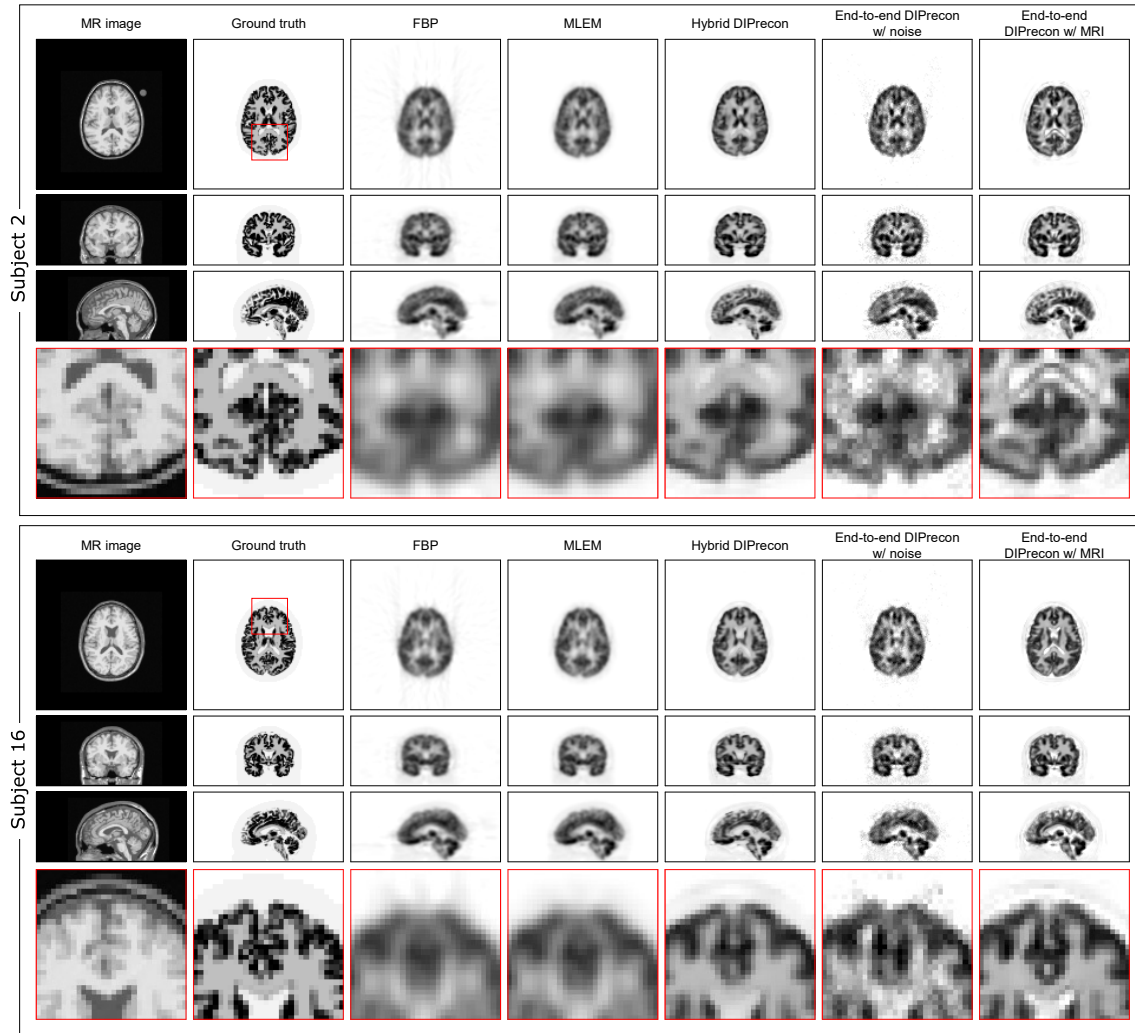


Figure 4.2: Three orthogonal slices of the reconstructed images and their magnified images of the red squared regions by different reconstruction algorithms in the computer simulation. The columns correspond to the MR images, ground truth, reconstructed images obtained using FBP, and MLEM, hybrid DIPrecon, and the proposed end-to-end DIPrecon with random noise and MRI input (left-to-right). The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE)

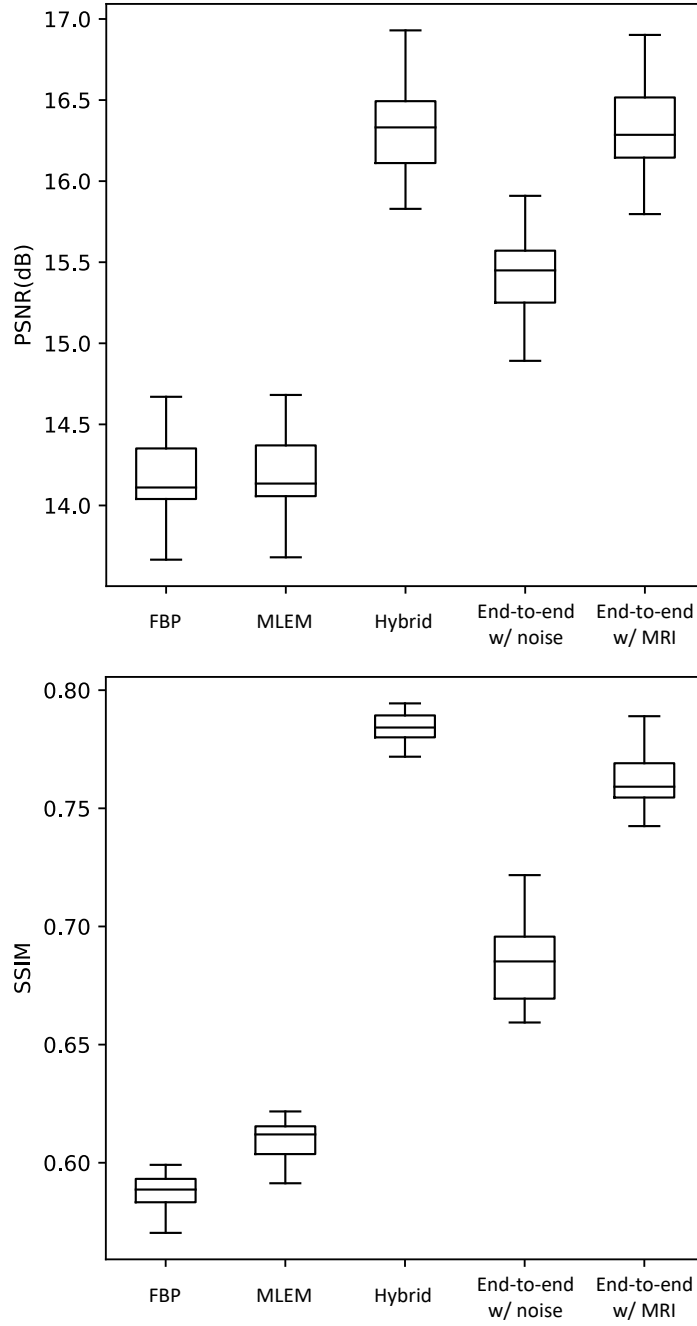


Figure 4.3: Quantitative results of the PSNR (top) and SSIM (bottom) for the computer simulation with different reconstruction algorithms: the MLEM, hybrid DIPrecon (Hybrid), end-to-end DIPrecon with random noise input (End-to-end w/ noise) and with MRI input (End-to-end w/ MRI). The line within the box represents the median value, and the upper and lower lines of the box represent the 75th and the 25th percentiles, respectively. The upper and lower whiskers represent the maximum and minimum values, respectively. The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE)

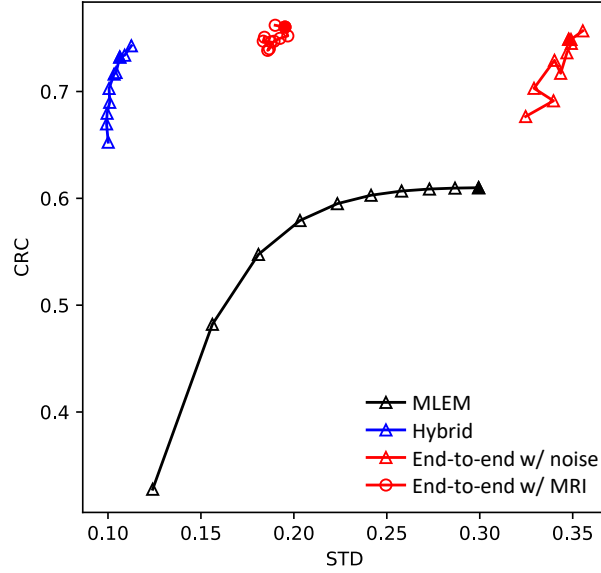


Figure 4.4: Tradeoffs between the CRC and STD at the gray matter region for the computer simulation with different reconstruction algorithms: the MLEM, hybrid DIPrecon (Hybrid), end-to-end DIPrecon with random noise input (End-to-end w/ noise) and with MRI input (End-to-end w/ MRI). Markers are plotted every ten iterations from ten to 100 in the MLEM, every iteration from one to ten in the hybrid DIPrecon, every three epochs from 18 to 45 in the end-to-end DIPrecon with the random noise input, and every epoch from 33 to 42 in the end-to-end DIPrecon with the MRI input. The reconstructed images in Figure 4.2 are labeled by filled makers. The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE)

4.4.2 Real preclinical PET data

Figure 4.5 displays the three orthogonal slices of the reconstructed images and their magnified images by different reconstruction algorithms in the real preclinical ^{18}F -FDG PET data of the conscious rhesus monkey brain. The proposed end-to-end DIPrecon algorithm provides high-quality PET images compared with both the FBP and MLEM algorithms.

Figure 4.6 illustrates the tradeoffs between the striatum uptake and STD by different reconstruction algorithms. The proposed end-to-end DIPrecon algorithm showed both the highest striatum uptake and lowest STD in the other reconstruction algorithms.

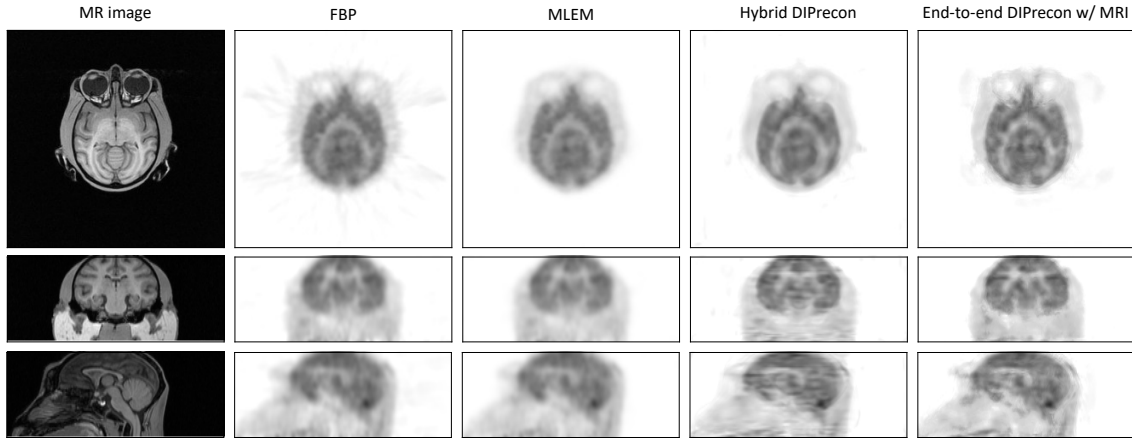


Figure 4.5: Three orthogonal slices of the reconstructed images by different reconstruction algorithms in the computer simulation. The columns correspond to the MR images, reconstructed images obtained using FBP, MLEM, hybrid DIPrecon, and the end-to-end DIPrecon with MRI input (left-to-right). The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE)

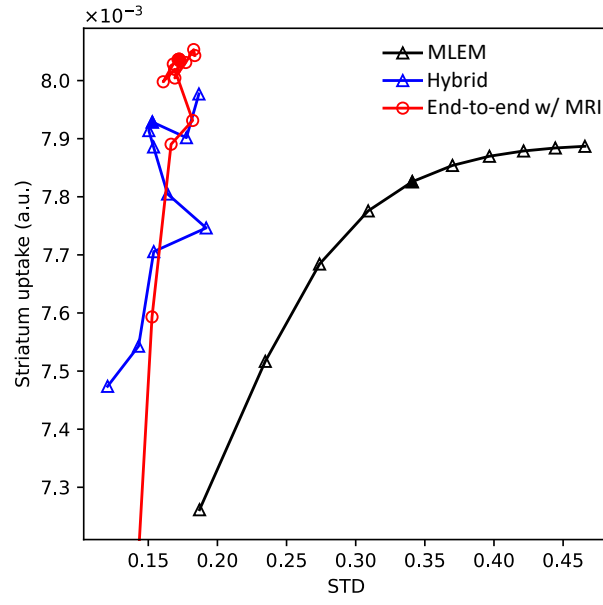


Figure 4.6: Tradeoffs between the striatum uptake and STD for real preclinical data with different reconstruction algorithms: the MLEM, hybrid DIPrecon (Hybrid), and end-to-end DIPrecon with MRI input (End-to-end w/ MRI). Markers are plotted every ten iterations from ten to 100 in MLEM, every iteration from one to ten in the hybrid DIPrecon, every three epochs from 9 to 18, and every epoch from 21 to 30 in the end-to-end DIPrecon with MRI input. The corresponding reconstructed images in Figure 4.5 are labeled by filled markers. The figure is reprinted from the work of Hashimoto et al. [113] (© 2022 IEEE)

4.5 Discussion

Similar to PET image denoising, various supervised deep learning algorithms have been developed for iterative PET image reconstruction [119-122]. For example, Gong et al. proposed a deep learning-based iterative PET image reconstruction algorithm incorporating a supervised CNN for image representation [119]. Xie et al. expanded the network in the above-mentioned work to generative adversarial networks [120]. Mehranian and Reader proposed MAP reconstruction to integrate residual networks by using a forward-backward splitting algorithm [121]. However, as mentioned in Chapter 3, there are still challenges in applying supervised deep learning reconstruction to unknown cases not covered in the training datasets, such as unfamiliar diseases, novel PET tracers, and different human races. The DIPrecon algorithms have the potential to open the doors to apply clinical use without any prior training datasets due to the synergy of the conditional DIP and iterative image reconstruction frameworks.

Our proposed end-to-end DIPrecon algorithm distinguishes itself from other deep learning-based PET reconstruction algorithms. Typically, supervised deep learning reconstruction requires a substantial number of reconstructed PET images for network training. In contrast, the DIPrecon requires only the measured data, attenuation correction information and additional prior information such as MR images. This can reduce the inductive biases in training datasets, such as artifacts introduced during the existing image reconstruction process.

The computation time for the proposed end-to-end DIPrecon and hybrid DIPrecon algorithms were 5.74 seconds and 226.76 seconds per iteration, which demonstrated that the proposed end-to-end DIPrecon approximately 8.8 times faster than the hybrid

DIPrecon in obtaining well-converged reconstructed images. The hybrid DIPrecon algorithm requires three calculation steps as described in (4.8) to (4.10): solving two subproblems iteratively and updating the Lagrangian multipliers. On the other hand, the proposed end-to-end DIPrecon algorithm transforms the constrained optimization problem described in (4.2) into the unconstrained optimization problem described in (4.3), and solves it without any penalty parameters in a single step using the neural network optimization. Thus, the proposed end-to-end DIPrecon is a more practical and faster algorithm for iterative PET image reconstruction without any complicated penalty parameter adjustments, unlike the hybrid DIP recon algorithm, which requires parameter tuning for the ADMM algorithm.

The proposed end-to-end DIPrecon algorithm has a drawback with respect to learning stability. An unstable behavior was observed during initial iterations because the proposed DIPrecon algorithm optimized from randomly initialized network parameters. To mitigate this behavior, we can use pre-trained network, the same as the implementation of the hybrid DIPrecon algorithm.

The limitation of this study is that the experiments were only on the simulated brain PET data with ^{18}F -FDG contrast and real preclinical brain PET data from a conscious rhesus monkey with ^{18}F -FDG. In the future, we plan to evaluate human clinical data, different PET scanners, PET tracers and other organs. The effects of the proposed DIPrecon algorithm in low-dose PET imaging should also be investigated in the future. In addition, we will investigate the impact of the proposed end-to-end DIPrecon algorithm on mismatched data between PET and MRI, such as lesions present on the MR images and not on the PET images.

In the experiments, we prepared 2D sinograms using the single-slice rebinning

method for the computer simulation and the Fourier rebinning method for the real preclinical data. However, most recent PET scanners do not use these methods because they inherently lack information on the measurement data, resulting in blurred images. In order to utilize all the information in the measurement data, it is necessary to handle fully 3D PET data. Therefore, we will introduce a practical algorithm for expanding the end-to-end DIPrecon to fully 3D PET data in the next chapter.

5 Expanding DIPrecon to 3D PET

5.1 Introduction

In the Chapter 4, we introduced the iterative PET image reconstruction algorithm incorporating the conditional DIP framework, called the end-to-end DIPrecon. Although the end-to-end DIPrecon provided accurate PET images while measuring consistency with measured emission data, it was limited to 2D sinograms.

The lack of GPU memory is the primary factor that restricts the expanding the end-to-end DIPrecon to fully 3D PET implementation. In this chapter, to address the challenge, we introduce the first attempt to implement an end-to-end DIPrecon algorithm for fully 3D PET data [123]. We modify the end-to-end DIPrecon algorithm to a block iteration and sequential learning of an ordered sequence of block sinograms. In addition, we implement the relative difference penalty (RDP) function [124] to the loss function of the network to further improve the quantitative accuracy of the reconstructed image. We evaluate the proposed 3D end-to-end DIPrecon algorithm with conventional iterative reconstruction algorithms using Monte Carlo simulation data and real preclinical data.

5.2 Methods

5.2.1 Block iterative algorithm for fully 3D PET data

The end-to-end DIPrecon can reconstruct PET images by minimizing the constrained optimization problem shown in (4.2). For 2D PET data, (4.2) can be solved by the unconstrained problem shown in (4.3) in one step. However, for the fully 3D PET data, the optimization in (4.3) cannot be computed on current GPU processors due to its huge data volume: the data includes the 3D projection data, network parameters, and network input such as MRI data, as well as calculation histories of the 3D forward projection for keeping gradients to optimize the network parameters through the back-propagation algorithm. In particular, it needs to allocate huge memory space for storing the calculation histories of 3D forward projection.

Inspired by a strategy of the OSEM algorithm which sequentially updates an ordered sequence of block sinograms, we apply block iteration-based optimization to (4.3) to reduce memory usage of calculation histories of the 3D forward projection for each optimization, as follows,

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \sum_{d=1}^D \|(\mathbf{A}^d f(\theta|\mathbf{g}) - \mathbf{y}_0^d) \odot \mathbf{m}^d\|, \quad (5.1)$$

where D and d represent the number of ordered subsets of the sinogram and the index of the subsets, respectively. $\mathbf{A}$ is a system matrix that denotes the probability of each voxel to each LOR, and f is a neural network with trainable parameters θ . $\mathbf{g}$ represents prior information as the network input. $\mathbf{y}_0$ is the measured projection data,

$\odot$ represents the Hadamard product, and $\mathbf{m}$ is the binary mask of the detector gaps. We sampled each subset to equally spaced angles to mitigate geometric correlations between subsets. In this implementation, the system matrix $\mathbf{A}$ stores as a sparse matrix in a coordinate list format to reduce memory usage.

This modification in (5.1) can realize a practical implementation of a fully 3D PET image reconstruction. However, it should be noted that the optimization in (4.3) and (5.1) cannot strictly converge to the same solutions. This situation is similar to the relationship between the MLEM and OSEM algorithms: the OSEM algorithm does not converge to the true ML solution and sometimes falls into the limit cycle problem [125,126].

5.2.2 Penalization for DIP optimization

As described in the Section 2.1.2, the MAP image reconstruction that integrates image priors has been used to achieve better image noise and contrast characteristics. In practical MAP reconstruction on commercial PET scanners, the RDP [127] has been introduced for the image prior [128], as follows,

$$RDP(\mathbf{x}) = \sum_j \sum_{k \in N_j} \frac{(x_j - x_k)^2}{(x_j + x_k) + \gamma |x_j - x_k|}, \quad (5.2)$$

where N_j represents a set of neighboring voxels for the j -th voxel, and γ is the shape of the RDP function. The RDP calculates the differences between neighboring voxels in the image domain. The RDP applies more smoothing in (relatively) low-activity regions and less smoothing in high-activity regions in the image domain. The MAP reconstruction with the RDP serves as a better penalty for PET imaging, achieving better contrast recovery and reduced background noise compared to the OSEM algorithm [129].

Inspired by MAP reconstruction, we introduce the RDP function into the loss function as a penalization of the end-to-end DIPrecon to improve the PET image quality and mitigate the overfitting problem in the DIP optimization, as follows,

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \sum_{d=1}^D \|(\mathbf{A}^d f(\theta|\mathbf{g}) - \mathbf{y}_0^d) \odot \mathbf{m}^d\| + \beta RDP(f(\theta|\mathbf{g})), \quad (5.3)$$

where β is the hyperparameter of the regularization.

5.2.3 Implementation

Figure 5.1 illustrates an implementation overview of the proposed fully 3D end-to-end DIPrecon algorithm. For the experiments, we used the same 3D U-Net architecture that was presented in Figure 4.1.

The proposed fully 3D end-to-end DIPrecon algorithm performs the following steps. (1) The prior information of the MR image is input into the network. (2) The PET image is obtained from the network output. (3) The blurred output image is obtained through a fixed convolution layer which calculates a shift-invariant blurring model of the PET scanner in the image-space domain. Please note that the fixed convolution layer is inspired by PSF reconstruction to improve the PET image quality. (4) The forward projection is sequentially calculated by the dot product of the blurred output image and the block sparse system matrix at each iteration to obtain each estimated block sinogram. (5) The loss is calculated with each measured block sinogram and estimated block sinogram. (5) The network parameters are sequentially updated in a mini-batch optimization manner through the back-propagation algorithm. We used the Siddon algorithm to implement 3D forward projection operation [130], and the Gaussian kernel with $\sigma = 0.5$ voxels was used for the kernel of the fixed convolution layer.

In the experiments, we used the stochastic L-BFGS algorithm [131] that extended to a mini-batch optimization. The experiments were conducted on a computer using the Ubuntu 20.04 with an NVIDIA A100 GPU with 80 GB of memory. The network models were implemented using PyTorch 1.12.1 (<https://pytorch.org/>).

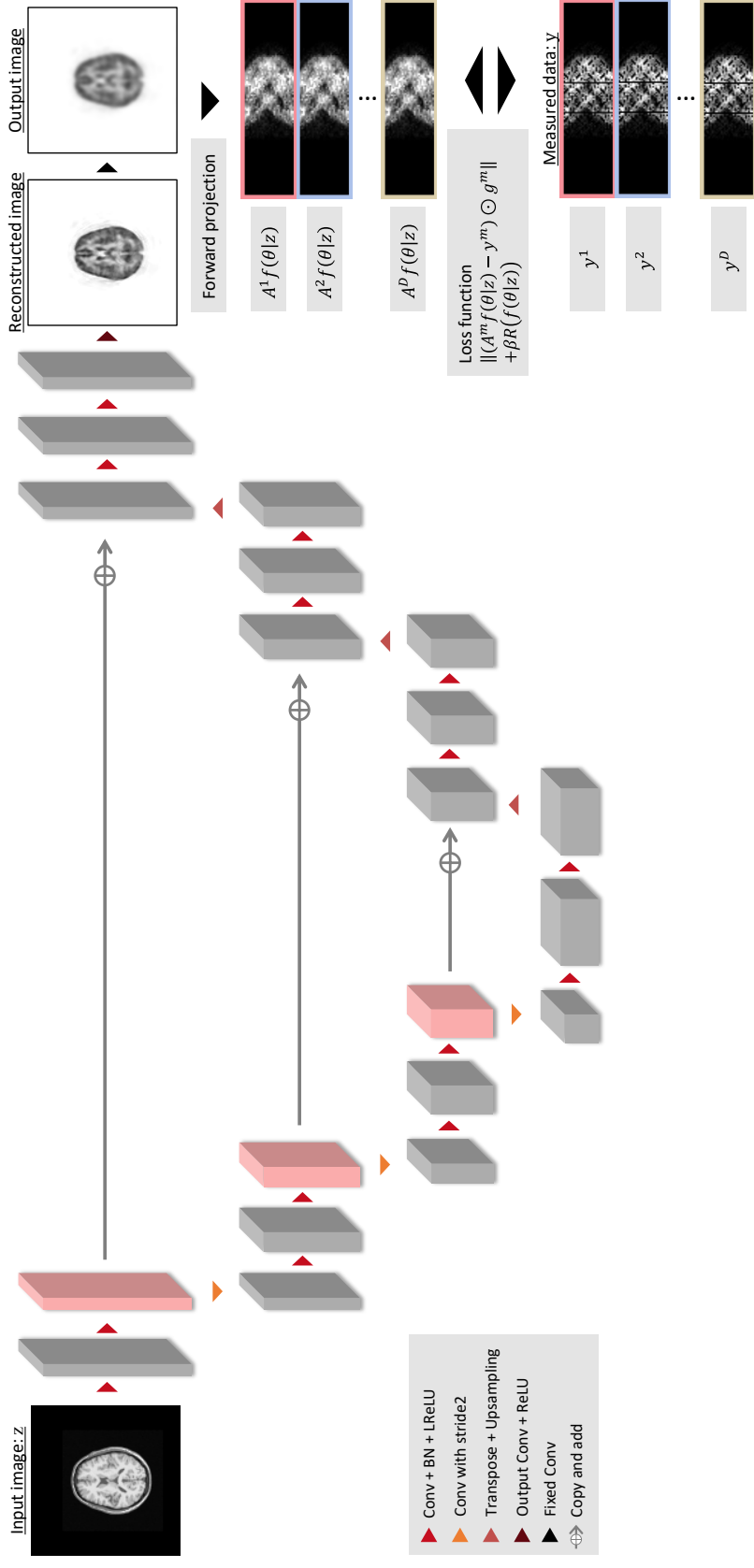


Figure 5.1: Implementation overview of the proposed fully 3D end-to-end PET image reconstruction using the conditional DIP framework. First, a PET image was calculated using the conditional DIP framework. Then, loss calculation was performed with each measured block sinogram and the estimated block sinogram from the DIP output. Finally, the network parameters were sequentially updated in a mini-batch optimization manner. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPENM)

5.3 Experimental setup

5.3.1 Computer simulation

We utilized 20 brain phantoms from the BrainWeb database [84,85] and generated projection data through Monte Carlo simulation. The radioactivity contrast ratios for gray matter, white matter, and cerebrospinal fluid were set to 1:0.25:0.05, reflecting the typical distribution of the ^{18}F -FDG. The attenuation coefficients were assigned to 0.00958 mm^{-1} for soft tissue and 0.0151 mm^{-1} for bone, respectively. In addition, three spherical tumors of 21, 15, and 12 mm diameters with contrasts of 1.1, 1.2, and 1.5 were inserted into white matter area, which cannot be identified on the MR image.

For the experiment, we assumed the same brain-dedicated PET scanner [111] as described in the Subsection 4.3.1. We excluded a positron range, angular deviation, and random events in the experiment. The scatter events were also excluded from the list-mode data to simplify the simulation. The 3D PET sinogram was created from the simulated list-mode data using nearest neighbor interpolation with the ring difference binned with a span of seven.

The sinogram and PET image sizes were $128 \text{ angles} \times 128 \text{ bins} \times 64 \text{ slices} \times 19 \text{ oblique angles}$, and $128 \times 128 \times 64 \text{ voxels}$ with $3.0 \times 3.0 \times 3.221 \text{ mm/voxel}$. The simulated sinogram had 2,921,540 counts. Before the reconstruction, both component-based normalization and attenuation correction were performed on the sinogram space.

For the quantitative evaluation, the PSNR and SSIM were used as shown in Equations (3.7) and (3.8). Additionally, we computed the CRC in (4.5) and the background STD in (4.6).

5.3.2 Real preclinical PET data

The real preclinical experiments were approved by the Animal Ethical Committee of the Central Research Laboratory, Hamamatsu Photonics K.K. (Approval number: HPK-2017-02).

We used the same real preclinical ^{18}F -FDG PET data as in the Subsections 3.3.2 and 4.3.2. The scatter correction was implemented through a convolution subtraction method, and attenuation correction was executed by reprojecting the transmission image into 3D sinogram space. Random correction was performed by subtracting delayed coincidence.

The low-dose PET data were simulated by periodic 1/20 downsampling of the measured list-mode emission data. The sinogram and PET image sizes were 128 angles $\times$ 128 bins $\times$ 64 slices $\times$ 19 oblique angles, and 128 $\times$ 128 $\times$ 64 voxels with 0.65 $\times$ 0.65 $\times$ 1.0167 mm/voxel. A T1-weighted MR image was scanned on a different day and was manually registered to PET image by two radiological technologists, as in the Subsection 4.3.2.

5.3.3 Comparison algorithms

In the experiments, we compared the proposed fully 3D end-to-end DIPrecon algorithm with the MLEM, MAPEM with the RDP function [124], and hybrid DIPrecon algorithms [112]. It should be noted that the image-space PSF reconstruction were used in these comparison algorithms for fair comparison.

We used 100 iterations for the MLEM and MAPEM algorithms, and 200 main iterations with two sub-iterations for the EM reconstruction and ten sub-iterations for the DIP optimization were used for the hybrid DIPrecon algorithm, which are the same settings for the hybrid DIPrecon algorithm as employed in the original paper [112].

5.4 Results

5.4.1 Computer simulation

Figure 5.2 displays the reconstructed results for the simulated data using different regularization parameters, and Figure 5.3 demonstrates the impact of the regularization parameter on the proposed fully 3D DIPrecon algorithm with two subsets. We achieved more detailed brain structures and more stable optimization with the regularization parameter of 5×10^{-9} . Therefore, we selected this value in the experiments.

Figure 5.4 presents the reconstructed results for the simulated data using different algorithms. The simulation results indicated that the proposed fully 3D end-to-end DIPrecon with the RDP enhanced the PET image quality by reducing the statistical noise and maintaining the brain structures and the inserted tumor contrast. The visual results are supported by the highest PSNR achieved with the proposed fully 3D end-to-end DIPrecon algorithm.

Figure 5.5 illustrates the tradeoffs between the tumor CRC and STD for different reconstruction algorithms. The trade-off curves are averaged over ten independent and identically distributed samples from the same digital brain phantom. The tradeoff curves demonstrate that the proposed end-to-end DIPrecon algorithm achieves competitive results with the hybrid DIPrecon algorithm in the CRC and the STD.

Figure 5.6 illustrates the influence of the RDP term on the PSNR for the proposed end-to-end DIPrecon algorithm. The RDP term into the loss function of the proposed algorithm significantly enhanced the reconstruction performance compared to that of the algorithm without the RDP. Additionally, the RDP term did not compromise the quantitative performance, even with an increased number of iterations.

We assessed the uncertainty and variability of the proposed DIPrecon algorithm by using ten independent and identically distributed samples from the same digital brain phantom. Figure 5.7 presents the voxel-wise mean and standard deviation images. We found that the variability of the proposed DIPrecon algorithm was less than or equal to that of the other algorithms. Figure 5.8 shows the PSNR results of ten independent and identically distributed simulation samples, which achieved the highest values for the proposed end-to-end DIPrecon algorithm with RDP.

Figure 5.9 represents the reconstructed results of the simulation data using different numbers of subsets, and Figure 5.10 demonstrates the impact of the numbers of subsets on the PSNR. It is important to note that the end-to-end DIPrecon algorithm is not feasible even on current best GPU boards such as NVIDIA A100 with 80GB memory, unless the number of subsets is greater than two due to the limitations of GPU memory. Similar curves were observed across these subsets, suggesting that the proposed block iteration algorithm effectively yielded high-quality images without any training dataset.

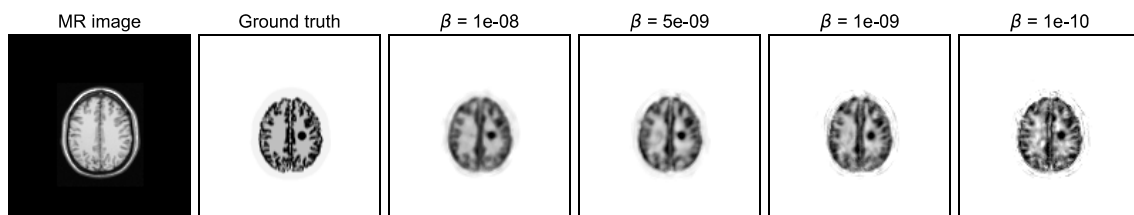


Figure 5.2: Reconstructed results of the human brain ^{18}F -FDG computer simulation with different regularization parameters β with two subsets. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPEM)

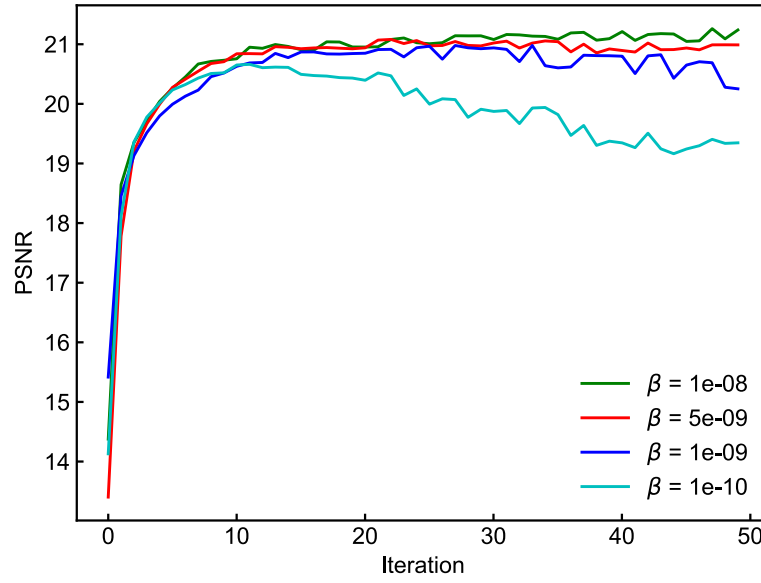


Figure 5.3: Impact of the settings of the regularization parameter β setting with two subsets, in the PSNR for the proposed fully 3D end-to-end DIPrecon algorithm. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPeM)

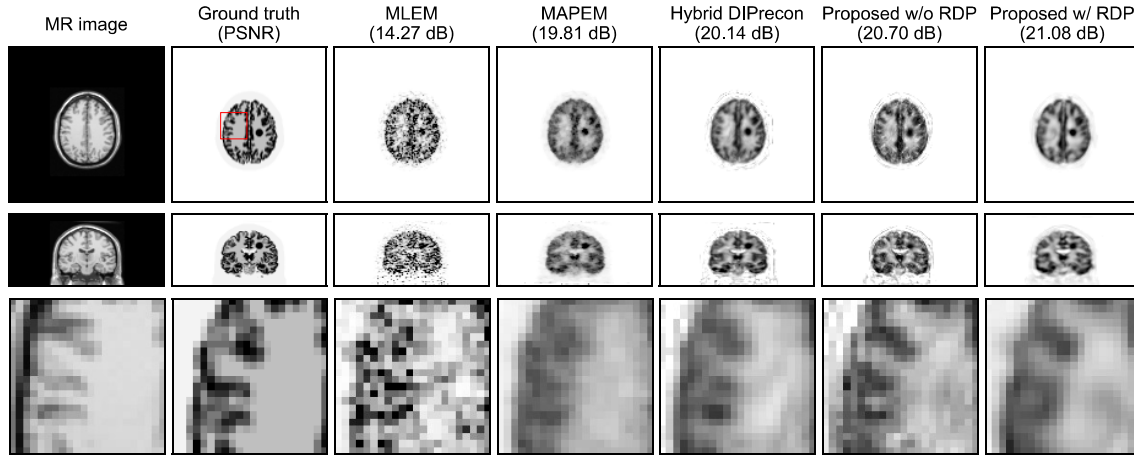


Figure 5.4: Transaxial and coronal slices of the reconstructed images and their magnified images of the red squared regions by different reconstruction algorithms in the computer simulation. The columns correspond to the MR images, ground truth, reconstructed images obtained using the MLEM, MAPEM, hybrid DIPrecon, and the proposed end-to-end DIPrecon without RDP (Proposed w/o RDP) and with RDP (Proposed w/ RDP) (left-to-right). The PSNR values for each algorithm are provided below the name of the algorithm. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPeM)

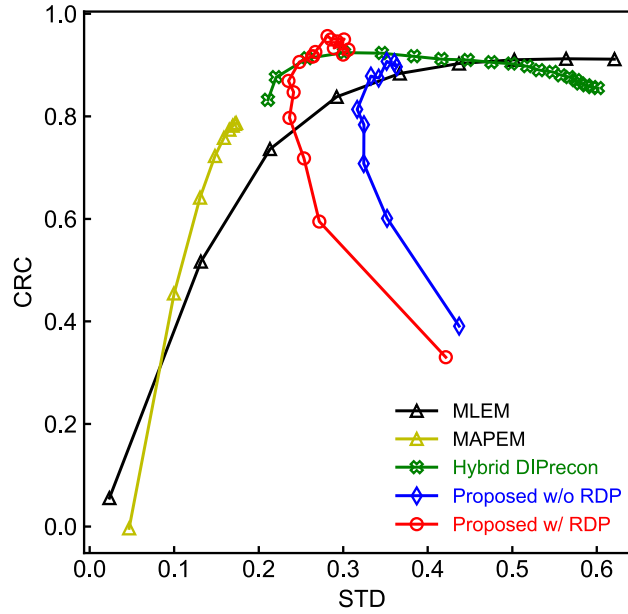


Figure 5.5: Mean tradeoffs between the tumor CRC and STD for the human brain ^{18}F -FDG computer simulation. Markers are plotted every ten iterations from one to 100 in the MLEM and MAPEM algorithms, every ten iterations from one to 200 in the hybrid DIPrecon algorithm, every two epochs from 1 to 20 in the proposed algorithm without RDP (Proposed w/o RDP), and every three epochs from 1 to 50 in the proposed algorithm with RDP (Proposed w/ RDP). The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM)

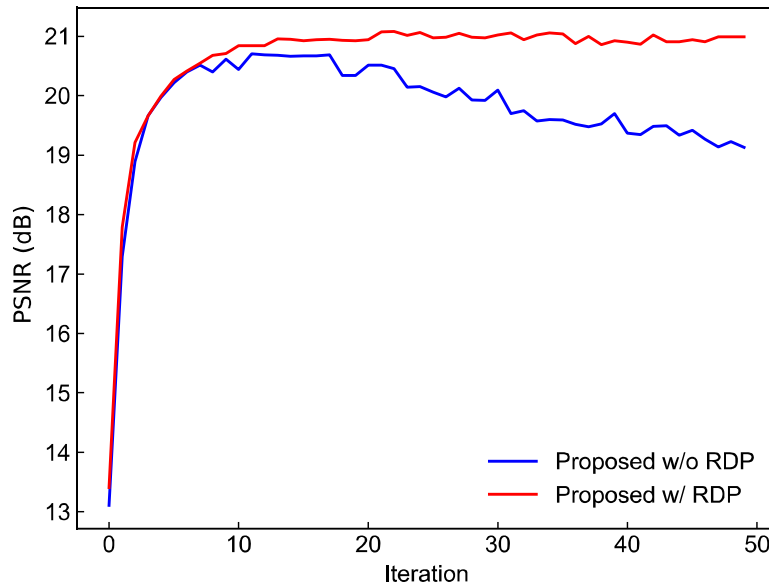


Figure 5.6: Impact of the RDP term on the PSNR for the proposed fully 3D end-to-end DIPrecon algorithm. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM)

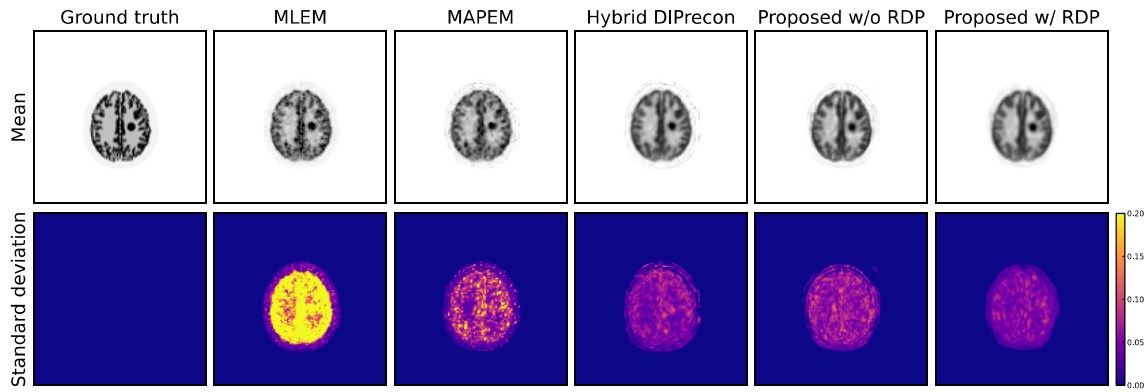


Figure 5.7: Mean and standard deviation PET images of ten independent and identically distributed samples from the same digital brain phantom for different reconstruction algorithms. From left to right, the MAPEM with the RDP, hybrid DIPrecon, proposed end-to-end DIPrecon without RDP (Proposed w/o RDP), and proposed end-to-end DIPrecon with RDP (Proposed w/RDP). The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPeM)

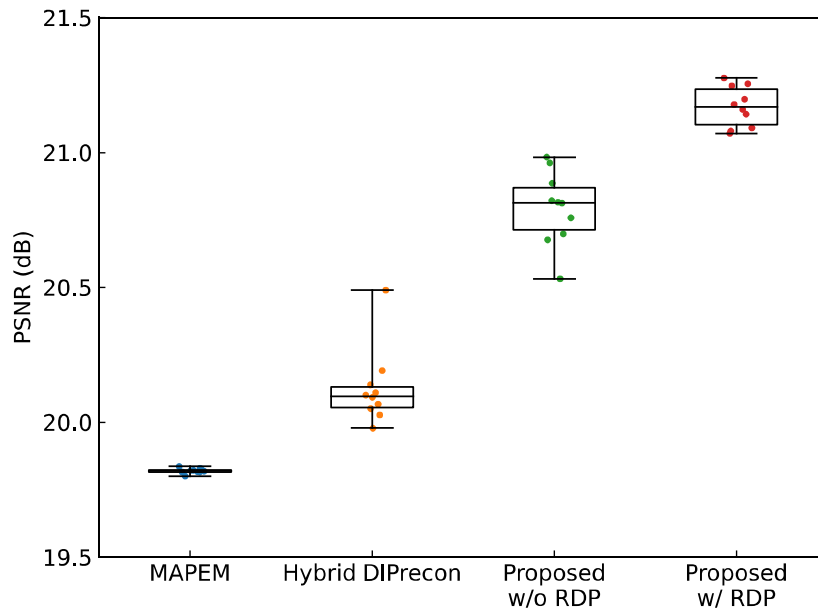


Figure 5.8: Quantitative results of ten independent and identically distributed samples from the same digital brain phantom for different reconstruction algorithms. From left to right, the MAPEM with the RDP, hybrid DIPrecon, proposed end-to-end DIPrecon without RDP (Proposed w/o RDP), and proposed end-to-end DIPrecon with RDP (Proposed w/RDP). The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPeM)

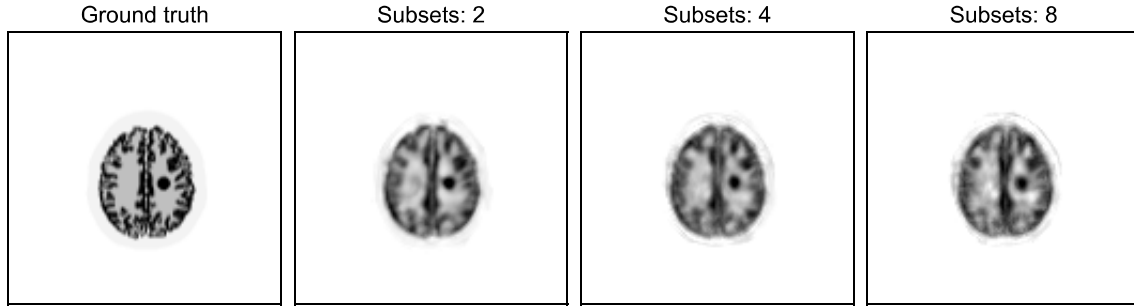


Figure 5.9: Reconstructed results of the human brain computer simulation with ^{18}F -FDG contrast for different number of subsets. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM)

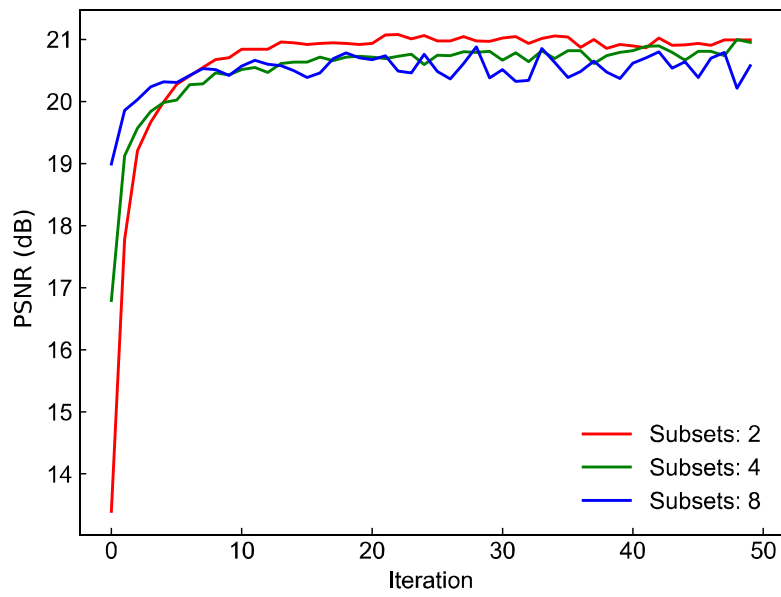


Figure 5.10: Impact of the number of subsets in terms of the PSNR for the proposed fully 3D end-to-end DIPrecon algorithm. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM)

5.4.2 Real preclinical PET data

Figure 5.11 represents the reconstructed results with subsets 2 for the low-dose real preclinical PET data using different algorithms. The proposed end-to-end DIPrecon algorithm with the RDP achieved the highest PSNR and offered more fine representations of putamen structures in low-dose PET imaging, compared with the other reconstruction algorithms.

Figure 5.12 displays the line profiles through the putamen regions using different algorithms, with the best recovery of the putamen uptake being observed for the proposed algorithm with the RDP. These results suggested that the proposed fully 3D end-to-end DIPrecon algorithm offers more precise quantitative reconstruction for low-dose PET imaging.

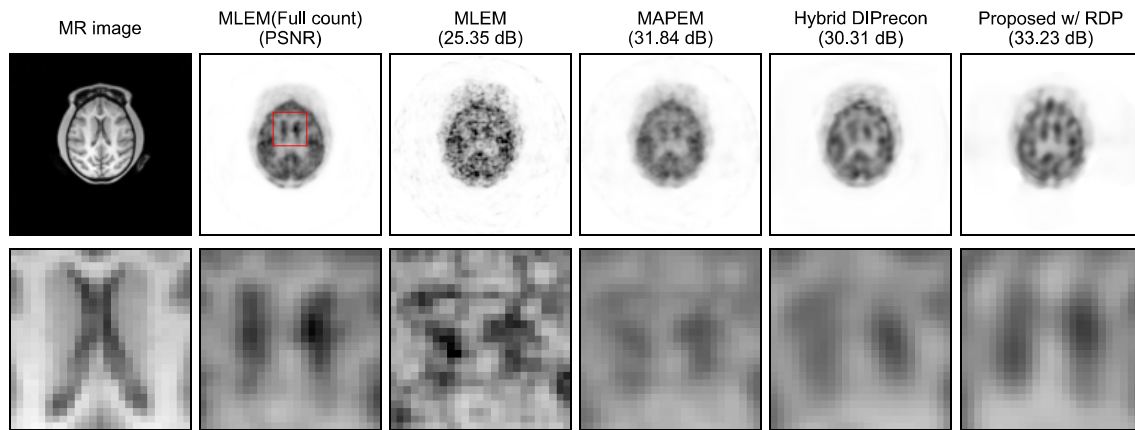


Figure 5.11: Reconstruction results of the reconstructed images and their magnified images of the red squared regions by different reconstruction algorithms for the real preclinical low-dose ^{18}F -FDG PET data. The columns correspond to the MR images, reconstructed images obtained using the MLEM with full count data and low-dose data, MAPEM, hybrid DIPrecon, and the proposed end-to-end DIPrecon with RDP (Proposed w/ RDP) (left-to-right). The PSNR values for each algorithm are provided below the name of the algorithm. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPeM)

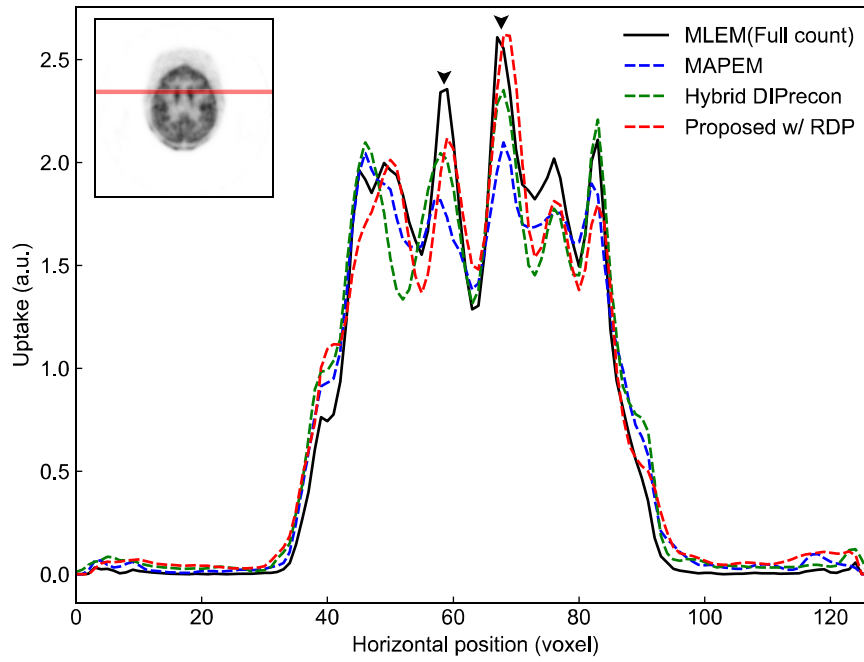


Figure 5.12: Line profiles for the real preclinical low-dose ^{18}F -FDG PET data by different reconstruction algorithms. The profile line is represented in the upper left. Triangular arrows indicate the putamen areas. The figure is reprinted from the work of Hashimoto et al. [123] (© 2023 IPPEM)

5.5 Discussion

In this Chapter, we introduced a practical implementation of a fully 3D PET image reconstruction using the DIP framework in an end-to-end manner. The proposed fully 3D end-to-end DIPrecon algorithm incorporated a 3D forward projection model into a loss function, and we modified the loss function to block iteration and sequential learning of an ordered sequence of block sinograms. Additionally, we integrated the RDP term in the loss function to improve the PET image quality and suppress the overfitting problem of the DIP optimization.

As shown in Figure 5.2 and Figure 5.3, the hyperparameter of the regularization β significantly influenced the smoothness of the reconstructed PET images. The results showed that β is an important hyperparameter that can adjust the contrast and noise characteristics. It should be noted that the scaling of β is substantially different from that used in the MAPEM [127], which is the original work in the RDP function. This is because the MSE was employed as the loss function in the proposed algorithm, whereas the MAPEM utilizes the negative log-likelihood as the objective function.

The proposed end-to-end DIPrecon algorithm demonstrated comparable performance to the hybrid DIPrecon algorithm, which is the same conditions as the proposed algorithm, as shown in Figure 5.5: it is consistent with the results for 2D PET data in Chapter 4. The proposed end-to-end DIPrecon algorithm with the RDP achieved superior PET image quality by suppressing image noise while preserving the brain structures and inserted tumors than the hybrid DIPrecon algorithm. It would be interesting to investigate the impacts of the different contrast PET tracers and tumors in the future, as the results were based on the largest inserted tumor.

The RDP term significantly enhanced the PET image quality, resulting in a stable optimization, as shown in Figure 5.6, where the graph exhibits a plateau at the highest PSNRs. The results showed that incorporating the RDP term into the loss function suppresses the overfitting problem of the DIP optimization and can eliminate the early stopping in the DIP framework. Therefore, strict monitoring of the optimization process is not required for the proposed end-to-end DIPrecon algorithm with the RDP.

The variability of the proposed algorithm with RDP is lower than or comparable to that of other reconstruction algorithms, as shown in the standard deviation images in Figure 5.7. In addition, neither the Hybrid DIPrecon algorithm nor the proposed end-to-end DIPrecon algorithm increased the variability around the inserted tumors. This result means that the proposed algorithm is stable in processing even when tumors cannot be identified on the MR image. The mean image of the proposed DIPrecon algorithm in Figure 5.7 obtained smooth white matter areas, suggesting that the heterogeneity in these areas in Figure 5.4 could be attributed to statistical variations.

In the real preclinical experiment, the proposed end-to-end DIPrecon algorithm with the RDP achieved superior performance compared to the other reconstruction algorithms in the PSNR. Additionally, the proposed DIPrecon algorithm successfully restored the putamen uptake and structures even when the 1/20 low-count PET data were used. In contrast, the hybrid DIPrecon algorithm did not restore the uptake and structures. These preclinical results demonstrated that the proposed DIPrecon algorithm has the potential to be used not only for the simulation data but also for real low-dose PET imaging. However, there were regions of insufficient recovery at tissue boundaries with the proposed DIPrecon algorithm, which is a future issue.

A block iterative algorithm is critical to practically implementing our fully 3D end-

to-end DIPrecon algorithm. The original DIP framework generally does not use a mini-batch training manner because it optimizes for a single target image [33,34]. On the other hand, we introduced a block iterative algorithm that implemented an ordered sequence of block sinograms, such as the OSEM algorithm, in a mini-batch training manner. The end-to-end fully 3D DIPrecon implementation does not even compute on current best GPU boards without the block iteration algorithm due to the limitations of the GPU memory. The proposed DIPrecon algorithm can reduce the GPU memory usage by increasing the number of subsets. For instance, the histories of the 3D forward projection, which consumed most of the GPU memory, can be reduced by a factor of 4 with 4 subsets. The experimental results shown in Figure 5.10 demonstrated that the number of subsets does not significantly affect the image quality in PSNR, suggesting that the proposed end-to-end DIPrecon algorithm is a practical and straightforward solution even for middle-range GPU boards.

In general, the optimization of the deep learning tends to converge on slightly different solutions depending on the size of the mini-batch [132]. For example, relatively small batches tend to converge to a flat minimum, while large batches tend to converge to a sharp minimum. Similar phenomena may occur with the proposed DIPrecon algorithm because the proposed algorithm implements block iteration in a mini-batch optimization manner. This suggests that the number of subsets influences the network convergence of the proposed DIPrecon algorithm, leading to variations in network parameters and PET images across different subsets. However, the impact of these phenomena on the PET image quality and characteristics remains unclear.

The proposed end-to-end DIPrecon algorithm computes the trainable parameters of the neural network in a one-step optimization using only deep learning frameworks. This

implementation differs from the hybrid DIPrecon algorithm, which optimizes two separate sub-problems, including the EM reconstruction and DIP optimization processes. Consequently, the proposed end-to-end DIPrecon algorithm is free from the settings of multiple hyperparameters: it requires only three hyperparameters of the number of iterations and subsets, as well as the learning rate. On the other hand, the hybrid DIPrecon algorithm requires five hyperparameters of the number of main iterations, two sets of sub-iterations, the regularization parameter, and the learning rate. In addition, the PET image quality can be easily adjusted with only one regularization parameter β , when the RDP term is used in the proposed DIPrecon algorithm. Thus, we concluded that the proposed DIPrecon algorithm is easy and straightforward implementation for fully 3D end-to-end PET image reconstruction.

The proposed DIPrecon algorithm employed the MR image as a network input to improve the PET image quality, drawing from previous studies that also used the conditional DIP framework[76-78,112,113]. Chapter 4 highlighted that the DIPrecon with random noise input had worse noise characteristics compared to the MR image input. Therefore, it is important to note that there is a potential degradation in noise characteristics with random noise input in the proposed DIPrecon algorithm.

The majority of deep learning-based PET image reconstruction algorithms are data-driven, utilizing trained neural networks that learn from large training datasets consisting of high- and low-quality PET images. The DIPrecon algorithms are not limited by the quality of training datasets because it does not face performance restrictions related to domain adaptation capabilities, such as variations in PET tracers, scanners, organs, and diseases. In addition, there is no theoretical upper limit to the quality of images that can be enhanced using this algorithm because of the absence of the above restrictions. Thus,

the proposed DIPrecon algorithm can enhance the performance of low-dose PET image reconstruction as well as can extend the capabilities of achieving high-quality PET image reconstruction beyond the current baselines.

The limitation of the study was the stability of the network training as the number of subsets increased, possibly due to the complexity of network optimization with randomly initialized parameters. The implementation of a pre-trained network, as recommended by Cui et al. [133] and Onishi et al. [134], might stabilize the network training.

Data corrections such as attenuation and scatter corrections are critical to the success of the end-to-end reconstruction framework. In the current implementation, we perform data corrections on the sinogram space prior to the reconstruction. In the future, a seamless integration of these corrections is needed.

Another area for improvement is its intensive computational cost. The current processing time requires approximately 13 minutes per epoch, positioning it as considerably more resource-intensive than other supervised reconstruction algorithms. Therefore, we need to develop a more practical algorithm for accelerating the processing time in clinical setting. We are currently working on accelerating the DIPrecon algorithm using a two-step optimization method [135].

The experiments of this study were limited to the Monte Carlo simulations of the human brain ^{18}F -FDG PET data and the real preclinical monkey brain ^{18}F -FDG PET data. Future research will expand to include clinical trials with various PET tracers, scanners, organs, and diseases.

6 Conclusion

In this thesis, we aimed at improving PET image quality in the frameworks of the PET image denoising and reconstruction, using the conditional DIP. First, we described the basic principles of the PET imaging, image reconstruction, and deep learning. Then, we introduced the conditional DIP framework for PET image denoising and expanded it to 4D dynamic PET imaging. We proposed the iterative PET image reconstruction incorporating the conditional DIP framework, called DIPrecon. Finally, we developed the DIPrecon for fully 3D PET data.

Chapter 3 first introduced the conditional DIP framework for PET image denoising, which does not require a prior training dataset, and then expanded the conditional DIP framework to 4D dynamic PET imaging. We evaluated the denoising performance of dynamic PET imaging using computer simulation and real preclinical data scanned by the animal PET scanner. The proposed conditional DIP denoising algorithm showed superior performance both visually and quantitatively in computer simulation and real data experiments compared with other denoising algorithms.

Chapter 4 proposed the iterative PET image reconstruction incorporating the conditional DIP framework, called DIPrecon. The DIPrecon incorporated the conditional DIP framework into iterative image reconstruction to achieve powerful noise reduction while measuring consistency with measured emission data, fundamentally solving the challenges of post-processing. We evaluated the proposed DIPrecon with conventional iterative reconstruction algorithms using Monte Carlo simulation data and real preclinical data. The proposed end-to-end DIPrecon algorithm showed superior performance in computer simulation and real preclinical data compared to other conventional

reconstruction algorithms. Furthermore, the proposed algorithm also demonstrated comparable performance and faster calculation time compared to the hybrid DIPrecon algorithm.

Chapter 5 expanded the DIPrecon to fully 3D PET data. We modified the end-to-end DIPrecon algorithm to a block iteration and sequential learning of an ordered sequence of block sinograms. In addition, we implemented the RDP function to the loss function of the network to further improve the quantitative accuracy of the PET image. We evaluated the proposed fully 3D end-to-end DIPrecon with conventional iterative reconstruction algorithms using Monte Carlo simulation data and real preclinical data and can produce high-quality images in computer simulation and real preclinical data.

In actual clinical situations, high-speed calculations that do not interfere with daily operations are required. The main advantages of the proposed conditional DIP denoising introduced in Chapter 3 are its fast computation time and easy implementation, which can potentially solve the above challenges. However, there is a risk of accidentally removing important image information, such as lesions, along with noise because the conditional DIP denoising is a post-processing method. The DIPrecon introduced in Chapter 4 can solve their challenges because this optimization is performed while measuring consistency with measurement data. As introduced in Chapter 5, we can push the DIPrecon algorithm to practical reconstruction by expanding it to fully 3D PET data.

Reference

- 1 M. E. Phelps, "PET: molecular imaging and its biological applications," New York: Springer; 2012. doi: 10.1007/978-0-387-22529-6
- 2 R. E. Carson, "Tracer Kinetic Modeling," in *Positron Emission Tomography*, London: Springer, 2005. doi: 10.1007/1-84628-007-9_6
- 3 H. Schoder, and M. Gonen, "Screening for cancer with PET and PET/CT: potential and limitations," *Journal of Nuclear Medicine*, vol. 48, suppl 1, pp. 4S-18S, 2007.
- 4 R. Minamimoto, M. Senda, K. Uno, S. Jinnouchi, T. Iinuma, et al., "Performance profile of FDG-PET and PET/CT for cancer screening on the basis of a Japanese Nationwide Survey," *Annals of Nuclear Medicine*, vol. 21, no. 9, pp 481-498, 2007.
- 5 L. Zhu, K. Ploessl, and H. F. Kung, "PET/SPECT imaging agents for neurodegenerative diseases," *Chemical Society Reviews*, vol. 43, no. 19, pp. 6683-6691, 2014.
- 6 H. Barthel, M. L. Schroeter, K. T. Hoffmann, and O. Sabri, "PET/MR in dementia and other neurodegenerative diseases," *Seminars in Nuclear Medicine*, vol. 45, no. 3, pp. 224-233, 2015.
- 7 T. Jones, and E. A. Rabiner, "The development, past achievements, and future directions of brain PET," *Journal of Cerebral Blood Flow and Metabolism*, vol. 32, no. 7, pp. 1426-1454, 2012.
- 8 A. R. De Pierro, "A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography," *IEEE Transactions on Medical Imaging*, vol. 14, no. 1, pp. 132-137, 1995.
- 9 M. A. Lodge, M. A. Chaudhry, and R. L. Wahl, "Noise Considerations for PET Quantification Using Maximum and Peak Standardized Uptake Value," *Journal of Nuclear Medicine*, vol. 53, no. 7, pp. 1041-1047, 2012.
- 10 G. Akamatsu, Y. Ikari, H. Nishida, T. Nishio, A. Ohnishi, et al., "Influence of Statistical Fluctuation on Reproducibility and Accuracy of SUV_{max} and SUV_{peak} : A Phantom Study," *Journal of Nuclear Medicine Technology*, vol. 43, no. 3, pp. 222-226, 2015.
- 11 O. Delcroix, D. Bourhis, N. Keromnes, P. Robin, P. Y. Le Roux, et al., "Assessment of Image Quality and Lesion Detectability With Digital PET/CT System," *Frontiers in Medicine*, vol. 8, 629096, 2021.
- 12 National Research Council (2006) Health risks from exposure to low levels of

- ionizing radiation: BEIR VII phase 2. The National Academies Press, Washington, DC. Available: <https://doi.org/10.17226/11340>
- 13 D. W. Townsend, R. A. Isoardi, and B. Bendriem, "Volume Imaging Tomographs," in *The Theory and Practice of 3D PET*, Dordrecht: Springer, 1998.
 - 14 T. Inubushi, M. Ito, Y. Mori, M. Futatsubashi, K. Sato, et al., "Neural correlates of head restraint: Unsolicited neuronal activation and dopamine release," *Neuroimage*, vol. 224, 117434, 2021.
 - 15 R. Ota, "Photon counting detectors and their applications ranging from particle physics experiments to environmental radiation monitoring and medical imaging," *Radiological Physics and Technology*, vol. 14, No. 2, pp. 134-148, 2021.
 - 16 S. Vandenberghe, E. Mikhaylova, E. D'Hoe, P. Mollet, and J. S. Karp, "Recent developments in time-of-flight PET," *EJNMMI Physics*, vol. 3, 3, 2016.
 - 17 S. R. Cherry, T. Jones, J. S. Karp, J. Qi, W. W. Moses, and R. D. Badawi, "Total-Body PET: Maximizing Sensitivity to Create New Opportunities for Clinical Research and Patient Care," *Journal of Nuclear Medicine*, vol. 59, no. 1, pp. 3-12, 2018.
 - 18 H. Tashima, and T. Yamaya, "Proposed helmet PET geometries with add-on detectors for high sensitivity brain imaging," *Physics in Medicine and Biology*, vol. 61, no. 19, 7205, 2016.
 - 19 K. Gong, S. Majewski, P. E. Kinahan, R. L. Harrison, B. F. Elston, et al., "Designing a compact high performance brain PET scanner—simulation study," *Physics in Medicine and Biology*, vol. 61, no. 10, 3681, 2016.
 - 20 G. Akamatsu, M. Takahashi, H. Tashima, Y. Iwao, E. Yoshida, et al., "Designing a compact high performance brain PET scanner—simulation study," *Physics in Medicine and Biology*, vol. 67, no. 22, 225011, 2022.
 - 21 A. J. Reader, G. Corda, A. Mehranian, C. da Costa-Luis, S. Ellis, and J. A. Schnabel, "Deep Learning for PET Image Reconstruction," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 1, pp. 1-25, 2021.
 - 22 F. Hashimoto, Y. Onishi, K. Ote, H. Tashima, A. J. Reader, and T. Yamaya, "Deep learning-based PET image denoising and reconstruction: a review," *Radiological Physics and Technology*, vol. 17, no. 1, pp. 24-46, 2024.
 - 23 A. Bousse, V. S. S. Kandarpa, K. Shi, K. Gong, J. S. Lee, et al., "A Review on Low-Dose Emission Tomography Post-Reconstruction Denoising with Neural Network Approaches," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 8,

- no. 4, pp. 333-347, 2024.
- 24 F. Hashimoto, H. Ohba, K. Ote, and H. Tsukada, "Denoising of Dynamic Sinogram by Image Guided Filtering for Positron Emission Tomography," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 2, no. 6, pp. 541-548, 2018.
 - 25 J. Dutta, R. M. Leahy, and Q. Li, "Non-local means denoising of dynamic PET images," *PLoS ONE*, vol. 8, no. 12, e81390, 2013.
 - 26 H. Arabi, and H. Zaidi, "Non-local mean denoising using multiple PET reconstructions," *Annals of Nuclear Medicine*, vol. 35, no., pp. 176-186, 2021.
 - 27 K. Ote, F. Hashimoto, A. Kakimoto, T. Isobe, T. Inubushi, et al., "Kinetics-Induced Block Matching and 5-D Transform Domain Filtering for Dynamic PET Image Denoising," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 4, no. 6, pp. 720-728, 2020.
 - 28 T. Hebert, and R. Leahy, "A generalized EM algorithm for 3-D Bayesian reconstruction from Poisson data using Gibbs priors," *IEEE Transactions on Medical Imaging*, vol. 8, no. 2, pp. 194-202, 1989.
 - 29 C. Comtat, P. E. Kinahan, J. A. Fessler, T. Beyer, D. W. Townsend, M. Defrise, and C. Michel, "Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels," *Physics in Medicine and Biology*, vol. 47, no. 1, pp. 1-20, 2001.
 - 30 J. Tang, and A. Rahmim, "Bayesian PET image reconstruction incorporating anatomic-functional joint entropy," *Physics in Medicine and Biology*, vol. no. 23, 7063, 2009.
 - 31 Y. LeCun, Y. Bengio, and G. Hinton G, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436-44, 2015.
 - 32 J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85-117, 2015.
 - 33 D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep Image Prior," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 9446-9454, doi: 10.1109/CVPR.2018.00984.
 - 34 D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," *International Journal of Computer Vision*, vol. 128, no. 7, pp. 1867-88, 2020.
 - 35 J. Pacák, Z. Točík, and M. Černý, "Synthesis of 2-deoxy-2-fluoro-D-glucose," *Journal of the Chemical Society D: Chemical Communications*, no. 2, p. 77, 1969.
 - 36 P. Som, H. L. Atkins, D. Bandyopadhyay, J. S. Fowler, R. R. MacGregor, et al., "A fluorinated glucose analog, 2-fluoro-2-deoxy-D-glucose (F-18): nontoxic tracer for

- rapid tumor detection,” *Journal of Nuclear Medicine*, vol. 21, no. 7, pp. 670-675, 1980.
- 37 G. J. Kelloff, J. M. Hoffman, B. Johnson, H. I. Scher, B. A. Siegel, et al., “Progress and promise of FDG-PET imaging for cancer patient management and oncologic drug development,” *Clinical Cancer Research*, vol. 11, no. 8, pp. 2785-2808, 2005.
 - 38 G. N. Ramachandran, and A. V. Lakshminarayanan, “Three-dimensional reconstruction from radiographs and electron micrographs: Application of convolutions instead of Fourier transforms,” *Proceedings of the National Academy of Sciences*, vol. 68, no. 9, pp. 2236-2240, 1971.
 - 39 L. A. Shepp, and B. F. Logan, “The Fourier reconstruction of a head section,” *IEEE Transactions on Nuclear Science*, vol. 21, no. 3, pp. 21-43, pp. 21-43, 1974.
 - 40 E. Tanaka, and T. Iinuma, “Correction functions for optimizing the reconstructed image in transverse section scan,” *Physics in Medicine and Biology*, vol. 20, no. 3, pp. 789-798, 1975.
 - 41 M. Defrise, and P. E. Kinahan, “Data acquisition and image reconstruction for 3D PET,” in *The Theory and Practice of 3D PET*, Dordrecht: Springer, 1998.
 - 42 J. Radon J, “On the determination of functions from their integral values along certain manifolds,” *IEEE Transactions on Medical Imaging*, vol. 5, no. 4, pp. 170-176, 1986.
 - 43 Y. Tsutsui, S. Awamoto, K. Himuro, Y. Umezu, S. Baba, and M. Sasaki, “Characteristics of smoothing filters to achieve the guideline recommended positron emission tomography image without harmonization.” *Asia Oceania Journal of Nuclear Medicine and Biology*, vol. 6, no. 1, pp. 15-23, 2018.
 - 44 L. A. Shepp, and Y. Varidi, “Maximum likelihood reconstruction for emission tomography,” *IEEE Transactions on Medical Imaging*, vol. 1., no. 2, pp. 113-122, 1982.
 - 45 K. Lange, and R. Carson, “EM reconstruction algorithm for emission and transmission tomography,” *Journal of Computer Assisted Tomography*,” vol. 8, no. 2, pp. 306-316, 1984.
 - 46 Y. Vardi, L. A. Shepp, and L. Kaufman, “A statistical model for positron emission tomography,” *Journal of the American Statistical Association*, vol. 80, no. 389, pp. 8-20, 1985.
 - 47 J. Qi, and R. M. Leahy, “Iterative reconstruction techniques in emission computed tomography,” *Physics in Medicine and Biology*, vol. 51, no. 15, pp. R541-R578, 2006.

- 48 H. M. Hudson, and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4, pp. 601-609, 1994.
- 49 J. Browne, and A. B. de Pierro, "A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography," *IEEE Transactions on Medical Imaging*, vol. 15, no.5, pp. 687-699, 1996.
- 50 E. Tanaka, and H. Kudo, "Subset-dependent relaxation in block-iterative algorithm for image reconstruction in emission tomography," *Physics in Medicine and Biology*, vol. 48, no. 10, pp. 1405-1422, 2003.
- 51 J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou, and T. H. Farquhar, "High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner," *Physics in Medicine and Biology*, vol. 43, no. 4, pp. 1001-1013, 1998.
- 52 A. J. Reade, S. Ally, F. Bakatselos, R. Manavaki, R. J. Walledge, et al., "One-pass list-mode EM algorithm for high-resolution 3-D PET image reconstruction into large arrays," *IEEE Transactions on Nuclear Science*, vol. 49, no. 3, pp. 693-699, 2002.
- 53 T. Yamaya, N. Hagiwara, T. Obi, M. Yamaguchi, N. Ohyama, et al., "Transaxial system models for jPET-D4 image reconstruction," *Physics in Medicine and Biology*, vol. 50, no. 22, pp. 5339-5355, 2005.
- 54 V. Y. Panin, F. Kehren, C. Michel, and M. Casey, "Fully 3-D PET reconstruction with system matrix derived from point source measurements," *IEEE Transactions on Medical Imaging*, vol. 25, no. 7, pp. 907-921, 2006.
- 55 D. L. Snyder, M. I. Miller, and D. G. Politte, "Noise and edge artifacts in maximum-likelihood reconstructions for emission tomography," *IEEE Transactions on Medical Imaging*, vol. 6, no. 3, pp. 228-238, 1987.
- 56 E. Levitan, G. T. Herman, "A maximum a posteriori probability expectation maximization algorithm for image reconstruction in emission tomography," *IEEE Transactions on Medical Imaging*, vol. 6, no. 3, pp. 185-192, 1987.
- 57 T. Herbert, R. Leachy, "A generalized EM algorithm for 3-D Bayesian reconstruction from projection data using Gibbs priors," *IEEE Transactions on Medical Imaging*, vol. 8, no. 2, pp. 194-202, 1989.
- 58 P. J. Green, "Bayesian reconstructions from emission tomography data using a modified EM algorithm," *IEEE Transactions on Medical Imaging*, vol. 9, no. 1, pp. 84-92, 1990.
- 59 A. R. De Pierro, M. E. B. Yamagishi, "Fast EM-like methods for maximum "a

- posteriori” estimates in emission tomography,” *IEEE Transactions on Medical Imaging*, vol. 20, no. 4, pp. 280-288, 2001.
- 60 I. Sutskever, J. Martens, G. Dahl, and G. Hinton, “On the importance of initialization and momentum in deep learning,” in *Proceedings of the 30th International Conference on Machine Learning*, vol. 28, no. 3, pp. 1139-1147, 2013.
 - 61 A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, “Deep learning for computer vision: A brief review,” *Computational intelligence and neuroscience*, 7068349, 2018.
 - 62 T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, et al., “Language models are few-shot learners,” in *Advances in Neural Information Processing Systems (NeurIPS 2020)*, 2020.
 - 63 N. O’Mahony, S. Campbel, A. Carvalho, S. Harapanahalli, G. V. Hernandez, et al., “Deep learning vs. Traditional computer vision,” in *Advances in Computer Vision*, Cham, Switzerland:Springer, vol. 943, pp. 128-144, 2020.
 - 64 J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, et al., “Recent advances in convolutional neural networks,” *Pattern recognition*, vol. 77, pp. 354-377, 2018.
 - 65 R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” *Insights into imaging*, vol. 9, no. 4, pp. 611-629, 2018.
 - 66 O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015. Lecture Notes in Computer Science*, vol 9351, 2015. doi: [10.1007/978-3-319-24574-4_28](https://doi.org/10.1007/978-3-319-24574-4_28).
 - 67 Ö Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation,” in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016. Lecture Notes in Computer Science*, vol 9901, 2016. doi: [10.1007/978-3-319-46723-8_49](https://doi.org/10.1007/978-3-319-46723-8_49).
 - 68 S. Kaji, and S. Kida, “Overview of image-to-image translation by use of deep neural networks: denoising, super-resolution, modality conversion, and reconstruction in medical imaging,” *Radiological Physics and Technology*, vol. 12, no. 3, pp. 235-348, 2019.
 - 69 N. S. Pun, and S. Agarwal, “Modality specific U-Net variants for biomedical image segmentation: a survey,” *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5845-5889, 2022.

- 70 F. Hashimoto, A. Kakimoto, N. Ota, S. Ito, and S. Nishizawa, "Automated segmentation of 2D low-dose CT images of the psoas-major muscle using deep convolutional neural networks," *Radiological Physics and Technology*, vol. 12, no. 2, pp. 210-215, 2019.
- 71 F. Hashimoto, K. Ote, T. Oida, A. Teramoto, and Y. Ouchi, "Compressed-sensing magnetic resonance image reconstruction using an iterative convolutional neural network approach," *Applied Sciences*, vol. 10, no. 6, 1902, 2020.
- 72 F. Hashimoto, and K. Ote, "ReconU-Net: a direct PET image reconstruction using U-Net architecture with back projection-induced skip connection," *Physics in Medicine and Biology*, 2024, doi: [10.1088/1361-6560/ad40f6](https://doi.org/10.1088/1361-6560/ad40f6).
- 73 B. Zhu, J. Z. Liu, S. F. Cauley, B. R. Rosen, M. S. Rosen, "Image reconstruction by domain-transform manifold learning," *Nature*, vol. 555, pp. 487-492, 2018.
- 74 I. Häggström, C. R. Schmidtlein, G. Campanella, T. J. Fuchs, "DeepPET: A deep encoder-decoder network for directly solving the PET image reconstruction inverse problem," *Medical Image Analysis*, vol. 54, pp. 253-262, 2019.
- 75 G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Transactions on Graphics*, vol. 23, no. 3, pp. 664-672, 2004.
- 76 J. Cui, K. Gong, N. Guo, C. Wu, X. Meng, K. Kim, et al., "PET image denoising using unsupervised deep learning," *European Journal of Nuclear Medicine and Molecular Imaging*, vol. 46, no. 13, pp. 2780-2789, 2019.
- 77 F. Hashimoto, H. Ohba, K. Ote, and A. Teramoto, "Unsupervised dynamic PET image denoising with anatomical information," *Medical Imaging and Information Sciences*, vol. 37, no. 3, pp. 58-61, 2020.
- 78 Y. Onishi, F. Hashimoto, K. Ote, H. Ohba, R. Ota, et al., "Anatomical-guided attention enhances unsupervised PET image denoising performance," *Medical Image Analysis*, vol. 74, 102226, 2021.
- 79 F. Hashimoto, H. Ohba, K. Ote, A. Teramoto, and H. Tsukada, "Dynamic PET image denoising using deep convolutional neural networks without prior training datasets," *IEEE Access*, vol. 7, pp. 96594-96603, 2019.
- 80 F. Hashimoto, H. Ohba, K. Ote, A. Kakimoto, H. Tsukada, and Y. Ouchi, "4D deep image prior: dynamic PET image denoising using an unsupervised four-dimensional branch convolutional neural network," *Physics in Medicine and Biology*, vol. 66, no. 1, 015006, 2021.

- 81 S. Ioffe, and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” in *Proceedings of the 32nd International Conference on Machine Learning, in Proceedings of Machine Learning Research*, vol. 37, pp. 448-456, 2015. Available : <https://proceedings.mlr.press/v37/ioffe15.html>
- 82 H. Fukui, T. Yamashita, Y. Kato, R. Matsui, T. Ogata, et al., “Multiple Facial Attributes Estimation Based on Weighted Heterogeneous Learning,” in *Computer Vision – ACCV 2016 Workshops. ACCV 2016*, vol. 10117, 2017. doi: [10.1007/978-3-319-54427-4_29](https://doi.org/10.1007/978-3-319-54427-4_29).
- 83 K. He, X. Zhang, S. Ren, and J. Sun, “Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification,” in *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1026-1034, 2015. doi: [10.1109/ICCV.2015.123](https://doi.org/10.1109/ICCV.2015.123).
- 84 B. Aubert-Broche, M. Griffin, G. B. Pike, A. C. Evans, and D. L. Collins, “Twenty new digital brain phantoms for creation of validation image data bases,” *IEEE Transactions on Medical Imaging*, vol. 25, no. 11, pp. 1410-1416, 2006.
- 85 BrainWeb: Simulated Brain Database. Available: <https://brainweb.bic.mni.mcgill.ca/>
- 86 D. Feng, K. P. Wong, C. M. Wu, and W. C. Siu, “A technique for extracting physiological parameters and the required input function simultaneously from PET image measurements: Theory and simulation study,” *IEEE Transactions on Information Technology in Biomedicine*, vol. 1, no. 4, pp. 243-254, 1997.
- 87 H. Watabe, Y. Ikoma, Y. Kimura, M. Naganawa, and M. Shidahara, “PET kinetic analysis—compartmental model,” *Annals of Nuclear Medicine*, vol. 20, no. 9, pp. 583-588, 2006.
- 88 C. S. Patlak, R. G. Blasberg, and J. D. Fenstermacher, “Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data,” *Journal of Cerebral Blood Flow and Metabolism*, vol. 3, no. 1, pp. 1-7, 1983.
- 89 PMOD Technologies LLC – PMOD Technologies. Available: <https://www.pmod.com/web/>
- 90 Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600-612, 2004.
- 91 L. Farde, E. Ehrin, L. Eriksson, T. Greitz, H. Hall, C. G. Hedström, et al., “Substituted benzamides as ligands for visualization of dopamine receptor binding in the human

- brain by positron emission tomography,” *Proceedings of the National Academy of Sciences*, vol. 82, no. 11, pp. 3863-3867, 1985.
- 92 Y. Tomonari, Y. Onishi, F. Hashimoto, K. Ote, T. Okamoto, and H. Ohba, “Animal PET scanner with a large field of view is suitable for high-throughput scanning of rodents,” *Annals of Nuclear Medicine*, 2024. (In press)
 - 93 E. Tanaka, and H. Kudo, “Optimal relaxation parameters of DRAMA (Dynamic RAMLA) aiming at one-pass image reconstruction for 3D-PET,” *Physics in Medicine and Biology*, vol. 55, no. 10, pp. 2917-2939, 2010.
 - 94 J. Logan, J. S. Fowler, N. D. Volkow, G. J. Wang, Y. S. Ding, and D. L. Alexoff, “Distribution volume ratios without blood sampling from graphical analysis of PET data,” *Journal of Cerebral Blood Flow and Metabolism*, vol. 16, no. 5, pp. 834-840, 1996.
 - 95 K. He, J. Sun, and X. Tang, “Guided image filtering,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1397-1409, 2012.
 - 96 K. Gong, J. Guan, C. C. Liu, and J. Qi, “PET image denoising using a deep neural network through fine tuning,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 153-161, 2018.
 - 97 K. Ote, A. Tokui, F. Hashimoto, T. Isobe, A. Saito, and T. Omura, “PET Image Noise Removal by Covolutional Neural Networks,” *Medical Imaging Technology*, vol.37, no. 1, pp. 35-45, 2019. [in Japanese]
 - 98 Y. Wang, B. Yu, L. Wang, C. Zu, D. S. Lalush, W. Lin, X. Wu, J. Zhou, D. Shen, and L. Zhou, “3D conditional generative adversarial networks for high-quality PET image estimation at low dose,” *Neuroimage*, vol. 174, pp. 550-562, 2018.
 - 99 Y. Lei, X. Dong, T. Wang, K. Higgins, T. Liu, W. J. Curran, H. Mao, J. A. Nye, and X. Yang, “Whole-body PET estimation from low count statistics using cycle-consistent generative adversarial networks,” *Physics in Medicine and Biology*, vol. 64, no. 21, 215017, 2019.
 - 100 S. Y. Yie, S. K. Kang, D. Hwang, and J. S. Lee, “Self-supervised PET denoising,” *Nuclear Medicine and Molecular Imaging*, vol. 54, np. 6, pp. 299-304, 2020.
 - 101 S. K. Kang, S. Y. Yie, J. S. Lee, “Noise2Noise Improved by Trainable Wavelet Coefficients for PET Denoising,” *Electronics*, vol. 10, no. 13, 1529, 2021.
 - 102 T. A. Song, F. Yang, J. Dutta, “Noise2Void: unsupervised denoising of PET images,” *Physics in Medicine and Biology*, vol. 66, 214002, 2021.
 - 103 K. Simonyan, and A. Zisserman, “Very deep convolutional networks for large-scale

- image recognition,” *arXiv preprint*, arXiv:1409.1556, 2014.
- 104 C. Chan, R. Fulton, R. Barnett, D. D. Feng, and S. Meikle, “Postreconstruction nonlocal means filtering of whole-body PET with an anatomical prior,” *IEEE Transactions on Medical Imaging*, vol. 33, no. 3, pp. 636-650, 2014.
 - 105 M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, “Non-local transform-domain filter for volumetric data denoising and reconstruction,” *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 119-133, 2013.
 - 106 R. Heckel, and P. Hand, “Deep decoder: Concise image representations from untrained non-convolutional networks,” *arXiv preprint*, arXiv:1810.03982, 2018.
 - 107 F. Hashimoto, H. Ohba, K. Ote, and A. Teramoto, “Unsupervised dynamic PET image denoising with anatomical information,” *Medical Imaging and Information Sciences*, vol. 37, no. 3, pp.58-61, 2020.
 - 108 D. Zhao, F. Zhao, and Y. Gan, “Reference-driven compressed sensing MR image reconstruction using deep convolutional neural networks without pre-training,” *Sensors*, vol. 20, no. 1, 308, 2020.
 - 109 A. Rahmim, K. Dinelle, J. C. Cheng, M. A. Shilov, W. P. Segars, S. C. Lidstone, S. Blinder, O. G. Rousset, H. Vajihollahi, B. M. W. Tsui, D. F. Wong, and V. Sossi, “Accurate event-driven motion compensation in high-resolution PET incorporating scattered and random events,” *IEEE Transactions on Medical Imaging*, vol. 27, no. 8, pp. 1018-1033, 2008.
 - 110 C. Catana, T. Benner, A. van der Kouwe, L. Byars, M. Hamm, D. B. Chonde, C. J. Michel, G. E. Fakhri, M. Schmand, and A. G. Sorensen, “MRI-assisted PET motion correction for neurologic studies in an integrated MR-PET scanner,” *Journal of Nuclear Medicine*, vol. 52, no. 1, pp. 154-161, 2011.
 - 111 Y. Onishi, T. Isobe, M. Ito, F. Hashimoto, T. Omura, and E. Yoshikawa, “Performance evaluation of dedicated brain PET scanner with motion correction system,” *Annals of Nuclear Medicine*, vol. 36, no. 8, pp. 746-755, 2022.
 - 112 K. Gong, C. Catana, J. Qi, and Q. Li, “PET image reconstruction using deep image prior,” *IEEE Transactions on Medical Imaging*, vol. 38, no. 7, pp. 1655-1665, 2019.
 - 113 F. Hashimoto, K. Ote, and Y. Onishi, “PET Image Reconstruction Incorporating Deep Image Prior and a Forward Projection Model,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 6, no. 8, pp. 841-846, 2022.
 - 114 E. V. R. Di Bella, A. B. Barclay, R. L. Eisner, and R. W. Schafer, “A comparison of rotation-based methods for iterative reconstruction algorithms,” *IEEE Transactions*

- on *Nuclear Science*, vol. 43, no. 6, pp. 3370-3376, 1996.
- 115 C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization," *ACM Transactions on Mathematical Software*, vol. 23, no. 4, pp. 550-560, 1997.
 - 116 Y. Nesterov, "A method of solving a convex programming problem with convergence rate $O(1/k^2)$," *Soviet Mathematics Doklady*, vol. 27, pp. 372-376, 1983.
 - 117 D. P. Kingma, and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint*, arXiv:1412.6980, 2014.
 - 118 S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1-122, 2011.
 - 119 K. Gong, J. Guan, K. Kim, X. Zhang, J. Yang, Y. Seo, G. E. Fakhri, J. Qi, and Q. Li, "Iterative PET image reconstruction using convolutional neural network representation," *IEEE Transactions on Medical Imaging*, vol. 38, no. 3, pp. 675-685, 2019.
 - 120 Z. Xie, R. Baikejiang, T. Li, X. Zhang, K. Gong, M. Zhang, W. Qi, E. Asma, and J. Qi, "Generative adversarial network based regularized image reconstruction for PET," *Physics in Medicine and Biology*, vol. 65, no. 12, 125016, 2020.
 - 121 A. Mehranian, and A. J. Reader, "Model-Based Deep Learning PET Image Reconstruction Using Forward-Backward Splitting Expectation-Maximization," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 5, no. 1, pp. 54-64, 2021.
 - 122 K. Kim, D. Wu, K. Gong, J. Dutta, J. H. Kim, Y. D. Son, H. K. Kim, G. E. Fakhri, and Q. Li, "Penalized PET reconstruction using deep learning prior and local linear fitting," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1478-1487, 2018.
 - 123 F. Hashimoto, Y. Onishi, K. Ote, H. Tashima, and T. Yamaya, "Fully 3D implementation of the end-to-end deep image prior-based PET image reconstruction using block iterative algorithm," *Physics in Medicine and Biology*, vol. 68, no. 15, 155009, 2023.
 - 124 J. Nuyts, D. Beque, P. Dupont, L. Mortelmans, "A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography," *IEEE Transactions on Nuclear Science*, vol. 49, no. 1, pp. 56-60, 2002.
 - 125 H. M. Hudson, and R. S. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Transactions on Medical Imaging*, vol. 13, no. 4,

- pp. 601-609, 1994.
- 126 S. R. Meikle, B. F. Hutton, D. L. Bailey, P. K. Hooper, and M. J. Fulham, "Accelerated EM reconstruction in total-body PET: potential for improving tumour detectability," *Physics in Medicine and Biology*, vol. 39, no. 10, 1689, 1994.
 - 127 J. Nuyts, D. Beque, P. Dupont, and L. Mortelmans, "A concave prior penalizing relative differences for maximum-a-posteriori reconstruction in emission tomography," *IEEE Transactions on Nuclear Science*, vol. 49, no. 1, pp. 56-60, 2002.
 - 128 K. Miwa, T. Yoshii, K. Wagatsuma, S. Nezu, Y. Kamitaka, T. Yamao, R. Kobayashi, S. Fukuda, Y. Yakushiji, N. Miyaji, and K. Ishii, "Impact of γ factor in the penalty function of Bayesian penalized likelihood reconstruction (Q. Clear) to achieve high-resolution PET images," *EJNMMI Physics*, vol. 10, no. 1, 4, 2023.
 - 129 E. J. Teoh, D. R. McGowan, R. E. Macpherson, K. M. Bradley, and F. V. Gleeson, "Phantom and clinical evaluation of the bayesian penalized likelihood reconstruction algorithm Q.clear on an LYSO PET/CT system," *Journal of Nuclear Medicine*, vol. 56, no. 9, pp. 1447-1452, 2015.
 - 130 R. L. Siddon, "Fast calculation of the exact radiological path for a 3-dimensional CT array," *Medical Physics*, vol. 12, no. 2, pp. 252-255, 1985.
 - 131 S. Yatawatta, L. De Clercq, H. Spreeuw, and F. Diblen, "A stochastic LBFGS algorithm for radio interferometric calibration," in *Proc. IEEE Data Science Workshop (DSW)*, pp 208-212, 2019.
 - 132 N. S. Keskar, D. Mudigere, J. Nocedal, M. Smelyanskiy, and P. T. P. Tang, "On large-batch training for deep learning: Generalization gap and sharp minima," *arXiv preprint*, arXiv:1609.04836, 2016.
 - 133 J. Cui, K. Gong, N. Guo, C. Wu, K. Kim, H. Liu, and Q. Li, "Populational and individual information based PET image denoising using conditional unsupervised learning," *Physics in Medicine and Biology*, vol. 66, no. 15, 155001, 2021.
 - 134 Y. Onishi, F. Hashimoto, K. Ote, K. Matsubara, and M. Ibaraki, "Self-supervised pre-training for deep image prior-based robust pet image denoising," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 8, no. 4, pp. 348-356, 2024.
 - 135 F. Hashimoto, Y. Onishi, K. Ote, H. Tashima, and T. Yamaya, "Two-step optimization for accelerating deep image prior-based PET image reconstruction," *Radiological Physics and Technology*, 2024. [Online]. Available: <https://doi.org/10.1007/s12194-024-00831-9>

Publication

Journal papers

1. Y. Iwao, K. Shiotsuki, F. Hashimoto, T. Ochiai, T. Kagawa, R. Nagata, M. Eto, Y. Hatanaka, Y. Yoshida, and Y. Asayama, “Explainable Deep Learning for Predicting Bone Mineral Density using Clavicle Features on Chest Radiographs: A Multi-task Approach with Regression and Segmentation,” *Radiological Physics and Technology*. (Under review)
2. K. Ote, F. Hashimoto, Y. Onishi, and Y. Ouchi, “List-Mode PET Image Reconstruction Using Dykstra-Like Splitting,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, 2024. [Online]. Available: <https://doi.org/10.1109/TRPMS.2024.3441526>
3. F. Hashimoto, Y. Onishi, K. Ote, H. Tashima, and T. Yamaya, “Two-step optimization for accelerating deep image prior-based PET image reconstruction,” *Radiological Physics and Technology*, 2024. [Online]. Available: <https://doi.org/10.1007/s12194-024-00831-9>
4. F. Hashimoto, and K. Ote, “ReconU-Net: a direct PET image reconstruction using U-Net architecture with back projection-induced skip connection,” *Physics in Medicine and Biology*, vol. 69, no. 10, 105022, 2024.
5. Y. Tomonari, Y. Onishi, F. Hashimoto, K. Ote, T. Okamoto, and H. Ohba, “Animal PET scanner with a large field of view is suitable for high-throughput scanning of rodents,” *Annals of Nuclear Medicine*, vol. 38, no. 7, pp. 544-552, 2024.
6. Y. Onishi*, F. Hashimoto, K. Ote, and R. Ota*, “Whole reconstruction-free system design for direct positron emission imaging from image generation to attenuation correction,” *IEEE Transactions on Medical Imaging*, vol. 43, no. 5, pp. 1654-1663, 2024. (* Equal contribution)
7. Y. Onishi, F. Hashimoto, K. Ote, K. Matsubara, and M. Ibaraki, “Self-Supervised Pre-Training for Deep Image Prior-Based Robust PET Image Denoising,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 8, no. 4, pp. 348-356, 2024.
8. A. Obana, K. Ote, Y. Gohto, H. Yamada, F. Hashimoto, S. Okazaki, and R. Asaoka, “Deep learning-based correction of cataract-induced influence on macular pigment optical density measurement by autofluorescence spectroscopy,” *PLoS ONE*, vol. 19,

no. 2, e0298132, 2024.

9. F. Hashimoto, Y. Onishi, K. Ote, H. Tashima, A. J. Reader, and T. Yamaya, “Deep learning-based PET image denoising and reconstruction: a review,” *Radiological Physics and Technology*, vol. 17, no. 1, pp. 24-46, 2024. (Invited review paper)
10. F. Hashimoto, Y. Onishi, K. Ote, H. Tashima, and T. Yamaya, “Fully 3D implementation of the end-to-end deep image prior-based PET image reconstruction using block iterative algorithm,” *Physics in Medicine and Biology*, vol. 68, no. 15, 155009, 2023.
11. K. Ote, F. Hashimoto, Y. Onishi, T. Isobe, and Y. Ouchi, “List-Mode PET Image Reconstruction Using Deep Image Prior,” *IEEE Transactions on Medical Imaging*, vol. 42, no. 6, pp. 1822-1834, 2023.
12. F. Hashimoto, K. Ote, and Y. Onishi, “PET Image Reconstruction Incorporating Deep Image Prior and a Forward Projection Model,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 6, no. 8, pp. 841-846, 2022.
13. Y. Onishi, T. Isobe, M. Ito, F. Hashimoto, T. Omura, and E. Yoshikawa, “Performance evaluation of dedicated brain PET scanner with motion correction system,” *Annals of Nuclear Medicine*, vol. 36, no. 8, pp. 746-755, 2022.
14. Y. Onishi, F. Hashimoto, K. Ote, and R. Ota, “Unbiased TOF estimation using leading-edge discriminator and convolutional neural network trained by single-source-position waveforms,” *Physics in Medicine and Biology*, vol. 67, no. 7, 04NT01, 2022.
15. K. Ote, and F. Hashimoto, “Deep-learning-based fast TOF-PET image reconstruction using direction information,” *Radiological Physics and Technology*, vol. 15, no. 1, pp. 72-82, 2022.
16. S. I. Kwon*, R. Ota*, E. Berg*, F. Hashimoto, K. Nakajima, I. Ogawa, Y. Tamagawa, T. Omura, T. Hasegawa, and S. R. Cherry, “Ultrafast timing enables reconstruction-free positron emission imaging,” *Nature Photonics*, vol. 15, no. 12, pp. 914-918, 2021. (* Equal contribution)
17. Y. Onishi*, F. Hashimoto*, K. Ote, H. Ohba, R. Ota, E. Yoshikawa, and Y. Ouchi, “Anatomical-Guided Attention Enhances Unsupervised PET Image Denoising Performance,” *Medical Image Analysis*, vol. 74, 102226, 2021. (* Equal contribution)
18. F. Hashimoto*, M. Ito*, K. Ote, T. Isobe, H. Okada, and Y. Ouchi, “Deep learning-based attenuation correction for brain PET with various radiotracers,” *Annals of*

Nuclear Medicine, vol. 35, no. 6, pp. 691-701, 2021. (* Equal contribution)

19. A. Obana, K. Ote, F. Hashimoto, R. Asaoka, Y. Gohto, S. Okazaki, and H. Yamada, "Correction for the Influence of Cataract on Macular Pigment Measurement by Autofluorescence Technique Using Deep Learning," *Translational Vision Science and Technology*, vol. 10, no. 2, 18, 2021.
20. F. Hashimoto, H. Ohba, K. Ote, A. Kakimoto, H. Tsukada, and Y. Ouchi, "4D deep image prior: dynamic PET image denoising using an unsupervised four-dimensional branch convolutional neural network," *Physics in Medicine and Biology*, vol. 66, no. 1, 015006, 2021.
21. J. Fang, H. Ohba, F. Hashimoto, H. Tsukada, F. Chen, and H. Liu, "Imaging mitochondrial complex I activation during a vibrotactile stimulation: A PET study using [^{18}F]BCPP-EF in the conscious monkey brain," *Journal of cerebral blood flow and metabolism*, vol. 40, no. 12, pp. 2521-2532, 2020.
22. K. Ote, R. Ota, F. Hashimoto, and T. Hasegawa, "Direct Annihilation Position Classification Based on Deep Learning Using Paired Cherenkov Detectors: A Monte Carlo Study," *Applied Sciences*, vol. 10, no. 22, 7957, 2020.
23. K. Ote, F. Hashimoto, A. Kakimoto, T. Isobe, T. Inubushi, et al., "Kinetics-Induced Block Matching and 5-D Transform Domain Filtering for Dynamic PET Image Denoising," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 4, no. 6, pp. 720-728, 2020.
24. F. Hashimoto, H. Ohba, K. Ote, and A. Teramoto, "Unsupervised dynamic PET image denoising with anatomical information," *Medical Imaging and Information Sciences*, vol. 37, no. 3, pp.58-61, 2020.
25. F. Hashimoto, K. Ote, T. Oida, A. Teramoto, and Y. Ouchi, "Compressed-sensing magnetic resonance image reconstruction using an iterative convolutional neural network approach," *Applied Sciences*, vol. 10, no. 6, 1902, 2020.
26. F. Hashimoto, H. Ohba, K. Ote, A. Teramoto, and H. Tsukada, "Dynamic PET image denoising using deep convolutional neural networks without prior training datasets," *IEEE Access*, vol. 7, pp. 96594-96603, 2019.
27. F. Hashimoto, A. Kakimoto, N. Ota, S. Ito, and S. Nishizawa, "Automated segmentation of 2D low-dose CT images of the psoas-major muscle using deep convolutional neural networks," *Radiological Physics and Technology*, vol. 12, no. 2, pp. 210-215, 2019.
28. F. Hashimoto, K. Ote, R. Ota, and T. Hasegawa, "A feasibility study on 3D

- interaction position estimation using deep neural network in Cherenkov-based detector: a Monte Carlo simulation study,” *Biomedical Physics and Engineering Express*, vol. 5, no. 3, 035001, 2019.
29. F. Hashimoto, H. Ohba, K. Ote, and H. Tsukada, “Denoising of Dynamic Sinogram by Image Guided Filtering for Positron Emission Tomography,” *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 2, no. 6, pp. 541-548, 2018.
 30. M. Kanazawa, S. Nishiyama, F. Hashimoto, T. Kakiuchi, and H. Tsukada, “Evaluation of D-isomers of 4-borono-2-¹⁸F-fluoro-phenylalanine and O-¹¹C-methyl-tyrosine as brain tumor imaging agents: a comparative PET study with their L-isomers in rat brain glioma,” *EJNMMI Research*, vol. 8, 47, 2018.
 31. F. Hashimoto, A. Teramoto, Y. Asada, S. Suzuki, and H. Fujita, “Dose reduction technique in diagnostic X-ray computed tomography by use of 6-channel multileaf collimators,” *Radiological Physics and Technology*, vol. 10, no. 1, pp. 60-67, 2017.
 32. 橋本二三生, 大西佑弥, 大手希望, “Deep learning PET 画像再構成への招待,” *Medical Imaging Technology*, vol. 41, no. 4-5, pp. 162-170, 2023. (招待レビュー論文)
- F. Hashimoto, Y. Onishi, and K. Ote, “Invitation to Deep Learning-based PET Image Reconstruction,” *Medical Imaging Technology*, vol. 41, no. 4-5, pp. 162-170, 2023. (Invited review paper, in Japanese)
33. 大手希望, 得居葵, 橋本二三生, 磯部卓志, 斉藤右典, 大村知秀, “畳み込みニューラルネットワークによる PET 画像ノイズ除去,” *Medical Imaging Technology*, vol. 37, no. 1, pp. 35-45, 2019.
- K. Ote, A. Tokui, F. Hashimoto, T. Isobe, A. Saito, and T. Omura, “PET Image Noise Removal by Convolutional Neural Networks,” *Medical Imaging Technology*, vol. 37, no. 1, pp. 35-45, 2019. (In Japanese)
34. 橋本二三生, 大手希望, “Image guided filter を用いたダイナミック PET 画像の MAP 画像再構成,” *医用画像情報学会雑誌*, vol. 35, no. 4, pp. 59-61, 2018.
- F. Hashimoto, and K. Ote, “MAP image reconstruction of dynamic PET images using image guided filter,” *Medical Imaging and Information Sciences*, vol. 35, no. 4, pp. 59-61, 2018. (In Japanese)
35. 橋本二三生, 寺本篤司, 浅田恭生, 鈴木昇一, 藤田広志, “二次元領域設定型 CT 装置の開発 -アクティブコリメータを搭載した実験装置の製作と基礎的評価-,” *Medical Imaging Technology*, vol. 34, no. 2, pp. 123-127, 2016.
- F. Hashimoto, A. Teramoto, Y. Asada, S. Suzuki, and H. Fujita, “Development of the

Two-dimensional Region-setting CT System: Development and Basic Evaluation of the Experimental System Using the Active Collimators,” *Medical Imaging Technology*, vol. 34, no. 2, pp. 123-127, 2016. (In Japanese)

36. 大澤慎也, 村田千佳, 橋本二三生, 寺本篤司, 藤田広志,”近赤外光 CT に関する基礎的検討 -ディジタル一眼レフカメラを用いた実験装置の開発-,” 医用画像情報学会雑誌, vol. 32, no. 2, pp. 44-47, 2015.

S. Osawa, C. Murata, F. Hashimoto, A. Teramoto, and H. Fujita, “ Basic study on the near-infrared light computed tomography - Development of the experimental system using a digital single - lens reflex camera -,” *Medical Imaging and Information Sciences*, vol. 32, no. 2, pp. 44-47, 2015. (In Japanese)

International conferences

1. T. Yamaya, T. Ishikawa, G. Akamatsu, H. Tashima, F. Nishikido, M. Takahashi, F. Hashimoto, R. Ota, S. I. Kwon, and S. R. Cherry, "Dual-panel PET system to be enabled by 30-ps super-fast detector: simulation study," *Society of Nuclear Medicine and Molecular Imaging (SNMMI) Annual Meeting 2024*, June, 2024.
2. T. Yamaya, T. Ishikawa, G. Akamatsu, H. Tashima, F. Nishikido, M. Takahashi, C. Toramatsu, Y. Iwao, F. Hashimoto, R. Ota, S. I. Kwon, and S. R. Cherry, "Dual-panel geometry for PET-guided therapy to be enabled by super-fast detector: simulation study," *The 10th Conference on PET, SPECT, and MR Multimodal Technologies, Total Body and Fast Timing in Medical Imaging*, May, 2024.
3. F. Hashimoto, K. Ote, H. Tashima, G. Akamatsu, Y. Iwao, M. Takahashi, and T. Yamaya, "Uncertainty-based mixture of a deep image prior and an original reconstructed images in PET," *The 3rd International Conference on Radiological Physics and Technology*, April, 2024.
4. T. Yamaya, T. Ishikawa, G. Akamatsu, H. Tashima, F. Nishikido, M. Takahashi, F. Hashimoto, and R. Ota, "Dual-panel PET system to be enabled by 30-ps super-fast detector: a preliminary simulation study," *The 3rd International Conference on Radiological Physics and Technology*, April, 2024.
5. F. Hashimoto, Y. Onishi, K. Ote, H. Tashima, and T. Yamaya, "Accelerated Deep Image Prior-based PET Image Reconstruction Using Two-Step Optimization," *The 2023 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov.

2023.

6. F. Hashimoto, K. Ote, and Y. Onishi, "ReconU-Net: Direct PET Image Reconstruction Using Back Projection-induced Skip Connection," *The 2023 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov. 2023.
7. Y. Onishi, F. Hashimoto, K. Ote, and R. Ota, "Proposal of Morphological Imaging for direct Positron Emission Imaging," *The 2023 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov. 2023.
8. F. Hashimoto, K. Ote, Y. Onishi, H. Tashima, and T. Yamaya, "End-to-end Unsupervised CNN-based PET Image Reconstruction with Relative Difference Penalty," *The 2nd International Conference on Radiological Physics and Technology*, April, 2023.
9. F. Hashimoto, K. Ote, Y. Onishi, H. Tashima, and T. Yamaya, "3D Implementation of the End-to-end Deep Image Prior-based PET Image Reconstruction," *The 2022 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov. 2022.
10. Y. Onishi, F. Hashimoto, K. Ote, K. Matsubara, and M. Ibaraki, "Using Self-Supervised Pretraining Model for Unsupervised PET Image Denoising," *The 2022 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov. 2022.
11. K. Ote, F. Hashimoto, Y. Onishi, and T. Isobe, "List-Mode PET Image Reconstruction Using Deep Image Prior," *The 2022 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov. 2022.
12. R. Ota, S. I. Kwon, E. Berg, F. Hashimoto, K. Nakajima, I. Ogawa, Y. Tamagawa, T. Omura, T. Hasegawa, and S. R. Cherry, "Reconstruction-free imaging of positron-emitting radionuclides using ultra-fast detectors," *The virtual 2021 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov. 2021.
13. Y. Onishi, F. Hashimoto, K. Ote, H. Ohba, R. Ota, E. Yoshikawa, and Y. Ouchi, "Unsupervised PET Image Denoising Using Attention-Guided Anatomical Information," *The virtual 2021 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov. 2021.
14. K. Ote, and F. Hashimoto, "Deep Learning-based Fast TOF-PET Image Reconstruction Using Direction Information," *The virtual 2021 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov. 2021.
15. F. Hashimoto, H. Ohba, K. Ote, A. Kakimoto, H. Tsukada, A. Teramoto, and Y. Ouchi, "Dynamic PET Image Denoising Using 4-dimensional Deep Image Prior," *The virtual 2020 IEEE Nuclear Science Symposium and Medical Imaging Conference*,

Nov. 2020.

16. K. Ote, R. Ota, F. Hashimoto, and T. Hasegawa, "Direct Annihilation Position Regression based on Deep Learning and Digital Offset using Pair of Cherenkov Detectors: Monte Carlo Study," *The virtual 2020 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov. 2020.
17. A. Obana, K. Ote, F. Hashimoto, Y. Gohto, S. Okazaki, and H. Yamada, "Correction of the influence of cataract on macular pigment measurement by autofluorescence technique using deep learning," *Association for Research in Vision and Ophthalmology (ARVO) 2020 Annual Meeting*, June, 2020.
18. F. Hashimoto, H. Ohba, K. Ote, A. Kakimoto, H. Tsukada, A. Teramoto, and Y. Ouchi, "End-to-end Dynamic PET Image Denoising Using 4-dimensional Unsupervised Convolutional Neural Network," *The 18th Conference of Peace through Mind/Brain Science*, Feb. 2020.
19. K. Ote, F. Hashimoto, A. Kakimoto, T. Isobe, T. Inubushi, A. Tokui, A. Saito, T. Omura, E. Yoshikawa, A. Teramoto, and Y. Ouchi, "Kinetics-Induced Block Matching and 5D Transform Domain Filtering for Dynamic PET Image Denoising," *The 18th Conference of Peace through Mind/Brain Science*, Feb. 2020.
20. F. Hashimoto, M. Ito, K. Ote, T. Isobe, H. Okada, Y. Ouchi, "Emis2Trans: Attenuation Correction for Brain PET With Many Types of PET Ligands Using Convolutional Neural Networks," *The 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Oct. 2019.
21. K. Ote, F. Hashimoto, A. Kakimoto, T. Isobe, A. Tokui, E. Yoshikawa, T. Omura, A. Teramoto, and Y. Ouchi, "Block Matching and 5D Filtering of Dynamic PET Images," *The 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Oct. 2019.
22. K. Ote, R. Ota, F. Hashimoto, and T. Hasegawa, "Direct Annihilation Position Classification based on Deep Learning using Pair of Cherenkov Detectors: Monte Carlo Study," *The 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Oct. 2019.
23. F. Hashimoto, K. Ote, and H. Tsukada, "Dynamic PET Image Denoising Using Deep Convolutional Neural Network Without Training Datasets," *Society of Nuclear Medicine and Molecular Imaging (SNMMI) Annual Meeting 2019*, June, 2019.
24. F. Hashimoto, K. Ote, R. Ota, R. Yamada, and T. Hasegawa, "Using deep learning to Estimate 3D Interaction Position in Cherenkov-based Detector: A Monte Carlo

- Simulation Study," *The 2018 IEEE Nuclear Science Symposium and Medical Imaging Conference*, Nov. 2018.
25. F. Hashimoto, H. Ohba, K. Ote, and H. Tsukada, "Dynamic PET Denoising by Image Guided Filtering: A Preliminary Study," *The 17th Conference of Peace through Mind/Brain Science*, Feb. 2018.
 26. F. Hashimoto, A. Teramoto, Y. Asada, S. Suzuki, and H. Fujita, "Novel Concept for Dose Reduction -Region-setting CT: Is Multileaf Collimator Also Valuable for Diagnostic CT?," *RSNA2015*, Dec. 2015.
 27. F. Hashimoto, A. Teramoto, S. Suzuki, and H. Fujita, "A preliminary study on the development of a region-setting CT system using multi-leaf active collimators," *The International Forum on Medical Imaging in Asia*, Jan. 2015.
 28. F. Hashimoto, C. Murata, A. Teramoto, S. Suzuki, and H. Fujita, "A basic study on region setting CT system: three-dimensional data collection and reconstruction using experimental system," *The 15th Asian Oceanian Congress of Radiology 2014*, Sep. 2014.
 29. A. Teramoto, T. Ohno, F. Hashimoto, C. Murata, K. Takahashi, and H. Fujita, "Basic study on the development of a high-resolution breast CT," *International Workshop on Breast Imaging*, June. 2014.
 30. C. Murata, F. Hashimoto, A. Teramoto, and H. Fujita, "Educational system of computed tomography using optical computed tomography," *Computer Assisted Radiology and Surgery 28th International Congress and Exhibition*, June. 2014.

Acknowledgement

本研究を進めるにあたりご指導、ご教示いただきました山谷泰賀客員教授に心より感謝いたします。山谷先生の強いリーダーシップや鋭い視点は大変勉強になりました。ありがとうございました。

また、本論文をまとめるにあたり有益なご助言、ご提言をいただきました羽石秀昭教授、兪文偉教授、菅幹生准教授に心より感謝いたします。

浜松ホトニクス株式会社中央研究所の原勉前所長をはじめ、塚田秀夫元 PET 研究グループ長、原田典弘企画推進グループ長、大庭弘行副第 5 研究室長には社会人として働きながら大学院進学をサポートしていただきました。大手希望主任部員、大田良亮部員、大西佑弥研究員とは日々の楽しいディスカッションの中で、多くの有益なアイデアが生まれました。ありがとうございました。これからもよろしくお願いします。

量子科学研究開発機構の田島英朗主幹研究員には、田島仕込みの画像再構成と野球を教えていただきました。ありがとうございました。

浜松医科大学の尾内康臣教授には、研究指導や臨床目線での助言だけでなく、大学院進学も相談にも乗っていただきました。ありがとうございました。

また、第 5 研究室のメンバー、イメージング物理研究グループのメンバーには公私にわたり大変お世話になりました。

最後に、これまでの人生や研究生活を応援し続けてくれた家族に心から感謝いたします。ありがとうございました。

2024 年 9 月 橋本二三生