Deep-learning-based fast TOF-PET image reconstruction using direction information

Kibo Ote1 · Fumio Hashimoto1

Received: 29 October 2021 / Revised: 26 January 2022 / Accepted: 27 January 2022 / Published online: 8 February 2022
© Japanese Society of Radiological Technology and Japan Society of Medical Physics 2022

Abstract
Although deep learning for application in positron emission tomography (PET) image reconstruction has attracted the attention of researchers, the image quality must be further improved. In this study, we propose a novel convolutional neural network (CNN)-based fast time-of-flight PET (TOF-PET) image reconstruction method to fully utilize the direction information of coincidence events. The proposed method inputs view-grouped histo-images into a 3D CNN as a multi-channel image to use the direction information of such events. We evaluated the proposed method using Monte Carlo simulation data obtained from a digital brain phantom. Compared with a case without direction information, the peak signal-to-noise ratio and structural similarity were improved by 1.2 dB and 0.02, respectively, at a coincidence time resolution of 300 ps. The calculation times of the proposed method were significantly lower than those of a conventional iterative reconstruction. These results indicate that the proposed method improves both the speed and image quality of a TOF-PET image reconstruction.

Keywords Positron emission tomography · Image reconstruction · Deep learning · Time-of-flight · Direction information

1 Introduction
Positron emission tomography (PET) is a functional imaging tool used in various medical applications, such as oncology, cardiology, and neurology [1]. It has the unique ability to quantitatively estimate radiotracer concentrations at as low as the picomolar scale; however, the radiotracer concentration cannot be directly imaged from a line of response measured through the coincidence detection of annihilation photons. Therefore, an image reconstruction process is required to estimate the distribution of the radiotracer concentration.

There are two main methods for achieving an image reconstruction: analytic and iterative [2]. An analytical method is simple and fast, although it is sensitive to statistics and prone to streak artifacts. An iterative method models the noise distribution and reconstructs an image through iterative updating. Although an iterative method improves the signal-to-noise ratio (SNR) of the reconstructed image better than an analytical approach, it is computationally expensive. Therefore, an image reconstruction method that improves both the speed and SNR is desired.

Deep learning has recently attracted attention for improving the image quality of PET [3–9], particularly for image reconstruction tasks [10–17]. In a pioneering study, Zhu et al. proposed an automated transform by manifold approximation (AUTOMAP) network that can directly reconstruct tomographic images from various sensing data [18]. The versatility of AUTOMAP is due to its unique network architecture, which first applies a fully connected (FC) layer to the sensing data, followed by a convolution layer to the feature maps, which are both repeated multiple times. However, AUTOMAP is limited to 2D reconstruction because the FC layer has a large number of parameters. Hägström et al. developed DeepPET, which is a deep convolutional encoder–decoder network for directly reconstructing PET images from PET sinogram data [19]. Although DeepPET provides less noisy images than those of the ordered subset expectation maximization algorithm (OSEM) [20], similar to AUTOMAP, it is limited to 2D reconstruction. Such limitations are due to the large network required for the direct mapping of sensing data to the image format.

To realize an actual 3D image reconstruction in near real time, Whiteley et al. proposed FastPET, a
deep-learning-based fast TOF-PET image reconstruction using direction information. FastPET first converts coincidence events into a histo-image using a 3D convolutional neural network (CNN). A histo-image is a more suitable format for a CNN than a sinogram, thus enabling near-real-time 3D image reconstruction. However, based on the theory of analytic TOF-PET image reconstruction, FastPET may be compromised in terms of the SNR because the histo-image does not preserve the direction information of coincidence events.

In this study, we propose a deep-learning-based fast TOF-PET image reconstruction method using directional information. The proposed method inputs view-grouped histo-images into a 3D CNN multichannel images. Using Monte Carlo simulation data of a digital brain phantom, we verified that the proposed method improves the SNR while maintaining a near real-time image reconstruction.

2 Materials and methods

2.1 Histo-image generation

The basis of an analytical TOF-PET image reconstruction is the application of a deconvolution after accumulation of coincidence events into a histo-image format using TOF information. This problem deals with the process of accumulating events in a histo-image. Accordingly, two methods have been considered, i.e., the most likely annihilation position (MLAP) and confidence weighting (CW).

In both methods, the MLAP of the coincident event connecting two detector elements whose coordinates are \( \vec{P}_1 \) and \( \vec{P}_2 \) is estimated using the TOF information.

\[
\vec{P}_{\text{MLAP}} = \frac{\vec{P}_1 + \vec{P}_2}{2} + c \frac{\Delta t}{2} \frac{\vec{P}_1 - \vec{P}_2}{\| \vec{P}_1 - \vec{P}_2 \|},
\]

where \( c \) and \( \Delta t \) denote the speed of light and the TOF information, respectively. The MLAP method simply accumulates an event in the voxel nearest to \( \vec{P}_{\text{MLAP}} \). The CW method accumulates the event as a line weighted by the TOF response function centered on \( \vec{P}_{\text{MLAP}} \), as illustrated in Fig. 1.

By mathematically modeling the variance of the voxel value of an analytically reconstructed image of a uniform disk phantom, it was concluded that the CW method is optimal in terms of the SNR. The MLAP method is suboptimal because the high-resolution information in the vertical direction of the coincidence event is lost after accumulating the events into the histo-image when using this approach. In other words, the direction of the coincidence event contains information regarding the resolution heterogeneity. Another reason for selecting the CW method is continuity with a conventional non-TOF-PET image reconstruction. In other words, the CW method reduces to the non-TOF-PET image reconstruction method, as the coincidence time resolution (CTR) increases toward infinity.

2.2 Angular-view grouping

Although the CW method is optimal from the perspective of the SNR, because it requires ray tracing, it is a time-consuming approach. In this study, we introduce angular-view grouping to implement the direction information. In this scheme, the events are divided into \( N \) groups, depending on the angle of coincidence. The group of events is separately accumulated in \( N \) histo-images using an MLAP method. Fig 2 shows an example of angular-view grouping at \( N = 8 \). Using angular-view grouping, we can preserve the direction likely annihilation position (a). Although the CW method is optimal in terms of the SNR, it is also time-consuming owing to the ray tracing required.
Fig. 2 Event accumulation of TOF-PET using angular-view grouping. Events are divided into \( N \) groups depending on the angle of coincidence. The groups of events are separately accumulated in \( N \) histo-images using the MLAP method. In this example, \( N = 8 \). View-grouped histo-images preserve the direction information of the coincidence event without ray tracing. Here, \( \varphi \) is the azimuthal angle of a coincidence event.

information of coincidence events as view-grouped histo-images without ray tracing. Note that the angular-view grouping in this study was applied at the azimuthal angle and not at the oblique angle.

The azimuthal angle \( \varphi \) and view group \( n \) can be calculated using the following equations:

\[
\varphi = \begin{cases} 
\cos^{-1} \frac{y_2 - y_1}{\sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}} & x_2 - x_1 \geq 0 \\
\pi - \cos^{-1} \frac{y_2 - y_1}{\sqrt{(x_2-x_1)^2 + (y_2-y_1)^2}} & \text{otherwise}
\end{cases},
\]

where \( (x_1, y_1), (x_2, y_2) \) are the 2D coordinates of \( \vec{P}_1 \) and \( \vec{P}_2 \), respectively, and \( \lfloor \cdot \rfloor \) denotes the operator of the round down.

2.3 TOF-PET image reconstruction using CNN

Figure 3 depicts the outline of the proposed method. Raw coincidence events are accumulated in \( N \) view-grouped histo-images using the above-mentioned angular-view grouping. Subsequently, the 3D CNN receives \( N \) view-grouped histo-images and one attenuation map as an input.
Deep-learning-based fast TOF-PET image reconstruction using direction information

(N+1) channel image and outputs the final image. A 3D CNN can use the direction information of coincidence events because view-grouped histo-images preserve the direction information of coincidence events. The view-grouped histo-images are normalized for variations in the detection efficiency. Then, the attenuation and scatter are corrected inside the 3D CNN. The 3D CNN learns the attenuation and scatter corrections through end-to-end learning of mapping of the view-grouped histo-images, which are not corrected for scatter and attenuation, to the phantom images of ground truth.

Figure 4 illustrates the network architecture of the 3D CNN used in this study. We employed a 3D U-net [26] architecture because it is suitable for medical image processing. To increase the speed of the inference, we reduced the trainable parameters of the proposed architecture to 3% of those of the FastPET architecture. The 3D U-net comprises an encoder, a decoder, and skip connections.

The encoder part extracts useful features for image reconstruction through a convolution, nonlinear activation, and down-sampling. The combination of a 3 × 3 × 3 3D convolution and a leaky rectified linear unit (LReLU) was repeated twice before down-sampling. Down-sampling was applied using a 4 × 4 × 3 3D convolution with stride (2, 2, 1), followed by LReLU. At each down-sampling stage, the x- and y-dimensions of the feature maps are halved, and the number of channels is doubled.

The decoder part reconstructs the final image from the feature maps through convolution, nonlinear activation, and up-sampling. The combination of 3 × 3 × 3 3D convolution, and LReLU was repeated twice before up-sampling. Up-sampling was applied using a 4 × 4 × 3 3D transpose convolution with stride (2, 2, 1), followed by LReLU. At each up-sampling stage, the x- and y-dimensions of the feature maps are doubled, and the number of channels is halved. The final image was reconstructed using a 3 × 3 × 3 3D convolution with a single channel output.

The feature maps of the encoder part before down-sampling are added to the feature maps of the decoder part after up-sampling through a skip connection.

2.4 Simulation dataset

Monte Carlo simulations were conducted using our Monte Carlo simulator to create a dataset. Segmented brain MRI images of 20 normal subjects downloaded from BrainWeb [27] were used to create digital brain phantoms. The contrast of radioactivity between the gray matter, white matter, and cerebrospinal fluid was set to 1: 0.25: 0.05, based on the [18F] FDG contrast. The attenuation coefficients of soft tissue and bone were set to 0.00895 and 0.0151 mm⁻¹, respectively.

In this study, we used the brain-dedicated PET scanner described in [28] as the detector arrangement for the simulation. A detector ring with a diameter of 486.83 mm was constructed using 28 and 4 detector units in the ring and axial directions, respectively. Each detector unit had a 16 × 16 array of cerium-doped lutetium–yttrium oxyorthosilicate (LYSO) crystals. The size of each LYSO crystal was 3.14 × 3.14 × 20 mm³. The image size had 128 × 128 × 70 voxels, with a voxel size of 3 × 3 × 3.221 mm³. An energy resolution of 15% and energy window of 400–650 keV were assumed. A total of 181.12 ± 6.08 M counts, including scatter events, were collected for each subject using a 3D acquisition. CTR values of 100, 300, and 600 ps were simulated. There were 72 rings, including the gap between the detector units in the axial direction. The maximum ring difference was set to ± 66.

We divided the 20 subjects into 15 and 5 subjects for training and testing, respectively. In addition, to monitor the...
validation loss during training, the training data were split into 12 data for real training and 3 data for validation.

To evaluate the robustness of the proposed method to abnormal cases, three simulated tumors were inserted into a digital brain phantom of the PET data. We downloaded a segmented MRI image of one normal subject from BrainWeb and embedded three tumors with contrast values of 1.1, 1.3, and 1.5.

2.5 Network training

The 3D CNN was trained for 500 epochs using the Adam optimizer with $\beta_1 = 0.5$. In this study, phantom images were used as the training labels. The mean-squared error was used as the loss function. We considered 64 updates using a mini-batch with a batch size of 32 as an epoch. We randomly cropped the $64 \times 64 \times 64$ voxel-sized sub-image from the original $128 \times 128 \times 70$ voxel-sized image during training for data augmentation. We monitored the loss of the validation data during training and selected the model with the minimum validation loss for testing. During the testing phase, the original $128 \times 128 \times 70$ voxel-sized images were used as the network input.

2.6 Evaluation

We reconstructed the five test datasets using the list-mode dynamic row-action maximum likelihood algorithm (List-DRAMA) [29, 30] and the proposed method of $N = 1, 2, 4, 8$, and 16. List-DRAMA was applied using 2 main iterations and 40 sub-iterations. List-DRAMA was calculated in parallel using eight cores of 3.33 GHz Intel Xeon X5680. The proposed method was calculated using an NVIDIA Quadro P6000 graphics board using Chainer 7.7.0 (https://chainer.org). For a quantitative evaluation, we measured the peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) of the reconstructed images, $\hat{x}$, using a phantom image, $x$, as a reference.

The PSNR is an indicator of the voxel-level image similarity and is calculated as

$$\text{PSNR} = 10 \log_{10} \left( \frac{\max_R x}{\sqrt{\sum_{j \in R} (\hat{x}_j - x_j)^2}} \right)^2 \text{dB},$$

where $j$, $R$ and $N_R$ denote the index of the voxel, the region of an entire brain, and the number of voxels inside the entire brain region, respectively.

In addition, the SSIM is an indicator of the image brightness, contrast, and structural similarities, and is calculated as

$$\text{SSIM} = \frac{1}{N_R} \sum_{j \in R} \frac{2 \mu_{\hat{x}_j} \mu_{x_j} + c_1 (2 \sigma_{\hat{x}_j x_j} + c_2)}{(\mu_{\hat{x}_j}^2 + \mu_{x_j}^2 + c_1)(\sigma_{\hat{x}_j}^2 + \sigma_{x_j}^2 + c_2)},$$

where $\mu$ and $\sigma$ are the mean and standard deviation of a small local region (patch) around the $j$-th voxel, respectively, and $\sigma_{\hat{x}_j x_j}$ is the covariance between the patches around the $j$-th voxel of the reconstructed and phantom images. In this study, the patch had $7 \times 7 \times 7$ voxels, $c_1 = (0.01L)^2$, and $c_2 = (0.03L)^2$, where $L$ is the dynamic range of the phantom image.

To evaluate the spatial distribution of the error, we confirmed the difference between the reconstructed and phantom images. We also measured the mean absolute deviation (MAD) between the reconstructed image and phantom image as follows:

$$\text{MAD} = \frac{1}{N_R} \sum_{j \in R} |\hat{x}_j - x_j|.$$  

Moreover, we measured the contrast recovery coefficients (CRCs) of the tumors as follows:

$$\text{CRC} = \frac{\sum_{x \in \text{tumor}} \hat{x}_j}{\sum_{x \in \text{tumor}} x_j},$$

where $R_{\text{tumor}}$ is a tumor region.

3 Results and discussion

Figure 5 shows the PSNR and SSIM relative to the number of views at CTRs of 100, 308, and 600 ps. The PSNR and SSIM increased as the number of views increased. This indicates that the direction information is beneficial for improving the SNR, even with a deep-learning-based image reconstruction method.

The PSNR at a CTR of 100 ps decreased slightly when the number of views reached 12 or more. It is known that the number of views required for an analytic TOF-PET image reconstruction was obtained as follows [25]:

$$N > 2 \pi \sigma_T / d_0,$$

where $\sigma_T$ is the standard deviation of the TOF response function, and $d_0$ denotes the required spatial resolution. If $\sigma_T = 6.37 \text{ mm}$, which corresponds to a CTR of 100 ps, and $d_0 = 4.5 \text{ mm}$, which corresponds to a 1.5 voxel width, then $N > 6$ is sufficient. These results are consistent with theory. Therefore, the optimal number of views for each TOF-PET scanner can easily be estimated using the above formula.
Figure 6 shows a comparison of the reconstructed images between the proposed and other methods at a CTR of 300 ps. The results of FastPET correspond to the results of the proposed method using a single view. The results of the proposed method are obtained from eight views. The proposed method improved both the PSNR and SSIM and provided sharper images than those of the other methods. The proposed method also recovered a finer structure than that of the other approaches, as shown by the magnified images of the red square. Fig 7 shows the differences between the reconstructed images and the phantom image at a CTR of 300 ps. The error distributions of the proposed method were similar to those of FastPET. The MAD of the proposed approach is smaller than those of the other methods. Fig 8 shows the results of the simulation data, including tumors, at a CTR of 300 ps. The mean and standard deviation of the CRCs of three tumors were 0.86±0.06, 0.78±0.09, and 0.80±0.06 for List-DRAMA, FastPET, and the proposed method with eight views, respectively. The differences between the reconstructed images and the phantom image in Fig 8 (b) indicate that the proposed method underestimated the value of the tumors. Although the CRC of the proposed method was lower than that of List-DRAMA, it was better than that of FastPET, which did not use direction information. In addition, the proposed method recovered tumor shapes more accurately than FastPET. These results indicate that the structural accuracy of the reconstructed image obtained using the proposed method was improved by the high-resolution information in the vertical direction of the coincidence events.

Table 1 shows a comparison of the calculation times, PSNRs, and SSIMs between the proposed method and the other methods for a CTR of 300 ps. The calculation time of the proposed method is slower than that of FastPET based on the calculation time required to generate view-grouped histo-images. In addition, the proposed method was one order of magnitude faster than List-DRAMA. These results indicate that the proposed method is capable of near real-time TOF-PET image reconstruction with a high image quality. Notably, the calculation time of histo-image generation is dependent on the number of events. For example, if we can generate view-grouped histo-images within 0.7 s for the list data of 1 s duration, we can reconstruct the images faster than acquisition because the reconstruction when applying the proposed method only required approximately 0.2 s.

One of the limitations of this study is that it was applied only to the simulation dataset. We plan to collect experimental data for training neural networks using a brain PET scanner. From this study, it appears that the findings of the theory of image reconstruction are also useful for deep-learning-based methods. For example, view-grouped histo-images can be beneficial for improving the PET image quality using an unsupervised CNN framework [6–8, 12].

In this study, we used angular-view grouping with MLAP, instead of CW. The CW method is expected to improve the SNR based on the principle of TOF-PET image reconstruction; however, it is impractical for near real-time image reconstruction from the perspective of the calculation cost because the strict calculation of CW requires event-by-event ray tracing. Angular-view grouping can be considered fast approximation of the CW method, and its performance is similar to that of the CW method as the number of views increases. In this study, we input a tuple of view-grouped histo-images and attenuation maps to the 3D CNN, similar to FastPET [21]. When no attenuation map is used for the CNN input, the accuracies of the scatter and attenuation corrections can be degraded because Compton scattering and photon attenuation are governed by the attenuation map. For a brain-dedicated PET scanner without an X-ray CT mechanism, such as that simulated in this study, the registration of CT images, taken separately from PET, to the PET image space is required for correction of the attenuation. In this case, image reconstruction without an attenuation map is conducted first for image registration, and image reconstruction with the attenuation map is then applied.
Fig. 6 Comparison of reconstructed images between the proposed method and the other methods at a CTR of 300 ps. From left to right, the phantom image, histo-image, List-DRAMA, and FastPET, which correspond to the proposed method using a single view, and the proposed method using eight views. The histo-image was made by accumulating all events into a single array using TOF information. The images were tagged with the PSNR and SSIM. The last row shows the zoomed-in images of a red square.
Fig. 7 Difference between reconstructed images and phantom image at a CTR of 300 ps. From left to right, List-DRAMA and FastPET, corresponding to the proposed method using a single view, and the proposed method using eight views. The images were tagged using MAD.
Fig. 8 Results of simulation data including tumors at CTR of 300 ps. (a) Comparison of reconstructed images. (b) Difference between reconstructed images and phantom image. From left to right, the phantom image, List-DRAMA and FastPET, which correspond to the proposed method using a single view, and the proposed method using eight views. The reconstructed images were tagged using the mean and standard deviation of the CRC of three tumor regions. Red arrows indicate tumors.
Table 1 Comparison of calculation times, PSNRs, and SSIMs between the proposed method and other methods at 300 ps CTR. The calculation times of both histo-image generation and reconstruction are shown. The calculation times of histo-image generation and List-DRAMA are dependent on the number of events.

| Reconstruction method | Number of views | Calculation time (s): | PSNR (dB) | SSIM |
|-----------------------|----------------|-----------------------|-----------|------|
|                       |                | Histo-image generation | Reconstruction |
| List-DRAMA            | –              | 7.92 ± 0.12           | 18.71 ± 0.41 | 0.818 ± 0.011 |
| FastPET               | 1              | 11.06 ± 0.42          | 20.44 ± 0.17 | 0.936 ± 0.001 |
| Proposed method       | 8              | 11.35 ± 0.22          | 20.95 ± 0.33 | 0.941 ± 0.002 |
|                       | 12             | 11.86 ± 0.28          | 20.98 ± 0.14 | 0.941 ± 0.001 |
|                       | 16             | 12.15 ± 0.19          | 21.00 ± 0.15 | 0.943 ± 0.002 |

4 Conclusion

We proposed a deep-learning-based fast TOF-PET image reconstruction method using direction information. We input the view-grouped histo-images into the 3D CNN to use the direction information. We then evaluated the proposed method using Monte Carlo simulation data from a digital brain phantom. The proposed method achieved better PSNR and SSIM results, recovered finer structures than those of the other methods, and except for the calculation time of the view-grouped histo-image generation, required a sub-second calculation time. These results indicate that the proposed method is beneficial for both the speed and image quality of TOF-PET image reconstruction.

Declarations

Conflict of interest The authors are employees of the Hamamatsu Photonics K.K. The company had no control over the interpretation, writing, or publication of this manuscript.

Ethical approval This study did not include research on human subjects or animals.

References

1. Phelps ME. PET: molecular imaging and its biological applications. New York: Springer; 2012.
2. Defrise M, Kinahan PE. Data acquisition and image reconstruction for 3D PET in The Theory and Practice of 3D PET. Dordrecht: Springer; 1998.
3. Wang Y, Yu B, Wang L, Zu C, Lalush DS, Lin W, et al. 3D conditional generative adversarial networks for high-quality PET image estimation at low dose. Neuroimage. 2018;174:550–62.
4. Chen KT, Gong E, de Carvalho MFB, Xu J, Boumis A, Khalighi M, et al. Ultra-low-dose 18F-Flurbetaen amyloid PET imaging using deep learning with multi-contrast MRI inputs. Radiology. 2019;290(3):649–56.
5. Gong K, Guan J, Liu CC, Qi J. PET image denoising using a deep neural network through fine tuning. IEEE Trans Radiat Plasma Med Sci. 2018;3(2):153–61.
6. Hashimoto F, Ohba H, Ote K, Teramoto A, Tsukada H. Dynamic PET image denoising using deep convolutional neural networks without prior training datasets. IEEE Access. 2019;7:96594–603.
7. Hashimoto F, Ohba H, Ote K, Kakimoto A, Tsukada H, Ouchi Y. 4D deep image prior: dynamic PET image denoising using an unsupervised four-dimensional branch convolutional neural network. Phys Med Biol. 2021;66(1):015006.
8. Hashimoto F, Ito M, Ote K, Isobe T, Okada H, Ouchi Y. Deep learning-based attenuation correction for brain PET with various radiotracers. Ann Nucl Med. 2021;35(6):691–701.
9. Sanaat A, Shiri I, Arabi H, Maitani J, Nkoulou R, Zaidi H. Deep learning-assisted ultra-fast/low-dose whole-body PET/CT imaging. Eur J Nucl Med Mol Imaging. 2021;48(8):2405–15.
10. Yang B, Ying L, Tang J. Artificial neural network enhanced Bayesian PET image reconstruction. IEEE Trans Med Imaging. 2018;37(6):1297–309.
11. Gong K, Guan J, Kim K, Zhang X, Yang J, Seco Y, et al. Iterative PET image reconstruction using convolutional neural network representation. IEEE Trans Med Imaging. 2019;38(3):675–85.
12. Gong K, Catana C, Qi J, Li Q. PET image reconstruction using deep image prior. IEEE Trans Med Imaging. 2019;38(7):1655–65.
13. Whiteley W, Luk WK, Gregor J. DirectPET: full-size neural network PET reconstruction from sinogram data. J Med Imaging. 2020;7(3):032503.
14. Reader AJ, Corda G, Mehranian A, da Costa-Luis C, Ellis S, Schnabel JA. Deep learning for PET image reconstruction. IEEE Trans Radiat Plasma Med Sci. 2021;5(1):1–25.
15. Hu Z, Xue H, Zhang Q, Gao J, Zhang N, Zou S, et al. DPIR-Net: direct PET image reconstruction based on the wasserstein generative adversarial network. IEEE Trans Radiat Plasma Med Sci. 2021;5(1):35–43.
16. Kandarpa VSS, Bousse A, Benoit D, Visvikis D. DUG-RECON: a framework for direct image reconstruction using convolutional generative networks. IEEE Trans Radiat Plasma Med Sci. 2021;5(1):44–53.
17. Mehranian A, Reader AJ. Model-based deep learning PET image reconstruction using forward-backward splitting expectation maximization. IEEE Trans Radiat Plasma Med Sci. 2021;5(1):54–64.
18. Zhu B, Liu JZ, Cauley SF, Rosen BR, Rosen MS. Image reconstruction by domain-transform manifold learning. Nature. 2018;555(7697):487–92.
19. Håggström I, Schmidtlein CR, Campanella G, Fuchs TJ. DeepPET: a deep encoder–decoder network for directly solving the PET image reconstruction inverse problem. Med Image Anal. 2019;54:253–62.
20. Hudson HM, Larkin RS. Accelerated image reconstruction using ordered subsets of projection data. IEEE Trans Med Imaging. 1994;13(4):601–9.
21. Whiteley W, Panin V, Zhou C, Cabello J, Bhrakhada D, Gregor J. FastPET: near real-time reconstruction of PET histo-image data using a neural network. IEEE Trans Radiat Plasma Med Sci. 2021;5(1):65–77.
22. Matej S, Suri S, Jayanthi S, Daube-Witherspoon ME, Lewitt RM, Karp JS. Efficient 3-D TOF PET reconstruction using
23. Snyder DL, Thomas LJ, Ter-Pogossian MM. A mathematical model for positron-emission tomography systems having time-of-flight measurements. IEEE Trans Nucl Sci. 1981;28(3):3575–83.

24. Tomitani T. Image reconstruction and noise evaluation in photon time-of-flight assisted positron emission tomography. IEEE Trans Nucl Sci. 1981;28(6):4581–9.

25. Tanaka E. Line-writing data acquisition and signal-to-noise ratio in time-of-flight positron emission tomography. IEEE Comput Soc. 1982;448:101–8.

26. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: Ourselin S, Joskowicz L, Sabuncu MR, Unal G, Wells W, editors. Medical Image Computing and Computer Assisted Intervention (MICCAI) LNCS, vol. 9901. Cham: Springer; 2016. p. 424–32.

27. Collins DL, Zijdenbos AP, Kollokian V, Sled JG, Kabani NJ, Holmes CJ, Evans AC. Design and construction of a realistic digital brain phantom. IEEE Trans Med Imaging. 1998;17(3):463–8.

28. Saito A, Yoshikawa E, Omura T, Yamanaka T, Ote K, Isobe T, et al. Development of a brain PET scanner with motion correction using motion capture technology. IEEE Nucl Sci Symp Med Imaging Conf 2018;M-07–146

29. Tanaka E, Kudo H. Subset-dependent relaxation in block-iterative algorithms for image reconstruction in emission tomography. Phys Med Biol. 2003;48(10):1405–22.

30. Nakayama T, Kudo H. Derivation and implementation of ordered-subsets algorithms for list-mode PET data. IEEE Nuc Sci Symp Med Imaging. Conf Rec 2005;3540-3

31. Vandenberghe S, Daube-Witherspoon ME, Lewitt RM, Karp JS. Fast reconstruction of 3D time-of-flight PET data by axial rebinning and transverse mashing. Phys Med Biol. 2006;51(6):1603–21.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.