


ORIGINAL RESEARCH

Open Access



Deep learning-based image quality improvement of ^{18}F -fluorodeoxyglucose positron emission tomography: a retrospective observational study

Junichi Tsuchiya^{1*} , Kota Yokoyama¹, Ken Yamagiwa¹, Ryosuke Watanabe¹, Koichiro Kimura¹, Mitsuhiro Kishino¹, Chung Chan², Evren Asma² and Ukihide Tateishi¹

* Correspondence: tuwu11@gmail.com

¹Department of Diagnostic Radiology and Nuclear Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan
Full list of author information is available at the end of the article

Abstract

Background: Deep learning (DL)-based image quality improvement is a novel technique based on convolutional neural networks. The aim of this study was to compare the clinical value of ^{18}F -fluorodeoxyglucose positron emission tomography (^{18}F -FDG PET) images obtained with the DL method with those obtained using a Gaussian filter.

Methods: Fifty patients with a mean age of 64.4 (range, 19–88) years who underwent ^{18}F -FDG PET/CT between April 2019 and May 2019 were included in the study. PET images were obtained with the DL method in addition to conventional images reconstructed with three-dimensional time of flight-ordered subset expectation maximization and filtered with a Gaussian filter as a baseline for comparison. The reconstructed images were reviewed by two nuclear medicine physicians and scored from 1 (poor) to 5 (excellent) for tumor delineation, overall image quality, and image noise. For the semi-quantitative analysis, standardized uptake values in tumors and healthy tissues were compared between images obtained using the DL method and those obtained with a Gaussian filter.

Results: Images acquired using the DL method scored significantly higher for tumor delineation, overall image quality, and image noise compared to baseline ($P < 0.001$). The Fleiss' kappa value for overall inter-reader agreement was 0.78. The standardized uptake values in tumor obtained by DL were significantly higher than those acquired using a Gaussian filter ($P < 0.001$).

Conclusions: Deep learning method improves the quality of PET images.

Keywords: Deep learning, ^{18}F -fluorodeoxyglucose positron emission tomography, Image quality

Background

Integrated positron emission tomography (PET) and computed tomography (CT) using ^{18}F -fluorodeoxyglucose (FDG) is a standard method used in oncology [1] and is also applied in other conditions, including infectious, ischemic, and degenerative diseases. In oncology, ^{18}F -FDG PET/CT is useful for differentiation between benign and malignant lesions, cancer staging, assessment of the response to treatment, and planning of radiation therapy. However, although ^{18}F -FDG PET is a promising modality, the images are noisy and their resolution is low [2, 3].

PET images are reconstructed by analytic methods such as filtered back projection [4]. However, reconstructions using analytic methods are challenging because noise statistics related to the emission of photons are difficult to model. Therefore, a statistical model in the maximum likelihood framework has been developed [5].

High-resolution PET images with a high signal-to-noise ratio (SNR) enable better visualization of precise anatomical structures, improving diagnostic accuracy and facilitating early diagnosis of disease and accurate staging. Several methods can be used to obtain high-quality PET images, including increasing the acquisition time, using a time of flight technique, and detection with a semiconductor [6, 7]. In clinical settings, increasing the data acquisition time is a common choice. However, an increased acquisition time leads to a longer examination time, which can be burdensome for patients. Denoising techniques have been used to improve the quality of low SNR images.

A convolutional neural network (CNN) has been applied in medical imaging, including CT and magnetic resonance imaging. Given that it is composed of several linear convolutional layers and nonlinear layers, CNN could reduce statistical noise in PET images without degrading image contrast. Such layers (linear and nonlinear) with a large number of parameters are optimized by training using large datasets. Parameters are optimized by training data extracted from PET/CT datasets. Deep learning approaches for image noise reduction have recently been reported for other modalities, including CT and magnetic resonance imaging [8–11]. Introduction of deep learning-based restoration in single-photon emission tomography (SPECT) images has also been reported. Dietze et al. used a CNN to upgrade images of technetium-99m macroaggregated albumin SPECT/CT pre-treatment, resulting in images comparable with Monte Carlo-based iterative reconstruction, which is known to render better quality images, but to be time-consuming [12]. Moreover, CNN can be applied to PET. CNN was applied for low-dose PET/CT and yielded high-quality images [13–18]. However, the applicability of CNN to clinical PET images has not been fully investigated.

Therefore, the main purpose of this study was to evaluate initial clinical experiences and to explore whether or not the newly developed deep learning-based method improves image quality in comparison with conventional images obtained with a Gaussian filter. We compared whole-body ^{18}F -FDG clinical images with the DL method and those with conventional reconstruction both visually and semi-quantitatively.

Methods

Patients

The institutional review board of the Tokyo Medical and Dental University approved the present study, and written informed consent was obtained from all patients. The

population of the testing studies consisted of 50 patients with a mean age of 64.4 (range, 19–88) years who underwent ^{18}F -FDG PET/CT between April 2019 and May 2019. The patient demographic and clinical data are summarized in Table 1.

PET/CT imaging

The patients fasted for at least 6 h before undergoing a PET/CT examination and their blood glucose levels were measured. Next, a 3.7-MBq/kg injection of ^{18}F -FDG was administered, and ^{18}F -FDG imaging was performed approximately 60 min later. Patients were scanned from the skull base to the mid-thigh region using a PET/CT scanner (Celesteion, Canon Medical, Tokyo, Japan). The CT parameters used for attenuation correction were as follows: 120 kV; field of view, 550 mm; pitch, 16.0; and slice thickness, 2.0 mm. PET emission datasets were obtained with 2 min in each bed position (for 16–18 min in total). The conventional PET images with Gaussian postfilter were reconstructed with three-dimensional ordered subset with 2 iterations and 10 subsets. Then, a Gaussian filter of size 6 mm was applied. The DL images were reconstructed with 4 iterations and 10 subsets, and DL-based noise reduction was applied. Next, for semi-quantitative assessment, voxels of interest (VOIs) were calculated. The ^{18}F -FDG PET/CT scans were analyzed using a commercially available dedicated Vox-base SP1000 workstation (J-MAC Systems, Sapporo, Japan).

Deep learning-based approach

Our DL-based approach is an application of deep convolutional neural network (DCNN), which comprises a network training step and a denoising step. In the training step, we prepared a large amount of low-quality and high-quality image pairs in which network parameters are optimized to map a noisy (low-quality) image to a low noise

Table 1 Patient demographic and clinical data ($n = 50$)

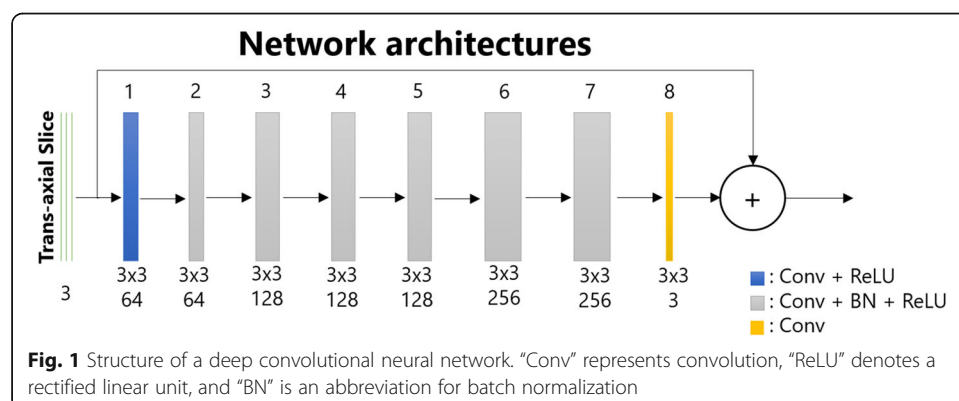
Age, years	64.9 ± 13.9
Sex	
Male	26
Female	24
Weight, kg	61.9 ± 10.9
Disease, n	
Malignancy	
Head and neck tumor	13
Lung cancer	10
Lymphoma	10
Breast cancer	6
Uterine cancer	2
Pancreatic cancer	2
Cancer of unknown primary	2
Others *1	5
Time delay, min	63.7 ± 6.7
Blood sugar level (mg/dl)	115.6 ± 16.0

*1, multiple myeloma, esophageal cancer, colon cancer, malignant melanoma, Takayasu aortitis

(high-quality) image. The low-quality and high-quality images are also referred to as training images and target images, respectively. The training dataset constituted of 6 patient studies of ^{18}F -FDG lung scans with 2 beds coverage (lung to upper abdomen) and $2 \times ^{18}\text{F}$ -FDG brain studies. The lung scans were acquired for 14 min/bed while the brain studies were scanned for 15 min/bed. These full scans were used as the target images. We then uniformly subsampled the listmode data into 8 noise levels as 30, 45, 60, 120, 180, 240, 300, and 420 s/bed as the noisy training samples. For example, for rebinning a 10-min listmode data into a count level equivalent to a 5-min acquisition, every other event is removed in the listmode data. In such case, these noisy samples share the same clean latent image but are corrupted with different magnitudes of noise. These scan durations cover count levels beyond the normal range seen in clinical practice. In the training process, all the inputs at different noise levels are paired with the same high-quality target [18]. The rationale behind this is that all the noisy samples share the same clean latent image but are corrupted with different levels of noise. The network filters are therefore optimized to estimate the noise residual from different noise levels in the input images. In such a way, the CNN can learn to adapt to different noise levels in the input image automatically and can always produce consistent high-quality images. The training images were reconstructed with an OSEM reconstruction algorithm incorporating time-of-flight (TOF), point spread function (PSF), attenuation, and scatter corrections. All the images were reconstructed with 4 iterations and 10 subsets. The reconstructed image dimension was $272 \times 272 \times 144$ for the lung studies and $120 \times 120 \times 81$ for the brain studies with 2-mm voxel dimension. Before training and testing, all the images were converted to a standardized uptake value (SUV), a quantity that is normalized based on the injected radioactivity and the subject's body weight. The SUV of normal tissue should be around 1, while a SUV of 2.5 or higher is generally indicative of malignant tissue. All these studies constitute 9234 2D training slices in total. The testing studies are reconstructed using the same protocol and converted to SUV prior to being fed into the network.

Deep CNN architecture

Our DL method consists of a DCNN with 8 layers as shown in Fig. 1. Here, we chose the residual network architecture that tries to estimate the noise \mathbf{n} instead of high-



quality image \mathbf{x} out of low-quality image $\mathbf{y} = \mathbf{x} + \mathbf{n}$, as proposed in the previous study [19]. The optimization process then tries to minimize the loss function L :

$$L(\Theta) = \frac{1}{N} \sum_{i \in N} \psi(\mathcal{F}(\mathbf{y}_i; \Theta) - (\mathbf{y}_i - \mathbf{x}_i)) \quad (1)$$

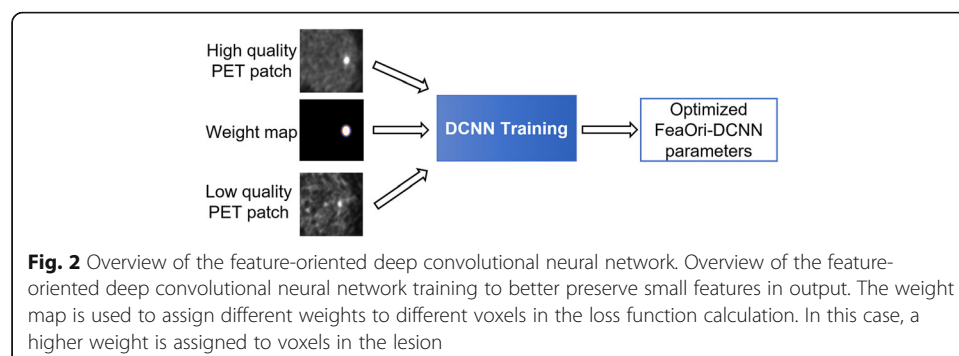
where Θ denotes the trainable weights, $\mathcal{F}(\mathbf{y}_i; \Theta)$ is the estimation of noise \mathbf{n} , ψ is the error function chosen as the mean square error, N represents the number of training images, \mathbf{y} denotes the training images, and \mathbf{x} denotes the target images.

In conventional DCNN training, the loss function equally weights all voxel-wise differences in image patches. We have proposed a feature-oriented approach in the DCNN training, which uses weight maps to steer the training toward contrast preservation for small features. Phantom and patient studies have demonstrated that this approach can effectively improve contrast recovery on small and low contrast lesions, when the number of high-quality training dataset with small lesions is limited [15].

As shown in Fig. 2, a weight map that assigns different weights to different voxels is provided in the network along with the low-quality input and high-quality target. The weight map forces the network to learn to preserve the desired small features while suppressing noise in the background. With the weight map, the loss function becomes:

$$L(\Theta) = \frac{1}{N} \sum_{i \in N} \psi((\mathcal{F}(\mathbf{y}_i; \Theta) - (\mathbf{y}_i - \mathbf{x}_i)) \cdot \mathbf{w}_i) \quad (2)$$

where \mathbf{w}_i denotes the weight map and the \cdot operator denotes voxel-wise multiplication. To generate the weight maps, we first segment the lesions in the target images in the training dataset using adaptive thresholding in order to create the lesion mask (Fig. 3). The weights of the background voxels are set to unity. The voxels in the lesion, which are more important, are set to 10. This value is determined empirically based on the tradeoff between preservation of lesion contrast versus noise reduction in our previous experiments. The weight map is then convolved with a Gaussian kernel to accommodate segmentation errors. The weight map is uniform for the patches that do not contain lesions. The trans-axial slices of sample training targets and the corresponding weight maps are shown in Fig. 3. All the lesions that used to generate the weight maps are located in the lung and liver. Please note the weight map is only needed in the training process that can be casted as a form of regularization in the loss function.



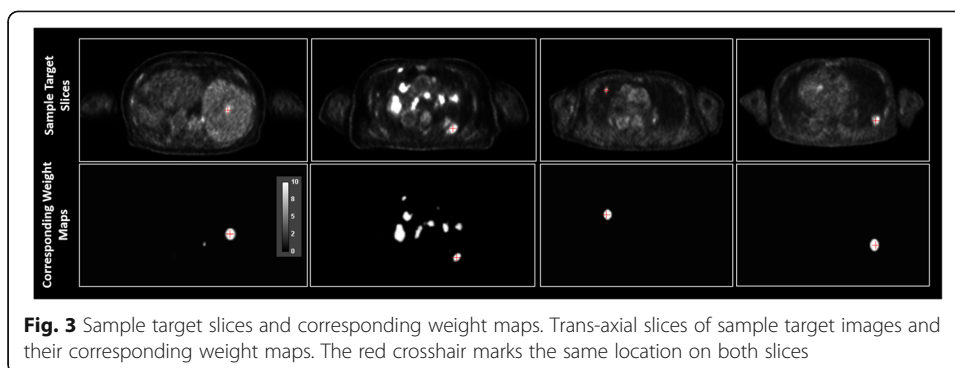


Fig. 3 Sample target slices and corresponding weight maps. Trans-axial slices of sample target images and their corresponding weight maps. The red crosshair marks the same location on both slices

Indeed, it prevents the network training overfitting to the noise in the input images. During inference, no weight map is needed.

The inputs and outputs of the networks are 3-channels (3 consecutive slices), which can be considered as a 2.5D convolution process. With stride 1 in the axial direction, some slices may be processed 2 or 3 times. The outputs slices are weight averaged in the axial redirection to form the final volume. During training, patches with a dimension of $32 \times 32 \times 3$ were extracted from the input images with 16, 16, and 2 pixels overlapping in the x , y , and z directions, respectively. During testing, the kernels were applied to the whole image slice. The convolutional kernel dimension was 3×3 in order to capture local noise distribution and small features instead of global features. The L2 norm was chosen as the loss function in order to account for high-intensity noise spikes and to obtain stable solution. The Adaptive Moment Estimation (Adam) algorithm optimizer was used to minimize the loss function. The network was trained for 450 total epochs with a gradually decreasing learning rate, i.e., 0.01 for 350 epochs, 0.001 for 50 epochs, and 0.0001 for the last 50 epochs. To validate the training results, we reserved 1% of the training samples for validation during training. These training samples were not used to train the network. Instead, they were used to monitor the performance of the network during training. We computed the loss function using the validation data set and compared to the loss function computed from the training dataset. The network was trained in a Matlab environment using the MatConvNet toolbox [20].

Qualitative analysis

The acquired images were independently reviewed and analyzed using the Vox-base SP1000 workstation. All PET images were blindly evaluated by two experienced nuclear medicine physicians (with 10 and 8 years of experience, respectively, in interpreting PET scans). The readers were not aware of the clinical indication for PET/CT.

Readers were allowed to manually adjust the standard window settings. Subsequently, based on 5-point scales, the following quality criteria were assessed: tumor delineation (ranging from 1 = lesion cannot be confirmed to 5 = excellent delineation of the lesion margin), overall image quality (ranging from 1 = poor overall image quality to 5 = excellent overall image quality), and image noise (ranging from 1 = enormous image noise to 5 = no perceivable image noise) [21, 22]. In the event of large differences in assessment between readers, the specific images were discussed in a consensus meeting. Readers were instructed to record when they detected any artifact or failure.

Quantitative analysis

For semi-quantitative analyses, one nuclear medicine physician placed 0.5-mL spherical VOIs in healthy tissues (aortic arch, semioval center [white matter], lung, left ventricle of the heart, parotid gland, quadriceps femoris muscle, and spleen) and also placed a 3-mL spherical VOI in the liver for reference tissue purposes. From these VOIs, the maximum, peak, and mean standardized uptake values (SUV_{max}, SUV_{mean}, and SUV_{peak}, respectively) were obtained. VOIs were also placed in the different lesions per patient (with an overall maximum of five lesions per patient, with a maximum of two in the same tissue type). From these measurements, the different SUV_{max}, SUV_{peak}, and SUV_{mean} values were compared between the two images. The normalized values of SUV_{max}, SUV_{peak}, and SUV_{mean} with reference to the aortic arch were calculated and compared between the two images.

Statistical analysis

Scores for the images acquired using the two methods were compared pairwise using a two-tailed paired samples *t*-test. For inter-reader agreement regarding tumor delineation, overall image quality, and image noise, the original 5-point scores were reassigned to 3-point scores (1 + 2 became 1, 3 became 2, and 4 + 5 became 3). Inter-reader agreement was subsequently evaluated using the kappa statistic. SUV parameters in healthy tissues and lesions were compared between the different systems using a two-tailed paired samples *t*-test. The statistical analysis was performed using SPSS for Windows (IBM Corp., Armonk, NY, USA). A *P*-value < 0.05 was considered statistically significant.

Results

Images using DL were scored significantly higher for tumor delineation, overall image quality, and image noise than at baseline (*P* < 0.001; Table 2). The Fleiss' kappa value for the overall inter-reader agreement was 0.78. In most of the healthy tissues, the SUVs measured with the DL method were higher than those measured with standard reconstruction (*P* = 0.456 to < 0.001; Table 3). We detected lesions in the region including the brain, parotid gland, thyroid, pharynx, lung, breast, liver, bile duct, pancreas, intestine, lymph nodes (cervical, subclavian, axillary, mediastinum, hepatic, paraaortic, and inguinal), subcutaneous area, and bones (rib, spine, and pelvic bones). The sizes of the lesions in the maximal axial plain ranged from 0.52 to 102.21 cm³. The difference between DL images and Gaussian filtered images was more significant in tumor tissues (*P* = 0.31 to < 0.001; Table 4). The difference in normalized values between DL images and Gaussian filtered images was also significant (*P* = 0.09 to < 0.001;

Table 2 Qualitative image analysis

	dPET	cPET	<i>P</i> -value
Delineation	4.06 ± 0.24	2.94 ± 0.24	< 0.0001
Noise	3.88 ± 0.48	2.40 ± 0.50	< 0.0001
Overall image quality	3.94 ± 0.44	2.98 ± 0.20	< 0.0001

The data are shown as the mean and standard deviation. cPET, conventional positron emission tomography; dPET, deep learning processed positron emission tomography

Table 3 SUVs in healthy organ tissues

Organs		cPET mean \pm SD	dPET mean \pm SD	P-value
Aortic arch	SUV max	1.63 \pm 0.26	1.72 \pm 0.26	< 0.0001
	SUV mean	1.34 \pm 0.20	1.44 \pm 0.20	< 0.0001
	SUV peak	1.48 \pm 0.21	1.57 \pm 0.21	< 0.0001
Semioval center	SUV max	2.28 \pm 0.43	2.19 \pm 0.39	0.04
	SUV mean	1.85 \pm 0.45	1.76 \pm 0.37	0.003
	SUV peak	2.28 \pm 0.43	2.37 \pm 0.47	< 0.0001
Liver	SUV max	2.25 \pm 0.43	2.09 \pm 0.27	< 0.0001
	SUV mean	1.76 \pm 0.33	1.81 \pm 0.30	0.456
	SUV peak	2.00 \pm 0.37	1.92 \pm 0.24	< 0.0001
Lung	SUV max	0.33 \pm 0.30	0.31 \pm 0.09	0.033
	SUV mean	0.26 \pm 0.23	0.24 \pm 0.07	< 0.0001
	SUV peak	0.31 \pm 0.27	0.29 \pm 0.08	0.006
Left ventricle	SUV max	1.62 \pm 0.29	1.64 \pm 0.27	0.19
	SUV mean	1.32 \pm 0.24	1.41 \pm 0.27	< 0.0001
	SUV peak	1.59 \pm 0.32	1.63 \pm 0.30	0.006
Parotid gland	SUV max	1.33 \pm 0.37	1.49 \pm 0.43	< 0.0001
	SUV mean	1.11 \pm 0.30	1.20 \pm 0.33	< 0.0001
	SUV peak	1.22 \pm 0.34	1.30 \pm 0.37	< 0.0001
Quadriceps muscle	SUV max	0.64 \pm 0.16	0.67 \pm 0.17	< 0.0001
	SUV mean	0.50 \pm 0.12	0.54 \pm 0.12	< 0.0001
	SUV peak	0.58 \pm 0.13	0.60 \pm 0.13	< 0.0001
Spleen	SUV max	1.87 \pm 0.30	1.92 \pm 0.31	0.023
	SUV mean	1.59 \pm 0.26	1.30 \pm 0.37	< 0.0001
	SUV peak	1.70 \pm 0.26	1.69 \pm 0.24	< 0.0001

The data are shown as the mean and standard deviation. Note: The current DCNN is trained to be used only for general whole-body studies but is not designed for the brain. *cPET*, conventional positron emission tomography; *dPET*, deep learning processed positron emission tomography; *SD*, standard deviation; *SUV*, standardized uptake value

Table 5). Representative cases are shown in Figs. 4 and 5. No artifact or failure was detected.

Discussion

This study evaluated the effect of the deep learning technique on the quality of FDG PET/CT images. In DL images, the image noise was lower, and the lesion delineation and image quality were superior compared to conventional reconstruction. DL enhanced image quality and reduced image noise.

A conventional spatial filter, such as the Gaussian filter, involves mixing of pixel values and the amount or ratio of mixture is determined simply by the distance

Table 4 SUVs in tumor lesions ($n = 108$)

		cPET mean \pm SD	dPET mean \pm SD	P-value
Lesions	SUV max	5.59 \pm 3.97	8.42 \pm 5.02	< 0.0001
	SUV mean	1.98 \pm 0.96	2.14 \pm 0.83	< 0.0001
	SUV peak	4.04 \pm 2.93	4.69 \pm 3.07	0.31

The data are shown as the mean and standard deviation. *cPET*, conventional positron emission tomography; *dPET*, deep learning processed positron emission tomography; *SD*, standard deviation

Table 5 Difference in SUV ratios compared to the background SUV in tumor lesions ($n = 108$)

		cPET Mean \pm SD	dPET Mean \pm SD	P-value
Lesions	SUV max ratio	3.42 \pm 2.32	4.95 \pm 3.26	< 0.0001
	SUV mean ratio	1.49 \pm 0.73	1.52 \pm 0.69	0.09
	SUV peak ratio	2.81 \pm 2.05	3.09 \pm 2.26	< 0.0001

The data are shown as the mean and standard deviation. cPET, conventional positron emission tomography; dPET, deep learning processed positron emission tomography; SD, standard deviation

between two pixels. This process results in unwanted blurring of organs and reduction of the SUV value of regions with small concentration in both tumor and healthy organs. The principle underlying the deep learning noise reduction process is different from the conventional convolution filter approach. When noisy data are used as the training input and high-quality data are presented as the training target, the network can learn to produce clean images from noise-contaminated images. Our training target dataset for DCNN comprised high-quality images with a longer acquisition time. DL imaging rendered a higher image quality than conventional imaging with Gaussian filtering. The difference between the conventional filter approach and DL approach results in a difference in the SUV parameters. For tumors and healthy tissues in small organs, Gaussian filtering shaves off pixel values in the images, while DL only removes noise. This difference results in a higher SUVmean in most concentrations. When we consider relatively larger organs like the liver, excluding its edge, both Gaussian filtering and DL do not change the overall SUV level. Thus, the lower SUVmax and SUVpeak of DL demonstrate that DL has stronger denoising capability than Gaussian filtering.

In the qualitative assessment, we compared DL PET images with conventional PET images. DL reduced the image noise and improved the image quality. Schaefferkoetter et al. demonstrated that the denoising method using CNN improved very noisy data.

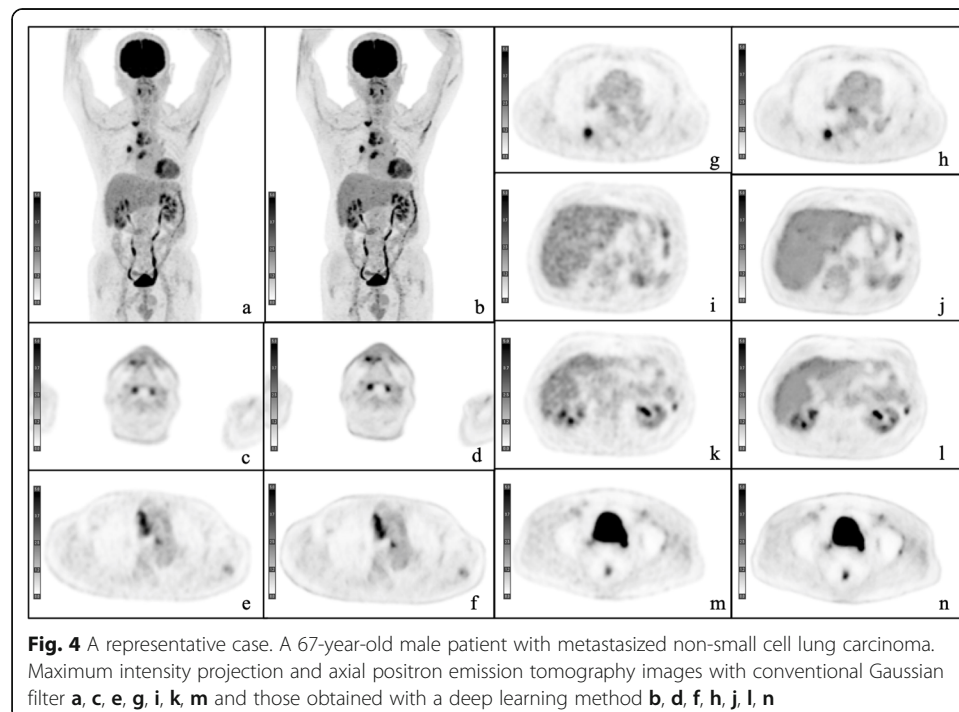
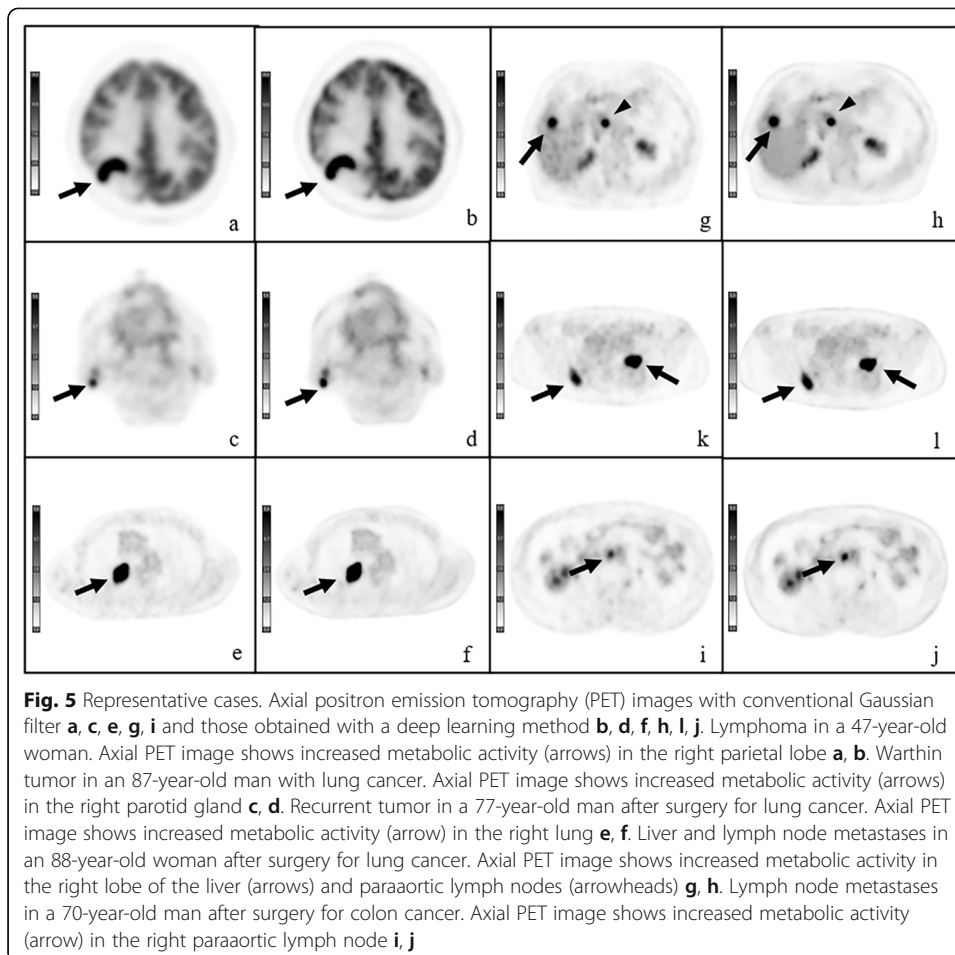


Fig. 4 A representative case. A 67-year-old male patient with metastasized non-small cell lung carcinoma. Maximum intensity projection and axial positron emission tomography images with conventional Gaussian filter **a, c, e, g, i, k, m** and those obtained with a deep learning method **b, d, f, h, j, l, n**



However, the lesser effect was appreciated by applying the technique to routinely encountered images [17]. The clinical impact of the DL method has to be investigated in future studies. Moreover, the application of the DL technique used in the current study may lead to a reduction in the radiation dose and a shorter acquisition time. Cheng et al. enhanced the image quality of 2 one-hundredth dose PET to that of standard-dose PET using the deep learning method [13]. Further studies are necessary to determine whether or not the diagnostic accuracy of imaging with DL methods used in this study is retained at reduced radiation doses.

PET images have a low SNR so need denoising. The commonly used technique is Gaussian filtering, which is sometimes implemented in iterative image reconstruction algorithms [23]. Gaussian filtering can increase the SNR. However, it also smooths the image and can produce a loss of resolution by averaging voxels together and blurring the distinction between two closely adjacent objects.

The network used in this study was specifically trained and applied on ^{18}F -FDG PET studies acquired on Canon Medical's Celesteion PET/CT scanner. The application of this network to other imaging tracers or scanners may not produce equivalent results as presented in this study, which requires further investigation. In addition, we have not investigated the impact of dose/scan duration reduction on the diagnostic quality of the network results, which also warrants future investigation.

Conclusions

PET imaging is already characterized by limited spatial resolution and associated partial volume effects. The DL methods used in this study enhanced the image quality by maintaining the values of the SUV parameters. In this regard, the DL technique is superior to conventional reconstruction.

In this initial study, we found that the DL method provides better perceived image quality than conventional imaging with a Gaussian filter; more sharply demarcated tumor lesions were seen, the overall image quality was higher, and a higher SNR was assessed visually. In terms of semi-quantitative image quality, the DL technique renders higher values for SUV parameters on imaging of tumors and healthy tissues in small organs, while Gaussian filtering decreases the SUV by blurring. Our results demonstrate that DL respects the tissue boundaries well and reduces the noise considerably without losing quantitative information on PET images, including SUV_{max} and SUV_{mean}, when compared with images using a Gaussian filter. However, improved quantitative performance may be feasible using clinically optimized reconstruction settings. Future studies that include groups of patients with more homogeneous oncologic disease are necessary to validate our findings and to assess the potential clinical impact of the DL method on PET imaging.

Abbreviations

CNN: Convolutional neural network; CT: Computed tomography; DL: Deep learning; DCNN: Deep convolutional neural network; FDG: Fluorodeoxyglucose; PET: Positron emission tomography; SNR: Signal-to-noise ratio; SPECT: Single-photon emission tomography; SUV: Standardized uptake value; VOI: Voxel of interest

Acknowledgements

Not applicable.

Authors' contributions

JT: conceptualization, formal analysis, and investigation; project administration; writing of the original draft; and visualization. KYo: validation and investigation. KYa: validation and investigation. RW: writing, review, and editing. KK: writing, review, and editing. CC: resources and data processing. EA: resources and data processing. MK: writing, review, and editing. UT: supervision. The authors read and approved the final manuscript.

Authors' information

Not applicable.

Funding

The present study was funded by Canon Medical Systems Corporation, Otawara, Tochigi, Japan.

Availability of data and materials

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The current study was approved by the Medical Ethics Committee of Tokyo Medical and Dental University in 2019. Patients consent to the use of their personal data for research purposes.

Consent for publication

Informed consent from all the patients was obtained.

Competing interests

Chung Chan and Evren Asma are employees of Canon Medical Research USA, Inc., Vernon Hills, IL, USA. The other authors declare that they have no conflicts of interest.

Author details

¹Department of Diagnostic Radiology and Nuclear Medicine, Tokyo Medical and Dental University, 1-5-45 Yushima, Bunkyo-ku, Tokyo 113-8510, Japan. ²Canon Medical Research USA, Inc., 706 N. Deerpath Drive, Vernon Hills, IL 60061, USA.

Received: 6 August 2020 Accepted: 12 March 2021

Published online: 25 March 2021

References

1. Heck RK, O'Malley AM, Kellum EL, Donovan TB, Ellzey A, Witte DA. Errors in the MRI evaluation of musculoskeletal tumors and tumorlike lesions. *Clin Orthop Relat Res.* 2007;459:28–33. <https://doi.org/10.1097/BLO.0b013e3180485681>.
2. Takamochi K, Yoshida J, Murakami K, Niho S, Ishii G, Nishimura M, Nishiwaki Y, Suzuki K, Nagai K, et al. Pitfalls in lymph node staging with positron emission tomography in non-small cell lung cancer patients. *Lung Cancer.* 2005;47(2):235–42. <https://doi.org/10.1016/j.lungcan.2004.08.004>.
3. van der Vos CS, Koopman D, Rijnsdorp S, Arends AJ, Boellaard R, van Dalen JA, Lubberink M, Willemsen ATM, Visser EP, et al. Quantification, improvement, and harmonization of small lesion detection with state-of-the-art PET. *Eur J Nucl Med Mol Imaging.* 2017;44(Suppl 1):4–16. <https://doi.org/10.1007/s00259-017-3727-z>.
4. Knesaurek K, Machac J, Vallabhajosula S, Buchsbaum MS. A new iterative reconstruction technique for attenuation correction in high-resolution positron emission tomography. *Eur J Nucl Med.* 1996;23(6):656–61. <https://doi.org/10.1007/BF00834527>.
5. Shepp LA, Vardi Y. Maximum likelihood reconstruction for emission tomography. *IEEE Trans Med Imaging.* 1982;1(2):113–22. <https://doi.org/10.1109/TMI.1982.4307558>.
6. Vandenberghe S, Mikhaylova E, D'Hoe E, Mollet P, Karp JS. Recent developments in time-of-flight PET. *EJNMMI Phys.* 2016;3(1):3. <https://doi.org/10.1186/s40658-016-0138-3>.
7. Slomka PJ, Pan T, Germano G. Recent advances and future progress in PET instrumentation. *Semin Nucl Med.* 2016;46(1):5–19. <https://doi.org/10.1053/j.semnuclmed.2015.09.006>.
8. Kidoh M, Shinoda K, Kitajima M, Isogawa K, Nambu M, Uetani H, Morita K, Nakaura T, Tateishi M, Yamashita Y, Yamashita Y, et al. Deep learning based noise reduction for brain MR imaging: tests on phantoms and healthy volunteers. *Magn Reson Med Sci.* 2019;19(3):195–206. <https://doi.org/10.2463/mrms.mp.2019-0018>.
9. Akagi M, Nakamura Y, Higaki T, Narita K, Honda Y, Zhou J, Yu Z, Akino N, Awai K. Deep learning reconstruction improves image quality of abdominal ultra-high-resolution CT. *Eur Radiol.* et al, 2019;29(11):6163–71. <https://doi.org/10.1007/s00330-019-06170-3>.
10. Tatsugami F, Higaki T, Nakamura Y, Yu Z, Zhou J, Lu Y, Fujioka C, Kitagawa T, Kihara Y, Iida M, Awai K, et al. Deep learning-based image restoration algorithm for coronary CT angiography. *Eur Radiol.* 2019;29(10):5322–9. <https://doi.org/10.1007/s00330-019-06183-y>.
11. Hauptmann A, Arridge S, Lucka F, Muthurangu V, Steeden JA. Real-time cardiovascular MR with spatio-temporal artifact suppression using deep learning-proof of concept in congenital heart disease. *Magn Reson Med.* 2019;81(2):1143–56. <https://doi.org/10.1002/mrm.27480>.
12. Dietze MMA, Branderhorst W, Kunnen B, Viergever MA, de Jong H. Accelerated SPECT image reconstruction with FBP and an image enhancement convolutional neural network. *EJNMMI Phys.* 2019;6(1):14. <https://doi.org/10.1186/s40658-019-0252-0>.
13. Chen KT, Gong E, de Carvalho Macruz FB, Xu J, Boumis A, Khalighi M, Poston KL, Sha SJ, Greicius MD, Mormino E, Pauly JM, Srinivas S, Zaharchuk G, et al. Ultra-low-dose (18)F-florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs. *Radiology.* 2019;290(3):649–56. <https://doi.org/10.1148/radiol.2018180940>.
14. Ouyang J, Chen KT, Gong E, Pauly J, Zaharchuk G. Ultra-low-dose PET reconstruction using generative adversarial network with feature matching and task-specific perceptual loss. *Med Physics.* 2019;46(8):3555–64. <https://doi.org/10.1002/mp.13626>.
15. 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.* 2019;3(2):153–61. <https://doi.org/10.1109/trpms.2018.2877644>.
16. Xu J GE, Pauly J, Zaharchuk. 200x low-dose PET reconstruction using deep learning. 2017.
17. Schaefferkoetter J, Yan J, Ortega C, Sertic A, Lechtman E, Eshet Y, Metser U, Veit-Haibach P. Convolutional neural networks for improving image quality with noisy PET data. *EJNMMI Res.* et al, 2020;10(1):105. <https://doi.org/10.1186/s13550-020-00695-1>.
18. Chan C, Zhou J, Yang L, Qi W, Kolthammer J, Asma E. Feature oriented deep convolutional neural network for PET image denoising. 2018 IEEE Nuclear Sci Symp Med Imag Conf Proceed. 2018:1–4.
19. Zhang K, Zuo W, Chen Y, Meng D, Zhang L. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising. *IEEE Trans Image Process.* 2017;26(7):3142–55. <https://doi.org/10.1109/tip.2017.2662206>.
20. Vedaldi A, Lenc K. MatConvNet: convolutional neural networks for MATLAB. Proceedings of the 23rd ACM international conference on Multimedia. Brisbane: Association for Computing Machinery; 2015. p. 689–92.
21. Sonni I, Baratto L, Park S, Hatami N, Srinivas S, Davidzon G, Gambhir SS, Iagaru A, et al. Initial experience with a SiPM-based PET/CT scanner: influence of acquisition time on image quality. *EJNMMI Physics.* 2018;5(1):9. <https://doi.org/10.1186/s40658-018-0207-x>.
22. Lenga L, Czwikla R, Wichmann JL, Leithner D, Albrecht MH, Booz C, Arendt CT, Yel I, D'Angelo T, Vogl TJ, Martin SS, et al. Dual-energy CT in patients with colorectal cancer: improved assessment of hypoattenuating liver metastases using noise-optimized virtual monoenergetic imaging. *Eur J Radiol.* 2018;106:184–91. <https://doi.org/10.1016/j.ejrad.2018.07.027>.
23. Alessio A, Kinahan P. PET image reconstruction. *Nucl Med.* 2006;1:1–22.

Publisher's Note

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