IPEM Institute of Physics and Engineering in Medicine

PAPER

Joint correction of attenuation and scatter in image space using deep convolutional neural networks for dedicated brain ^{18}F-FDG PET

View the article online for updates and enhancements.

Physics in Medicine & Biology

IPEM Institute of Physics and Engineering in Medicine

CrossMark

PAPER

Joint correction of attenuation and scatter in image space using deep convolutional neural networks for dedicated brain ^{18}F-FDG PET

Jaewon Yang[1,3,4], Dookun Park[2], Grant T Gullberg[1] and Youngho Seo[1]

1 Physics Research Laboratory, Department of Radiology and Biomedical Imaging, University of California San Francisco, San Francisco, CA, United States of America
2 Microsoft, Bellevue, WA, United States of America
3 UCSF Physics Research Laboratory, 185 Berry Street, Suite 350, San Francisco, CA 94143-0946, United States of America
4 Author to whom any correspondence should be addressed.

E-mail: jaewon.yang@ucsf.edud

Abstract

Dedicated brain positron emission tomography (PET) devices can provide higher-resolution images with much lower doses compared to conventional whole-body PET systems, which is important to support PET neuroimaging and particularly useful for the diagnosis of neurodegenerative diseases. However, when a dedicated brain PET scanner does not come with a combined CT or transmission source, there is no direct solution for accurate attenuation and scatter correction, both of which are critical for quantitative PET. To address this problem, we propose joint attenuation and scatter correction (ASC) in *image space* for non-corrected PET (PET$_{NC}$) using deep convolutional neural networks (DCNNs). This approach is a one-step process, distinct from conventional methods that rely on generating attenuation maps first that are then applied to iterative scatter simulation in *sinogram space*. For training and validation, time-of-flight PET/MR scans and additional helical CTs were performed for 35 subjects (25/10 split for training and test dataset). A DCNN model was proposed and trained to convert PET$_{NC}$ to DCNN-based ASC PET (PET$_{DCNN}$) directly in *image space*. For quantitative evaluation, uptake differences between PET$_{DCNN}$ and reference CT-based ASC PET (PET$_{CT-ASC}$) were computed for 116 automated anatomical labels (AALs) across 10 test subjects (1160 regions in total). MR-based ASC PET (PET$_{MR-ASC}$), a current clinical protocol in PET/MR imaging, was another reference for comparison. Statistical significance was assessed using a paired t test. The performance of PET$_{DCNN}$ was comparable to that of PET$_{MR-ASC}$, in comparison to reference PET$_{CT-ASC}$. The mean SUV differences (mean $\pm$ SD) from PET$_{CT-ASC}$ were 4.0% $\pm$ 15.4% ($P < 0.001$) and $-4.2\% \pm 4.3\%$ ($P < 0.001$) for PET$_{DCNN}$ and PET$_{MR-ASC}$, respectively. The overall larger variation of PET$_{DCNN}$ (15.4%) was prone to the subject with the highest mean difference (48.5% $\pm$ 10.4%). The mean difference of PET$_{DCNN}$ excluding the subject was substantially improved to $-0.8\% \pm 5.2\%$ ($P < 0.001$), which was lower than that of PET$_{MR-ASC}$ ($-5.07\% \pm 3.60\%, P < 0.001$). In conclusion, we demonstrated the feasibility of directly producing PET images corrected for attenuation and scatter using a DCNN (PET$_{DCNN}$) from PET$_{NC}$ in *image space* without requiring conventional attenuation map generation and time-consuming scatter correction. Additionally, our DCNN-based method provides a possible alternative to MR-ASC for simultaneous PET/MRI.

1. Introduction

Positron emission tomography (PET) enables an understanding of biochemical changes in the brain at early stages of disease prior to structural changes or clinical symptoms. PET neuroimaging is particularly useful in the diagnosis of neurodegenerative diseases (e.g. Alzheimer's disease, frontotemporal dementia, dementia with Lewy bodies, Parkinson's disease, and Huntington's disease) such as diseases associated with significant changes in brain metabolism (Silverman and Alavi 2005). The most commonly used tracer in clinical brain PET imaging is ^{18}F-fluorodeoxyglucose (FDG) but the development of other radiotracers such as dopaminergic, amyloid plaque and tau imaging agents is still an active development area (Brown *et al* 2014).

To keep pace with the expanding demand for PET neuroimaging, dedicated high-resolution brain PET systems such as high-resolution research tomographs (HRRTs) (Wienhard *et al* 2002) and CareMiBrain (Oncovision) have been introduced in the market for human brain imaging, supporting significantly improved high-resolution images typically with a lower dose compared to a conventional whole-body PET system. Additionally, wearable PET scanners such as ambulatory micro-dose PET (AM-PET) (Melroy *et al* 2017) and helmet-PET (Tashima and Yamaya 2016) have been developed and are being evaluated for functional brain research. However, these dedicated or wearable PET systems do not provide direct solutions for attenuation and scatter correction (ASC). These systems are not combined with computed tomography (CT) or magnetic resonance imaging (MRI) that can provide attenuation maps used for ASC. Thus, a challenge for dedicated brain PET systems is to develop a practical and robust method for accurate ASC without an additional imaging modality, such as CT or MRI.

ASC is critical for quantitative accuracy as well as image quality in PET (Meikle and Badawi 2005). Attenuated and scattered events occur due to photoelectric effects and Compton scattering induced by the presence of dense material along lines of response (LORs). Without attenuation correction, regions near the skin appear darker (emitting more photons) and regions surrounding brain tissues appear brighter (emitting less photons). The scatter fraction can reach 50% to 60% of LORs recorded in whole-body 3D PET and, without scatter correction, LORs recorded outside an object boundary due to scatter contribute noise in image reconstruction. Therefore, it is important to compensate for attenuation and scatter for quantitative PET. In a hybrid PET/CT or PET/MR imaging, the current implementation for attenuation correction is to transform CT (Lonn *et al* 2003) or MR-derived pseudo-CT images into attenuation maps (Wollenweber *et al* 2013, Berker and Li 2016, Yang *et al* 2017b); while the current implementation for scatter correction is to estimate scatter iteratively by a 3D-model-based simulation using down-sampled attenuation and emission images (Watson *et al* 2004, Iatrou *et al* 2006, Zaidi and Montandon 2007). Both ASCs are separately performed due to the difference in photoelectric effects and Compton scattering (Meikle and Badawi 2005) in *sinogram space* where LORs are conventionally recorded as counts according to their locations and orientations and then reconstructed to PET images by an ordered subset expectation maximization (OSEM) algorithm (Defrise *et al* 2005).
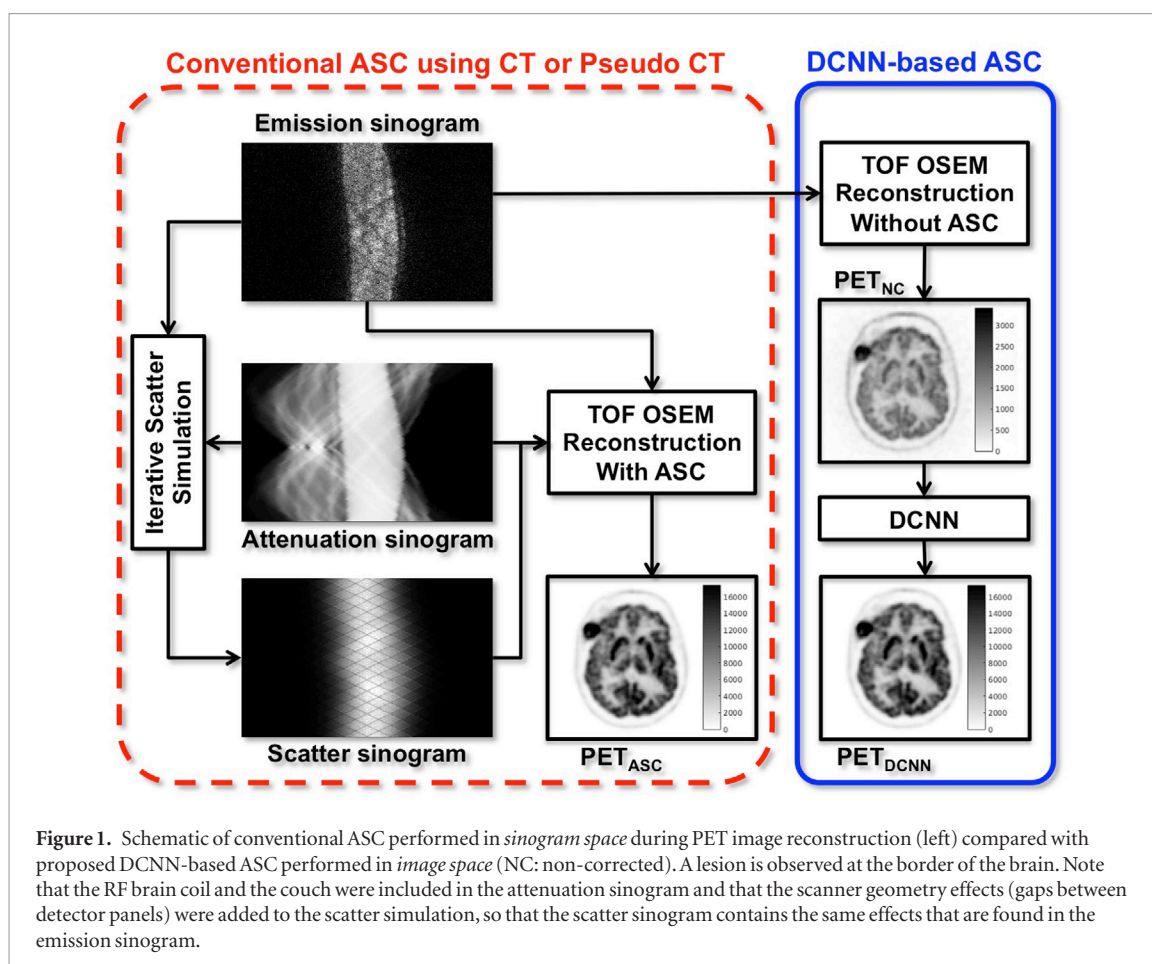
Recently, deep convolutional neural networks (DCNNs) have been widely applied to medical imaging based on the success of deep learning to computer vision tasks (Ronneberger *et al* 2015). DCNN demonstrated the direct conversion to pseudo-CT from T1-weighted MR (Liu *et al* 2018a), Dixon and Zero-TE MR (Gong *et al* 2018) or non-attenuation-corrected PET (Liu *et al* 2018b) for attenuation correction in PET neuroimaging. Also, DCNNs have been demonstrated to improve the quality of noisy attenuation maps generated by simultaneous maximum-likelihood reconstruction of activity and attenuation by time-of-flight (TOF) information (Hwang *et al* 2018).

In this paper, we propose a new approach for joint ASC in *image space* using only non-corrected PET images without depending on another imaging modality and performing a scatter simulation. Since brain tissues and their boundaries (e.g. white and gray matters, skin, bone, etc) are perceptible in non-corrected PET, DCNNs can extract important patterns successfully for joint ASC in PET neuroimaging. The proposed joint ASC is a one-step process, distinct from conventional methods that rely on generating attenuation maps first that are then applied to iterative scatter simulation in *sinogram space*.

2. Methods

2.1. Patient information

The patient study was approved by the Institutional Review Board, and all patients signed an informed consent form before the examinations. Thirty-five patients (16 male and 19 female) underwent whole-body ^{18}F-FDG PET/MRI and helical CT scans. The average patient age was 57.7 ± 11.5 y (range 29–76 years), the average weight was 73.7 ± 17.4 kg (range 39.5–109.8 kg), and the average administered dose of ^{18}F-FDG was 308.5 ± 74.6 MBq (range 170.2–468.1 MBq). The average scan duration of the whole brain was 227.2 ± 137.5 s (range 135–900 s), and the average time difference between injection and scan was 150.8 ± 24.5 min (range 111.0–190.1 min). A tumor was observed in the head for only one subject who was included in the test set (figure 1).

Figure 1. Schematic of conventional ASC performed in *sinogram space* during PET image reconstruction (left) compared with proposed DCNN-based ASC performed in *image space* (NC: non-corrected). A lesion is observed at the border of the brain. Note that the RF brain coil and the couch were included in the attenuation sinogram and that the scanner geometry effects (gaps between detector panels) were added to the scatter simulation, so that the scatter sinogram contains the same effects that are found in the emission sinogram.
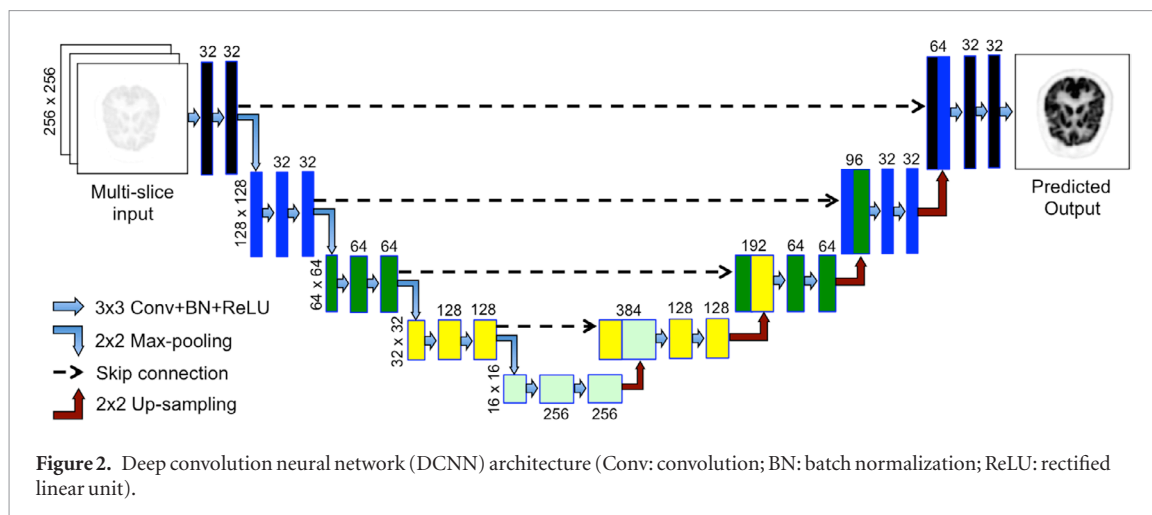
2.2. PET/MRI and CT data acquisition

TOF PET/MRI examinations were performed on a SIGNA PET/MR scanner (GE Healthcare). PET had a 600 mm transaxial field of view (FOV) and 250 mm axial FOV, with a TOF timing resolution of approximately 400 ps and average measured sensitivity of 22.65 cps kBq^{-1}. While PET data were acquired, Dixon MR (FOV 500 × 500 × 312 mm; resolution 1.95 × 1.95 mm; slice thickness 5.2 mm; slice spacing 2.6 mm; scan time 18 s) sequences were acquired for MR-based attenuation and scatter correction (MR-ASC), using the head and neck coil array. Helical CT images of the patients were acquired on a Discovery PET/CT (GE Healthcare) or Biograph HiRez 16 (Siemens Healthcare) scanner with variable parameter settings (120 kVp; 105–599 mA; rotation time 0.5 s; pitch 0.98 and 0.75; rotation 39.37 and 34.45 mm; axial FOV 700 and 500 mm; slice thickness 3.75 and 5.00 mm; matrix size 512 × 512; and voxel sizes 2.73 × 2.73 × 3.75 and 1.95 × 1.95 × 5.00 mm^3 for GE and Siemens systems, respectively) for CT-based attenuation and scatter correction (CT-ASC). The methodology described in our previous work (Yang *et al* 2017a, 2017b) was used for CT image preprocessing and coregistration to MR images: MR and CT image pairs were coregistered using the vendor-developed registration toolkit or Advanced Normalization Tools (Avants *et al* 2009) based on the Insight Segmentation and Registration Toolkit (Kitware, Clifton Park, NY).

2.3. PET image reconstruction

As depicted in figure 1, non-corrected PET (PET$_{NC}$) and CT-based attenuation/scatter-corrected PET (PET$_{CT-ASC}$) images were reconstructed by a TOF OSEM algorithm (four iterations; 28 subsets; axial FOV 350 mm; matrix size 256 × 256 × 89; voxel size 1.37 × 1.37 × 2.78 mm^3; 4.0 mm in-plane Gaussian filter followed by axial 3-slice 1:4:1 filtering) in the offline PET/MR toolbox (REL_1_28, GE Healthcare). Also, MR-based attenuation/scatter-corrected PET (PET$_{MR-ASC}$) images were reconstructed as a silver standard for evaluation, providing the bottom line of a clinically acceptable performance limit. Corrections including normalization, dead time, decay, point-spread function, and randoms were applied during the reconstruction.

2.4. Deep convolutional neural networks (DCNNs)

The aim of this work is to develop a DCNN model that can transform PET$_{NC}$ to PET$_{CT-ASC}$ directly in *image space* (PET$_{DCNN}$), without generating attenuation maps and performing an iterative scatter simulation.

Figure 2. Deep convolution neural network (DCNN) architecture (Conv: convolution; BN: batch normalization; ReLU: rectified linear unit).

2.4.1. DCNN architecture

The proposed DCNN consists of five encoder-decoder stages symmetrically concatenated with skip connections (figure 2) based on the U-Net (Ronneberger *et al* 2015). In each stage, convolution (Conv) with 3×3 kernels, batch normalization (BN) (Ioffe and Szegedy 2015), and rectified linear unit (ReLU) is sequentially performed twice. Between stages, the downsampling and upsampling are done by 2×2 max pooling and bilinear interpolation (Xu *et al* 2017), respectively. In order to preserve local information and resolution of the image, skip connections transfer the 2nd convolution layer of the encoder, performed prior to the BN and ReLU activation, to the decoder after upsampling at the same stage (Liu *et al* 2018b).

2.4.2. Preprocessing

PET_{NC} and $PET_{CT\text{-}ASC}$ images were utilized as paired input and output for training/testing our proposed DCNN architecture. For each training dataset, PET raw values (Bq/ml) were scaled down (to kBq/ml) to reduce the dynamic range of input/output values, and PET slices above the top of the head and below the cerebellum were removed to focus on the brain. Also, activities out of the head were considered as noise and removed by binary masking.

2.4.3. Model training

The DCNN model was trained with a three-slice input to provide volumetric information due to attenuation and scatter in the axial direction. Training multi-slice inputs can provide higher efficiency with fewer parameters than training with depth-wise operation of 3D convolution (Xu *et al* 2017). Before being fed into the model, the input was randomly rotated (-10–$10°$), flipped horizontally, and translated horizontally (<50 pixels in FOV) for data augmentation to simulate a larger dataset and to avoid overfitting. Model training was performed with the mean squared error (or L2 loss) and RMSprop optimizer (Hinton *et al*) with a learning rate initialized by 0.001, which halved automatically if the loss did not decrease in ten epochs. Weights for convolution were initialized with truncated Gaussian distributions with zero mean and standard deviation of 0.02. All biases were initialized with zero. A mini-batch of 32 input/output patches (1375 patches in total) was used for training and the loss was converged in 140 epochs.

2.4.4. Computation

Training and testing our proposed DCNN were performed on a Ubuntu server with a single Tesla P100 (NVIDIA) graphics processing unit. The proposed model was implemented using Tensorflow (version 1.9.0 with CUDA 9.1) and Keras libraries (version 2.2.0). Model training takes approximately 160 min to reach stability, which occurred at approximately 5880 iterations (140 epochs $\times$ 42 iterations per epoch). At that point, the training was stopped. After training the model, it took only 0.4 s on average to generate PET_{DCNN} volumetric images (89 slices) with the single Tesla P100.

2.5. Evaluation metrics

2.5.1. Quantitative analysis

All PET images were spatially registered to a brain template with 116 automated anatomical labels (AALs) for a generalized regional analysis. Dixon MRAC T1 images simultaneously acquired with PET were registered to the T1 brain template provided by the Montreal Neurological Institute (MNI) (Tzourio-Mazoyer *et al* 2002) using Advanced Normalization Tools (Avants *et al* 2009). The derived registration parameters were applied to deform PET images to the template. Absolute and relative (%) differences (mean $\pm$ SD) of standardized uptake values

Table 1. Uptake differences (SUV and %) of PET_{DCNN} and PET_{MR-ASC} from reference PET_{CT-ASC} across 116 regions × ten test subjects (1160 regions in total).

	Difference		
	SUV	%	*P*
PET_{DCNN}	0.20 ± 0.92	3.98 ± 15.42	<0.001
PET_{MR-ASC}	-0.31 ± 0.31	-4.24 ± 4.29	<0.001

Differences are mean ± SD. A paired *t* test was performed for the pair of PET_{CT-ASC} and PET_{DCNN}.

(SUV = image-derived uptake [MBq/ml]/injection dose [MBq] × patient's weight [g]) between PET_{DCNN} and reference PET_{CT-ASC} were computed for 116 AAL regions across ten test subjects (1160 regions in total) as follows:

$$\text{Difference (SUV)} = \text{mean (voxels of } PET_{DCNN} \text{ in AAL\#)}$$
$$-\text{mean(voxels of } PET_{CT-ASC} \text{ in AAL\#)}$$
$$\text{Difference (\%)} = [\text{mean (voxels of } PET_{DCNN} \text{ in AAL\#)}$$
$$-\text{mean(voxels of } PET_{CT-ASC} \text{ in AAL\#)}]$$
$$/\text{mean(voxels of } PET_{CT-ASC} \text{ in AAL\#)} \times 100.$$

The distributions of the differences were presented in Bland–Altman plots. A difference between PET_{MR-ASC} and PET_{CT-ASC} was also calculated as a reference for comparison. Statistical significance was assessed using a paired *t* test and a *P* value < 0.05 was deemed statistically significant.

Additionally, to evaluate a performance variation of DCNN between subjects, subject-specific SUV differences across 116 AAL regions were computed separately. Also, to assess a performance variation of DCNN in the brain regions, SUV differences were calculated in eight merged regions: the cerebellum, temporal lobes, occipital lobes, parietal lobes, frontal lobes, thalamus, putamen, and caudate nucleus.

2.5.2. Qualitative analysis
PET_{CT-ASC}, PET_{DCNN} and their difference images were illustrated for three representative test subjects: subject 1 with the longest scan duration (900 s) selected for presenting the effect of a scan time variation (mean scan duration: 227.2 ± 137.5 s); subject 2 with a tumor in the head chosen for presenting the effect of a pathological variation; and subject 3 with outliers selected for presenting the visual effect of a large difference.

2.5.3. Voxelwise analysis
A joint histogram was used to show the distribution of voxel-based PET uptake correlation between PET_{DCNN} and reference PET_{CT-ASC} within the SUV range of 0.5–20.0 (g ml^{-1}). Also, an error histogram was used to show the distribution of voxel-based PET uptake differences within the same SUV range.

All the processing and analyses above were performed in MATLAB (MathWorks).

3. Results

The average difference between PET_{DCNN} and PET_{CT-ASC} was 0.20 ± 0.92 ($3.98\% \pm 15.42\%$, $P < 0.001$) and the average difference between PET_{MR-ASC} and PET_{CT-ASC} was -0.31 ± 0.31 ($-4.24\% \pm 4.29\%$, $P < 0.001$) (table 1). PET_{DCNN} was slightly overestimated (4.0%; range, $-13.4 \sim +63.1\%$) with a larger variation (15.4%), while PET_{MR-ASC} was slightly underestimated (-4.2%; range, $-28.8 \sim +9.3\%$) with a smaller variation (4.3%). In the Bland–Altman plot (figure 3), most of the differences (circles) with DCNN are positioned within $\pm 10\%$, whereas most of the differences (triangles) with MR-ASC were prone to negative areas within $\pm 10\%$. Box plots of 116 AAL regions across test subjects are available in supplementary figure 1 (stacks.iop.org/PMB/64/075019/mmedia).

Because the overall larger variation of DCNN (15.4%) was mostly due to the outliers between 40%–60% differences in the Bland–Altman plot (figure 3), we calculated subject-specific differences across 116 regions to evaluate a performance variation of DCNN across subjects. Figure 4(a) demonstrates that subject 3 with the highest mean difference contributed the outliers to the Bland–Altman plot of DCNN. In order to derive a more generalized result without the outliers, the average differences of PET_{DCNN} and PET_{MR-ASC} from PET_{CT-ASC} were recalculated after excluding subject 3 and the updated results are summarized in table 2. The average difference of DCNN was substantially reduced from 4.0% to -0.8% without the outliers, which is much smaller than the average difference of MR-ASC (-5.1%). However, the SD of DCNN ($\pm 5.2\%$) was still slightly higher than the SD of MR-ASC ($\pm 3.6\%$).

Also, the average differences were calculated in eight merged regions to assess a regional performance variation of DCNN. Figures 4(b) and (c) compare the generalized regional differences with and without subject 3
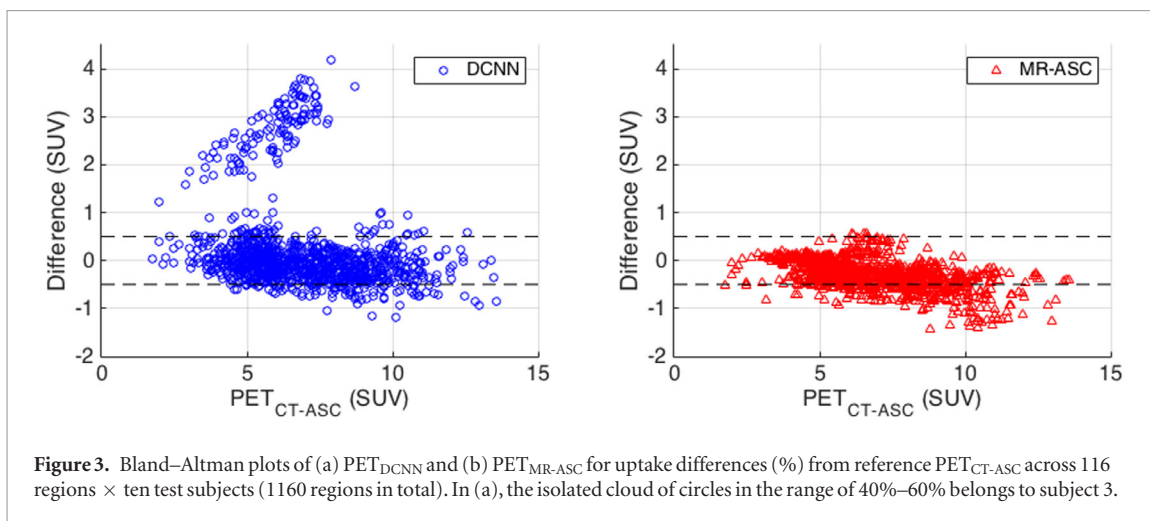
Figure 3. Bland–Altman plots of (a) PET$_{DCNN}$ and (b) PET$_{MR-ASC}$ for uptake differences (%) from reference PET$_{CT-ASC}$ across 116 regions × ten test subjects (1160 regions in total). In (a), the isolated cloud of circles in the range of 40%–60% belongs to subject 3.
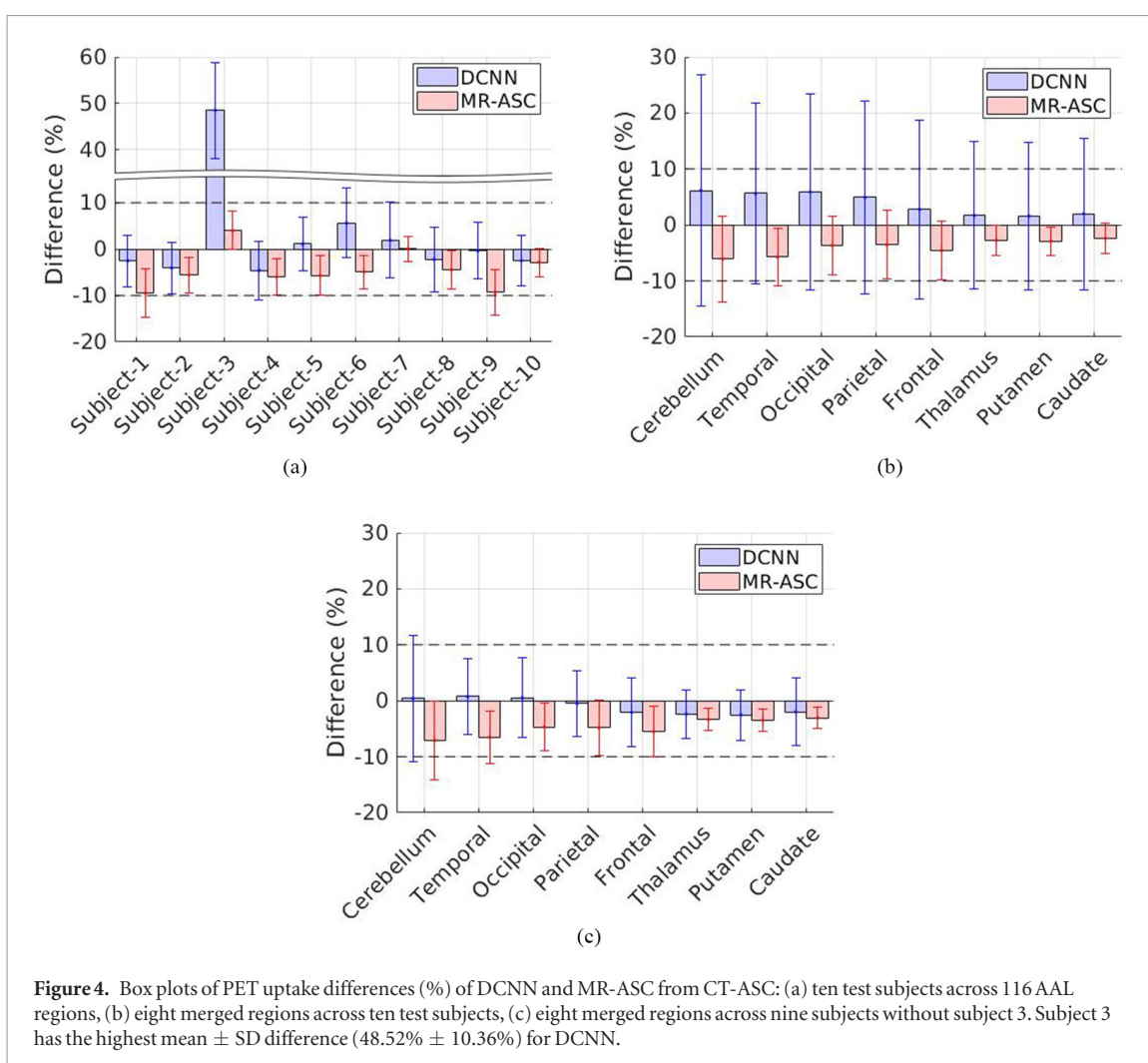
Figure 4. Box plots of PET uptake differences (%) of DCNN and MR-ASC from CT-ASC: (a) ten test subjects across 116 AAL regions, (b) eight merged regions across ten test subjects, (c) eight merged regions across nine subjects without subject 3. Subject 3 has the highest mean ± SD difference (48.52% ± 10.36%) for DCNN.

(outliers), which was also consistent with the overall result above. For example, the average difference of DCNN was substantially reduced from 6.2% to 0.4% for the cerebellum after excluding subject 3. The reduced difference of DCNN (0.4%) was much smaller than the recalculated difference of MR-ASC (−7.1%), while the SD of DCNN (±11.3%) was still higher than the SD of MR-ASC (±7.1%).

Next, figure 5 illustrates the qualitative differences of PET$_{DCNN}$ and PET$_{MR-ASC}$ from reference PET$_{CT-ASC}$ with examples of three representative subjects (subject 1, 2, and 3). The examples show the overall similarity between PET images with different ASC methods (reference CT-ASC, DCNN, and MR-ASC) and the voxel-based difference patterns of DCNN and MR-ASC from CT-ASC. In figure 5(a) (subject 1), the major difference between DCNN and MR-ASC is that DCNN differences are randomly distributed with a mixture of over- and

Table 2. PET uptake differences (SUV and %) of DCNN and MR-ASC from reference PET$_{CT-ASC}$ across 116 regions × nine subjects (excluding subject 3).

	Difference		
	SUV	%	*P*
PET$_{DCNN}$	-0.08 ± 0.33	-0.83 ± 5.20	<0.001
PET$_{MR-ASC}$	-0.36 ± 0.27	-5.07 ± 3.60	<0.001

Differences are mean ± SD. A paired *t* test was performed for the pair of PET$_{CT-ASC}$ and PET$_{DCNN}$.

under-estimated patterns, while MR-ASC differences depict overall slight underestimation across the brain but a strong underestimation near the skull. In figure 5(b) (subject 2), the SUV$_{max}$ of the tumor with DCNN was underestimated by -13.5%, while the SUV$_{max}$ of the tumor with MR-ASC was underestimated by -6.7%, compared to the SUV$_{max}$ with CT-ASC. In figure 5(c) (subject 3), both PET$_{DCNN}$ (structural similarity index (SSIM) (Wang *et al* 2004) $= 0.9670$) and PET$_{MR-ASC}$ (SSIM $= 0.9968$) are qualitatively similar to PET$_{CT-ASC}$, though PET$_{DCNN}$ was substantially overestimated across the brain in the result of figure 4(a) ($48.5\% \pm 10.4\%$ versus $4.0\% \pm 4.1\%$).

Further voxelwise evaluation was performed by the joint histogram (figure 6) and error histogram (figure 7). The analysis shows the voxel-wise similarity between PET$_{DCNN}$ and reference PET$_{CT-ASC}$ with the slope of 1.01 and R^2 of 0.98 (figure 6(a)). PET$_{DCNN}$ achieved higher accuracy for lower uptake voxels with smaller variation but lower accuracy for larger uptake voxels with larger variation, whereas PET$_{MR-ASC}$ achieved smaller variation with a trend of underestimation with the slope of 0.97 and R^2 of 0.99 (figure 6(b)), which was consistent with the result of the error histograms of PET$_{DCNN}$ and PET$_{MR-ASC}$ (figure 7).

4. Discussion

Our results demonstrate that a DCNN can achieve joint ASC that converts PET$_{NC}$ to PET$_{ASC}$ directly in *image space* without depending on a conventional approach that performs ASC separately in *sinogram space*. To our knowledge, this is the first work to investigate the feasibility of performing ASC simultaneously in *image space* through a DCNN. In this study, the DCNN demonstrated comparable quantitative ($-0.83\% \pm 5.20\%$ difference without outliers) and qualitative results compared to conventional CT-ASC or MR-ASC performed in *sinogram space* using CT or MR-derived attenuation maps and time-consuming iterative scatter correction.

Encouragingly, the results demonstrate the great potential of DCNN-based joint ASC to promote the clinical feasibility for a dedicated brain PET system such as a small-FOV PET or a wearable PET (Tashima and Yamaya 2016, Melroy *et al* 2017). Since it is technically difficult and not practical to combine such small PET systems with an additional CT or MR system, it is important to devise a practical and robust method for ASC without requiring additional anatomical imaging. Additionally, it is encouraging to provide a possible alternative to MR-based attenuation correction (MRAC) in simultaneous PET/MRI. Although the solutions for MRAC of simultaneous brain PET/MRI have been improved (Ladefoged *et al* 2017), these solutions are still not regarded as on the same level as CT-based attenuation correction. The comparison of PET$_{DCNN}$ and PET$_{MR-ASC}$ (-0.83% versus -5.07% difference without subject 3) demonstrate that DCNN could be a potential alternative for MRAC, which was the motivation of using PET/MR data instead of PET/CT data in this study.

The DCNN-based ASC approach is fully automated and fully data-driven: PET$_{NC}$ is directly converted to PET$_{ASC}$ in *image space* almost in real time (0.4 s on average for 89 volumetric images) without requiring additional anatomical information or time-consuming scatter correction. In the offline toolbox, scatter simulation took approximately 150 s and TOF-PET reconstruction with ASC took 147.8 min – 2.2 times longer than the processing time without ASC (67.8 min) in our workstation (Intel i7 CPU with 3.40 GHz × 8 cores and 31.4 GB memory). Scatter correction increased overall processing time substantially in the course of image reconstruction, but we can overcome this issue by applying our DCNN to reconstructed images without ASC. In general, DCNN applications for PET are more challenging than those for MR and CT due to the low resolution and noise characteristics of PET. Nevertheless, the success of low-dose FDG PET reconstruction demonstrated the capability of DCNN to deal with noisy FDG PET data (Xu *et al* 2017) and the feasibility study of DCNN-based conversion from FDG PET$_{NC}$ to pseudo-CT demonstrated the perceptibility of brain tissues and their boundaries in FDG PET$_{NC}$ (Liu *et al* 2018b). Therefore, it is not surprising to expect the successful result of our proposed DCNN-based joint ASC for FDG PET. In our study, the pattern differences were derived from how ASC changes uncorrected image patterns simultaneously, which enabled our proposed DCNN to predict attenuation/scatter-correction patterns for corrected images.
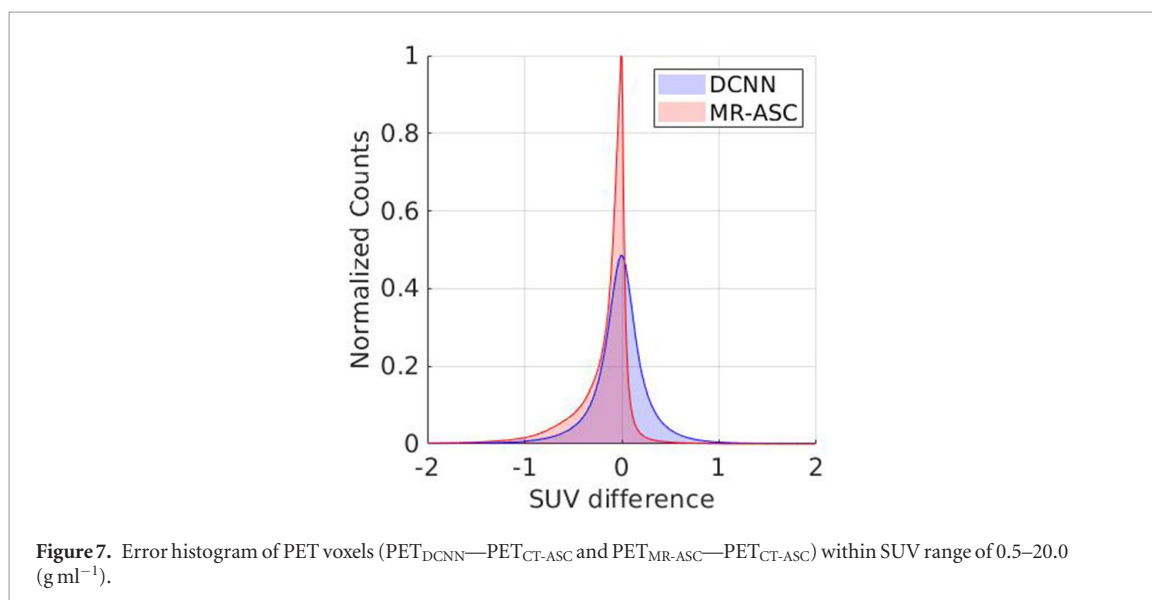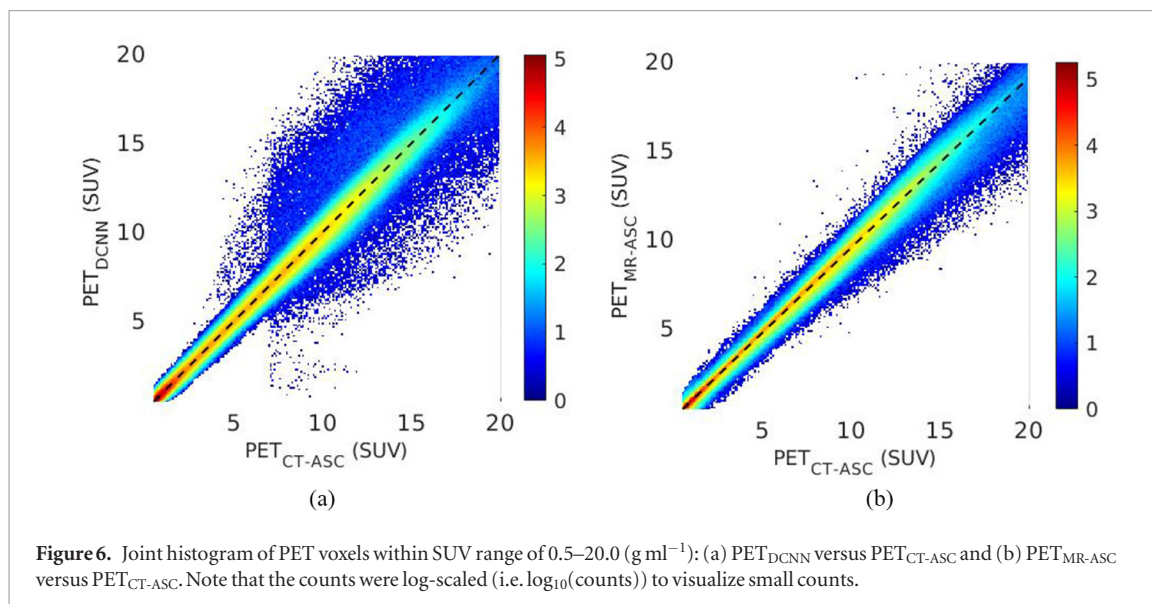
For the model training, we did not consider the information about the table couch and external head coils that should be always included in attenuation maps derived from CT or MR for accurate attenuation correction
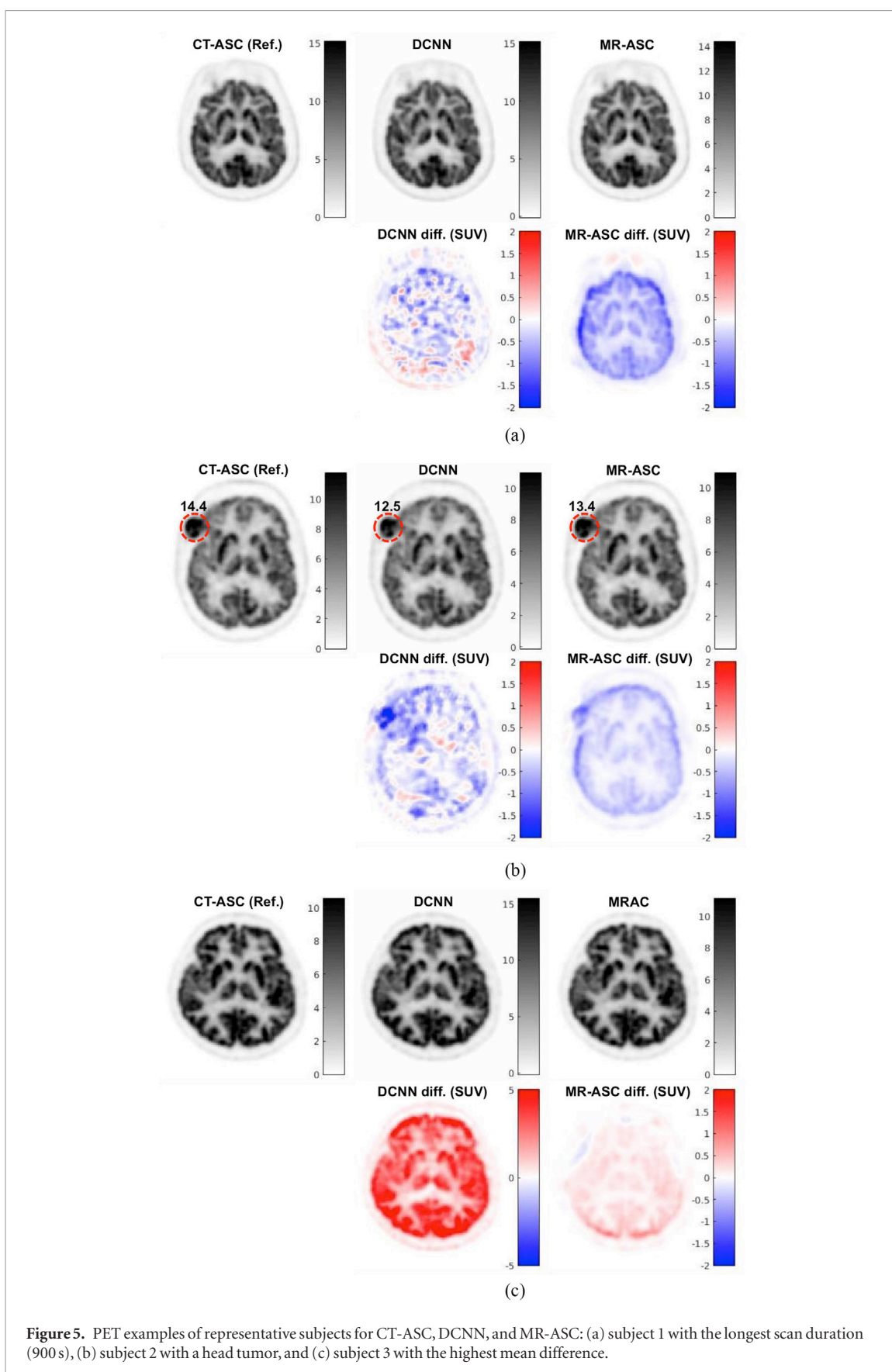
Figure 5. PET examples of representative subjects for CT-ASC, DCNN, and MR-ASC: (a) subject 1 with the longest scan duration (900 s), (b) subject 2 with a head tumor, and (c) subject 3 with the highest mean difference.

and scatter simulation. Surprisingly, however, the omitted information was not problematic since the attenuation information caused by the external materials could be embedded in training images themselves. For this reason, it is important not to perform vertical flipping for data augmentation because the back of the head was always positioned near the couch in the field of view and brain tissues near the couch were more attenuated in this setting.

Figure 6. Joint histogram of PET voxels within SUV range of 0.5–20.0 (g ml^{-1}): (a) PET$_{DCNN}$ versus PET$_{CT\text{-}ASC}$ and (b) PET$_{MR\text{-}ASC}$ versus PET$_{CT\text{-}ASC}$. Note that the counts were log-scaled (i.e. log$_{10}$(counts)) to visualize small counts.

Figure 7. Error histogram of PET voxels (PET$_{DCNN}$—PET$_{CT\text{-}ASC}$ and PET$_{MR\text{-}ASC}$—PET$_{CT\text{-}ASC}$) within SUV range of 0.5–20.0 (g ml^{-1}).

The proposed DCNN achieved the high accuracy (-0.83% difference) on average except for one test subject (subject 3) who seemed to be not represented in the training cohort. To clarify, since the average skull density is 685.6 ± 61.1 HU (min: 569.6 HU, max: 805.1 HU) in our data set (34 subjects except subject 3), such a low skull density (e.g. subject 3: 475.1 HU) is not common. The low bone density of subject 3 was qualitatively and quantitatively presented in additional supplementary figures 2 and 3. Although the quantitative difference (48.5% difference) of this subject was clinically not acceptable, we expect that this overall overestimation problem might not be problematic for diagnostic purposes because the relative contrast was consistent for both DCNN and standard approaches for subject 3 (figure 5). Therefore, it may be acceptable to use our DCNN for quantifying pre- and follow-up scans if a DCNN is consistently used for an outlier such as subject 3, since the values would be consistently overestimated with an acceptable contrast.

Despite the promising results, our study has several limitations. First, the amount of training and test data (25/10 data split) may not be enough to derive a fully generalized optimal DCNN model. Since the generalization power of DCNN is largely dependent on the amount of data, we performed data augmentation to simulate a larger dataset and to avoid overfitting. Nevertheless, as shown in the result of subject 3, the model substantially overestimated PET$_{DCNN}$ compared to reference PET$_{CT\text{-}ASC}$ (figure 4(a)), though the qualitative similarity between them was still high (figure 5(c)). Additionally, the training set did not include pathological subjects, which resulted in the relatively larger difference (-13.5%) of DCNN than the difference of MR-ASC (-6.7%) for the tumor uptake in subject 2. Therefore, in order to deal with any pathological pattern in our DCNN model, it is very important to include the variety of pathological patterns that may include a similar pattern for a target pathology in our extended training set. With the addition of an increased number of subjects with various pathologies in the training dataset, we expect that our model should be robust to a potential source of bias or

variation. Second, the model has not been tested for other radiotracers with different mean/maximum positron ranges (^{18}F, 0.66/2.63 mm; ^{11}C, 1.27/4.46 mm; ^{13}N, 1.73/5.57 mm; ^{15}O, 2.97/9.13 mm; ^{68}Ga, 3.56/10.27 mm; ^{82}Rb, 7.49/18.6 mm in water (Champion and Le Loirec 2007, Le Loirec and Champion 2007)) or biological distributions that can change image resolution and visually recognized brain patterns. Radiotracer-specific model training is likely necessary to evaluate the radiotracer-specific performance variation of our DCNN. Third, we have not investigated the performance of our DCNN for dynamic scans. Neuroreceptor studies potentially done with a dedicated brain PET system may be performed as dynamic scans, and the image patterns of such a study change continuously from the perfusion phase to the final image of neuroreceptor binding. Basically, since the theory of ASC is consistent for static and dynamic PET, the DCNN model could derive dynamic-imaging-specific ASC patterns if enough dynamic brain PET data were used to train the model. As discussed above, the flexibility of our DCNN model is substantially dependent on the amount of pattern information in the training data as well as the network architecture, so we may need to tailor our model architecture and retrain the model, considering data diversity such as inter-patient variation (pathological and anatomical difference) and inter-radiotracer variation (positron-range and biological distribution). Finally, there were large variations between PET protocols for the brain (e.g. scan duration, post-injection scan time, etc) since the brain PET scan was taken from a whole-body scan. However, DCNN-based joint ASC performed considerably well in spite of the variation, considering the result of subject 1 with the longest scan duration (900 s for subject 1, 227.2 ± 137.5 s for the others).

In future work, substantially increasing training data is our most important task to derive a fully generalized model that can interpret inter-patient and inter-radiotracer variations. Since many clinical and research PET/CT and PET/MR scans are ongoing in our department, we will access archived data and prospectively acquired data to increase our training/validation data set. We can increase the training/validation data set using PET/CT and PET/MR data simultaneously; however, it may be necessary to train the model according to PET/CT and PET/MR data separately and to compare their results, since PET system resolution and reconstruction details are different between PET/CT and PET/MR systems. For future validation, it will be worthwhile to use cross validation to perform a more generalized evaluation in a large cohort study. Second, although subjects with a very low bone density such as subject 3 are generally not likely, it is still important to control this potential problem to avoid an unexpectedly large quantitative error. To address this problem, it is probably feasible to develop a peripheral DCNN that might be able to derive a mean skull density for preventing unacceptable over/under estimation by training bone-specific voxel pairs of noncorrected PET and corrected PET, which is not a voxel-by-voxel conversion but a regression that might be resolved by DCNN. Third, we may increase the confidential level of our DCNN using conditional generative adversarial networks (c-GANs). The c-GANs include a generator network and a discriminator network which are trained simultaneously in order to guarantee the predicted output is close to the ground truth (Wang *et al* 2018). Finally, we plan to investigate a trade-off between image-based (our DCNN) and sinogram-based (Liu *et al* 2018b) approaches using only PET data in terms of robustness and practice in a large cohort study.

5. Conclusion

We demonstrated the feasibility of directly producing PET images corrected for attenuation and scatter using DCNN (PET$_{DCNN}$) from noncorrected PET (PET$_{NC}$) in *image space* without requiring additional anatomical imaging and time-consuming scatter correction. This approach is a one-step process, distinct from conventional methods that rely on generating attenuation maps first that are then applied to iterative scatter simulation in *sinogram space*. In particular, DCNN-based joint ASC has great potential to promote the clinical feasibility of a dedicated brain PET system that needs a practical and robust way for ASC without requiring anatomical imaging. Additionally, our DCNN-based method provides a possible alternative to MR-ASC for simultaneous PET/MRI.

Acknowledgments

Conflict of interest

There are no conflicts of interest or industry support for the project.

Financial support

ORCID iDs

Jaewon Yang ⬤ https://orcid.org/0000-0001-7637-0436

References

Avants B B, Tustison N J, Stauffer M, Song G, Wu B and Gee J C 2014 The Insight ToolKit image registration framework *Front Neuroinform* **8** 44

Berker Y and Li Y 2016 Attenuation correction in emission tomography using the emission data—a review *Med. Phys.* **43** 807–32

Brown R K, Bohnen N I, Wong K K, Minoshima S and Frey K A 2014 Brain PET in suspected dementia: patterns of altered FDG metabolism *Radiographics* **34** 684–701

Champion C and Le Loirec C 2007 Positron follow-up in liquid water: II. Spatial and energetic study for the most important radioisotopes used in PET *Phys. Med. Biol.* **52** 6605–25

Defrise M, Kinahan P E and Michel C J 2005 *Positron Emission Tomography* (*Image Reconstruction Algorithms in PET*) (Berlin: Springer) ch 4, pp 63–91

Gong K, Yang J, Kim K, El Fakhri G, Seo Y and Li Q 2018 Attenuation correction for brain PET imaging using deep neural network based on Dixon and ZTE MR images *Phys. Med. Biol.* **63** 125011

Hinton G, Srivastava N and Swersky K 2012 Neural networks for machine learning lecture 6e rmsprop: divide the gradient by a running average of its recent magnitude (https://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf)

Hwang D, Kim K Y, Kang S K, Seo S, Paeng J C, Lee D S and Lee J S 2018 Improving the accuracy of simultaneously reconstructed activity and attenuation maps using deep learning *J. Nucl. Med.* **59** 1624–9

Iatrou M, Manjeshwar R M, Ross S G, Thielemans K and Stearns C W 2006 3D implementation of scatter estimation in 3D PET *2006 IEEE Nuclear Science Symp. Conf. Record* vol 1–6 pp 2142–5

Ioffe S and Szegedy C 2015 Batch normalization: accelerating deep network training by reducing internal covariate shift (arXiv:1502.03167)

Ladefoged C N *et al* 2017 A multi-centre evaluation of eleven clinically feasible brain PET/MRI attenuation correction techniques using a large cohort of patients *NeuroImage* **147** 346–59

Le Loirec C and Champion C 2007 Track structure simulation for positron emitters of medical interest. Part I: the case of the allowed decay isotopes *Nucl. Instrum. Methods Phys. Res.* A **582** 644–53

Liu F, Jang H, Kijowski R, Bradshaw T and McMillan A B 2018a Deep learning MR imaging-based attenuation correction for PET/MR imaging *Radiology* **286** 676–84

Liu F, Jang H, Kijowski R, Zhao G, Bradshaw T and McMillan A B 2018b A deep learning approach for ^{18}F-FDG PET attenuation correction *Eur. J. Nucl. Med. Mol. Imaging* **5** 24

Lonn A H, Hsieh J, Stearns C W, Chao E H and Grekowicz B 2003 Evaluation of extension of the CT attenuation map in PET/CT *J. Nucl. Med.* **44** 39p

Meikle S R and Badawi R D 2005 *Positron Emission Tomography* (*Quantitative Techniques in PET*) (Berlin: Springer) ch 5, 93–124

Melroy S, Bauer C, McHugh M, Carden G, Stolin A, Majewski S, Brefczynski-Lewis J and Wuest T 2017 Development and design of next-generation head-mounted ambulatory microdose positron-emission tomography (AM-PET) system *Sensors* **17** 1164

Ronneberger O, Fischer P and Brox T 2015 U-Net: convolutional networks for biomedical image segmentation *Lect. Notes Comput. Sci.* **9351** 234–41

Silverman D H and Alavi A 2005 PET imaging in the assessment of normal and impaired cognitive function *Radiol. Clin. North Am.* **43** 67–77

Tashima H and Yamaya T 2016 Proposed helmet PET geometries with add-on detectors for high sensitivity brain imaging *Phys. Med. Biol.* **61** 7205–20

Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, Mazoyer B and Joliot M 2002 Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain *NeuroImage* **15** 273–89

Wang Y, Yu B, Wang L, Zu C, Lalush D S, Lin W, Wu X, Zhou J, Shen D and Zhou L 2018 3D conditional generative adversarial networks for high-quality PET image estimation at low dose *NeuroImage* **174** 550–62

Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: from error visibility to structural similarity *IEEE Trans. Image Process.* **13** 600–12

Watson C C, Casey M E, Michel C and Bendriem B 2004 Advances in scatter correction for 3D PET/CT *2004 IEEE Nuclear Science Symp. Conf.* vol 1–7 pp 3008–12

Wienhard K *et al* 2002 The ECAT HRRT: performance and first clinical application of the new high resolution research tomograph *IEEE Trans. Nucl. Sci.* **49** 104–10

Wollenweber S D, Ambwani S, Delso G, Lonn A H R, Mullick R, Wiesinger F, Piti Z, Tari A, Novak G and Fidrich M 2013 Evaluation of an atlas-based PET head attenuation correction using PET/CT & MR patient data *IEEE Trans. Nucl. Sci.* **60** 3383–90

Xu J, Gong E, Pauly J and Zaharchuk G 2017 200× low-dose PET reconstruction using deep learning (arXiv:1712.04119)

Yang J, Jian Y, Jenkins N, Behr S C, Hope T A, Larson P E Z, Vigneron D and Seo Y 2017a Quantitative evaluation of atlas-based attenuation correction for brain PET in an integrated time-of-flight PET/MR imaging system *Radiology* **284** 169–79

Yang J, Wiesinger F, Kaushik S, Shanbhag D, Hope T A, Larson P E Z and Seo Y 2017b Evaluation of sinus/edge-corrected zero-echo-time-based attenuation correction in brain PET/MRI *J. Nucl. Med.* **58** 1873–9

Zaidi H and Montandon M L 2007 Scatter compensation techniques in PET *PET Clin.* **2** 219–34