



Deep-JASC: joint attenuation and scatter correction in whole-body ^{18}F -FDG PET using a deep residual network

Isaac Shiri¹ · Hossein Arabi¹ · Parham Geramifar² · Ghasem Hajianfar³ · Pardis Ghafarian^{4,5} · Arman Rahmim^{6,7} · Mohammad Reza Ay^{8,9} · Habib Zaidi^{1,10,11,12} 

Received: 7 November 2019 / Accepted: 1 May 2020 / Published online: 15 May 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract

Objective We demonstrate the feasibility of direct generation of attenuation and scatter-corrected images from uncorrected images (PET-nonASC) using deep residual networks in whole-body ^{18}F -FDG PET imaging.

Methods Two- and three-dimensional deep residual networks using 2D successive slices (DL-2DS), 3D slices (DL-3DS) and 3D patches (DL-3DP) as input were constructed to perform joint attenuation and scatter correction on uncorrected whole-body images in an end-to-end fashion. We included 1150 clinical whole-body ^{18}F -FDG PET/CT studies, among which 900, 100 and 150 patients were randomly partitioned into training, validation and independent validation sets, respectively. The images generated by the proposed approach were assessed using various evaluation metrics, including the root-mean-squared-error (RMSE) and absolute relative error (ARE %) using CT-based attenuation and scatter-corrected (CTAC) PET images as reference. PET image quantification variability was also assessed through voxel-wise standardized uptake value (SUV) bias calculation in different regions of the body (head, neck, chest, liver-lung, abdomen and pelvis).

Results Our proposed attenuation and scatter correction (Deep-JASC) algorithm provided good image quality, comparable with those produced by CTAC. Across the 150 patients of the independent external validation set, the voxel-wise REs (%) were $-1.72 \pm 4.22\%$, $3.75 \pm 6.91\%$ and $-3.08 \pm 5.64\%$ for DL-2DS, DL-3DS and DL-3DP, respectively. Overall, the DL-2DS approach led to superior performance compared with the other two 3D approaches. The brain and neck regions had the highest and lowest RMSE values between Deep-JASC and CTAC images, respectively. However, the largest ARE was observed in the chest (15.16

This article is part of the Topical Collection on Advanced Image Analyses (Radiomics and Artificial Intelligence)

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00259-020-04852-5>) contains supplementary material, which is available to authorized users.

✉ Habib Zaidi
habib.zaidi@hcuge.ch

¹ Division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211 Geneva 4, Switzerland

² Research Center for Nuclear Medicine, Shariati Hospital, Tehran University of Medical Sciences, Tehran, Iran

³ Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Science, Tehran, Iran

⁴ Chronic Respiratory Diseases Research Center, National Research Institute of Tuberculosis and Lung Diseases (NRITLD), Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁵ PET/CT and Cyclotron Center, Masih Daneshvari Hospital, Shahid Beheshti University of Medical Sciences, Tehran, Iran

⁶ Departments of Radiology and Physics, University of British Columbia, Vancouver, BC, Canada

⁷ Department of Integrative Oncology, BC Cancer Research Centre, Vancouver, BC, Canada

⁸ Research Center for Molecular and Cellular Imaging, Tehran University of Medical Sciences, Tehran, Iran

⁹ Department of Medical Physics and Biomedical Engineering, Tehran University of Medical Sciences, Tehran, Iran

¹⁰ Geneva University Neurocenter, Geneva University, Geneva, Switzerland

¹¹ Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, Groningen, Netherlands

¹² Department of Nuclear Medicine, University of Southern Denmark, Odense, Denmark

$\pm 3.96\%$) and liver/lung ($11.18 \pm 3.23\%$) regions for DL-2DS. DL-3DS and DL-3DP performed slightly better in the chest region, leading to AREs of $11.16 \pm 3.42\%$ and $11.69 \pm 2.71\%$, respectively (p value < 0.05). The joint histogram analysis resulted in correlation coefficients of 0.985, 0.980 and 0.981 for DL-2DS, DL-3DS and DL-3DP approaches, respectively.

Conclusion This work demonstrated the feasibility of direct attenuation and scatter correction of whole-body ^{18}F -FDG PET images using emission-only data via a deep residual network. The proposed approach achieved accurate attenuation and scatter correction without the need for anatomical images, such as CT and MRI. The technique is applicable in a clinical setting on standalone PET or PET/MRI systems. Nevertheless, Deep-JASC showing promising quantitative accuracy, vulnerability to noise was observed, leading to pseudo hot/cold spots and/or poor organ boundary definition in the resulting PET images.

Keywords PET/CT · Whole-body · Scatter correction · Attenuation correction · Deep learning

Introduction

^{18}F -FDG PET imaging is commonly used for clinical diagnosis, staging, restaging and monitoring of response to treatment in clinical oncology [1]. Quantitative and semi-quantitative image-derived PET metrics, such as the standardized uptake value (SUV), are used for non-invasive quantification of physiological processes, providing valuable information for diagnosis and therapy monitoring of malignant diseases [2].

Good image quality is essential for pertinent qualitative visual interpretation and accurate quantitative analysis of ^{18}F -FDG PET images [3]. Compton scatter of one or both annihilation photons, which undergoes interaction within dense materials in the gantry (patient body, instrument bed, detector and electronic components), results in a scatter fraction of 30–35% and 50–60% of all recorded events in 3D brain and whole-body PET data collection, respectively. Scattered photons can be rejected by falling outside the predetermined acceptance energy window (e.g. [435–650 keV] on the considered PET scanner). These photons are no longer relevant for scatter correction but still contribute to attenuation events [4]. However, the wrong lines of response (LORs) assigned following path change of scattered photons (even LORs recorded outside the patient's body) within the acceptance energy window require scatter correction. Attenuation and scatter events result in local decrease and increase in the number of detected counts, which lead to underestimation and overestimation of the tracer uptake, respectively, resulting in loss of image contrast and quantification errors. The distribution and magnitude of scattered photons in non-attenuation and scatter corrected (PET-nonSAC) images are to some extent dependent on the energy resolution, sensitivity and time-of-flight capability of PET scanners. Correction of physical degrading factors, such as Compton scatter (SC) [4] and photon attenuation correction (AC) [5], which can result in mislocation of loss of detected events, is crucial. On commercial hybrid PET/CT scanners, SC can be performed by modelling the scatter component based on attenuation and uncorrected emission PET images. AC is also performed by converting Hounsfield units in CT images to

linear attenuation coefficients at 511 keV. AC in PET/MRI is still an open area of research since MRI provides information about proton density which cannot be directly converted to electron density [6]. To address this, various strategies have been devised, including bulk segmentation of tissues from different MR sequences into a predefined number of classes [7, 8], atlas-based methods [9] and emission-based methods that are capable of direct estimation of attenuation maps from emission data [10, 11]. While some of these methods proved promising for AC in PET/MRI, segmentation and atlas-based techniques are limited by tissue misclassification, intra-subject co-registration between CT and MR images, truncation or metal-induced artefacts and by the presence of anatomical abnormalities. As such, AC remains a significant and nontrivial challenge in PET/MRI [6].

Deep learning algorithms have been developed for various medical image analysis applications [12]. Pioneering efforts have successfully applied deep convolutional networks to various tasks, including image denoising [13], resolution recovery [14] and image reconstruction [15]. Multiple studies explored the suitability of deep learning approaches for cross-modality transformation from MRI to CT images [16] and vice-versa [17], as well as in AC of PET data [18–20]. Recent works focused on the generation of pseudo-CT images from T1-weighted [16, 21, 22], ultra-short echo time (UTE) [23], zero echo time (ZTE) [24] and Dixon [25] MR sequences for AC of ^{18}F -FDG PET images in the brain and pelvic regions. Two closely related but independent works reported on the direct conversion of non-attenuation corrected brain ^{18}F -FDG PET images to the attenuation corrected image using convolutional encoder-decoder neural networks [20, 26].

Interest in simultaneous maximum likelihood reconstruction of attenuation and activity (MLAA) based on emission data was revived with the introduction of time-of-flight (TOF) PET [10, 27]. However, these algorithms are limited by insufficient coincidence time resolution, which leads to positional uncertainty [28], ill-posedness of the problem and high noise levels, which result in slow convergence and inaccurate estimations. Recent studies proposed to generate pseudo-CT AC

maps through the application of deep learning to the output of joint activity and attenuation estimation in brain and whole-body PET imaging in an attempt to address the limitation of the MLAA [29, 30]. In addition, recently proposed methods that generate pseudo-CT images from MRI are limited by the need for co-registered pairs of CT and MRI images for the training phase. Such methods also require cross-modality conversion in the image space (e.g. from a proton density (MRI) to an electron density (CT)) [6]. However, more recently MRI to CT transformation using unpaired data was also reported [31].

Previous works aiming at direct attenuation and scatter correction in image space focused mostly on brain imaging resulting in a promising performance with clinically tolerable errors owing to the rigidity of structures within the head and minor inter-patient variability [20, 26]. More recent independent work focused on application of this approach in whole-body imaging demonstrated promising reliability and quantitative accuracy using a limited number of training/test datasets [32]. However, the application of this approach in whole-body imaging warranted a thorough investigation of potential outliers owing to large inter-patient anatomical and posture variability and physiological motion. As such, this study sets out to investigate the potential of this approach using a large sample size through the assessment of outliers and to provide some insight into its potential to handle respiratory motion artefacts. The objective of this work is to demonstrate the feasibility of the direct generation of attenuation and scatter corrected images from non-attenuation and scatter corrected (PET-nonASC) images in whole-body ^{18}F -FDG PET imaging without the use of anatomical information using an unprecedentedly large sample size for training, test and validation. This is achieved using a framework involving the use of a deep residual convolutional neural network for joint attenuation and scatter correction (Deep-JASC).

Materials and methods

PET/CT data acquisition

This retrospective study was approved by our institutional review board and informed consent was obtained from all patients (ethic number REC.1397.095, *Tehran University of Medical Sciences*). Clinical whole-body ^{18}F -FDG PET/CT studies of 1388 patients acquired between 2016 and 2018 on a Siemens Biograph 6 True point scanner. Two hundred thirty-eight patients were excluded from this study owing to technical or logistic issues (artefacts, missing CT and/or PET-nonASC images). The patients were injected with an activity of 370 ± 49 MBq of ^{18}F -FDG and scanned 60 ± 13 min post-injection. Detailed patient demographics are given in Table 1 for training, validation and external validation datasets. A

low-dose CT scan (110 kVp, 145 mAs) was performed prior to PET data acquisition for attenuation correction. Scatter correction was performed only on PET-CTAC images using the single-scatter simulation (SSS) algorithm with two iterations [33]. In Siemens Biograph 6 True Point, the energy resolution is $\sim 12\%$ at 511 keV and the energy window set to 425–650 keV. The SSS algorithm estimates the scatter component for each LOR only for single Compton scattering [33]. Analytical calculations using the Klein-Nishina equation provide an estimation of the contribution of scattered photons to each LOR where the total amount (distribution) of single scattering is determined from the superposition of the estimations. The total scatter fraction is determined through tail fitting of the estimated scatter distribution to scattered photons outside the body contour. Random, dead-time, decay, normalization corrections were applied prior to point-spread function (PSF)-based image reconstruction for both PET-nonASC and CTAC images. PET images were reconstructed using the ordinary Poisson ordered subsets-expectation maximization (OP-OSEM) algorithm with 2 iterations and 21 subsets followed by with 5-mm FWHM Gaussian post-reconstruction smoothing. All images were reconstructed into a 168×168 matrix and cropped to 154×154 matrix (excluding empty regions on both sides) with a voxel size of $4.073 \times 4.073 \times 3$ mm². Both reference CT, CT-based attenuation/scatter-corrected (PET-CTAC) and PET-nonASC images were archived for all patients.

Data preparation

The 1150 whole-body ^{18}F -FDG PET studies included in this work were randomly divided into training (900), validation (100) and independent validation (150) datasets. Evaluation of the proposed algorithm was performed using the unseen external validation set. These images were not used within the training or evaluation of the network. In the first step, the voxel intensities in the whole dataset consisting of 1150 PET-CTAC and PET-nonASC images were converted to standardized uptake values (SUVs) to reduce the dynamic range of the intensity of PET images. Moreover, to further reduce the dynamic range of the voxel intensities for the sake of effective training of the network, PET-CTAC images were normalized with an empiric factor equal to 9 determined from the images. Likewise, PET-nonASC images were normalized with a factor equal to 3.

Deep-JASC network architecture

The residual network (*ResNet*) [34] proposed in the Python library of Niftynet pipeline [35], built upon TensorFlow (version 1.12.) [36], was used to implement Deep-JASC algorithm. *ResNet* is composed of 20 convolutional layers where a $3 \times 3 \times 3$ -voxel convolution kernel is used in the first seven

Table 1 Patient demographics of clinical whole-body PET/CT studies enrolled in this study protocol

Data set	Parameters	No.	Mean	std	Min	25%	50%	75%	Max
Training	Injected activity (MBq)	900	373	48.80	123	348	377	400	555
Validation		100	372	52.52	139	352	374	400	481
External validation		150	368	40.29	170	352	370	392	470
Training	Weight (kg)	900	73	15.56	27	62	72	82	145
Validation		100	75	16.06	42	65	73	82	138
External Validation		150	73	13.31	25	63	74	81	113
Training	Time post-injection (min)	900	61	6.61	46	57	60	65	89
Validation		100	60	6.31	48	56	59	64	86
External Validation		150	62	5.92	49	58	61	65	77
Training	Age (year)	900	50	16.55	10	36	52	62	87
Validation		100	47	16.56	13	34	48	62	83
External Validation		150	48	15.61	11	35	49	62	78
Training	Male	446	Lymphoma (36.8%), colorectal (18.6%), lung (14.4%), head and neck (12.9%), other (12.6%), no metabolic abnormalities (4.7%)						
	Female	454							
Validation	Male	51	Lymphoma (35.1%), colorectal (12.7%), lung (18.4%), head and neck (13.7%), other (11.2%), no metabolic abnormalities (8.9%)						
	Female	49							
External validation	Male	69	Lymphoma (37.1%), colorectal (14.1%), lung (13.4%), head and neck (13.7%), other (13.6%), no metabolic abnormalities (8.1%)						
	Female	81							

layers. These layers extract low-level features, for instance edges, from the input data. The next seven layers employ the dilated convolution kernel by a factor of two to encode the medium-level features from the input. The last six layers also increase the dilation of the convolution kernel by 4, in addition to the previous layers, to enable capturing high-level features (Fig. 1A). A residual connection links every two convolutional layers. A batch normalization and element-wise rectified linear unit (ReLU) is connected to the convolutional layers located in the residual blocks.

Implementation details

The training of the Deep-JASC was performed using 900 pairs of PET-nonASC and PET-CTAC images as input and output, respectively. To this end, the Deep-JASC model training repeated three times in 2D mode using successive slices (DL-2DS), anisotropic 3D mode extended across several slices (DL-3DS) and isotropic 3D mode using a fixed patch size of $64 \times 64 \times 64$ (DL-3DP) (Fig. 1B). For DL-2DS training, a 2-dimensional (2D) spatial window equal to 154×154 voxels and batch size of 30 were set. For the DL-3DS mode, a spatial window of $154 \times 154 \times 32$ voxels and a batch size of 2 were selected. For the DL-3DP mode, a spatial window equal to $64 \times 64 \times 64$ voxels and a batch size of 4 were set. The same

settings were used for the three training modes as follows: learning rate = 0.001, sample per volume = 1, optimizer = Adam, loss function = L2norm and decay = 0.0001. During the training process, 10% of the training dataset was devoted to evaluation within the training of the model to verify the possibility of overfitting. Insignificant differences between the evaluation and training losses (errors) were observed for the three modes of training, which indicates no risk of overfitting. Decay factor refers to the regularization imposed on the trainable parameters. To this end, a ridge regression scheme using the following formula was utilized to prevent overfitting.

$$\text{Loss (regularized)} = \text{Loss (original)} + \lambda \times \sum (w)^2 \quad (1)$$

where λ is the decay factor ($\lambda = 0.0001$) and w is an array containing the entire trainable parameters. In addition, the training was started with a learning rate equal to 0.005 and for the last 2 epochs, the learning rate was reduced to 0.

Evaluation strategy

Qualitative and quantitative evaluation of the proposed framework was performed using an independent external validation set consisted of 150 patients. The image quality of the

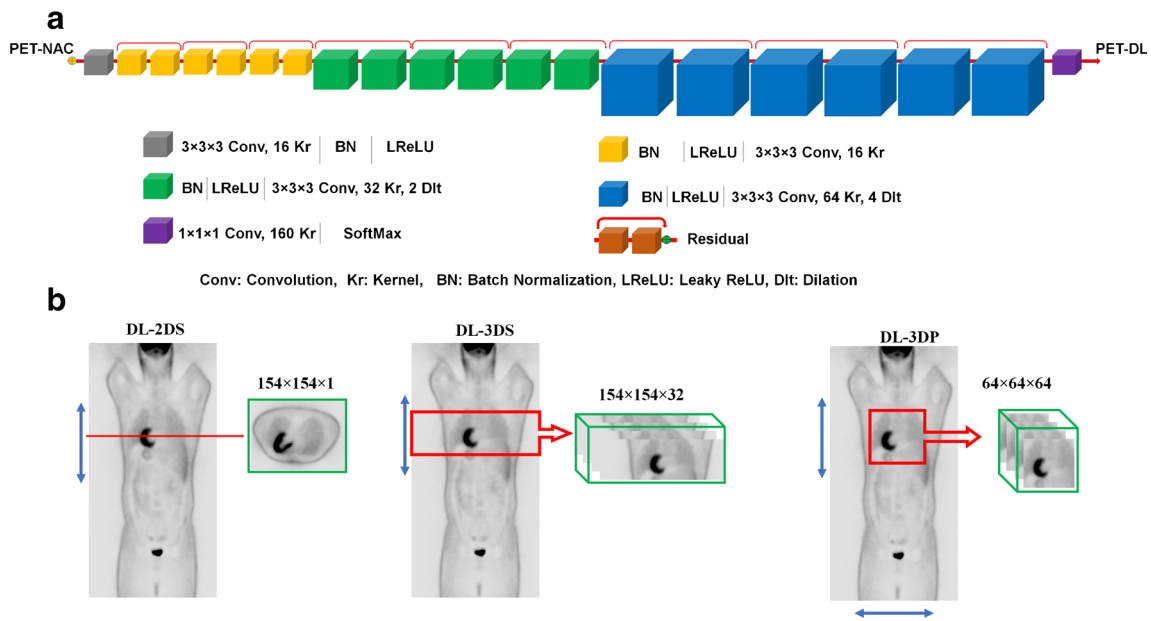


Fig. 1 (A) Architecture of the deep residual network (Deep-JASC) and implementation framework employed in the present study. (B) The training of the network was repeated three times: (left) in 2D slice-by-slice

mode (DL-2DS), (middle) anisotropic 3D mode extended across several slices (DL-3DS) and (right) isotropic 3D mode using a fixed patch size of 64 × 64 × 64 (DL-3DP)

generated images was assessed using the following metrics using PET-CTAC images as a reference for evaluation.

Voxel-wise mean error (ME), mean absolute error (MAE), relative error (RE%) and absolute relative error (ARE%) were calculated between reference PET-CTAC and predicted PET images for the three training modes, namely DL-2DS, DL-3DS and DL-3DP.

$$ME = \frac{1}{vxl} \sum_{v=1}^{vxl} PET_{Predicted}(v) - PET_{ref}(v) \quad (2)$$

$$MAE = \frac{1}{vxl} \sum_{v=1}^{vxl} |PET_{predicted}(v) - PET_{ref}(v)| \quad (3)$$

$$RE(\%) = \frac{1}{vxl} \sum_{v=1}^{vxl} \frac{(PET_{predicted})_v - (PET_{ref})_v}{(PET_{ref})_v} \times 100\% \quad (4)$$

$$ARE(\%) = \frac{1}{vxl} \sum_{v=1}^{vxl} \left| \frac{(PET_{predicted})_v - (PET_{ref})_v}{(PET_{ref})_v} \right| \times 100\% \quad (5)$$

Here, $PET_{predicted}$ stands for PET images corrected for attenuation/scatter using Deep-JASC whereas PET_{ref} stands for the reference PET-CTAC images. vxl and v denote the total number of voxels and voxel index, respectively.

Moreover, the root mean square error (RMSE), structural similarity index (SSIM) [37] and peak signal-to-noise ratio (PSNR) were calculated to assess the quality of the predicted PET images using Eqs. 6–8.

$$RMSE = \sqrt{\frac{1}{vxl} \sum_{v=1}^{vxl} (PET_{predicted}(i) - PET_{ref}(i))^2} \quad (6)$$

$$PSNR(dB) = 10 \log_{10} \left(\frac{Peak^2}{MSE} \right) \quad (7)$$

$$SSIM = \frac{(2Ave_{ref}Ave_{predicted} + C_1)(2\delta_{ref,predicted} + C_2)}{(Ave_{ref}^2 + Ave_{predicted}^2 + C_1)(\delta_{ref}^2 + \delta_{predicted}^2 + C_2)} \quad (8)$$

In Eq. (7), peak denotes the maximum intensity of either PET_{ref} or $PET_{predicted}$ whereas MSE indicates the mean squared error. In Eq. (8), Ave_{ref} and $Ave_{predicted}$ stand for the mean value of PET_{ref} and $PET_{predicted}$, respectively. δ_{ref} and $\delta_{predicted}$ denote the variances and $\delta_{ref,predicted}$ the covariances of PET_{ref} and $PET_{predicted}$ images, respectively. The constants ($C_1 = 0.01$ and $C_2 = 0.02$) were set to avoid division by very small values. The quantitative assessment of the proposed techniques was performed separately for the whole-body and for different regions of the body, namely the head, neck, chest, lung/liver, abdomen and pelvis as (Supplemental Figure 1). The lung/liver region contains slices from the liver dome until the end of the lungs. This region was reported separately since most respiratory motion errors/artefacts occur in this region.

Moreover, the performance of Deep-JASC was further investigated through the analysis of different types of malignant lesions delineated manually on 150 patients in the test dataset. To this end, volumes of interests (VOIs) were manually defined on lesions depicted on PET-CTAC images and copied on the corresponding predicted PET images. Conventional PET quantitative parameters, including SUV_{max} and SUV_{mean} as well as histogram-based, second- and high-order texture features were extracted from the VOIs defined on the lesions. For patients lacking metabolic abnormalities, a 3 cm² VOI was defined on the liver to extract the abovementioned

features. Overall, 171 VOIs from 150 patients (external validation) were extracted for further analysis. A detailed description of the radiomics features is provided in Supplemental Table 1. The relative errors (REs%) were calculated between features extracted from reference PET-CTAC and Deep-JASC images using Eq. 9.

$$RE(\%) = \frac{(\text{PET}_{\text{predicted}})_{\text{VOI}} - (\text{PET}_{\text{ref}})_{\text{VOI}}}{(\text{PET}_{\text{ref}})_{\text{VOI}}} \times 100 \quad (9)$$

Paired-sample *t* test was used for statistical analysis to compare the image-derived metrics between the different training modes. The significance level of *p* values was set to 0.05 for all comparisons. Moreover, a joint histogram analysis (with Pearson correlation) was performed to display the distribution of voxel-wise PET SUV correlation between the Deep-JASC approaches and reference PET-CTAC for a SUV range of 0.1 to 18 in 200 bins.

Results

Figure 2 provides head-to-head comparison of a representative clinical example of PET-nonASC, PET-CTAC, PET-DL-2DS, PET-DL-3DS and PET-DL-3DP images as well as difference SUV maps with respect to reference PET-CTAC image. As shown in Fig. 2, Deep-JASC provides image quality comparable with PET-CTAC. Quantification errors can be observed in the chest region and in regions with high heterogeneity and complexity, such as the liver, bones, soft tissues and air. The DL-2DS method provided overly better image quality and quantitative accuracy compared with the two 3D methods. A comparison of horizontal profiles drawn through a lesion in the lung between PET-CTAC (reference) and the three synthesized PET images is also shown. The profiles of the PET-CTAC and Deep-JASC images match each other in both high and low activity areas of the chest region of lung cancer patients. Supplemental Figure 2 shows maximum intensity projections of PET-CTAC and the three different Deep-JASC algorithms of ten randomly selected patients.

The comparisons of image quality metrics computed on SUV images (mean \pm SD) including ME, MAE, voxel-wise RE (%), voxel-wise ARE (%), SSIM, PSNR, and RMSE between the different methods (DL-2DS, DL-3DS and DL-3DP) for different regions in the body are summarized in Table 2. Supplemental Tables 2–4 provide the *p* values for comparisons of the three Deep-JASC techniques in different body regions. Across the 150 patients of the independent external validation set, the voxel-wise REs (%) were $-1.72 \pm 4.22\%$, $3.75 \pm 6.91\%$ and -3.08 ± 5.64 for DL-2DS, DL-3DS and DL-3DP, respectively.

The DL-2DS method led to lower RE (%) compared with both DL-3DS and DL-3DP approaches (*p* values = 0.01 and

0.03, respectively). In the DL-2DS technique, the lowest and highest REs (%) were $0.5 \pm 5.04\%$ in the neck region and $-4.7 \pm 4.02\%$ in the lung region, respectively. For DL-3DS, the lowest and highest REs (%) were 0.41 ± 6.86 in the lung-liver and 5.81 ± 4.85 in the pelvic region, respectively. The RE (%) decreased significantly for DL-2DS ($1.25 \pm 2.51\%$) in the pelvis region (*p* value < 0.05).

The average ME and ARE (%) for DL-2DS attenuation/scatter corrected images were -0.02 ± 0.06 and $11.61 \pm 4.25\%$, respectively. The DL-3DP, DL-3DS, and DL-2DS techniques ranked from the highest to lowest in terms of ARE (%) and MAE metrics. The DL-2DS, DL-3DS and DL-3DP approaches led to PSNR of 34.59 ± 3.28 , 34.24 ± 3.44 and 34.09 ± 3.32 and SSIM of 0.96 ± 0.03 , 0.95 ± 0.03 , and 0.95 ± 0.03 , respectively. However, the differences between these results were not statistically significant (Table 2).

The quantitative evaluation of PET images was conducted by voxel-wise comparisons of tracer uptake (SUV) within the body contour. Figure 3 compares the box plots of ARE, RE, MAE and RMSE, SSIM, PSNR, respectively, of the various methods within different regions of the body (box plot of ME is provided in Supplemental Figure 3). The lowest RMSE (0.22 ± 0.08) and MAE (0.22 ± 0.09) were achieved by DL-2DS.

Figure 4 depicts the heat map of REs (%) calculated on 17 radiomics features for different types of cancer in Deep-JASC PET images. Table 3 summarizes the mean \pm SD of radiomics features for different Deep-JASC approaches, wherein REs of SUV_{mean} of $-2.54 \pm 2.36\%$, $-1.75 \pm 4.02\%$ and $-5.03 \pm 3.05\%$ were obtained for 2DS, 3DS and 3DP approaches, respectively. 2DS and 3DS approaches resulted in mean REs of less than 5% for all radiomics features, while 3DP led to REs of $-5.03 \pm 3.05\%$, $-5.27 \pm 3.22\%$ and $-5.24 \pm 2.98\%$ for SUV_{mean}, Q1 and Q2 from histogram features, respectively. Supplemental Tables 5–7 present details of RE% (mean \pm SD) calculated for different radiomics features and different cancer types in Deep-JASC PET images.

The joint histogram analysis exhibited voxel-wise similarity between reference CTAC and Deep-JASC attenuation and scatter correction techniques with slopes of 0.98, 1.05 and 1.02 for DL-2DS, DL-3DS and DL-3DP, respectively (Fig. 5). Moreover, the DL-2DS approach showed higher correlation ($R^2 = 0.985$) compared with DL-3DS and DL-3DP ($R^2 = 0.980$ and $R^2 = 0.981$, respectively). Overall, DL-2DS exhibited smaller variation across the whole range of SUVs compared to the two 3D methods. However, for lower uptake voxels, DL-3DP showed comparable performance (variation) to the DL-2DS approach.

Figure 6 shows an example of PET-CTAC affected by noticeable respiratory motion owing to the local mismatch between PET-nonASC and CT images in the liver dome region. This mismatch led to large discrepancies between PET-CTAC and PET-nonASC that was reasonably well recovered

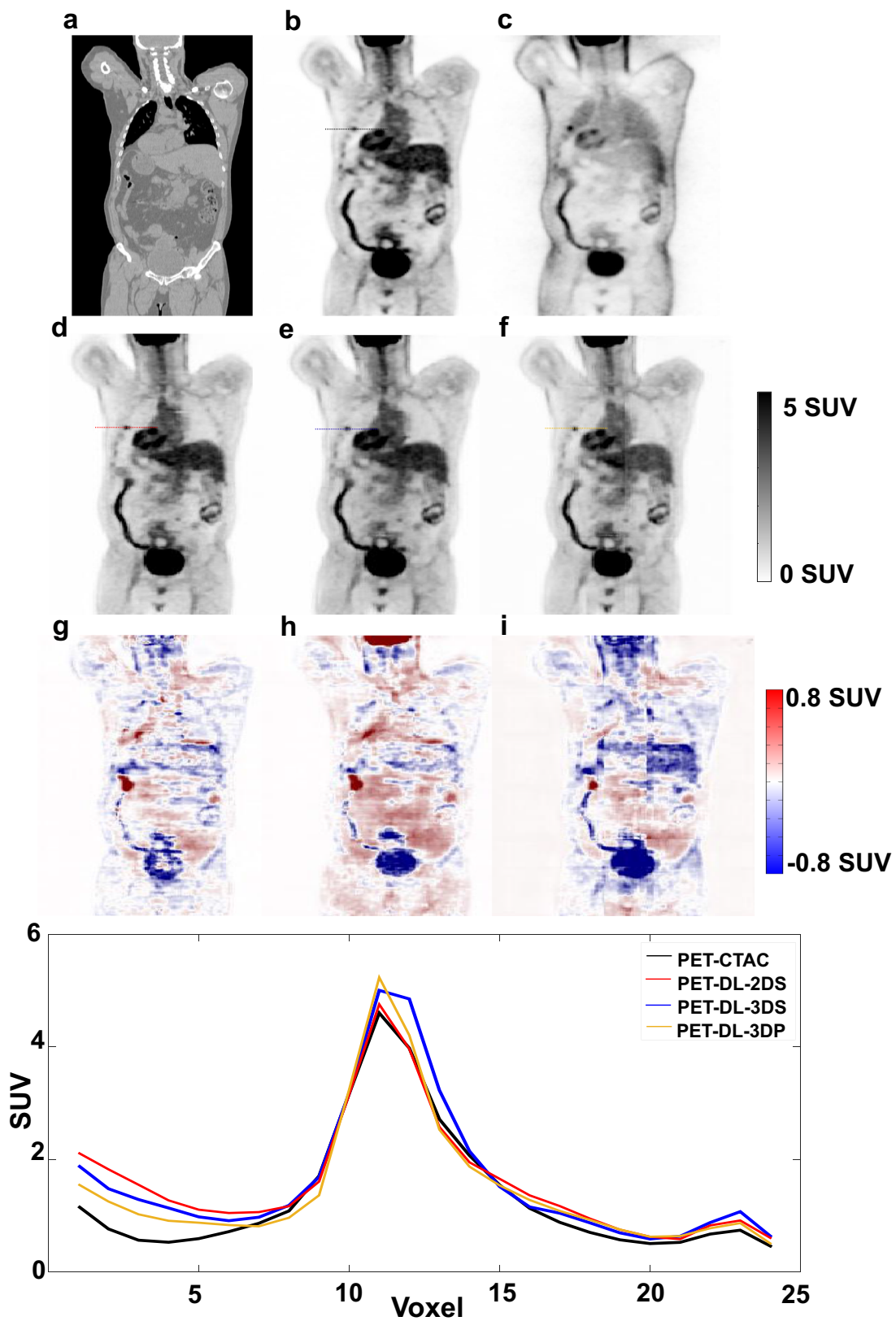


Fig. 2 Comparison of coronal views of the PET images corrected for attenuation/scatter using Deep-JASC approach. (A) CT, (B) PET-CTAC, (C) PET-nonASC, (D) PET-DL-2DS, (E) PET-DL-3DS, (F) PET-DL-3DP and the difference bias maps (G) PET-DL-2DS – PET-

CTAC, (H) PET-DL-3DS – PET-CTAC, (I) PET-DL-3DP – PET-CTAC. The plot shows SUV profiles drawn through a lung lesion on the four PET images

Table 2 Image quality metrics reflecting quantitative accuracy of the estimated tracer uptake within different regions of the body for the images generated using DL-2DS, DL-3DS and DL-3DP SC/AC approaches

Regions	ME	MAE	RMSE	RE (%)	ARE (%)	SSIM	PSNR
3DP_Brain	-0.04 ± 0.16	0.53 ± 0.08	0.31 ± 0.09	-0.62 ± 7.02	17.61 ± 5.15	0.95 ± 0.03	35.03 ± 2.99
3DS_Brain	0.11 ± 0.20	0.45 ± 0.14	0.28 ± 0.13	3.09 ± 9.31	14.47 ± 7.37	0.96 ± 0.03	35.30 ± 3.00
2DS_Brain	-0.01 ± 0.04	0.34 ± 0.09	0.27 ± 0.10	-1.37 ± 4.29	13.09 ± 4.47	0.97 ± 0.02	36.00 ± 2.93
3DP_Neck	0.08 ± 0.07	0.31 ± 0.04	0.16 ± 0.04	-1.23 ± 6.17	13.5 ± 4.92	0.95 ± 0.03	35.10 ± 2.33
3DS_Neck	0.08 ± 0.08	0.19 ± 0.05	0.18 ± 0.05	5.14 ± 6.64	12.3 ± 5.82	0.96 ± 0.03	35.21 ± 2.52
2DS_Neck	0.00 ± 0.04	0.17 ± 0.04	0.17 ± 0.05	0.50 ± 5.04	11.18 ± 5.20	0.97 ± 0.02	35.90 ± 2.63
3DP_Chest	0.02 ± 0.08	0.16 ± 0.04	0.20 ± 0.07	-1.01 ± 3.03	11.69 ± 2.71	0.92 ± 0.03	31.13 ± 2.69
3DS_Chest	0.03 ± 0.09	0.18 ± 0.05	0.21 ± 0.07	2.54 ± 5.61	11.16 ± 3.42	0.92 ± 0.03	31.20 ± 2.79
2DS_Chest	-0.01 ± 0.09	0.29 ± 0.05	0.22 ± 0.07	-4.70 ± 4.02	15.16 ± 3.96	0.93 ± 0.03	31.70 ± 2.52
3DP_Lung/Liver	-0.10 ± 0.11	0.43 ± 0.08	0.25 ± 0.08	-8.1 ± 5.05	14.68 ± 3.26	0.93 ± 0.03	31.80 ± 2.73
3DS_Lung/Liver	-0.03 ± 0.14	0.26 ± 0.07	0.24 ± 0.09	0.41 ± 6.86	12.22 ± 3.64	0.94 ± 0.03	32.01 ± 3.31
2DS_Lung/Liver	-0.06 ± 0.08	0.23 ± 0.06	0.24 ± 0.07	-2.02 ± 4.27	11.18 ± 3.23	0.94 ± 0.03	31.91 ± 2.49
3DP_Abdomen	0.03 ± 0.08	0.31 ± 0.06	0.22 ± 0.06	-4.42 ± 3.75	10.96 ± 2.37	0.97 ± 0.02	35.61 ± 2.70
3DS_Abdomen	0.07 ± 0.10	0.20 ± 0.05	0.23 ± 0.06	5.51 ± 5.69	10.36 ± 2.68	0.97 ± 0.02	35.70 ± 2.61
2DS_Abdomen	-0.02 ± 0.05	0.16 ± 0.04	0.22 ± 0.05	-1.49 ± 2.89	8.72 ± 2.06	0.97 ± 0.02	35.91 ± 2.54
3DP_Pelvis	0.06 ± 0.07	0.28 ± 0.05	0.24 ± 0.08	-3.10 ± 3.82	10.22 ± 1.80	0.97 ± 0.02	35.90 ± 2.91
3DS_Pelvis	0.09 ± 0.08	0.17 ± 0.04	0.27 ± 0.09	5.81 ± 4.85	10.31 ± 2.52	0.97 ± 0.02	36.00 ± 2.97
2DS_Pelvis	0.00 ± 0.04	0.14 ± 0.03	0.23 ± 0.07	-1.25 ± 2.51	7.54 ± 1.41	0.97 ± 0.02	36.10 ± 2.67
3DP_WB	0.01 ± 0.12	0.34 ± 0.13	0.23 ± 0.08	-3.08 ± 5.64	13.11 ± 4.38	0.95 ± 0.03	34.09 ± 3.32
3DS_WB	0.06 ± 0.13	0.24 ± 0.12	0.23 ± 0.09	3.75 ± 6.91	11.8 ± 4.80	0.95 ± 0.03	34.24 ± 3.44
2DS_WB	-0.02 ± 0.06	0.22 ± 0.09	0.22 ± 0.08	-1.72 ± 4.22	11.61 ± 4.25	0.96 ± 0.03	34.59 ± 3.28

by the Deep-JASC approach. Overall, despite the promising results achieved by the Deep-JASC approach, pseudo hot and cold spots were observed in Deep-JASC PET images in some of the cases in the independent validation dataset (150 patients). One such example is illustrated in Fig. 7 showing CT, PET-nonASC, PET-CTAC, PET-DL-2DS, PET-DL-3DS and PET-DL-3DP images as well as different SUV maps relative to PET-CTAC. A similar pseudo hot uptake appeared in the abdomen for the three Deep-JASC images where the line profiles drawn on the hot spot show the level of local SUV overestimation in Deep-JASC images compared with the reference PET-CTAC. Supplemental Figures 4–6 show representative examples wherein Deep-JASC failed to identify underlying structures (Supplemental Figure 4), to pinpoint lesions (Supplemental Figure 5) due to high levels of noise in PET-nonASC images, or cause artefact-like pseudo structures (Supplemental Figure 6).

Among the different Deep-JASC approaches, DL-2DS resulted in the lowest SUV error metrics (RMSE = 0.22 ± 0.08, RE = -1.72 ± 4.22% and SSIM = 0.96 ± 0.03). The highest error metrics (ARE = 15.16 ± 3.96%, lowest PSNR = 31.7 ± 2.52 and SSIM = 0.93 ± 0.03) were observed in the chest region. Regarding 3D implementations of Deep-JASC, ARE decreased significantly for DL-3DS (11.16 ± 3.42%, *p* value < 0.05) and DL-3DP (11.69 ± 2.71, *p* value < 0.05) approaches, respectively.

Discussion

The aim of the present work was to demonstrate the feasibility of the direct generation of attenuation/scatter-corrected whole-body ¹⁸F-FDG PET images from uncorrected images without using anatomical information. The evaluation of the proposed methods was performed on an independent validation set consisting of 150 patients using established quantitative imaging metrics. SUV quantification in the generated SC/AC images was highly reproducible.

Attenuation patterns in the chest region are different from other regions in PET-nonASC images. Lung tissue appears brighter in contrast to other regions of the body, whereas areas closer to the skin appear brighter and areas closer to the center appear darker in PET-nonASC images. The bias maps showed the highest error in the chest and liver regions owing to their higher heterogeneity (soft tissue, lung, air, bone and ribs). In addition, slight respiratory motion, causing local miss-match between PET and CT images, added to the complexity of this region which led to a higher relative error compared with other regions, including the brain, pelvis and abdomen. Prior to training, pairs of PET images suffering from low quality were removed from the training/test datasets. The causes of poor image quality were mostly gross patient motion within PET images (blurring) and miss-match between CT and PET

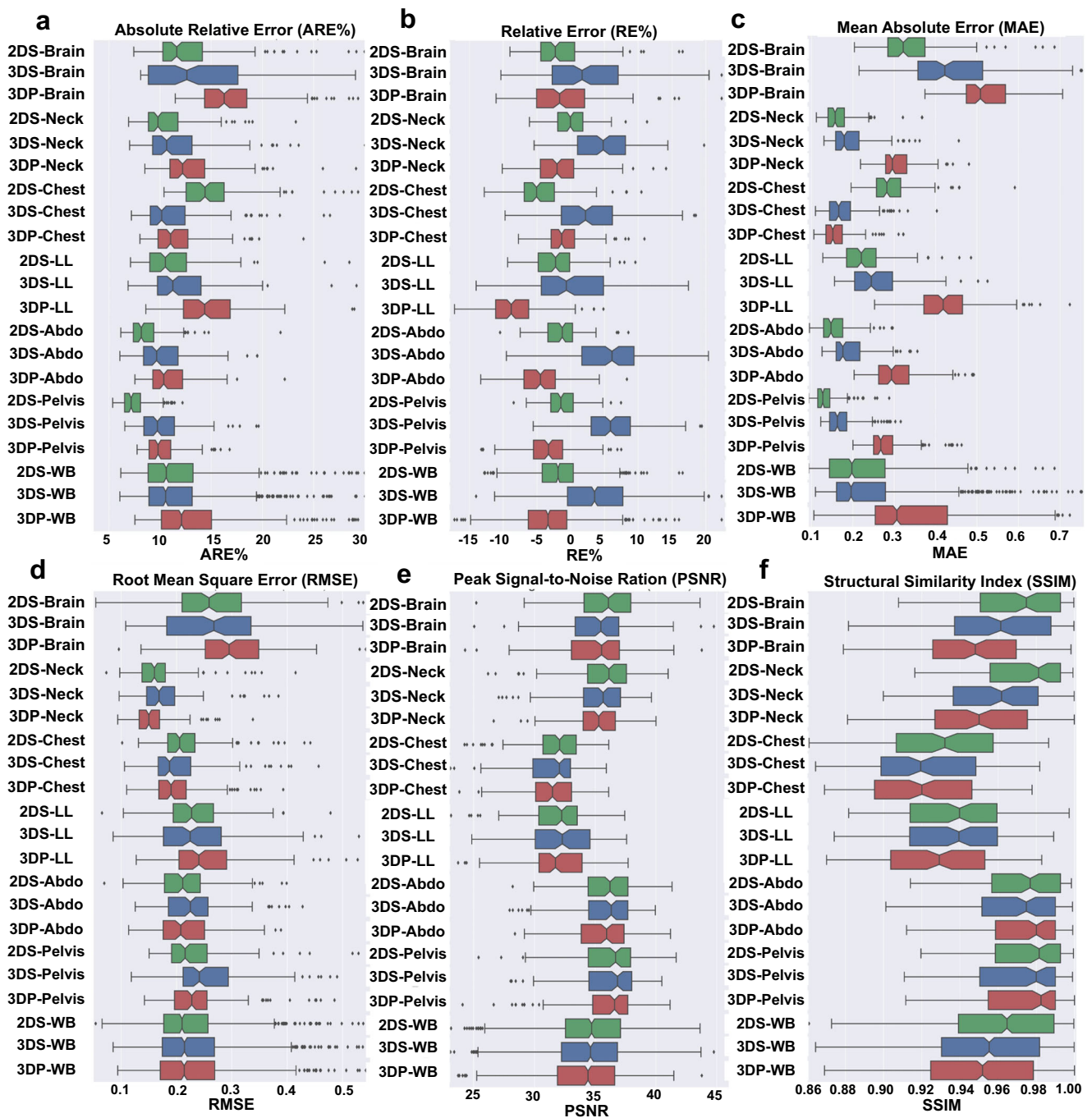


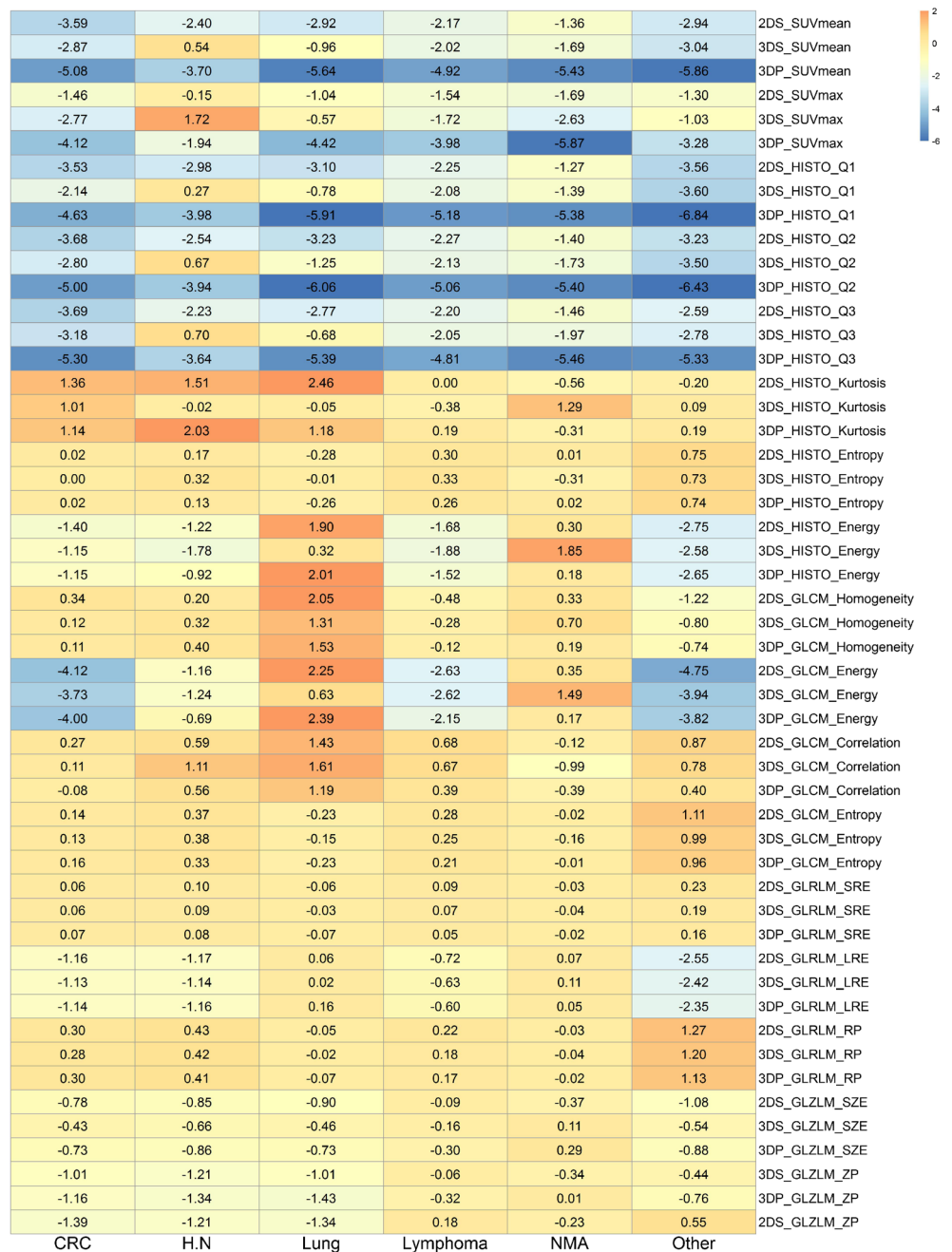
Fig. 3 Comparison of (A) absolute relative errors (ARE), (B) relative errors (RE%), (C) mean absolute errors (MAE), (D) root mean-square errors (RMSE), (E) peak signal to noise ratio (PSNR) and (F) structural similarity index (SSIM) obtained in different body regions when using

DL-2DS, DL-3DS and DL-3DP SC/AC approaches. The box plots display the intraquartile range (IQR), minimum (Q1–1.5 × IQR), first quartile (Q1), median, third quartile (Q3), maximum (Q3 + 1.5 × IQR) and outliers

images. In some cases, belonging to the external validation set (Fig. 6), one can observe a large difference between PET images generated using Deep-JASC and CTAC in the lung region and the liver dome. The mismatch between PET and CT images induced by respiratory motion in the lung and upper abdomen cause artefacts in PET-CTAC images. In Deep-JASC, attenuation and scatter corrections are performed

using only emission data without the use of anatomical images, such as CT and MRI. Therefore, Deep-JASC is less vulnerable to mismatch artefacts in the chest and upper abdomen, which partly explains the difference between Deep-JASC and CTAC PET images in these regions. Even though large ARE (%) and RE (%) were observed in the chest region, partly due to the division by very small SUV values, the

Fig. 4 Heat map depicting relative errors of 17 radiomics features extracted from 171 VOIs of 150 patients (external validation) across different cancer types for PET-DL-2DS, PET-DL-3DS and PET-DL-3DP images versus reference PET-CT images



absolute SUV differences between the predicted PET images and reference PET-CTAC images are comparable with those obtained in other regions. Previous whole-body PET AC studies showed SUV bias in lung lesions up to 15% when using atlas-based methods and over 20% when using segmentation-based [7, 38].

Despite the large anatomical intra- and inter-patient variability, the spatial distribution of the scatter component in PET images follows commonly a smooth variation across slices and/or regions. As such, the scatter patterns are not much affected in the chest region. Moreover, the scatter fraction, estimated via the tail fitting approach, is commonly

regularized over the axial coverage to ensure smooth scatter pattern changes. Therefore, although large anatomical variations exist in this region, the scatter patterns do not change drastically across the patients and through the different regions.

The results demonstrated that training in 2D mode led to superior (or at least slightly better) performance compared with the two 3D training modes. Three different methods, namely 2D slices (low computation burden, large data size with high variability, AC/SC performed in 2D), 3D slices (considering the 3D volume and include different information from sagittal and coronal planes, which allows a better

Table 3 Summary of the REs (%) of radiomics features extracted from 171 VOIs of 150 patients (external validation) for the different deep learning–based approaches with respect to PET-CTAC

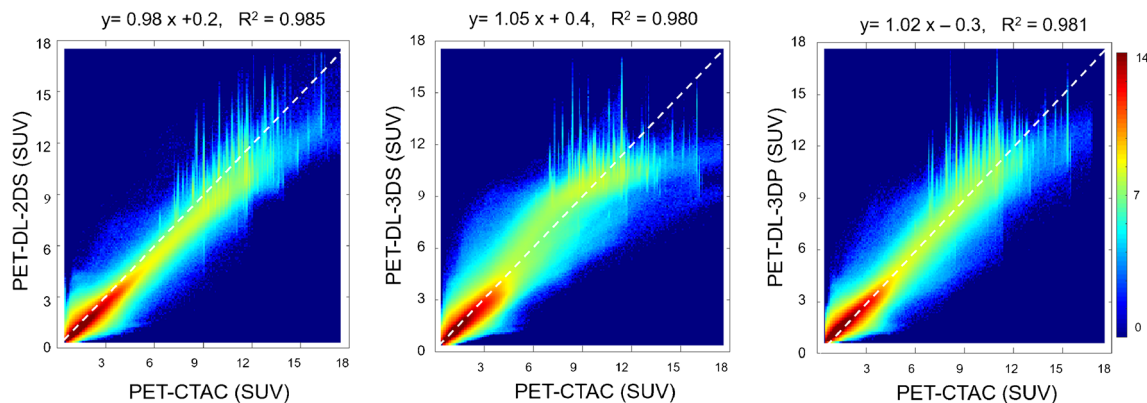
Quantitative parameters	2DS	3DS	3DP
SUV _{mean}	-2.54 ± 2.36	-1.75 ± 4.02	-5.03 ± 3.05
SUV _{max}	-1.25 ± 4.28	-1.21 ± 5.84	-3.83 ± 4.90
HISTO_Q1	-2.74 ± 2.71	-1.73 ± 4.06	-5.27 ± 3.22
HISTO_Q2	-2.69 ± 2.47	-1.86 ± 4.08	-5.24 ± 2.98
HISTO_Q3	-2.48 ± 2.58	-1.73 ± 4.33	-4.91 ± 3.21
HISTO_kurtosis	0.66 ± 4.53	0.11 ± 3.97	0.67 ± 4.63
HISTO_entropy	0.20 ± 1.37	0.24 ± 1.31	0.18 ± 1.31
HISTO_energy	-1.08 ± 5.51	-1.27 ± 5.16	-0.93 ± 5.52
GLCM_homogeneity	0.03 ± 2.98	0.08 ± 2.62	0.15 ± 2.75
GLCM_energy	-2.03 ± 8.78	-2.0 ± 7.92	-1.64 ± 8.68
GLCM_correlation	0.67 ± 1.96	0.66 ± 1.71	0.39 ± 1.57
GLCM_entropy	0.29 ± 1.64	0.27 ± 1.54	0.24 ± 1.53
GLRLM_SRE	0.08 ± 0.36	0.07 ± 0.33	0.05 ± 0.32
GLRLM_LRE	-0.92 ± 3.26	-0.87 ± 3.18	-0.83 ± 3.2
GLRLM_RP	0.34 ± 1.66	0.32 ± 1.58	0.3 ± 1.53
GLZLM_SZE	-0.56 ± 1.46	-0.34 ± 1.36	-0.53 ± 1.29
GLZLM_ZP	-0.42 ± 2.44	-0.55 ± 2.0	-0.76 ± 2.08

continuity of image intensity in the axial direction) and 3D patches (seeing the different structures in different FOVs), were considered for implementation of Deep-JASC. Nevertheless, the 3D approaches provided better results in highly heterogeneous regions, such as the lung region owing to the inclusion of information from neighbouring slices. The training dataset in 2D mode is markedly larger than in 3D mode, which might partly explain its better performance. In addition, 3D training mode may also complicate the learning/convergence process of regression and consequently results in sub-optimal performance. 3D training dramatically increases the number of trainable parameters, subsequently complicating the optimization process to converge to a global minimum loss. It should be noted that the training of datasets for a

specific body region may improve the quality of the outcome. Dividing the body into a number of sub-regions for training is an option to consider to enhance the performance of the Deep-JASC approach. In addition, hybrid 3D and 2D training might result in overall enhanced performance, for instance in the lung region where 3D implementations outperformed the 2D one.

Deep learning techniques were previously reported for cross-modality transformation from MRI to CT for MRI-only guided radiation therapy and AC in PET/MRI. Dedicated (UTE/ZTE) MR sequences require long acquisition time to generate an appropriate attenuation map, rendering these methods impractical for whole-body imaging. Hwang et al. [30] proposed a deep learning–based synthetic CT generation which the input to this network is the attenuation map estimated from MLAA reconstruction in brain imaging. They reported the average voxel-wise RE $12.82\% \pm 2.45\%$, $5.61\% \pm 0.68\%$ and $2.05\% \pm 1.51\%$ for MLAA based AC, four segment-based AC and deep learning–based AC, respectively. Shi et al. [39] proposed a line-integral projection loss (LIP-loss) function that incorporates the physics of PET attenuation into attenuation map generation. They reported a MAE of 21.4%, 7.8%, 9.4%, 9.2% and 11.26% for MLAA; 3.5%, 4.2%, 4.8%, 4.4% and 4.1% for the deep learning method; and 3.2%, 3.7%, 4.1%, 3.6% and 3.6% for Deep-LIP in the head, neck/chest, abdomen, pelvis and whole-body regions, respectively. Despite the promising results, pseudo-CT synthesis from MRI for AC faces a number of challenges [24, 40, 41], including mismatch between MRI and PET images, internal organ displacements and the lack of direct relationship between proton density and electron density [6]. To address these challenges, our proposed method performs direct AC from emission data only. This approach is less vulnerable to the challenges of pseudo-CT synthesis and is applicable to standalone PET scanners and PET/MRI systems.

Van Hemmen et al. [42] applied a deep convolutional encoder-decoder (CED) network using a U-Net architecture for whole-body PET attenuation correction and reported an ARE of

**Fig. 5** Joint histogram analysis displaying the correlation between activity concentration in PET- DL-2DS, PET- DL-3DS and PET- DL-3DP images and versus reference PET-CT image. Note that a logarithmic scale was used to display the SUV levels

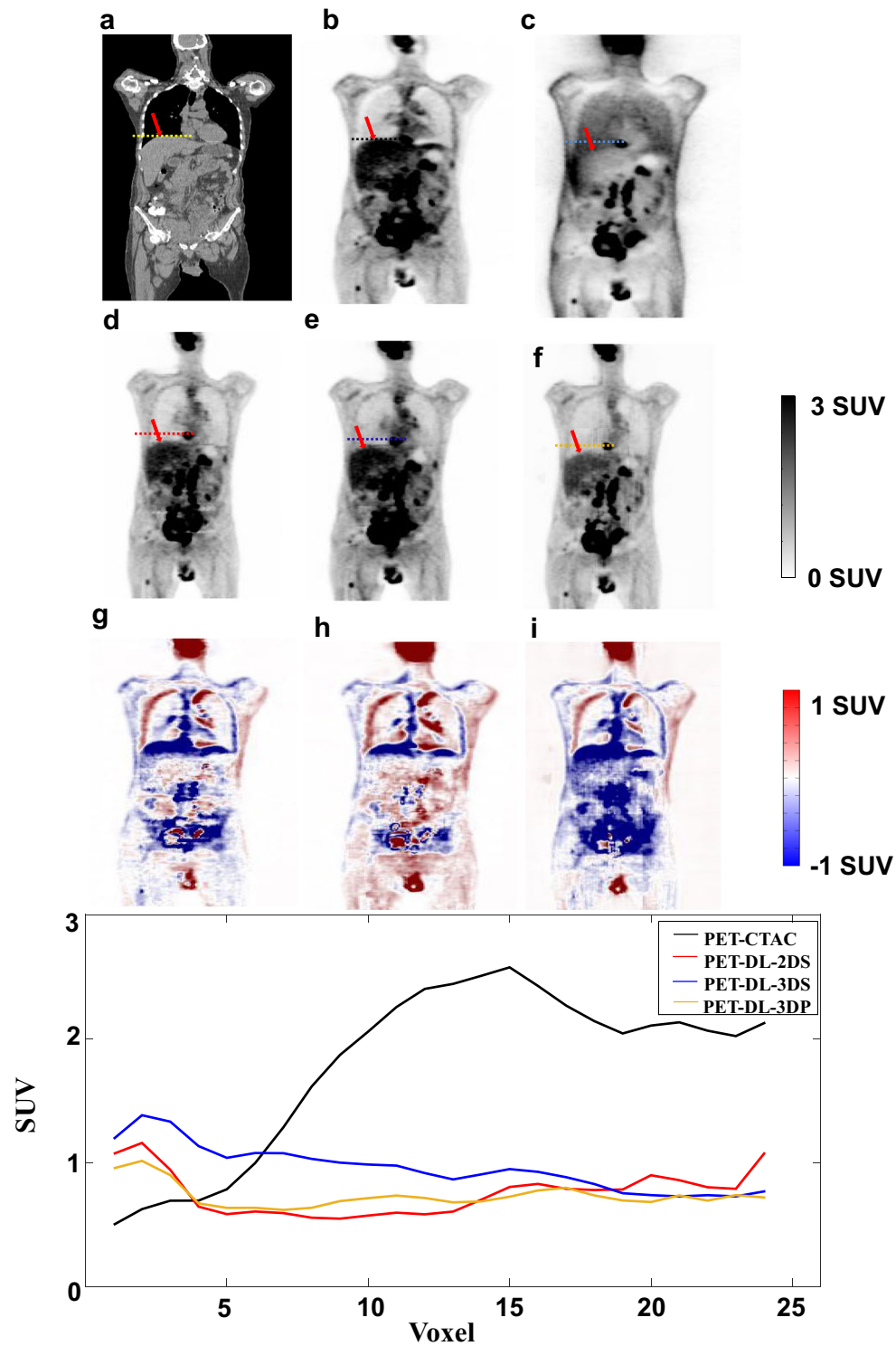


Fig. 6 Case report of a clinical PET/CT study suffering from mismatch between PET and CT images owing to respiratory motion reflected in the PET-CTAC images that was partially compensated by PET images corrected using deep learning-based SC/AC approach. (A) CT, (B) PET-CTAC, (C) PET-nonASC, (D) PET-DL-2DS, (E) PET-DL-3DS,

(F) PET-DL-3DP and the difference bias maps (G) PET-DL-2DS – PET-CTAC, (H) PET-DL-3DS – PET-CTAC and (I) PET-DL-3DP – PET-CTAC. The plot shows SUV profiles drawn through the pseudo hot spot

~30% for 10 subjects. Yang et al. [43] applied 3D generative adversarial networks for attenuation and scatter correction of whole-body PET images reporting an average ME and MAE in whole-body imaging of $2.49\% \pm 7.98\%$ and $16.55\% \pm 4.43\%$, respectively, in a population of 30 patients. More recently, Dong et al. [32] proposed a deep learning approach for whole-body PET AC in the absence of structural information. They reported an average ME of $13.84\% \pm 10.11\%$, $13.42\% \pm 10.13\%$ and $-17.02\% \pm 11.98\%$ in the lung and $2.05\% \pm 2.21\%$, $2.25\% \pm 1.93\%$ and $0.62\% \pm 1.26\%$ in whole-body imaging for U-Net, GAN and cycle-GAN, respectively. Reliable assessment of deep learning algorithms requires a large dataset for training and validation, allowing for proper interpretation of the results and robustness to outliers. In the present study, we used an unprecedentedly large sample size of training, test and validation dataset for reliable assessment of Deep-JASC, which enabled the reporting of potential outliers and deficiencies of the proposed approach.

In an MRI-guided AC in whole-body PET imaging study, Arabi et al. [44] reported REs (%) of 15.8 ± 10.1 , -1.0 ± 6.8 and 3.9 ± 11.7 in the lung and -9.6 ± 8.2 , -3.4 ± 5.2 and -4.9 ± 6.7 in soft-tissue when using segmentation-based, atlas-based and Hofmann's pseudo CT generation approaches, respectively. Mehranian et al. [45] performed a clinical assessment of MLAA-guided MRI-constrained Gaussian mixture model (GMM) algorithm and 4-class MRI segmentation-based (MRAC) attenuation correction in whole-body time TOF PET/MR imaging. They reported an average SUV RE % (and RMSE) of -5.4 ± 12.0 (13.1) and -3.5 ± 6.6 (7.5) in the lung, -7.4 ± 1.8 (7.6) and -5.4 ± 3.2 (6.3) in the liver, and -9.2 ± 6.0 (11.0) and -3.1 ± 6.8 (7.5) in the myocardium for MRAC and MLAA-GMM methods, respectively. Hofmann et al. [46] evaluated segmentation and atlas-pattern recognition (AT&PR)-based whole-body AC in PET/MRI and reported REs (%) of 14.1 ± 10.2 and 7.7 ± 8.4 for the segmentation and the AT&PR approaches, respectively. Segmentation-based techniques require substantial pre-processing steps and have limited performance in the lung and bone regions because of weak signals produced by both tissues. Different MR sequences address the signal generation problem in conventional MR sequences. However, they are limited by image artefacts, high levels of noise, limited diagnostic value and long acquisition time. Atlas-based AC methods are affected by registration accuracy, which is limited by inter/intra-patients variability.

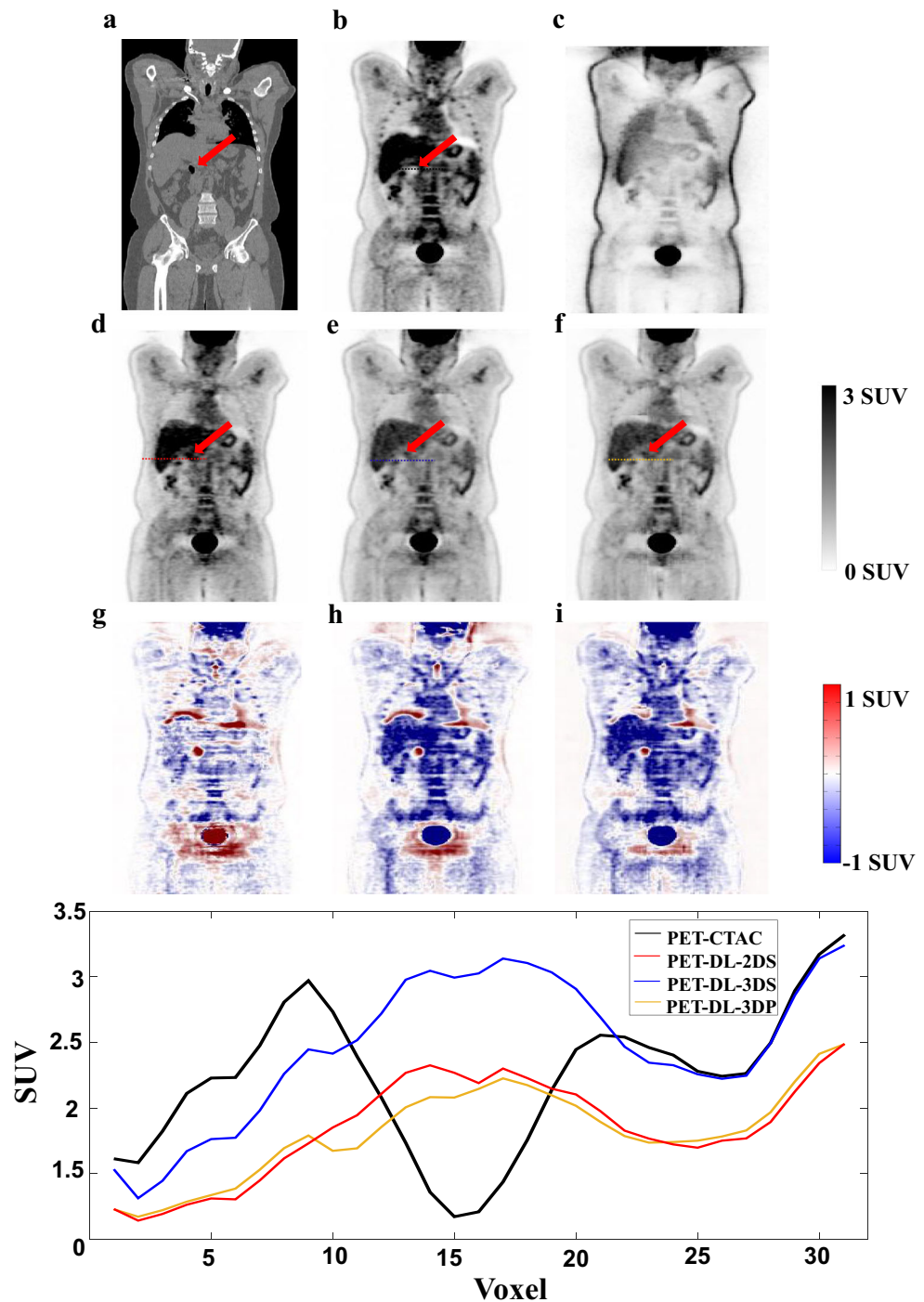
Chang et al. [47] developed a PET-nonASC-based segmentation using an active contour algorithm for PET AC and reported SUV bias of $6\% \pm 7\%$ in whole-body imaging. More recently, Rezaei et al. [11] applied joint reconstruction of activity and attenuation in TOF PET imaging, achieving a RE (%) of -7.5 ± 4.6 in the whole-body. Likewise, Salomon et al. [48] exploited a different framework for implementation of the joint activity and attenuation correction and reported averaged REs

of -10.3 , -2.9 , -2.0 and -5.7 for bone, soft-tissue, fat and lung regions, respectively. In an investigation of the impact of TOF reconstruction on PET quantification, Mehranian et al. [49] reported REs of -3.4 ± 11.5 for the lung and -21.8 ± 2.9 for bone using a non-TOF MRI-guided AC method. In this study, the voxel-wise REs across the 150 patients of the independent external validation set were $-1.72 \pm 4.22\%$, $3.75 \pm 6.91\%$ and $-3.08 \pm 5.64\%$ for DL-2DS, DL-3DS and DL-3DP methods, respectively. The DL-2DS method led to lower RE compared with both DL-3DS (p value < 0.02) and DL-3DP (p value < 0.05) approaches. In the DL-2DS technique, the lowest and highest REs were $0.5 \pm 5.04\%$ in the neck region and $-4.7 \pm 4.02\%$ in the lung region. For DL-3DS, the lowest and highest REs were 0.41 ± 6.86 in the lung-liver and 5.81 ± 4.85 in the pelvic region. The REs in the lung region were $-4.7 \pm 4.02\%$, $2.54 \pm 5.61\%$ and $-1.01 \pm 3.03\%$ for DL-2DS, DL-3DS and DL-3DP, respectively. Among the different Deep-JASC approaches, DL-2DS resulted in the lowest SUV error metrics (RMSE = 0.22 ± 0.08 , RE = $-1.72 \pm 4.22\%$ and SSIM = 0.96 ± 0.03). The highest error metrics (ARE = $15.16 \pm 3.96\%$, lowest PSNR = 31.7 ± 2.52 and SSIM = 0.93 ± 0.03) were observed in the chest region.

Overall, the 2DS approach exhibited superior performance over the different body regions, except the chest region (Fig. 3). Smaller ARE% and MAE were observed in the brain, neck, lung/liver, abdomen and pelvis regions when using the 2DS approach compared with 3D modes. This trend was reversed in the chest region where 3D approaches led to significantly smaller ARE% and MAE compared with the 2DS approach. This observation could be partly attributed to the complexity and irregularity of attenuation correction in this region wherein insufficient structural/underlying uptake information in 2D slices led to the inferior performance of the 2DS approach. Despite the higher ARE% and MAE, 2DS exhibited better performance with respect to PSNR and SSIM metrics, which may be due to the patchy appearance of the images in the 3D approaches.

The average errors calculated over either whole-body or regional body regions reflected mostly the technical performance of Deep-JASC approaches. However, a detailed analysis of tracer uptake in malignant lesions clarified the relevant clinical performance of these approaches. Both 2DS and 3DS approaches resulted in less than 4% errors in SUV_{mean} and SUV_{max} estimation (outperforming the 3DP approach with more than 5% REs). Moreover, the significantly lower standard deviations achieved by 2DS confirmed the better performance of this approach. The high tracer uptake, signal-to-noise ratio and contrast of malignant lesions make them easy to resolve in PET-nonASC images. As such, Deep-JASC performed relatively better on malignant lesions in terms of tracer uptake quantification compared with other body regions containing mixtures of tissues/activity concentration levels.

Fig. 7 Case report of an outlier showing: A pseudo hot spot appeared in the PET images corrected for attenuation/scatter using the different deep learning-based approaches. (A) CT, (B) PET-CTAC, (C) PET-nonASC, (D) PET-DL-2DS, (E) PET-DL-3DS, (F) PET-DL-3DP and the difference bias maps (G) PET-DL-2DS – PET-CTAC, (H) PET-DL-3DS – PET-CTAC and (I) PET-DL-3DP – PET-CTAC. The plot shows SUV profiles drawn through the pseudo hot spot



A major limitation of this approach is its vulnerability to noise beyond the normal/average levels of noise or signal-to-noise ratio (SNR) in PET-nonASC images existing in the training dataset. In such a situation, the deep learning-based technique would fail to discriminate noise from the underlying structures or distinguish genuine organ/lesion boundaries (supplemental Figures 4 and 5), thus potentially leading to noise-induced pseudo structures and/or loss of signal pertinent to underlying structures

(Supplemental Figure 6). The PET-nonASC image shown in Supplemental figure 4C suffers from low SNR reflected by poor organ boundaries particularly in the lung and shoulders regions where the deep learning-based approach failed to recover the actual activity distribution, leading to gross errors in the chest region. However, in other patients (e.g. Fig. 2), where PET-nonASC images bear relatively good SNR, the proposed approach succeeded to recover the actual activity distribution.

Although direct attenuation and scatter correction in the image domain is immune to mismatch between PET and CT images and potential respiratory motion-related errors, this approach warrants the implementation of thorough quality control measures of the resulting PET images to ensure accurate lesion conspicuity and quantitative accuracy. This is mainly due to the fact that this approach is still in an embryonic stage and requires further research and development efforts to shed light on the potential capabilities and pitfalls of this approach. Moreover, this approach performs the correction in an end-to-end fashion, and as such, failure of the deep learning approach would directly translate into errors in the resulting PET images.

We included disease-free and pathological patients with various indications, such as age, weight and disease type (for training and independent validation sets) resulting in a heterogeneous database. This study naturally bears a number of limitations. First, the training and independent validation processes were performed on only ^{18}F -FDG as a tracer. For different radiotracer distributions, the network will need to be retrained on PET images using other radiotracers. However, the network trained on ^{18}F -FDG PET images could be used to initialize networks trained on PET images acquired with other radiotracers via transfer learning [50]. This will also help to address the issue of having a limited training dataset. Second, all images were acquired on a PET/CT scanner. Hence, the network has not yet been validated for PET/MRI applications since this hybrid imaging uses rigid and surface MR coils (invisible in PET images) in the field-of-view. This will add challenges to the Deep-JASC approach, in particular for non-rigid surface coils as they may not impact PET-nonASC images similarly/repeatability across the subjects. The network should be validated for PET/MRI. Yet, PET/CT images included the CT couch, yet the network performed reasonably well.

Conclusion

This work demonstrated the feasibility of joint attenuation and scatter correction of whole-body ^{18}F -FDG PET images in the image space using a residual convolutional neural network (Deep-JASC) using an unprecedentedly large sample size. The proposed algorithm does not require the use of anatomical images (CT or MRI), thus providing a sensible solution in a clinical setting for standalone PET scanners and PET/MRI systems. Nevertheless, despite the promising results obtained, the technique is vulnerable to and noise-induced artefacts, which resulted in pseudo hot/cold spots or noticeable quantification errors across the organ boundaries, in particular between the lung and liver. Further research is guaranteed to improve the algorithm's performance.

Funding information This work was supported by the Swiss National Science Foundation under grant SNRF 320030_176052 and the Swiss Cancer Research Foundation under Grant KFS-3855-02-2016.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Research involving human participants All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Ben-Haim S, Ell P. ^{18}F -FDG PET and PET/CT in the evaluation of cancer treatment response. *J Nucl Med.* 2009;50:88–99.
2. Zaidi H, Karakatsanis N. Towards enhanced PET quantification in clinical oncology. *Br J Radiol.* 2018;91:20170508.
3. Boellaard R. Standards for PET image acquisition and quantitative data analysis. *J Nucl Med.* 2009;50:11S–20.
4. Zaidi H, Koral KF. Scatter modelling and compensation in emission tomography. *Eur J Nucl Med Mol Imaging.* 2004;31:761–82.
5. Kinahan PE, Hasegawa BH, Beyer T. X-ray-based attenuation correction for positron emission tomography/computed tomography scanners. *Semin Nucl Med.* 2003;33:166–79.
6. Mehranian A, Arabi H, Zaidi H. Vision 20/20: magnetic resonance imaging-guided attenuation correction in PET/MRI: challenges, solutions, and opportunities. *Med Phys.* 2016;43:1130–55.
7. Arabi H, Rager O, Alem A, Varoquaux A, Becker M, Zaidi H. Clinical assessment of MR-guided 3-class and 4-class attenuation correction in PET/MR. *Mol Imaging Biol.* 2015;17:264–76.
8. Martinez-Moller A, Souvatzoglou M, Delso G, Bundschuh RA, Chefd'hotel C, Ziegler SI, et al. Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: evaluation with PET/CT data. *J Nucl Med.* 2009;50:520–6.
9. Arabi H, Koutsouvelis N, Rouzaud M, Miralbell R, Zaidi H. Atlas-guided generation of pseudo-CT images for MRI-only and hybrid PET–MRI-guided radiotherapy treatment planning. *Phys Med Biol.* 2016;61:6531–52.
10. Mehranian A, Zaidi H. Joint estimation of activity and attenuation in whole-body TOF PET/MRI using constrained Gaussian mixture models. *IEEE Trans Med Imaging.* 2015;34:1808–21.
11. Rezaei A, Deroose CM, Vahle T, Boada F, Nuyts J. Joint reconstruction of activity and attenuation in time-of-flight PET: a quantitative analysis. *J Nucl Med.* 2018;59:1630–5.
12. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* 2017;42:60–88.
13. Sanaat A, Arabi H, Mainta I, Garibotto V, Zaidi H. Projection-space implementation of deep learning-guided low-dose brain PET imaging improves performance over implementation in image-space. *J Nucl Med.* 2020; *in press.*
14. Shiri I, Leung K, Geramifar P, Ghafarian P, Oveisi M, Ay MR, et al. PSFNET: ultrafast generation of PSF-modelled-like PET images using deep convolutional neural network [abstract]. *J Nucl Med.* 2019;60:1369.

15. Haggstrom I, Schmidlein 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.
16. Han X. MR-based synthetic CT generation using a deep convolutional neural network method. *Med Phys.* 2017;44:1408–19.
17. Jin CB, Kim H, Liu M, Jung W, Joo S, Park E, et al. Deep CT to MR synthesis using paired and unpaired data. *Sensors.* 2019;19:2361.
18. Bortolin K, Arabi H, Zaidi H. Deep learning-guided attenuation and scatter correction in brain PET/MRI without using anatomical images. *IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC).* Manchester, UK; 2019.
19. Arabi H, Zeng G, Zheng G, Zaidi H. Novel adversarial semantic structure deep learning for MRI-guided attenuation correction in brain PET/MRI. *Eur J Nucl Med Mol Imaging.* 2019;46:2746–59.
20. Yang J, Park D, Gullberg GT, Seo Y. Joint correction of attenuation and scatter in image space using deep convolutional neural networks for dedicated brain (18)F-FDG PET. *Phys Med Biol.* 2019;64:075019.
21. Spuhler KD, Gardus J 3rd, Gao Y, DeLorenzo C, Parsey R, Huang C. Synthesis of patient-specific transmission image for PET attenuation correction for PET/MR imaging of the brain using a convolutional neural network. *J Nucl Med.* 2019;60:555–60.
22. Arabi H, Dowling JA, Burgos N, Han X, Greer PB, Koutsouvelis N, et al. Comparative study of algorithms for synthetic CT generation from MRI: consequences for MRI-guided radiation planning in the pelvic region. *Med Phys.* 2018;45:5218–33.
23. Jang H, Liu F, Zhao G, Bradshaw T, McMillan AB. Technical note: deep learning based MRAC using rapid ultra-short echo time imaging. *Med Phys.* 2018;45:3697–704.
24. Gong K, Yang J, Kim K, El Fakhri G, Seo Y, Li Q. Attenuation correction for brain PET imaging using deep neural network based on Dixon and ZTE MR images. *Phys Med Biol.* 2018;63:125011.
25. Leynes AP, Yang J, Wiesinger F, Kaushik SS, Shanbhag DD, Seo Y, et al. Zero-echo-time and Dixon deep pseudo-CT (ZeDD CT): direct generation of pseudo-CT images for pelvic PET/MRI attenuation correction using deep convolutional neural networks with multiparametric MRI. *J Nucl Med.* 2018;59:852–8.
26. Shiri I, Ghafarian P, Geramifar P, Leung KH, Ghelichoghli M, Oveisi M, et al. Direct attenuation correction of brain PET images using only emission data via a deep convolutional encoder-decoder (deep-DAC). *Eur Radiol.* 2019;29:6867–79.
27. Rezaei A, Defrise M, Bal G, Michel C, Conti M, Watson C, et al. Simultaneous reconstruction of activity and attenuation in time-of-flight PET. *IEEE Trans Med Imaging.* 2012;31:2224–33.
28. Mehranian A, Arabi H, Zaidi H. Quantitative analysis of MRI-guided attenuation correction techniques in time-of-flight brain PET/MRI. *Neuroimage.* 2016;130:123–33.
29. Hwang D, Kim KY, Kang SK, Seo S, Paeng JC, Lee DS, et al. Improving the accuracy of simultaneously reconstructed activity and attenuation maps using deep learning. *J Nucl Med.* 2018;59:1624–9.
30. Hwang D, Kang SK, Kim KY, Seo S, Paeng JC, Lee DS, et al. Generation of PET attenuation map for whole-body time-of-flight (18)F-FDG PET/MRI using a deep neural network trained with simultaneously reconstructed activity and attenuation maps. *J Nucl Med.* 2019;60:1183–9.
31. Wolterink JM, Dinkla AM, Savenije MHF, Seevinck PR, van den Berg CAT, Išgum I. Deep MR to CT synthesis using unpaired data. *International Workshop on Simulation and Synthesis in Medical Imaging SASHIMI.* Cham: Springer International Publishing; 2017. p. 14–23.
32. Dong X, Lei Y, Wang T, Higgins K, Liu T, Curran WJ, et al. Deep learning-based attenuation correction in the absence of structural information for whole-body PET imaging. *Phys Med Biol.* 2020;65(5):055011.
33. Watson CC. New, faster, image-based scatter correction for 3D PET. *IEEE Trans Nucl Sci.* 2000;47:1587–94.
34. Li W, Wang G, Fidon L, Ourselin S, Cardoso MJ, Vercauteren T. On the compactness, efficiency, and representation of 3D convolutional networks: brain parcellation as a pretext task. In: Niethammer M, Styner M, Aylward S, Zhu H, Oguz I, Yap P-T, et al., editors. *Information Processing in Medical Imaging.* Cham: Springer International Publishing; 2017. p. 348–60.
35. Gibson E, Li W, Sudre C, Fidon L, Shakir DI, Wang G, et al. NiftyNet: a deep-learning platform for medical imaging. *Comput Methods Prog Biomed.* 2018;158:113–22.
36. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, et al. TensorFlow: a system for large-scale machine learning. *Proceedings of the 12th USENIX conference on Operating Systems Design and Implementation.* Savannah, GA, USA: USENIX Association; 2016. p. 265–83.
37. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Proc.* 2004;13:600–12.
38. Arabi H, Zaidi H. One registration multi-atlas-based pseudo-CT generation for attenuation correction in PET/MRI. *Eur J Nucl Med Mol Imaging.* 2016;43:2021–35.
39. Shi L, Onofrey JA, Revilla EM, Toyonaga T, Menard D, Ankrah J, et al. A novel loss function incorporating imaging acquisition physics for PET attenuation map generation using deep learning. *International Conference on Medical Image Computing and Computer-Assisted Intervention: Springer;* 2019. p. 723–31.
40. Liu F, Jang H, Kijowski R, Bradshaw T, McMillan AB. Deep learning MR imaging-based attenuation correction for PET/MR imaging. *Radiology.* 2017;286:676–84.
41. Liu F, Jang H, Kijowski R, Zhao G, Bradshaw T, McMillan AB. A deep learning approach for (18)F-FDG PET attenuation correction. *EJNMMI Phys.* 2018;5:24.
42. Van Hemmen H, Massa H, Hurley S, Cho S, Bradshaw T, McMillan A. A deep learning-based approach for direct whole-body PET attenuation correction [abstract]. *J Nucl Med.* 2019;60:569.
43. Yang X, Lei Y, Dong X, Wang T, Higgins K, Liu T, et al. Attenuation and scatter correction for whole-body PET using 3D generative adversarial networks [abstract]. *J Nucl Med.* 2019;60:174.
44. Arabi H, Zaidi H. Magnetic resonance imaging-guided attenuation correction in whole-body PET/MRI using a sorted atlas approach. *Med Image Anal.* 2016;31:1–15.
45. Mehranian A, Zaidi H. Clinical assessment of emission-and segmentation-based MR-guided attenuation correction in whole-body time-of-flight PET/MR imaging. *J Nucl Med.* 2015;56:877–83.
46. Hofmann M, Bezrukov I, Mantlik F, Aschoff P, Steinke F, Beyer T, et al. MRI-based attenuation correction for whole-body PET/MRI: quantitative evaluation of segmentation-and atlas-based methods. *J Nucl Med.* 2011;52:1392–9.
47. Chang T, Diab RH, Clark JW Jr, Mawlawi OR. Investigating the use of nonattenuation corrected PET images for the attenuation correction of PET data. *Med Phys.* 2013;40:082508.
48. Salomon A, Goedicke A, Schweizer B, Aach T, Schulz V. Simultaneous reconstruction of activity and attenuation for PET/MR. *IEEE Trans Med Imaging.* 2011;30:804–13.
49. Mehranian A, Zaidi H. Impact of time-of-flight PET on quantification errors in MR imaging-based attenuation correction. *J Nucl Med.* 2015;56:635–41.
50. Mesnil G, Dauphin Y, Glorot X, Rifai S, Bengio Y, Goodfellow I, et al. Unsupervised and transfer learning challenge: a deep learning approach. *Proceedings of the 2011 International Conference on Unsupervised and Transfer Learning workshop-Volume 27: JMLR. org;* 2011. p. 97–111.

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