Deep learning-guided estimation of attenuation correction factors from time-of-flight PET emission data

Hossein Arabi[a], Habib Zaidi[a,b,c,d,*]

[a] Division of Nuclear Medicine and Molecular Imaging, Department of Medical Imaging, Geneva University Hospital, CH-1211 Geneva 4, Switzerland
[b] Geneva Neuroscience Center, Geneva University, CH-1205 Geneva, Switzerland
[c] Department of Nuclear Medicine and Molecular Imaging, University of Groningen, University Medical Center Groningen, 9700 RB Groningen, Netherlands
[d] Department of Nuclear Medicine, University of Southern Denmark, DK-500 Odense, Denmark

ARTICLE INFO

ABSTRACT

Purpose: Attenuation correction (AC) is essential for quantitative PET imaging. In the absence of concurrent CT scanning, for instance on hybrid PET/MRI systems or dedicated brain PET scanners, an accurate approach for synthetic CT generation is highly desired. In this work, a novel framework is proposed wherein attenuation correction factors (ACF) are estimated from time-of-flight (TOF) PET emission data using deep learning.

Methods: In this approach, referred to as called DL-EM), the different TOF sinogram bins pertinent to the same slice are fed into a multi-input channel deep convolutional network to estimate a single ACF sinogram associated with the same slice. The clinical evaluation of the proposed DL-EM approach consisted of 68 clinical brain TOF PET/CT studies, where CT-based attenuation correction (CTAC) served as reference. A two-tissue class consisting of background-air and soft-tissue segmentation of the TOF PET non-AC images (SEG) as a proxy of the technique used in the clinic was also included in the comparative evaluation. Qualitative and quantitative PET analysis was performed through SUV bias maps quantification in 63 different brain regions.

Results: The DL-EM approach resulted in $6.1 \pm 9.7\%$ relative mean absolute error (RMAE) in bony structures compared to SEG AC method with RMAE of $16.1 \pm 8.2\%$ (*p*-value <0.001). Considering the entire head region, DL-EM led to a root mean square error (RMSE) of 0.3 ± 0.01 outperforming the SEG method with RMSE of 0.8 ± 0.02 SUV (*p*-value <0.001). The region-wise analysis of brain PET studies revealed less than 7% absolute SUV bias for the DL-EM approach, whereas the SEG method resulted in more than 14% absolute SUV bias (*p*-value <0.05).

Conclusions: Qualitative assessment and quantitative PET analysis demonstrated the superior performance of the DL-EM approach over the segmentation-based technique with clinically acceptable SUV bias. The results obtained using the DL-EM approach are comparable to state-of-the-art MRI-guided AC methods. Yet, this approach enables the extraction of interesting features about patient-specific attenuation which could be employed not only as a stand-alone AC approach but also as complementary/prior information in other AC algorithms.

1. Introduction

Positron emission tomography (PET) has emerged as a unique molecular imaging modality enabling the non-invasive assessment of biological processes at the cellular level as well as the characterization of brain diseases. Brain PET scanning using ^{18}F-fluorodeoxyglucose (FDG) is commonly performed in clinical set-ting. Quantitative PET imaging plays a key role in the characterization as well as the early detection of brain diseases. Accurate correction for attenuated and scattered photons is an essential step towards dependable and clinically relevant quantitative PET imaging (Zaidi et al., 2007).

Computed tomography (CT)-based derivation of attenuation correction (AC) maps provided by commercial hybrid PET/CT scanners used in the clinic is considered the gold standard technique since it provides adequate patient-specific modeling of photon attenuation. However, systems lacking transmission scanning, such

as stand-alone dedicated brain PET and hybrid PET/MRI scanners still face the challenge of accurate attenuation map generation. For the latter, MRI signal intensity of conventional sequences lacks direct correlation with attenuation characteristics of biological tissues (Zaidi et al., 2003). Attenuation and scatter corrections are not only crucial for quantitative PET analysis, but also have tremendous impact on PET image quality and hence visual clinical interpretation. In the absence of transmission scanning, commonly used strategies to estimate AC maps relying on structural MRI can be divided into three generic approaches (Mehranian et al., 2016b): (i) Segmentation-based approaches, which classify MR images into a number of tissue classes with distinct attenuation coefficients. Thereafter, predefined attenuation coefficients are assigned to each tissue class (Arabi et al., 2015; Martinez-Moller et al., 2009; Schulz et al., 2011). Despite the implementation of these approaches on commercial units and their use in the clinic, their performance is quite limited. Dedicated MR sequences enabling to discriminate between air and bone could potentially provide fairly accurate AC maps (Keereman et al., 2010; Sekine et al., 2016). (ii) Atlas-based approaches rely on a number of co-registered MR and CT pairs and a transformation rule enabling to predict synthetic CT images for the target subject (Wollenweber et al., 2013). This is achieved through using an established mapping function between MR and CT data (Arabi et al., 2016; Arabi and Zaidi, 2016b) or a nonlinear learning algorithm to estimate AC maps from target MR images (Arabi and Zaidi, 2016a; Hofmann et al., 2011). (iii) Joint estimation of activity/attenuation distribution from PET emission data using time-of-flight (TOF) information enables attenuation map estimation even without the use of structural imaging (Defrise et al., 2012). However, exploitation of anatomical information, for instance MR images, could potentially improve the joint attenuation/activity estimation process (Mehranian et al., 2017).

The popularity and promising results achieved by machine learning and particularly deep learning approaches in various image analysis applications, including CT synthesis from MR images stimulated tremendous growth in the field (Han, 2017; Liu et al., 2017). Deep learning approaches provide an efficient and versatile framework for accurate CT synthesis from a single or multiple MR sequences (Emami et al., 2018; Gong et al., 2018; Han, 2017; Liu et al., 2018) in the brain (Arabi et al., 2019; Dinkla et al., 2018) and pelvis (Leynes et al., 2018) for either MRI-guided PET AC or MRI-only treatment planning in radiation therapy. Deep learning algorithms have also been adopted to improve the accuracy of joint reconstruction of activity and AC maps in brain PET imaging (Hwang et al., 2018).

Deep learning algorithms were initially utilized to predict synthetic CT images from structural MRI for PET attenuation correction prior to image reconstruction. However, correction for attenuation and scatter has been recently incorporated directly in the image space without the need for anatomical images (Bortolin et al., 2019; Shiri et al., 2019; Yang et al., 2019). In this approach, deep learning is exploited to learn an end-to-end regression from non-attenuation corrected PET (PET-nonAC) to CT-based attenuation corrected PET (PET-CTAC) images in the image domain.

PET emission data intrinsically bears sort of subject-specific information about the attenuating medium. As illustrated in Fig. 1A, in the object space, the activity distribution is independent of the attenuation map. Yet, the events recorded by the PET detectors bear a strong correlation with the attenuating medium. The counts recorded by the PET detectors convey not only information about the actual activity distribution in the object space but also object-specific information about the attenuated photons along the different lines of response (LORs). Nevertheless, this information would be lost during conventional image reconstruction (owing to the averaging effect of the reconstruction algorithm) to generate PET-nonAC images (Fig. 1B). Hence, deep learning-based approaches

proposed to correct PET images for attenuation and scatter in the image domain would not lead to unique subject-specific solutions. Yet, the information pertinent to the attenuation medium is reflected in the sinogram space (Fig. 1C), and as such, attenuation map estimation from the sinogram or through direct attenuation correction in the sinogram space could potentially result in subject-specific solutions. Exploiting the latent information in the projection domain has also been investigated in a hybrid approach benefiting from maximum likelihood reconstruction of activity and attenuation (MLAA) and convolutional neural networks to yield a patient-specific attenuation map from the PET data. To this end, a novel loss function based on line integral projection modeling the physics of PET attenuation was introduced into convolutional neural networks (Shi et al., 2019).

In this work, a novel approach is proposed to predict patient-specific attenuation correction factors (ACFs) from the PET emission data using deep learning approaches (DL-EM). In this regard, the PET emission sinogram data are used as input for training of a deep convolutional neural network to estimate the corresponding attenuation map or, more precisely, the ACF sinograms. The ACF sinograms derived from reference CT images were used as reference based on which the deep convolutional network was trained and evaluated. The predicted ACF sinograms were used to perform attenuation correction within PET image reconstruction and evaluated against reference CT-based attenuation correction. The performance of the proposed approach was compared to a segmentation-based method as proxy to techniques used in the clinic to provide the bottom line of clinically relevant performance.

2. Materials and methods

2.1. PET/CT data acquisition

Time-of-flight (TOF) PET/CT data of 68 clinical brain studies were retrospectively used for training and evaluation the proposed DL-EM approach. Brain scans were performed on the Biograph mCT PET/CT scanner (Siemens Healthcare, Erlangen, Germany) following intravenous injection of 205 ± 15 MBq of ^{18}F-FDG. Table 1 summarizes the patient demographics of the clinical ^{18}F-FDG PET brain studies. PET data acquisition started 31 ± 5 min post-injection and lasted 20 min. Low-dose CT scanning for attenuation correction (kVp = 120, mAs = 20, rotation speed = 0.3 s and voxel size = $1 \times 1 \times 2.5$ mm^3) preceded PET data acquisition. PET raw data were saved in list-mode format to allow offline reconstruction of PET images using synthetic CT images and ACF sinograms. The offline Siemens e7 reconstruction tool was employed to generate emission sinograms from PET raw data and ACF matrices from reference CT-based AC maps for the training of the DL-EM algorithm. The mCT scanner equipped with TOF capability with time resolutions of 530 ps generates emission sinograms in 13 TOF bins (312 ps wide).

2.2. ACF estimation from PET emission data

As shown in Fig. 1, PET emission data represented in sinogram format conveys patient-specific information about the attenuating medium in the object space. The objective of the DL-EM approach is to extract the latent information from the emission data to predict ACFs which represents a unique solution for activity/emission correspondence problem. To this end, two distinct strategies were followed: (i) A single sinogram matrix containing non-TOF PET emission data (segment = 9 and matrix size = $400 \times 168 \times 621$) was considered as input to DCNN learning to find a unique solution for the estimation of ACFs. Fig. 2A illustrates this procedure where the ACF matrices obtained from CT-based AC were considered as reference to train the DCNN. It should be noted that all brain PET
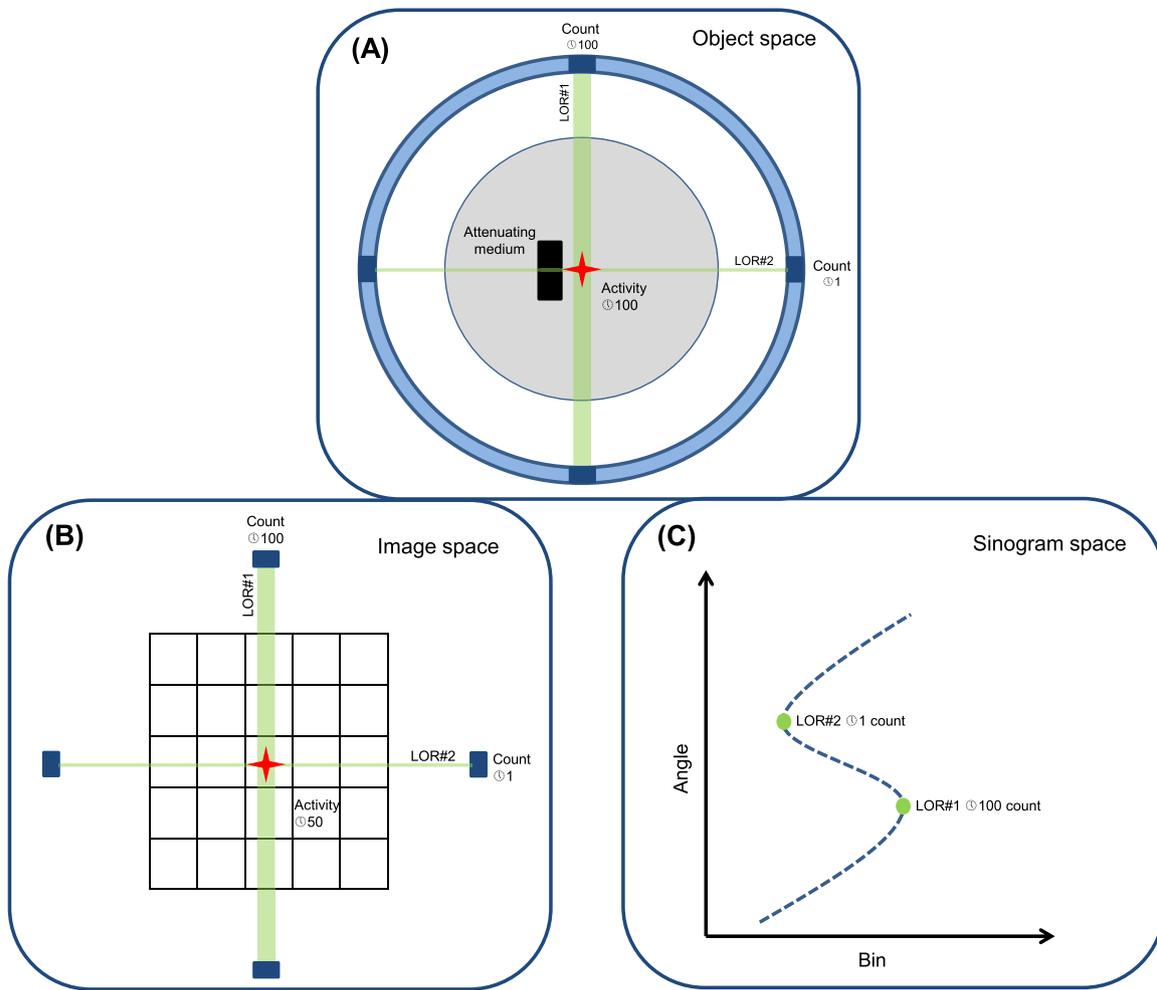
Fig. 1. A) Illustration of the impact of the attenuating medium on the distribution of acquired counts across PET detectors. B) The information about the attenuation medium is lost in image space (PET-nonAC images) due to the averaging nature of the image reconstruction process. C) The subject-specific information about the attenuation map, which intrinsically exists in emission PET data is reflected in the sinogram space.

Table 1
Patient demographics and clinical characteristics of the ^{18}F-FDG PET brain studies.

	Age (mean, range)	Gender	Weight	Injected activity	Indication/Diagnosis
68 PET/CT scans	61.3 (50–82) yrs	32F/36M	67.5 kg	205±15 MBq	Cognitive symptoms of possible neurodegenerative etiology

scans were performed with TOF capability; however, in this procedure, the 13-bin TOF sinograms generated using e7 tool were summed up to generate a single non-TOF sinogram. The ACFs estimation from emission data was examined without using TOF information since this procedure is memory efficient given the large size of the 4D sinogram data, which increase by 13-fold when TOF information is used. Using non-TOF emission sinogram facilitated the training and evaluation of the DL-EM approach; however, remarkable overfitting and sub-optimal performance were observed. (ii) Given the non-satisfactory results obtained from non-TOF data, the second procedure involved the use of TOF information as displayed in Fig. 2B. In this method, the TOF sinograms pertinent to the same acquisition plane were used as input to DCNN in parallel to predict a single ACF matrix corresponding to the same acquisition plane. In this regard, a multi-channel DCNN was employed to simultaneously process the different TOF bin sinograms. The rationale behind this strategy is that the different TOF bin sino-

grams generated from the same acquisition plane passed through the same attenuating medium, and as such, they convey information pertinent to ACFs. In other words, TOF bin sinograms contain different representations of the attenuation map encoded latently in the distribution of emission counts. Hence, the features that the DCNN extracts from the TOF PET emission data could rise by up to 13-fold (according to the TOF resolution of the mCT scanner) compared to non-TOF PET acquisition.

As indicated in Fig. 2B, only 7 out of 13 TOF bins (−3 to +3) were utilized for the training/evaluation of the DL-EM approach, owing to the relatively small size of head with respect to the transaxial FOV of the PET scanner (7 TOF bins cover the entire head region while the remaining 6 TOF bins are almost empty and practically convey no information disregarding scattered radiation). This training of the DCNN was computationally intensive when the TOF information was incorporated, but enabled to achieve promising results.
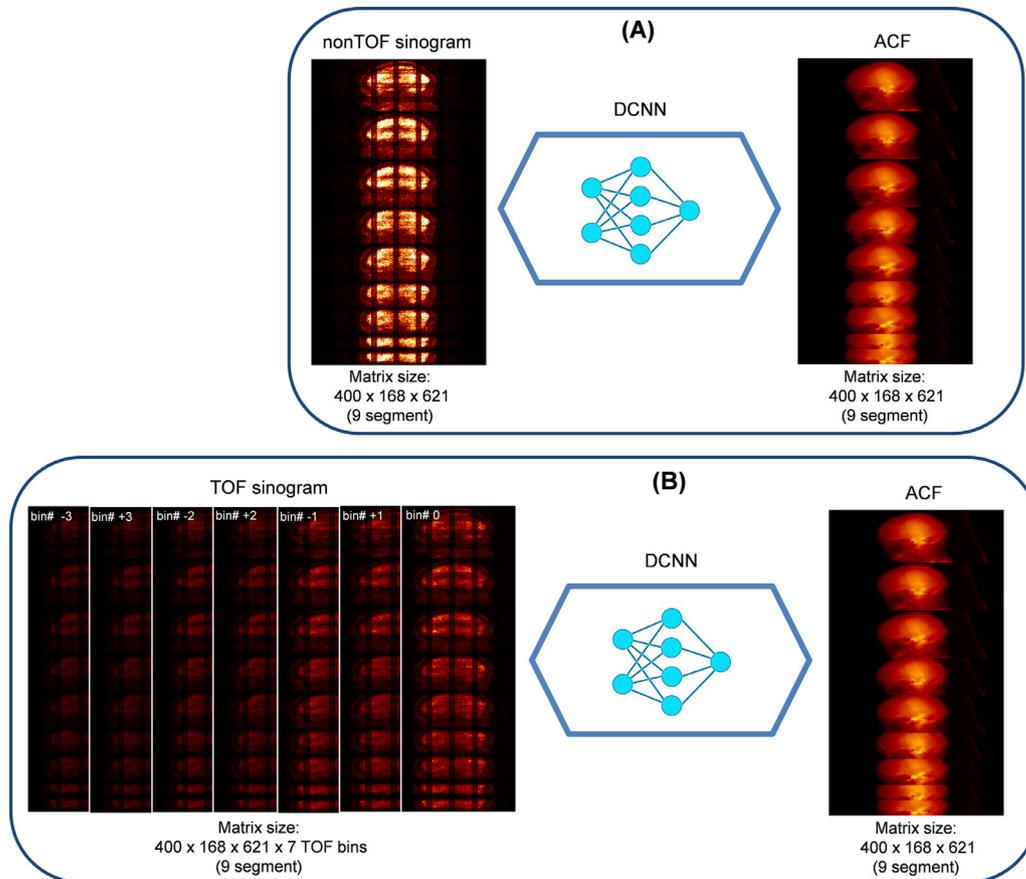
Fig. 2. Attenuation correction factors (ACFs) estimation from PET emission data using deep convolutional neural network (DCNN). A) ACFs prediction from non-TOF PET emission sinogram and B) ACFs prediction from TOF PET emission sinogram using 7 TOF bins (−3 to 3) projected in 9 segments.

2.3. Deep learning framework

Niftynet, an open–source pipeline for implementation of deep convolutional neural networks, was adopted to implement the DL-EM approach to predict ACFs from PET emission data. This infrastructure provides a modular deep-learning platform for the accomplishment of common medical image analysis applications, such as image segmentation (Gibson et al., 2018). Niftynet, built upon TensorFlow module in Python environment, allows for rapid and efficient implementation of deep learning algorithms either from scratch or using state-of-the-art deep learning architectures.

The DL-EM approach was implemented using the regression module embedded within the Niftynet platform accessible through net_regress application command. A compact DCNN, referred to as highresnet, was trained to estimate ACFs from the PET emission data. This DCNN, originally developed for volumetric image segmentation, consists of twenty convolutional layers with convolution kernels of $3 \times 3 \times 3$ voxel (Li et al., 2017). The low level image features are extracted by the first seven layers, while the following seven and six layers capture medium and high level image features using dilated convolutional kernels by factors of two and four, respectively. A residual connection links every two consecutive convolutional layers followed by a batch normalization layer.

2.3.1. Implementation details

Siemens e7 reconstruction tool allows the generation of the following data from the PET list mode files and CT scan: a) TOF PET sinograms containing the prompts (true and random events) in 13 TOF bins and 9 segments (matrix size = $13 \times 400 \times 168 \times 621$),

Table 2
The emission PET data in the form of sinograms was used for training of the DL-EM approach with and without correction for randoms, normalization and scatter. Five different scenarios were examined through combining correction of randoms, normalization and scatter obtained from SEG (Scatter-SEG) and CT-based (Scatter-CT) AC methods to the emission sinogram.

	Normalization	Randoms	Scatter-SEG	Scatter-CT
Scenario #1	×		×	×
		×		
Scenario #2	✔	×	×	×
Scenario #3	✔	✔	×	×
Scenario #4	✔	✔	✔	×
Scenario #5	✔	✔	×	✔

b) Random sinograms in a single matrix for all TOF bins (matrix size = $400 \times 168 \times 621$), c) ACF sinogram in 9 segments (matrix size = $400 \times 168 \times 621$), d) Normalization factors (matrix size = $400 \times 168 \times 621$) and scatter sinograms (matrix size = $400 \times 168 \times 109$). It should be noted that the above values correspond to a single bed position for a brain scan resulting in PET images with a matrix size of $200 \times 200 \times 109$ voxels.

As shown in Fig. 2B, the DL-EM approach employs only 7 out of 13 TOF bin emission sinograms to predict the ACFs (the remaining 6 TOF bins are almost empty). Before training the DL-EM approach, the TOF PET sinograms are corrected for randoms, scatter and normalization. In this regard, five different scenarios were investigated to determine the most efficient setting for implementation of the DL-EM approach. Table 2 summarizes five potential scenarios

where in the first scenario the training of the DL-EM approach was performed using the 7 TOF bin emission sinograms without applying any correction for randoms, normalization and scatter. Scenario #2 involves only the normalization factors whereas scenario #3 involves both randoms correction and normalization factors. Scatter events in the PET sinogram would have an adverse impact on the performance of the DL-EM approach since they provide wrong information about the activity concentration and attenuation correction factors. To account for scattered events, the TOF scatter sinogram can be estimated from a simple synthetic CT image derived from TOF PET-nonAC images. To this end, the TOF PET-nonAC images, which bear strong signal at the boundary of body/head and background air, underwent head contour segmentation to separate background air from the head region. Subsequently, attenuation coefficients of air and soft-tissue were assigned to voxels outside and inside the head contour, respectively, to generate SEG attenuation map (SEG AC map). Given the SEG AC map, PET data were reconstructed using the Siemens e7 reconstruction tool with the default clinical parameters to obtain the TOF scatter sinogram. Siemens e7 tool uses the single-scatter simulation (SSS) and tail fitting algorithm implemented on the Biograph mCT scanner to estimate TOF scatter sinograms. This scatter sinogram was used to correct the PET sinogram data prior to application of the DL-EM approach to reduce the adverse impact of scattered events. Scenario #4 was examined wherein the scattered photons were estimated from SEG AC map (Scatter-SEG) and the PET data were corrected for normalization, random and scattered events (Scatter-SEG). Slight improvement was observed when scenario #4 was examined compared to scenario #3. Hence, the investigation was further extended to include reference CT-based scatter estimation (Scatter-CT) in scenario #5. The only difference between scenarios #3 and #4 is the correction for scattered events estimated from the SEG AC map. To obtain Scatter-CT, CT-based reconstruction of the PET data was performed and the resulting TOF scatter sinogram was used to correct PET data prior to application of the DL-EM. This scenario, though not realistic, was investigated to put into perspective the extent to which the scattered events would impact the performance of the DL-EM approach.

All TOF PET sinogram counts, after applying the random and scatter correction as well as normalization, were normalized to a range of 0–1 by using a fixed intensity value for all subjects. Similarly, the inverse of the ACF matrices, where the intensities ranged from almost 0 to 1, were used for the training of the DL-EM approach.

The highresnet convolutional network, implemented in Niftynet, was trained using the 7 PET emission sinograms corresponding to 7 TOF bins as input and the corresponding ACFs as the output. The following hyper parameter values were set for the training/evaluation: spatial window = 168 × 200, learning rate = 0.002 and 0.001 (following an optimization procedure proposed in (Smith, 2018)), optimizer = Adam, loss function = L2norm, decay = 0.0001 and 0 (following an optimization procedure proposed in (Smith, 2018)), batch size = 70 and sample per subject = 2. The details of the hyperparameters (config file) used for the training of the DL-EM approach are provided in the Supplemental materials.

The Biograph mCT scanner generates TOF sinograms in 621 oblique and direct planes for each bed position. The training of the DL-EM approach was carried out plane-wise in such a way that each iteration deals with the 7 TOF bin sinograms (7 TOF bins × 400×168 voxels) of that plane (either direct or oblique) as input to learn the corresponding ACF sinogram (400×168 voxels). Therefore, the DL-EM model takes into account all direct and oblique planes to generate the ACF matrix (621× 400×168 voxel) of a bed position in a plane-wise manner.

The training and generation of the ACFs (inference) were performed in a two-dimensional manner wherein all the emission and attenuation data in the sinogram space corresponding to a single acquisition plane (in the projection space) was considered as one training sample. Prior to training and inference, the emission sinograms (7 two-dimensional sinograms corresponding to 7 meaningful TOF bins) belonging to a single plane were corrected for randoms, normalization, scattered events (considering scenario #4 in Table 2). Randoms smoothing and normalization are readily provided by the reconstruction software; however, the scatter distribution should be calculated from the SEG AC map prior to training or inference. The ACF matrix (readily provided by the scanner reconstruction software) does not undergo any major preprocessing and is employed as reference within the training of the DL-EM approach. In the inference phase, the estimated ACF matrix can also be directly incorporated within the PET image reconstruction process. In Fig. 2, all sinograms and ACF matrices associated with a single bed acquisition (in 9 segments) are shown as input and output of the DL-EM approach. However, in practice, each single row of the 7 emission sinogram matrices and ACF matrix are used as input and output of the network, respectively.

The training and evaluation of the DL-EM approach were performed based on a four-fold cross-validation framework. Therefore, the quantitative analysis is reported for the entire dataset.

2.4. Evaluation

The performance evaluation of the DL-EM approach included validation against the reference CT-based AC and comparison to a segmentation-based method generating a two-tissue class AC map (SEG). SEG AC maps obtained from segmentation of the head contour from the TOF PET-nonAC images followed by the assignment of 0.1 cm^{-1} ($\approx$ 0 HU) and 0.0 cm^{-1} (≈ -1000 HU) attenuation coefficients to voxels inside and outside the head contour, respectively. SEG approach was included in the comparative assessment as a representative approach similar to the commercial segmentation-based algorithm implemented on the Philips TF PET/MRI scanner (Zaidi et al., 2011).

Given the ACFs obtained from the DL-EM approach, SEG and reference CT AC maps, PET raw data was reconstructed three times resulting in PET-DL-EM, PET-SEG and reference PET-CT images. An ordinary Poisson ordered subsets-expectation maximization (OP-OSEM) algorithm with 5 iteration and 21 subsets was used for PET image reconstruction followed by post-reconstruction Gaussian filtering with 2 mm FWHM. The reconstruction included scatter, decay, random and dead-time corrections.

Taking PET-CT images as reference, standardized uptake value (SUV) bias was calculated for PET-DL-EM and PET-SEG images within different tissue types. To this end, CT images were segmented into bone, soft-tissue and air cavities using the following intensity threshold levels: >160 HU for bone, between −400 and 160 HU for soft-tissue and <−400 HU inside the head contour for air cavities. The relative mean error (RME) and relative mean absolute error (RMAE) were computed for DL-EM and SEG approaches against reference CT-based AC using Eqs. (1) and 2, respectively.

$$RME\ (\%) = \frac{1}{P} \sum_{p=1}^{P} \left(100 \times \frac{PET_{tst}(p) - PET_{ct}(p)}{PET_{ct}(p)} \right) \quad (1)$$

$$RMAE\ (\%) = \frac{1}{P} \sum_{p=1}^{P} \left| 100 \times \frac{PET_{tst}(p) - PET_{ct}(p)}{PET_{ct}(p)} \right| \quad (2)$$

where P indicates the total number of voxels inside the head regions and $PET_{tst}(p)$ represents the SUV for the voxel p within either PET-DL-EM or PET-SEG images.
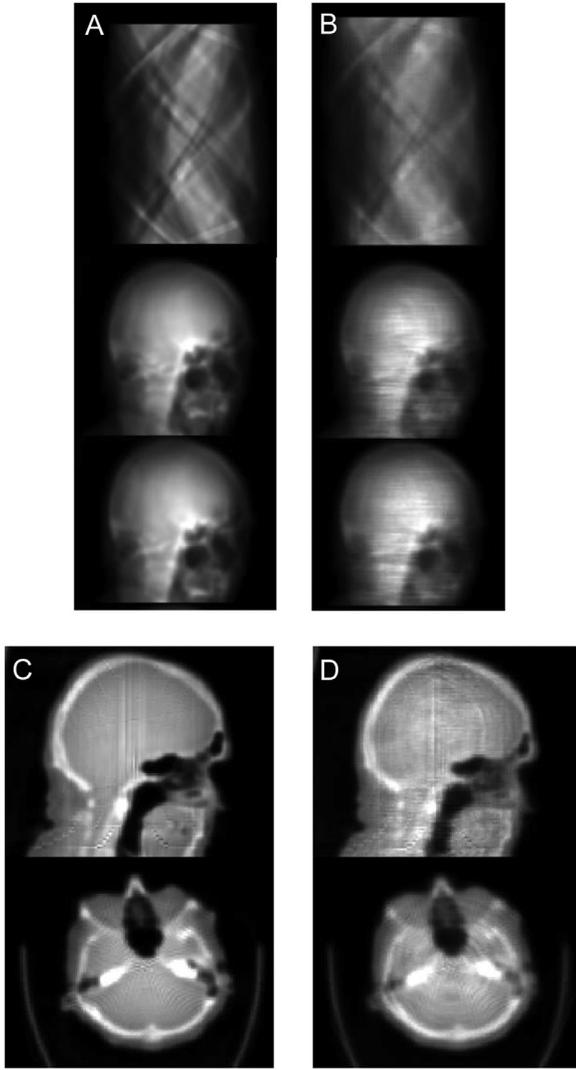
Fig. 3. A) ACF sinograms obtained from CT-based AC. B) ACF sinograms predicted by the DL-EM approach. C) PET AC map reconstructed from the reference CT ACF sinograms. D) PET AC map reconstructed from the predicted ACF sinograms using the DL-EM approach.

Moreover, the root mean square error (RMSE), peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were calculated using the Eqs. (3), 4 and 5, respectively.

$$RMSE = \sqrt{\frac{1}{P} \sum_{p=1}^{P} (PET_{tst}(p) - PET_{ct}(p))^2} \qquad (3)$$

$$PSNR(dB) = 10\log_{10}\left(\frac{Mval^2}{MSE}\right) \qquad (4)$$

$$SSIM = \frac{(2 \times mean\ (PET_{ct}) \times mean\ (PET_{tst}) + K_1)(2\delta_{PETct,PETtst} + K_2)}{\left(mean\ (PET_{ct})^2 + mean\ (PET_{tst})^2 + K_1\right)\left(\delta^2_{PETct} + \delta^2_{PETtst} + K_2\right)} \qquad (5)$$

In Eq. (4), *Mval* denotes the maximum intensity value of PET_{ct} or PET_{tst} images whereas *MSE* indicates the mean squared error. In Eq. (5), *mean()* returns the mean value of the image whereas δ_{PETct} and δ_{PETtst} stand for variances of PET_{ct} and PET_{tst} images, respectively, and $\delta_{PETct,PETtst}$ is their covariance. K_1 and K_2 (K_1=0.01 and K_2=0.02) are variables were used to avoid division by zero or very small values.
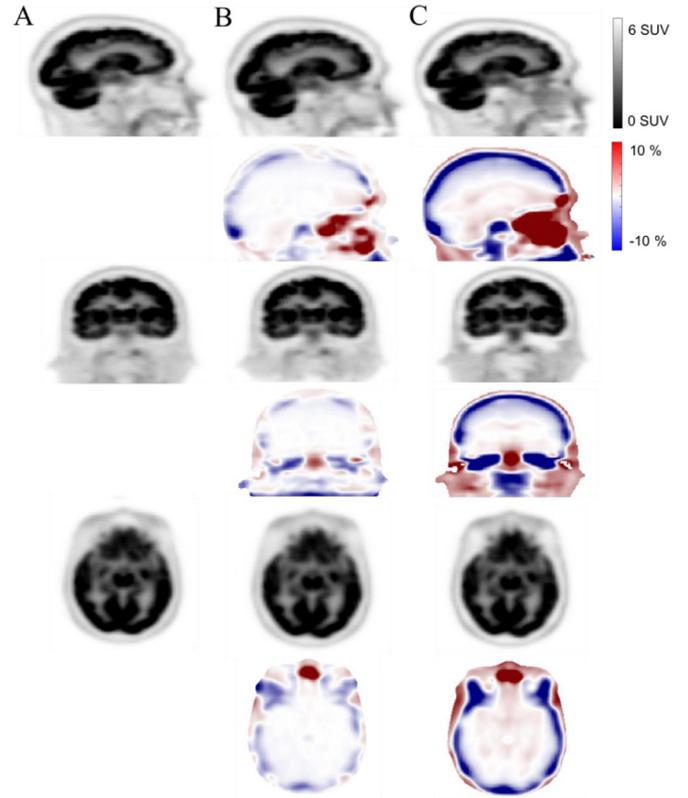

Fig. 4. PET attenuation corrected images using reference CT-based, DL-EM and SEG approaches. A) PET-CT. B) PET-DL-EM. C) PET-SEG. The SUB bias maps for DL-EM and SEG approaches are also shown.

The quantitative evaluation of DL-EM and SEG AC approaches was performed through SUV bias calculation in 63 brain regions. To this end, the Hermes BRASS brain analysis tool (Hermes Medical Solutions AB, Sweden), providing standard PET and MRI templates to transform test images into a common spatial coordinate, was utilized to align test PET images to the template followed by region-of interest analysis. The relative SUV bias (*RB(%)*) and absolute relative bias (*ARB(%)*) were calculated for each region using Eqs. (6) and 7, respectively. *R* indicates the brain region index.

$$RB(\%) = \frac{(PET_{tst})_R - (PET_{ct})_R}{(PET_{ct})_R} \times 100\% \qquad (6)$$

$$ARB(\%) = \left| \frac{(PET_{tst})_R - (PET_{ct})_R}{(PET_{ct})_R} \right| \times 100\% \qquad (7)$$

Moreover, joint histogram analysis was carried out to depict correlations between voxel-wise tracer distribution in PET-DL-EM and PET-SEG versus reference PET-CT images. Statistical significance of the difference between the results obtained in this work was assessed using paired *t*-test analysis where a *p*-value less than 0.05 indicates statistical significance.

3. Results

The DL-EM approach predicts ACF sinograms in an end-to-end fashion from the emission sinograms corrected for randoms, scatter and normalization. Fig. 3B depicts a representative ACF sinogram generated using the DL-EM method along with the reference CT-based ACFs shown in Fig. 3A. The AFCs were estimated for all 9 segments (only two segments are shown in Fig. 3) simultaneously, thus leading to a sinogram matrix of 400 × 168 × 621 elements. The ACF matrices depicted in Figs. 3A and 3B are presented
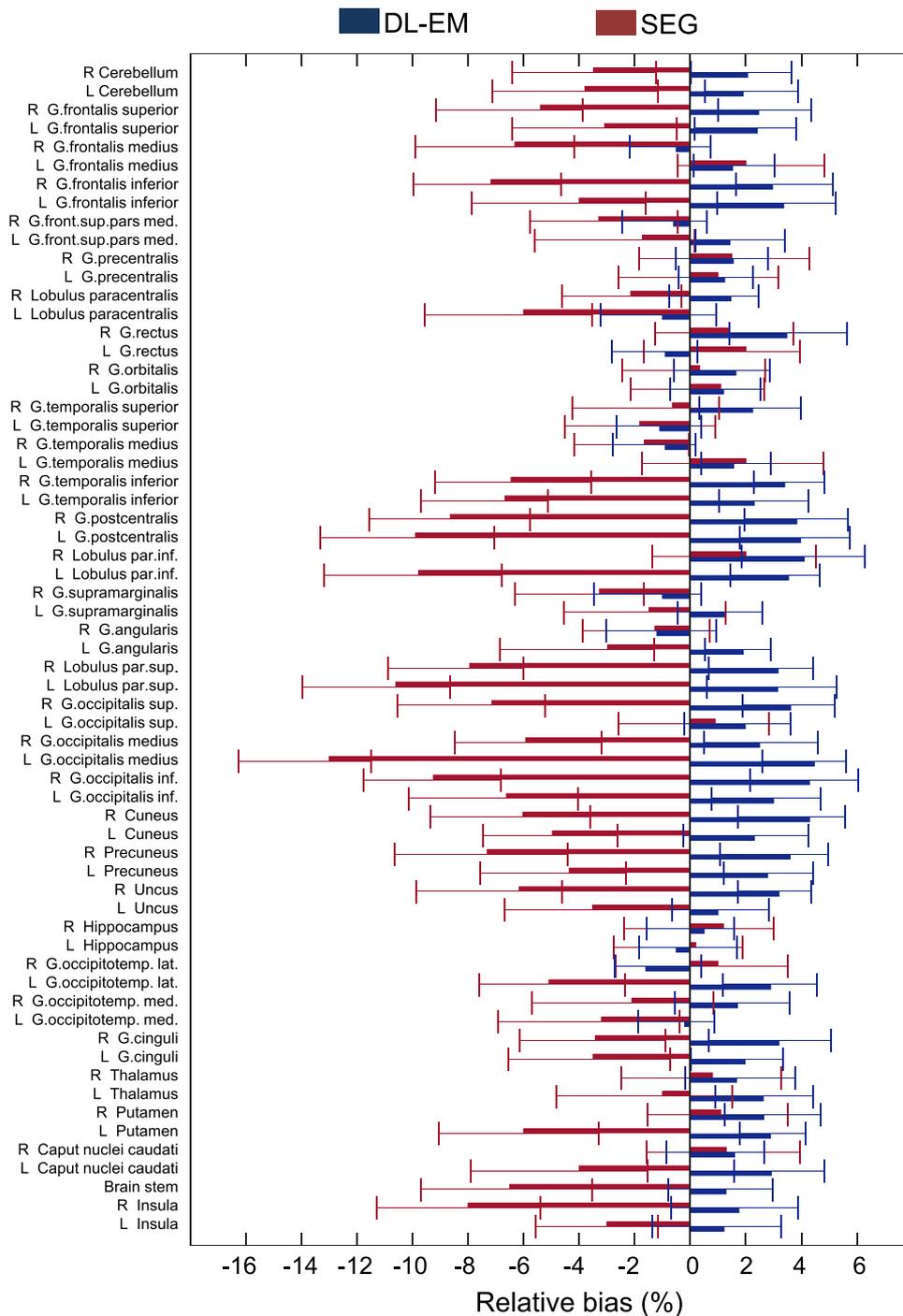
Fig. 5. Mean relative SUV bias measured in 63 brain regions for the DL-EM and SEG attenuation correction approaches over 68 patients.

in the projection space according to the mCT PET scanner geometry. These ACF matrices were reconstructed using the default PET reconstruction algorithm used in clinical setting to convert the ACF information from projection space to image space solely for the purpose of visualizing the anatomical details. Since the PET reconstruction algorithm assumes Poisson noise in the projection space and uses the PET system matrix and correction factors for the reconstruction of the ACF matrices, some artifacts appeared in the image domain (Fig. 3C and 3D). These artifacts are the result of mismatch between the ACF matrices and the PET reconstruction algorithm, which can be dealt with through the use of a dedicated

reconstruction algorithm. It is worth noting that attenuation correction of PET images was performed directly using the ACF sinograms.

The quantitative analysis of PET images corrected for attenuation using SEG and DL-EM approaches is presented in Table 3 for different tissue types considering CT-based AC as reference. Regarding quantification bias within bone tissue, the DL-EM approach led to RME and RMAE of 1.2% and 6.1%, respectively, while the SEG method led to RME of 15.0% and RMAE of 16.1%. Table 3 summarizes the results associated with scenario #4 in Table 2 as this scenario led to the best performance. The results of the alterna-

Table 3

PET quantification errors measured in air cavities, bone and soft-tissue as well as the entire head region for DL-EM and SEG AC approaches in terms of relative mean error (RME) and relative mean absolute error (RMAE).

	Soft-tissue RME ± SD (RMAE ± SD)	Bone RME ± SD (RMAE ± SD)	Air cavity RME ± SD (RMAE ± SD)	Head RME ± SD (RMAE ± SD)
DL-EM	2.0 ± 10.6 (4.1 ± 9.3)	1.2 ± 10.2 (6.1 ± 9.7)	5.3 ± 13.7 (5.8 ± 13.2)	2.9 ± 3.1 (3.9 ± 8.5)
SEG	−2.8 ± 11.1 (5.8 ± 9.0)	−15.0 ± 9.1 (16.1 ± 8.2)	43.3 ± 11.8 (45.9 ± 9.2)	−5.9 ± 4.7 (8.1 ± 9.9)
p-value	0.05 (0.002)	<0.001 (<0.001)	<0.001 (<0.001)	0.002 (0.001)

Table 4

SSIM, PSNR (dB) and RMSE (SUV) calculated between PET-SEG and PET-DL-EM images and reference PET-CT images.

	RMSE (SUV)	PSNR (dB)	SSIM
PET-DL-EM	0.3 ± 0.01	33.6 ± 1.4	0.93±0.02
PET-SEG	0.8 ± 0.02	30.9 ± 1.1	0.90±0.02
p-value	<0.001	0.001	0.001

tive versions of the DL-EM approach, namely non-TOF scenario #1, #2, #3 and #5 are provided in Supplemental Table 1. Fig. 4 shows representative views of PET images corrected for attenuation using reference CT-based AC, DL-EM and SEG approaches together with relative bias maps for PET-DL-EM and PET-SEG images versus PET-CT. The bias maps clearly show the overall reduced quantification error achieved by the DL-EM approach.

Moreover, the SSIM, PSNR and RMSE metrics calculated between PET-DL-EM and PET-SEG versus PET-CT images are reported in Table 4. The quantitative evaluation metrics demonstrated the superior performance of the DL-EM approach compared to the segmentation-based technique. The SSIM, PSNR and RMSE metrics associated with alternative versions of the DL-EM approach are provided in Supplemental Table 2.

The accuracy of tracer uptake quantification using DL-EM and SEG approaches was evaluated in different brain regions through mapping PET images to the BRASS template. The region-wise analysis of PET images is reflected in Fig. 5 showing the mean SUV bias measured in 63 brain regions for DL-EM and SEG approaches. The average SUV bias observed across the 68 patients is reported for each brain region. The differences between the results presented in Fig. 5 were all statistically significant ($p<0.05$) except for the Precentralis, Orbitalis, Angularis, Puntamen and Hippocampus regions (p-values = 0.07, 0.08, 0.08, 0.06 and 0.09, respectively). Likewise, the absolute mean SUV bias for the different brain regions is portrayed in Fig. 6 where the absolute SUV bias for the DL-EM approach was less than 7% for all brain regions, while the SEG method led to more than 14% SUV bias ($p<0.05$). The voxel-wise mean SUV bias along with the standard deviation of the bias maps is shown in Fig. 7. The SUV bias maps are calculated across PET images of 68 patients mapped to the BRASS template. The voxel-wise correlation plots between the tracer uptake in PET-DL-EM and PET-SEG images versus PET-CT images are portrayed in Fig. 8. The linear regression analysis was performed over the 68 patients resulting in a high correlation coefficient for the DL-EM approach (R^2=0.99). The linear regression analysis demonstrated underestimation of tracer uptake and a lower correlation coefficient (R^2=0.982) when using the SEG technique. It should be emphasized that the results presented in this section were achieved for scenario #4 in Table 2 as this setting led to superior performance.
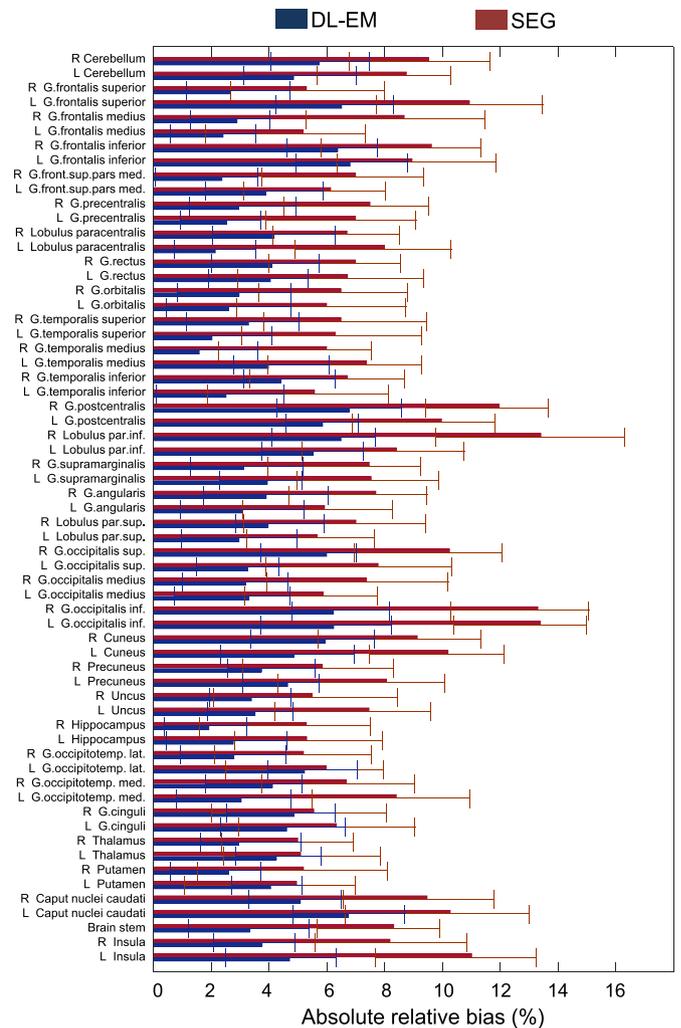

Fig. 6. Mean absolute relative SUV bias measured in 63 brain regions for the DL-EM and SEG attenuation correction approaches over 68 patients.

4. Discussion

The foundation of the proposed deep learning (DL-EM) approach lies in the fact that PET emission data convey object-specific information about the photon attenuating medium within the FOV of the PET acquisition. Though this information would be lost within the image reconstruction process (in image space), emission sinogram encodes this data in the form of relative count changes across different detector bins and angles. The DL-EM approach was intended to decode this information from the PET emission data aiming at synthesizing an object-specific ACF sinogram or PET attenuation map. In this regard, the DL-EM approach was initially considered with the non-TOF emission data (Fig. 2A) which resulted in sub-optimal performance in terms of ACF estimation showing evident signs of over-fitting. Inclusion of the TOF information resolved the over-fitting issue and led to an acceptable ACF estimation. Training of the DL-EM approach involved only seven out of thirteen TOF emission bins, which contained essential information for brain PET studies given the active transaxial FOV and TOF resolution of the Biograph mCT scanner. To this end, the convolutional neural network was modified to have seven input parallel channels to simultaneously process the TOF emission bins (Fig. 2B). Different TOF emission bins belonging to a specific acquisition plane convey different information about the activity distribution, but the same information about the attenuating medium
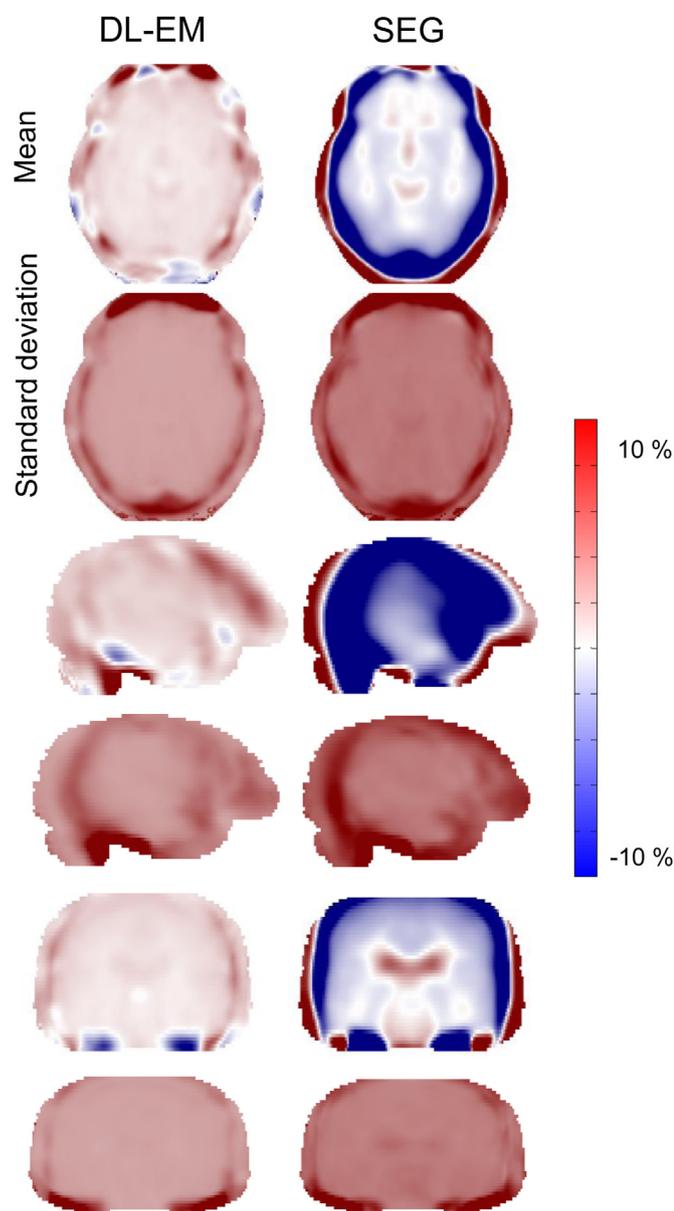
Fig. 7. SUV bias map, mean and standard deviation, in transaxial, coronal and sagittal views calculated over 68 patients for the DL-EM and SEG attenuation correction approaches.

(Surti, 2015). As such, employing the TOF data assisted the DL-EM approach to decode the latent information about photon attenuation from emission sinograms.

The emission sinograms obtained from the Biography mCT TOF PET/CT scanner contained prompt (sum of trues and randoms) coincidences; and as such, different scenarios were examined to find the most efficient setting for the training of the DL-EM approach (Table 2). Correction for normalization, scatter and randoms prior to the training of the DL-EM approach (scenario #4) led to the least test, validation and training loss. However, correction for scatter obtained from CT-based AC did not impact the quality of DL-EM outcome. Undoubtedly, emission sinograms correction using scatter coincidences estimated from CT-based AC is not a practical option. Nevertheless, this scenario was examined to investigate the impact of scatter correction on the performance of the DL-EM approach, wherein no substantial improvement was observed. This observation may only be valid for this specific dataset or brain

imaging where the scatter fraction is relatively low (estimated to be within about 20% for CT-based AC using the single scatter simulation method followed by tail fitting scaling). For pelvis or whole-body PET imaging, this assumption might not be valid, wherein a more accurate scatter correction prior to training of the DL-EM approach may be crucial. This issue warrants further investigation in future studies.

The DL-EM approach learns to estimate the whole ACF sinogram involving 9 segments (for the Biograph mCT scanner) simultaneously. However, as a matter of fact, the different sinogram segments contain highly correlated information about the ACFs. This existing correlation can be utilized in favor of more efficient training of the DL-EM approach. For instance, the training could be performed in an iterative way where the estimated ACFs from direct planes in the first iteration are fed into later iterations as a prior information for estimation of the ACFs in oblique planes and vice versa.

The quantitative analysis of PET images demonstrated the superior performance of the DL-EM approach over SEG method, leading to reduced SUV bias in all brain regions. The SEG technique was included in this evaluation as it provides a bottom line of clinically tolerable inaccuracy (Varoquaux et al., 2014). Moreover, the visual inspection of the ACFs obtained from the DL-EM approach (Fig. 3) revealed that important features of patient-specific anatomical details could be extracted by this method, such as air cavities, small bony structures and sharp tissue boundaries, particularly between bone and air. This is considered as an outstanding achievement of this approach, which enables dependable estimation of anatomical structures based solely on PET emission data. Besides, the DL-EM approach, resulted in an SUV bias of less than 8%, while SEG AC approach led to SUV bias of up to 14%.

The overall positive bias reported in Table 3 is observed when using this approach. Despite the overall slight positive SUV bias, the relatively large standard deviation of the quantities reported in Table 3 demonstrates a certain range of both positive and negative SUV biases imposed by the DL-EM approach among the different subjects. Moreover, Fig. 5 demonstrates that there are certain regions with overall negative SUV bias across all subjects. Hence, the DL-EM approach doesn't result in systematic positive SUV bias wherein depending on the anatomical region or subject under study, negative SUV bias may be observed. Besides, It should be noted that the DL-EM approach tend to slightly overestimate the ACFs with very small values, such as air canals or sinus cavities, which would justify the overall slight positive bias observed when using this method.

Even though the quantitative PET assessment demonstrated promising performance of the DL-EM approach, comparative studies involving joint estimation of activity/attenuation, deep learning-based CT synthesis from MR images and more importantly atlas-based methods would shed light on the strengths and pitfalls of the DL-EM approach. Despite the promising results achieved by deep learning-based approaches in the area of CT synthesis from MR images (Fu et al., 2018; Xiang et al., 2018) and direct attenuation and scatter correction in the image domain (Yang et al., 2019), comparison studies have shown comparable performance of atlas-based methods (Arabi et al., 2018, 2019; Mehranian et al., 2016a), wherein the deep learning-based methods exhibited higher vulnerability to outliers, leading to gross errors as reported in (Arabi et al., 2018, 2019).

Lastly, our work didn't allow to draw any conclusions regarding the performance of the DL-EM approach with respect to other existing techniques, including atlas-based or MRI-guided synthetic CT generation, or to claim that this approach is the ultimate solution for PET AC in the absence of CT or transmission scanning. One significant finding of this study is that this approach enables the extraction of appealing patient-specific attenuation features, which
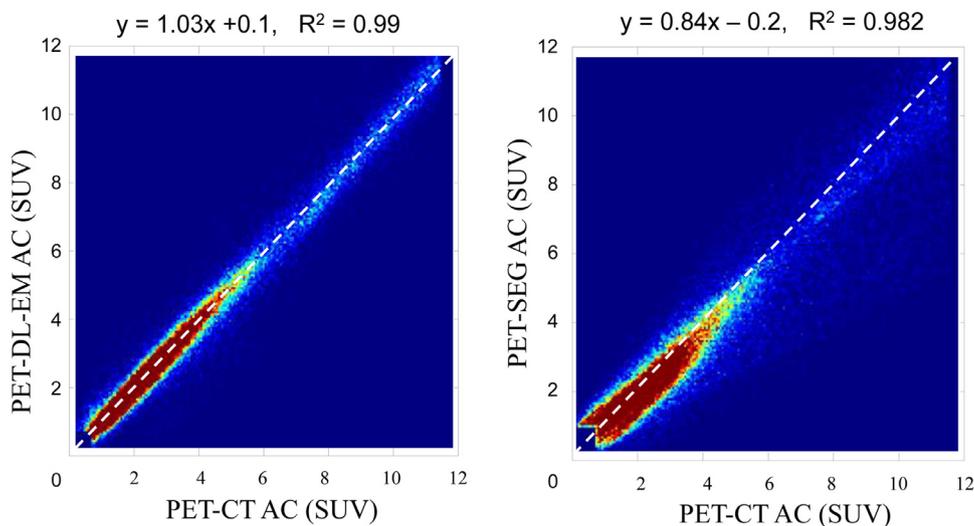
Fig. 8. Joint histograms of SUVs measured on PET-DL-EM (left) and PET-SEG (right) images versus reference PET-CT images.

could be utilized as complementary/prior information in other algorithms to enhance the overall accuracy of PET attenuation correction.

5. Conclusion

In the present work, the DL-EM approach was introduced, which enables estimation of the ACFs from TOF PET emission data. This approach relies on a deep learning algorithm to extract latent patient-specific information about the attenuating medium from TOF PET emission data in projection space (sinogram). Quantitative PET analysis demonstrated the superior performance of the DL-EM approach over the segmentation-based technique representative of algorithms used in the clinic. The DL-EM approach led to absolute SUV bias of less than 8% in all brain regions while the SEG method resulted in up to 14% absolute SUV bias. The results achieved by the DL-EM approach are comparable to other state-of-the-art MRI-guided AC methods without requiring anatomical information. However, future improvements in TOF PET temporal resolution would directly enhance the efficiency of this approach.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Hossein Arabi: Conceptualization, Methodology, Software, Writing - original draft. **Habib Zaidi:** Data curation, Investigation, Supervision, Writing - review & editing.

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.media.2020.101718.

References

Arabi, H., Dowling, J.A., Burgos, N., Han, X., Greer, P.B., Koutsouvelis, N., Zaidi, H., 2018. Comparative study of algorithms for synthetic CT generation from MRI: consequences for MRI-guided radiation planning in the pelvic region. Med. Phys. 45, 5218–5233.

Arabi, H., Koutsouvelis, N., Rouzaud, M., Miralbell, R., Zaidi, H., 2016. Atlas-guided generation of pseudo-CT images for MRI-only and hybrid PET–MRI-guided radiotherapy treatment planning. Phys. Med. Biol. 61, 6531–6552.

Arabi, H., Rager, O., Alem, A., Varoquaux, A., Becker, M., Zaidi, H., 2015. Clinical assessment of MR-guided 3-class and 4-class attenuation correction in PET/MR. Mol. Imaging Biol. 17, 264–276.

Arabi, H., Zaidi, H., 2016a. Magnetic resonance imaging-guided attenuation correction in whole-body PET/MRI using a sorted atlas approach. Med. Image Anal. 31, 1–15.

Arabi, H., Zaidi, H., 2016b. One registration multi-atlas-based pseudo-CT generation for attenuation correction in PET/MRI. Eur. J. Nucl. Med. Mol. Imaging 43 (11), 2021–2035.

Arabi, H., Zeng, G., Zheng, G., Zaidi, H., 2019. Novel adversarial semantic structure deep learning for MRI-guided attenuation correction in brain PET/MRI. Eur. J. Nucl. Med. Mol. Imaging 46 (13), 2746–2759.

Bortolin, K., Arabi, H., Zaidi, H., 2019. 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, pp. 1-3.

Defrise, M., Rezaei, A., Nuyts, J., 2012. Time-of-flight PET data determine the attenuation sinogram up to a constant. Phys. Med. Biol. 57, 885–899.

Dinkla, A.M., Wolterink, J.M., Maspero, M., Savenije, M.H., Verhoeff, J.J., Seravalli, E., Išgum, I., Seevinck, P.R., van den Berg, C.A., 2018. MR-only brain radiotherapy: dosimetric evaluation of synthetic CTs generated by a dilated convolutional neural network. Int. J. Radiat. Oncol. Biol. Phys. 102 (4), 801–812.

Emami, H., Dong, M., Nejad-Davarani, S.P., Glide-Hurst, C., 2018. Generating synthetic CT s from magnetic resonance images using generative adversarial networks. Med. Phys. 45 (8), 3627–3636.

Fu, J., Yang, Y., Singhrao, K., Ruan, D., Low, D.A., Lewis, J.H., 2018. Male pelvic synthetic CT generation from T1-weighted MRI using 2D and 3D convolutional neural networks. arXiv preprint arXiv:1803.00131.

Gibson, E., Li, W., Sudre, C., Fidon, L., Shakir, D.I., Wang, G., Eaton-Rosen, Z., Gray, R., Doel, T., Hu, Y., Whyntie, T., Nachev, P., Modat, M., Barratt, D.C., Ourselin, S., Cardoso, M.J., Vercauteren, T., 2018. NiftyNet: a deep-learning platform for medical imaging. Comput. Methods Programs Biomed. 158, 113–122.

Gong, K., Yang, J., Kim, K., El Fakhri, G., Seo, Y., Li, Q., 2018. Attenuation correction for brain PET imaging using deep neural network based on Dixon and ZTE MR images. Phys. Med. Biol. 63 (12), 125011–125013.

Han, X., 2017. MR-based synthetic CT generation using a deep convolutional neural network method. Med. Phys. 44, 1408–1419.

Hofmann, M., Bezrukov, I., Mantlik, F., Aschoff, P., Steinke, F., Beyer, T., Pichler, B.J., Scholkopf, B., 2011. MRI-based attenuation correction for whole-body PET/MRI: quantitative evaluation of segmentation- and Atlas-based methods. J. Nucl. Med. 52, 1392–1399.

Hwang, D., Kim, K.Y., Kang, S.K., Seo, S., Paeng, J.C., Lee, D.S., Lee, J.S., 2018. Improving the accuracy of simultaneously reconstructed activity and attenuation maps using deep learning. J. Nucl. Med. 59, 1624–1629.

Keereman, V., Fierens, Y., Broux, T., De Deene, Y., Lonneux, M., Vandenberghe, S., 2010. MRI-based attenuation correction for PET/MRI using ultrashort echo time sequences. J. Nucl. Med. 51, 812–818.

Leynes, A.P., Yang, J., Wiesinger, F., Kaushik, S.S., Shanbhag, D.D., Seo, Y., Hope, T.A., Larson, P.E., 2018. 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. 59, 852–858.

Li, W., Wang, G., Fidon, L., Ourselin, S., Cardoso, M.J., Vercauteren, T., 2017. On the compactness, efficiency, and representation of 3D convolutional networks: brain parcellation as a pretext task. In: International Conference on Information Processing in Medical Imaging. Springer, pp. 348–360.

Liu, F., Jang, H., Kijowski, R., Bradshaw, T., McMillan, A.B., 2017. Deep learning MR imaging–based attenuation correction for PET/MR imaging. Radiology 286, 676–684.

Liu, F., Jang, H., Kijowski, R., Zhao, G., Bradshaw, T., McMillan, A.B., 2018. A deep learning approach for 18 F-FDG PET attenuation correction. EJNMMI Phys. 5 ((1):24), 1–15.

Martinez-Moller, A., Souvatzoglou, M., Delso, G., Bundschuh, R.A., Chefd'hotel, C., Ziegler, S.I., Navab, N., Schwaiger, M., Nekolla, S.G., 2009. Tissue classification as a potential approach for attenuation correction in whole-body PET/MRI: evaluation with PET/CT data. J. Nucl. Med. 50, 520–526.

Mehranian, A., Arabi, H., Zaidi, H., 2016a. Quantitative analysis of MRI-guided attenuation correction techniques in time-of-flight brain PET/MRI. Neuroimage 130, 123–133.

Mehranian, A., Arabi, H., Zaidi, H., 2016b. Vision 20/20: magnetic resonance imaging-guided attenuation correction in PET/MRI: challenges, solutions, and opportunities. Med. Phys. 43, 1130–1155.

Mehranian, A., Zaidi, H., Reader, A.J., 2017. MR-guided joint reconstruction of activity and attenuation in brain PET-MR. Neuroimage 162, 276–288.

Schulz, V., Torres-Espallardo, I., Renisch, S., Hu, Z., Ojha, N., Börnert, P., Perkuhn, M., Niendorf, T., Schäfer, W., Brockmann, H., Krohn, T., Buhl, A., Günther, R., Mottaghy, F., Krombach, G., 2011. Automatic, three-segment, MR-based attenuation correction for whole-body PET/MR data. Eur. J. Nucl. Med. Mol. Imaging 38, 138–152.

Sekine, T., Ter Voert, E.E., Warnock, G., Buck, A., Huellner, .M..W, Veit-Haibach, P., Delso, G., 2016. Clinical evaluation of ZTE attenuation correction for brain FDG-PET/MR imaging-comparison with atlas attenuation correction. J. Nucl. Med. 57, 1927–1932.

Shi, L., Onofrey, J.A., Revilla, E.M., Toyonaga, T., Menard, D., Ankrah, J., Carson, R.E., Liu, C., Lu, Y., 2019. A novel loss function incorporating imaging acquisition physics for PET attenuation map generation using deep learning. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, pp. 723–731.

Shiri, I., Ghafarian, P., Geramifar, P., Leung, K.H.-Y., Ghelichoghli, M., Oveisi, M., Rahmim, A., Ay, M.R., 2019. Direct attenuation correction of brain PET images using only emission data via a deep convolutional encoder-decoder (Deep-DAC). Eur. Radiol. 21, 1–13.

Smith, L.N., 2018. A disciplined approach to neural network hyper-parameters: Part 1–learning rate, batch size, momentum, and weight decay. arXiv:1803.09820.

Surti, S., 2015. Update on time-of-flight PET imaging. J. Nucl. Med. 56, 98–105.

Varoquaux, A., Rager, O., Poncet, A., Delattre, B.M., Ratib, O., Becker, C.D., Dulguerov, P., Dulguerov, N., Zaidi, H., Becker, M., 2014. Detection and quantification of focal uptake in head and neck tumours: (18)F-FDG PET/MR versus PET/CT. Eur. J. Nucl. Med. Mol. Imaging 41, 462–475.

Wollenweber, S., Ambwani, S., Delso, G., Lonn, A., Mullick, R., Wiesinger, F., Piti, Z., Tari, A., Novak, G., Fidrich, M., 2013. Evaluation of an atlas-based PET head attenuation correction using PET/CT & MR patient data. IEEE Trans. Nucl. Sci. 60, 3383–3390.

Xiang, L., Wang, Q., Nie, D., Zhang, L., Jin, X., Qiao, Y., Shen, D., 2018. Deep embedding convolutional neural network for synthesizing CT image from T1-Weighted MR image. Med. Image Anal. 47, 31–44.

Yang, J., Park, D., Gullberg, G.T., Seo, Y., 2019. Joint correction of attenuation and scatter in image space using deep convolutional neural networks for dedicated brain (18)F-FDG PET. Phys. Med. Biol. 64 (7), 075019 11p.

Zaidi, H., Montandon, M.-.L., Meikle, S., 2007. Strategies for attenuation compensation in neurological PET studies. Neuroimage 34, 518–541.

Zaidi, H., Montandon, M.-.L., Slosman, D.O., 2003. Magnetic resonance imaging-guided attenuation and scatter corrections in three-dimensional brain positron emission tomography. Med. Phys. 30, 937–948.

Zaidi, H., Ojha, N., Morich, M., Griesmer, J., Hu, Z., Maniawski, P., Ratib, O., Izquierdo-Garcia, D., Fayad, Z.A., Shao, L., 2011. Design and performance evaluation of a whole-body Ingenuity TF PET-MRI system. Phys. Med. Biol. 56, 3091–3106.