Am J Nucl Med Mol Imaging 2026;16(1):44-54
www.ajnmmi.us /ISSN:2160-8407/ajnmmi0166625

Original Article

Using multimodal PET+MR data as conditional generative adversarial network inputs improves pseudo-CT and attenuation correction estimates for brain PET/MR

Jonathan Fisher[1,2], Emily Anaya[1,2], Garry Chinn[2], Craig S Levin[1,2,3,4]

[1]Department of Electrical Engineering, Stanford University, Stanford, CA, USA; [2]Department of Radiology, Stanford University, Stanford, CA, USA; [3]Department of Physics, Stanford University, Stanford, CA, USA; [4]Department of Bioengineering, Stanford University, Stanford, CA, USA

Received June 7, 2025; Accepted January 28, 2026; Epub February 15, 2026; Published February 28, 2026

Abstract: To obtain qualitatively and quantitatively accurate positron emission tomography (PET) images, the recorded PET emission data must be corrected for photon attenuation. Attenuation correction (AC) factors are typically estimated from X-ray computed tomography (CT) data acquired during an integrated PET/CT study. Estimating these factors from magnetic resonance (MR) data in an integrated PET/MR scanner is challenging, as MR images don't provide direct information about annihilation photon attenuation. Conditional generative adversarial networks (cGANs) have shown promising results for both emission-based and MR-based AC. This study explored whether combining these approaches could further improve brain PET AC accuracy. Thirty-five patients who received same-day whole-body PET/MR and PET/CT scans participated in this study. The non-attenuation-corrected and non-scatter-corrected (NASC) PET, MR, and CT reconstructed head regions were cropped and automatically co-registered. Four networks were trained to translate NASC PET and MR images into pseudo-CTs. Three used single-modality input, and the fourth used multi-modality. The multi-modality cGAN produced significantly better pseudo-CTs vs. the single-modality cGANs, with an average structural similarity index (SSIM) and dice similarity coefficients for bone, soft-tissue, and air of 0.865±0.001, 0.715±0.002, 0.915±0.001, and 0.567±0.004, respectively, vs. 0.841±0.001, 0.660±0.003, 0.894±0.001, and 0.524±0.005, for the single-modality cGANs with the best results. When comparing the AC PET reconstructed images, all cGANs outperformed the clinical atlas-based method used in commercially available PET/MR systems, and, as expected, the multi-modal cGAN achieved the highest quality results with average SSIM, and peak signal-to-noise ratio of 0.9987±0.0001, and 50.0±0.4, respectively, vs. 0.9913±0.0024, and 44.3±0.3 for the atlas method.

Keywords: PET, PET/MR, attenuation correction (AC), emission-based attenuation correction (EBAC), MR-based attenuation correction (MRAC), deep learning (DL), generative adversarial network (GAN)

Introduction

PET is a molecular imaging technique commonly used in oncology to detect and assess malignant lesions [1] and in neuroscience to classify and stage neurodegenerative diseases [2]. Most PET scanners are integrated with CT or magnetic resonance imaging (MRI), which generate complementary data, such as anatomical information. In PET/CT systems, CT data is also commonly used to derive linear attenuation coefficients (LAC) for PET attenuation correction (AC).

Several corrections must be made to the acquired emission data to obtain accurate PET images. These corrections include attenuation, scatter, randoms, and normalization. AC compensates for the reduction in the number of coincidence events detected by each pair of detectors due to the intersection of one or both annihilation photons with the patient's body. The most common types of interactions are Compton scatter and photoelectric effect. The photon attenuation depends on the effective atomic number (Z) and density of the tissues traversed before reaching the PET detectors. Tissues with higher Z and density, such as bone, attenuate more photons, leading to more significant attenuation artifacts in PET images. A μ-map, representing the LAC of the different tissues in the body at 511 keV, is used to correct the PET images for attenuation during image reconstruction. Since the CT signal is directly related to tissue density, applying a simple bilinear conversion on the CT data is used to generate a μ-map [3]. However, certain artifacts can appear in the attenuation-corrected PET images when using CT for AC. These artifacts can occur due to PET, and CT registration mismatch and propagation of artifacts present in CT, such as truncation and metal artifacts [4].

In standalone PET and PET/MR scanners, the lack of CT data makes AC more challenging. This is primarily because the MR signal is dependent on proton density and tissue relaxation properties, which do not exhibit a direct relationship with the attenuation coefficients [5].

Different solutions have been previously used to solve this AC problem. Most of the solutions can be roughly divided into five categories:

https://doi.org/10.62347/BONK5634

1. Contour fitting method [6]: Uses the emission sinograms to find the border between activity and scatter background. This border/contour is used to estimate the attenuation map. The method was originally used for the human head and provided a good first-order approximation of the μ-map by assuming that the entire head (defined by the contour) is made of soft tissue, with or without a skull (constant thickness of bone at the contour edge).

2. Sequence-based approaches [7, 8]: These approaches are only relevant for PET/MR scanners. Special MRI sequences, such as ultrashort echo time (UTE) and zero echo time (ZTE), were developed for MRI-based AC (MRAC). Unlike most MRI sequences, these sequences can distinguish between bone and air-filled cavities, which results in a more accurate estimation of μ-maps.

3. Template/atlas-based methods [9-11]: The template/atlas uses a reference μ-map with a co-registered PET and MRI image(s). After the new subject is scanned, a nonrigid registration algorithm is used to evaluate the deformation field between the reference and the new subject's MR/PET image(s). Applying this deformation to the reference μ-map generates a subject-registered μ-map.

4. Maximum-likelihood reconstruction of activity and attenuation (MLAA) algorithm [12-14]: enables the joint estimation of both the PET tracer distribution (activity) and the attenuation map directly from the emission data. The first MLAA solution, which didn't use timing information, lacks a unique solution and suffers from cross-talk between activity and attenuation estimates. To address this, a time-of-flight (TOF) MLAA was developed, using additional timing information to improve separation and reduce cross-talk errors.

5. Image-to-image translation using deep learning (DL) methods: This category can be divided into several subcategories: a. MRI to pseudo-CT (pCT) translation [15-18]: The input image(s) can either be an MR image(s) of diagnostic sequences (T1, T2), or non-diagnostic sequences (Dixon, UTE, ZTE). The MR-to-pCT translation method has also been used for other applications, such as MRI-only treatment planning [19, 20]. b. Uncorrected PET to pCT translation [21-23]: The input image(s) can either be non-attenuated corrected (NAC) PET images or non-attenuated and non-scatter corrected (NASC) images. c. Uncorrected PET to corrected PET translation [24-26]: NAC/NASC PET images are translated to fully corrected PET images.

To date, no previously published research has used both NASC PET and MR images as multimodal input to image-to-image DL translation model for pCT estimation. We hypothesize that using multimodal input provides additional information, enhancing the quality of the generated pCT compared to using only single-modality images. In this work, we explore and validate this idea.

Materials and methods

Dataset

This study included 36 patients. The patients' ages ranged from 25 to 79 years old. There were 16 females and 19 males, and one patient's sex was not listed. As a unique and significant aspect of this project, each patient received both a full-body PET/MR scan and a full-body PET/CT scan on the same day at Stanford Hospital.

The PET/MR scans were performed on a GE Healthcare Signa scanner. Each PET/MR scan acquired a 2-point Dixon sequence with a resolution of 1.95×1.95×5.2 mm^3 for Signa's scanner standard AC. Unfortunately, due to archiving issues, PET data from the PET/MR scan was unavailable, so we used PET data from the PET/CT studies in this work.

The PET/CT scans were performed on a GE Discovery 600 or 690 scanner. The CT scans were acquired at 140 kVp with resolutions of either 0.98×0.98×3.75 mm^3 or 1.37×1.37×3.75 mm^3, and the 18F-FDG PET emission images were acquired with a resolution of 3.65×3.65×3.27 mm^3.

Data preprocessing

Image registration is a crucial preprocessing step for generating accurate training pairs for image-to-image translation models [27]. To achieve this, we co-registered the head regions of the PET, MR, and CT volumes. Since the PET and CT volumes were acquired using the same scanner, they were inherently aligned, leaving only the MR volumes to be registered to the CT volumes.

Initially, we isolated the brain region from the PET/MR and PET/CT scans. In the majority of cases, patients positioned their arms alongside their bodies, allowing us to easily select the relevant slices. However, in some of the PET/CT scans, patients had their arms raised above their heads. For these instances, where the arms were raised above the head, we manually defined a bounding box that prioritizes minimizing impact on the head region, while removing as much of the arms as possible.

Once the PET/CT and PET/MR head volumes were obtained, we co-registered each pair of volumes, using SimpleITK [28], in the following way: 1. The MR volume was resampled to have the same resolution as the CT volume. 2. The MR volume was roughly aligned with the CT volume. 3. The roughly aligned MR was then registered to the fixed CT reference, utilizing the automatic 3D rigid registration using a multi-resolution approach to achieve a robust registration of MR and CT scans with efficient execution, which was previously developed in our group [29]. 4. The registered MR was subsequently registered to the fixed CT reference using a single-resolution affine registration step to refine the transformation parameters further.

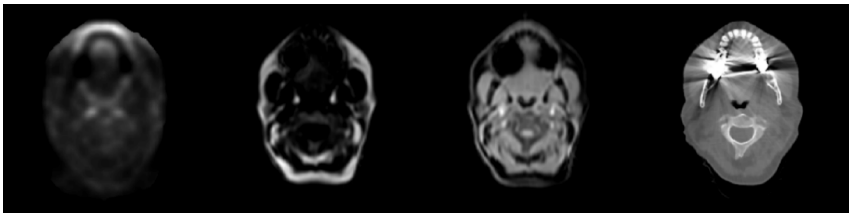

Figure 1. Left to right: Example of registered NASC PET images, MR Dixon Fat, MR Dixon Water, and CT axial slices with visible metal artifact in the MR and CT images.

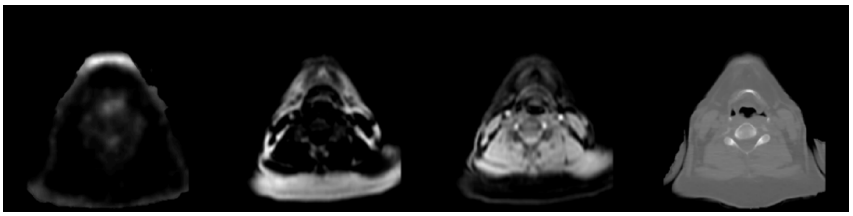

Figure 2. Left to right: Example of pixels corresponding to remainder of subtracted arms present in: NASC PET images, MR Dixon Fat, MR Dixon Water, and CT.

One patient's registration visibly failed and was discarded. For the remaining patients, we visually verified that the registration was performed successfully. The registration of all patients took approximately 414 seconds in total (~12 seconds per patient, for each of the 35 in this study).

After registering the PET, MR, and CT volumes, we obtained 2D grayscale images (single channel) from each modality, including MR Fat, MR Water, NASC PET, and CT images. We generated 100 axial slices for each patient with a resolution of 256×256 pixels. These single-modality 2D slices were used as inputs for the single-channel to single-channel image-to-image translational models. Additionally, we generated three-channel images by assigning MR Fat, MR Water, and NASC PET to separate image channels. These multi-modality, three-channel 2D slices were used as inputs for the multi-channel to single-channel image-to-image translational model.

It is important to note that some CT and MR images used for training contained metal artifacts, as illustrated in **Figure 1**, which can negatively impact the performance of the image translation model. Beyond metal artifacts, some images also contained residual arm structures. To address this, after registering the three modalities, an MR-based mask was applied to CT and NASC PET images to remove any remaining arm pixels. However, in some cases, the arms were positioned close to the head region, making complete removal challenging. **Figure 2** illustrates this issue, highlighting the areas where arm structures are still visible in the CT images.

Network structure

The pix2pix model described in [30] was used, and its implementation was downloaded from a GitHub repository [31]. The cGAN comprises a generator and discriminator. An eight-layer U-Net architecture with skip connections was used for the generator, which takes single or multiple imaging modalities (MR fat, MR water, NASC PET) as input(s) and generates a single channel pCT output. For the discriminator, we used a 70×70 PatchGAN. The discriminator gets pairs of images that include the input data with either the pCT or the real CT, and determines which pair contains the fake (generated) pCT and which contains the real CT.

Hyperparameter tuning, training, and testing

We tuned, trained, and tested four different networks, three performed single-channel to single-channel image-to-image translation, estimating the pCT using NASC PET, MR Water, or MR Fat individually, while the fourth performed multi-channel to single-channel translation using NASC PET, MR Water, and MR Fat as inputs.

Given the relatively small dataset of 35 patients, we applied 5-fold nested cross-validation (CV), as described in [32], by randomly splitting the data into five groups of seven patients each. For every fold, one group was held out for testing, while the remaining four were used for hyperparameter tuning with 4-fold CV. Hyperparameters selected for each of the 5 folds were those consistently associated with the best performance across all nested folds and all metrics presented in the network performance analysis section below. Best performance metrics include those who produced the top results or those not statistically different from it. Once tuned, the network was retrained on all four groups and tested on the held out fifth group.

The hyperparameters optimized in this process included the Adam optimizer momentum term (β_1), the initial learning rate, the weighting factor λ balancing the cGAN and L^1 loss objectives, and the number of epochs.

Network performance analysis

To quantitatively analyze the different networks' performance, the following metrics were used:

1. Mean squared error (MSE): $\frac{1}{N}\sum_{i=1}^{N}(CT_i - pCT_i)^2$, where CT is the ground truth CT image, pCT is the CT estimate generated by the network, N is the number of pixels in the 2D image, and $i \in [1...N]$ is the pixel index number. The MSE is measured in Hounsfield Units (HU). A smaller MSE $\in [0, \infty)$ indicates a smaller L^2 distance between the estimation and the ground truth, which means better consistency between model prediction and ground truth.

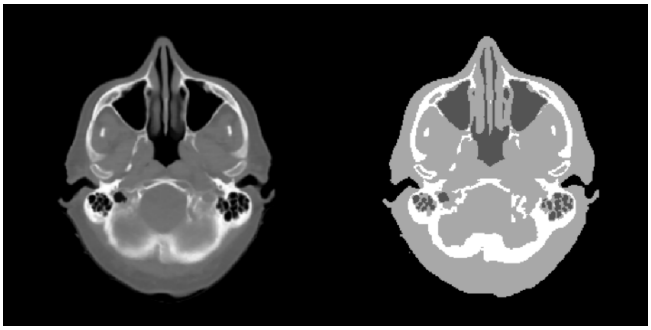

Figure 3. Left to right: Example of a single slice CT segmentation into background, air cavities in the head, soft tissue, and bone (dark to bright respectively).

2. Mean absolute error (MAE): $\frac{1}{N}\Sigma_{i=1}^{N}|CT_i - pCT_i|$, where CT, pCT, *N*, and *i* are the same as in 1. The MAE is measured in HU. A smaller MAE ∈ [0, ∞), indicates a smaller L^1 distance between the estimation and the ground truth, which means better consistency between model prediction and ground truth.

3. Dice similarity coefficient (DSC): $\frac{2\cdot|X\cap Y|}{|X|+|Y|}$, where *X*, *Y* are either the bone, soft tissue, or air cavities in the patient's head calculated for the CT and pCT respectively (for example, the DSC of bone is $DSC_{Bone}=\frac{2\cdot|Bone_{CT}\cap Bone_{pCT}|}{|Bone_{CT}|+|Bone_{pCT}|}$). Higher DSC ∈ [0, 1], means better consistency between model prediction and ground truth. Segmentation was performed using HU-based thresholding: bone (300, ∞), soft tissue [-400, 300], and air (-∞, -400) [33]. Air cavities were identified as segmented air regions overlapping with a mask of the patient's head, defined as fill (bone ∪ soft tissue), as shown in **Figure 3**.

4. Masked MSE and masked MAE: MSE and MAE calculated within the mask of the head, considering only the relevant FOV (excluding the background). A smaller MSE/MAE indicates a smaller L^2/L^1 distance between the estimation and the ground truth, which means better consistency between model prediction and ground truth. The masks were derived from the filled binary thresholded images of bone and soft tissue (as in 3). The final mask was obtained by overlapping the masks from the CT and pCT images.

5. Peak signal-to-noise ratio (PSNR): $10\cdot\log_{10}\left(\frac{MAX_I^2}{MSE}\right)$, where MAX_I is the maximum possible pixel value of the image. Higher PSNR ∈ [0, ∞), means better consistency between model prediction and ground truth.

6. Structural similarity index measure (SSIM): $\frac{(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)}{(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)}$, where μ_x is the average of *x*, μ_y is the average of *y*, σ_x^2 is the variance of x, σ_y^2 is the variance of y, σ_{xy} is the covariance between *X* and *Y*, C_1, and C_2 are variables to stabilize the division with weak denominator. Higher SSIM ∈ [-1, 1], means better consistency between model prediction and ground truth.

The quantitative analysis was performed using 4-fold cross-validation.

Image reconstruction

We reconstructed AC PET images for two patients using an offline GE image reconstruction tool by applying the 3D ordered-subsets expectation-maximization algorithm (OSEM) with two iterations. Our dataset had limitations, as many patients were scanned with their arms positioned beside their heads, which can contribute to attenuation. Since our model did not account for the arms, we conducted reconstructions for two patients whose arms were placed alongside their bodies. We applied AC using our different cGAN generated pCT images and compared our method to the uncorrected PET data, the clinical atlas-based approach [34], and the gold standard CTAC approach. The reconstructions produced 89 axial slices, each with a resolution of 192×192 pixels, where each pixel measures 3.125 mm^2. The images were also corrected for scatter, randoms, and decay during the image reconstruction.

Results

Hyper parameter tuning and cross-validation

Table 1 summarizes the cross-validation results. The multi-modal network outperformed the single-modality networks in 94.5% of all folds and metrics, while in the remaining 5.5% of cases (3 out of 54) the differences were not statistically significant. Notably, all instances where the multi-modal network did not significantly outperform the single-modality network occurred for the masked MSE and masked MAE in fold 3, as well as for the masked MAE in the validation set.

Pseudo-CT

The qualitative and quantitative comparisons of the pCT generated by the four networks are presented in **Figure 4** and **Tables 2**, **3** respectively. The multi-modal network outperforms the single-modality networks across all metrics, except masked MAE, where the results are better, but the difference in accuracy is not statistically significant. PSNR is better by 6.5-9.1%, SSIM by 2.8-4.2%, and the Dice coefficients for air, soft tissue, and bone are better by 7.6-25.7%, 2.3-3.5%, and 7.7-15%, respectively.

Tables 4 and **5** present a detailed assessment of the effects of residual arm structures and metal-induced artifacts on the accuracy of the pCT estimation. According to **Table 4**, the presence of metal artifacts degrades performance only in the single-modality model that uses NASC PET as input, while it appears to improve the results for the other models across most evaluation metrics. The influence of residual arm structures is less straightfor-

Table 1. Quantitative analysis of the different models' average performances on the validation set and all cross-validation folds (standard errors shown in parentheses)

NASC PET + MR input						
	validation	fold 1	fold 2	fold 3	fold 4	fold 5
MSE	13558 (84)	12772 (244)	13489 (296)	11772 (186)	13461 (281)	11912 (285)
MSE masked	46185 (2866)	40985 (1086)	42856 (1204)	61100 (23911)	45417 (918)	39816 (4181)
MAE	37.9 (0.2)	37.1 (0.5)	37.6 (0.5)	36.0 (0.4)	36.5 (0.5)	35.8 (0.7)
MAE masked	137 (15)	122 (4)	119 (2)	235 (121)	125 (2)	115 (6)
PSNR	23.96 (0.04)	24.23 (0.12)	23.88 (0.09)	24.48 (0.13)	24.34 (0.16)	24.82 (0.14)
DSC air	0.554 (0.003)	0.600 (0.008)	0.529 (0.010)	0.612 (0.009)	0.538 (0.010)	0.554 (0.011)
DSC ST	0.913 (0.001)	0.917 (0.002)	0.915 (0.003)	0.918 (0.003)	0.908 (0.003)	0.916 (0.004)
DSC bone	0.695 (0.002)	0.702 (0.005)	0.698 (0.006)	0.734 (0.005)	0.705 (0.006)	0.734 (0.005)
SSIM	0.860 (0.001)	0.862 (0.002)	0.859 (0.002)	0.868 (0.002)	0.870 (0.002)	0.868 (0.003)
NASC PET input						
	validation	fold 1	fold 2	fold 3	fold 4	fold 5
MSE	22220 (336)	21781 (446)	24604 (1599)	17391 (315)	17066 (345)	19381 (424)
MSE masked	69693 (1728)	84910 (4660)	81488 (23492)	80674 (25319)	58815 (1141)	74475 (13016)
MAE	49.9 (0.3)	50.4 (0.6)	51.2 (1.4)	45.2 (0.6)	43.3 (0.6)	48.2 (0.8)
MAE masked	164 (3)	187 (9)	247 (102)	250 (107)	152 (2)	171 (14)
PSNR	22.11 (0.04)	21.74 (0.09)	22.20 (0.14)	22.81 (0.11)	23.10 (0.16)	22.57 (0.13)
DSC air	0.395 (0.003)	0.365 (0.010)	0.448 (0.010)	0.471 (0.009)	0.413 (0.011)	0.408 (0.011)
DSC ST	0.877 (0.002)	0.877 (0.004)	0.874 (0.007)	0.896 (0.004)	0.890 (0.003)	0.881 (0.005)
DSC bone	0.577 (0.002)	0.550 (0.007)	0.586 (0.007)	0.642 (0.006)	0.631 (0.006)	0.630 (0.006)
SSIM	0.824 (0.001)	0.815 (0.002)	0.823 (0.002)	0.833 (0.002)	0.849 (0.002)	0.829 (0.003)
MR water input						
	validation	fold 1	fold 2	fold 3	fold 4	fold 5
MSE	18619 (106)	18733 (345)	17225 (252)	16111 (240)	19895 (420)	17073 (348)
MSE masked	61764 (551)	62207 (1475)	55314 (1195)	53141 (2886)	70880 (1803)	60139 (5672)
MAE	45.5 (0.2)	45.7 (0.6)	44.4 (0.5)	42.8 (0.5)	45.2 (0.7)	44.2 (0.7)
MAE masked	153 (2)	157 (5)	144 (4)	174 (32)	161 (3)	154 (13)
PSNR	22.52 (0.04)	22.51 (0.14)	22.69 (0.11)	23.05 (0.11)	22.48 (0.15)	23.04 (0.12)
DSC air	0.511 (0.003)	0.565 (0.010)	0.487 (0.010)	0.559 (0.010)	0.498 (0.010)	0.506 (0.011)
DSC ST	0.893 (0.001)	0.893 (0.002)	0.902 (0.002)	0.905 (0.003)	0.878 (0.004)	0.893 (0.004)
DSC bone	0.643 (0.002)	0.641 (0.006)	0.644 (0.006)	0.699 (0.005)	0.643 (0.006)	0.671 (0.006)
SSIM	0.837 (0.001)	0.836 (0.002)	0.833 (0.002)	0.846 (0.002)	0.848 (0.002)	0.840 (0.003)
MR fat input						
	validation	fold 1	fold 2	fold 3	fold 4	fold 5
MSE	21032 (112)	21507 (377)	19824 (292)	18296 (282)	21549 (423)	21134 (414)
MSE masked	69883 (485)	73546 (1703)	62904 (1332)	82201 (26298)	77753 (2220)	72207 (4141)
MAE	49.0 (0.2)	49.9 (0.7)	47.9 (0.6)	46.1 (0.6)	48.4 (0.7)	49.6 (0.8)
MAE masked	165 (1)	173 (5)	152 (3)	276 (131)	178 (5)	168 (10)
PSNR	21.92 (0.04)	21.80 (0.10)	22.02 (0.09)	22.55 (0.12)	22.17 (0.18)	22.05 (0.13)
DSC air	0.433 (0.003)	0.468 (0.010)	0.462 (0.010)	0.529 (0.010)	0.413 (0.012)	0.404 (0.012)
DSC ST	0.882 (0.001)	0.879 (0.003)	0.887 (0.004)	0.893 (0.003)	0.874 (0.003)	0.879 (0.004)
DSC bone	0.596 (0.002)	0.589 (0.006)	0.625 (0.005)	0.672 (0.005)	0.606 (0.007)	0.612 (0.005)
SSIM	0.825 (0.001)	0.823 (0.002)	0.827 (0.002)	0.834 (0.002)	0.835 (0.002)	0.825 (0.003)

The significantly better results are seen throughout the multi-modal input section of the Table (the top).

ward, as no consistent trend is observed across all metrics. However, the results suggest that input images without residual arm structures generally leads to higher accuracy. Finally, **Table 5** indicates that the absence of both metal artifacts and residual arms structures yields the best overall performance especially when using NASC PET as input.

As mentioned in the *Image Reconstruction* section, only two patients underwent PET/CT scanning with their arms

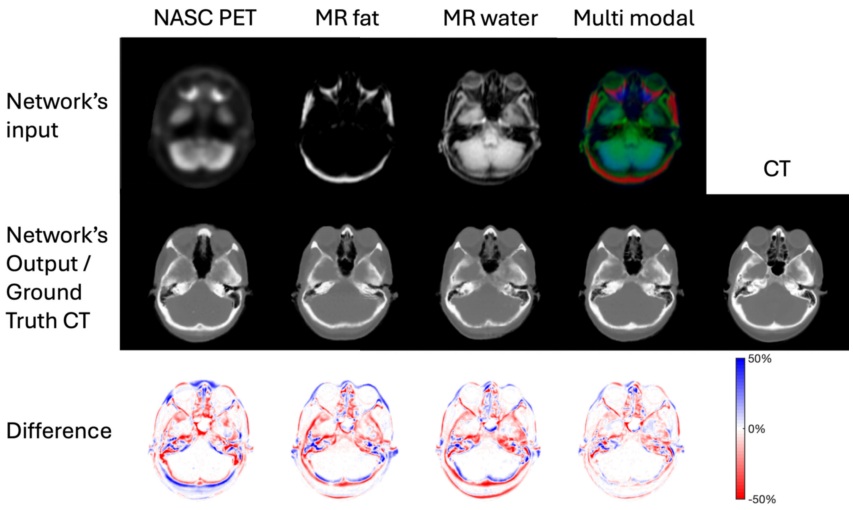

Figure 4. First row: the different network inputs, left to right - NASC PET, MR fat, MR water, multi-modal (MR + NASC PET). Second row: the different networks' output vs the ground truth CT, left to right - output of the network translating MR fat, MR water, uncorrected PET, multi-modal input to pCT, and the ground truth CT. Third row: the difference between the network's output and the reference CT, expressed as a percentage of the ground truth dynamic range.

Table 2. Quantitative analysis of the different models' average performances (standard errors shown in parentheses)

Metric	NASC PET + MR	MR Water	MR Fat	NASC PET
MSE	12674 (117)	17786 (147)	20440 (163)	20066 (361)
MSE masked	46141 (5019)	60200 (1365)	73738 (5532)	76260 (7521)
MAE	36.6 (0.2)	44.5 (0.3)	48.3 (0.3)	47.7 (0.4)
MAE masked	144 (25)	158 (7)	190 (27)	202 (30)
PSNR	24.35 (0.06)	22.76 (0.06)	22.12 (0.06)	22.48 (0.06)
DSC air	0.567 (0.004)	0.524 (0.005)	0.456 (0.005)	0.421 (0.005)
DSCS T	0.915 (0.001)	0.894 (0.001)	0.883 (0.001)	0.884 (0.002)
DSC bone	0.715 (0.002)	0.660 (0.003)	0.621 (0.003)	0.607 (0.003)
SSIM	0.865 (0.001)	0.841 (0.001)	0.828 (0.001)	0.830 (0.001)

The significantly better results are in the multi-modal results (column1).

Table 3. Quantitative analysis of the different models' median performances (standard errors shown in parentheses)

Metric	NASC PET + MR	MR Water	MR Fat	NASC PET
MSE	11754 (117)	16497 (147)	19015 (163)	17438 (361)
MSE masked	35766 (5019)	51986 (1365)	58773 (5532)	53428 (7521)
MAE	35.7 (0.2)	43.6 (0.3)	46.4 (0.3)	44.7 (0.4)
MAE masked	110 (25)	135 (7)	144 (27)	141 (30)
PSNR	23.91 (0.06)	22.43 (0.06)	21.82 (0.06)	22.19 (0.06)
DSC air	0.635 (0.004)	0.593 (0.005)	0.519 (0.005)	0.483 (0.005)
DSC ST	0.926 (0.001)	0.909 (0.001)	0.900 (0.001)	0.908 (0.002)
DSC bone	0.734 (0.002)	0.679 (0.003)	0.643 (0.003)	0.632 (0.003)
SSIM	0.864 (0.001)	0.838 (0.001)	0.827 (0.001)	0.828 (0.001)

The significantly better results are in the multi-modal results (column1).

positioned alongside their bodies, whereas the rest were scanned with their arms raised above their heads. This difference can introduce a bias, as the arms contributes additional photon attenuation. **Table 6** evaluates how this attenuation effect influences pCT accuracy for the models that use NASC PET as input. According to the results, no consistent trend was observed across all metrics; however, the pCT obtained using NASC PET without arms attenuation were significantly more accurate for most of the evaluated metrics.

AC-PET

A qualitative and quantitative comparison between NASC PET, the different methods of AC PET studied in this paper, and the ground truth CTAC PET is presented in **Figure 5** and **Table 7**, respectively. The cGAN-based approach outperforms the atlas-based method, showing statistically significant better average results, with average PSNR improvement of 7.7-9.6% with single-modality input and 12.7% with multi-modality input. Similarly, the average SSIM improves by 0.5-0.7% and 0.7% for single- and multi-modality inputs, respectively. The multi-modal configuration demonstrates significantly superior performance among the different networks, with 2.8-4.7% better average PSNR.

Discussion

Despite the presence of metal artifacts in both the CT and MR images, as shown in **Figure 1**, residual arm structures, as shown in **Figure 2**, and the introduction of attenuation correction errors due to the cropping of the arms, the results demonstrate a significant improvement in accuracy when using the proposed cGAN method over the method currently used in commercial PET/MR scanners, especially when a multimodal input is employed.

When comparing the different cGAN networks' performance, the multimodal network produced the most realistic images as can be seen by comparing the pCT images in the second row of **Figure 4** to the ground truth. **Tables 1-3** also demonstrate that the multimodal network produced the best quality pCT images with the most accurate average and median results according to all metrics compared to the single modality input cases. This is expected since the additional input will either

Table 4. Quantitative analysis of the influence of residual arm structure and metal artifacts (standard errors shown in parentheses)

	Residual arm structures		Metal artifacts	
	With n=17	Without n=18	With n=27	Without n=8
NASC PET + MR input				
MSE	13988 (172)	**11372 (152)**	**12508 (135)**	13229 (233)
MSE masked	50624 (9899)	41669 (1737)	47446 (6511)	41760 (748)
MAE	40.4 (0.3)	**32.8 (0.3)**	**36.3 (0.3)**	37.5 (0.5)
MAE masked	168 (50)	120 (3)	151 (32)	120 (2)
PSNR	23.70 (0.07)	**24.98 (0.09)**	24.37 (0.06)	24.26 (0.15)
DSC air	**0.581 (0.006)**	0.552 (0.006)	0.566 (0.005)	0.571 (0.009)
DSC ST	**0.917 (0.002)**	0.912 (0.002)	0.916 (0.002)	0.912 (0.003)
DSC bone	0.701 (0.004)	**0.729 (0.003)**	0.717 (0.003)	0.708 (0.006)
SSIM	0.848 (0.001)	**0.882 (0.001)**	**0.866 (0.001)**	0.861 (0.002)
NASC PET input				
MSE	23031 (674)	**17127 (245)**	20728 (457)	**17855 (345)**
MSE masked	78989 (14090)	73571 (5492)	75357 (7712)	79265 (20022)
MAE	53.1 (0.7)	**42.3 (0.4)**	48.4 (0.5)	**45.4 (0.6)**
MAE masked	234 (60)	171 (7)	191 (29)	238 (87)
PSNR	21.87 (0.07)	**23.08 (0.09)**	22.37 (0.06)	**22.83 (0.14)**
DSC air	**0.450 (0.006)**	0.389 (0.007)	0.406 (0.005)	**0.467 (0.009)**
DSC ST	0.886 (0.003)	0.881 (0.003)	0.882 (0.003)	**0.890 (0.003)**
DSC bone	0.595 (0.004)	**0.620 (0.004)**	0.602 (0.003)	**0.627 (0.007)**
SSIM	0.809 (0.001)	**0.850 (0.001)**	0.828 (0.001)	**0.835 (0.002)**
MR water input				
MSE	18821 (193)	**16760 (219)**	**17401 (163)**	19068 (328)
MSE masked	**56547 (892)**	63846 (2575)	**59194 (1724)**	63570 (1371)
MAE	48.1 (0.4)	**40.8 (0.4)**	**43.8 (0.3)**	46.5 (0.6)
MAE masked	157 (13)	159 (7)	159 (9)	156 (3)
PSNR	22.31 (0.06)	**23.20 (0.09)**	**22.83 (0.06)**	22.51 (0.13)
DSC air	**0.533 (0.006)**	0.514 (0.007)	0.524 (0.005)	0.523 (0.010)
DSC ST	0.901 (0.001)	0.887 (0.003)	**0.896 (0.002)**	0.887 (0.003)
DSC bone	0.650 (0.004)	**0.670 (0.004)**	**0.664 (0.003)**	0.645 (0.006)
SSIM	0.823 (0.001)	**0.858 (0.001)**	**0.842 (0.001)**	0.835 (0.002)
MR fat input				
MSE	21258 (211)	**19629 (246)**	**20151 (186)**	21401 (332)
MSE masked	74185 (10912)	73291 (1911)	74319 (7130)	71787 (2852)
MAE	51.5 (0.4)	**45.2 (0.4)**	**47.8 (0.3)**	50.3 (0.6)
MAE masked	209 (54)	172 (5)	196 (35)	171 (8)
PSNR	21.75 (0.06)	**22.48 (0.10)**	22.13 (0.06)	22.07 (0.16)
DSC air	**0.485 (0.006)**	0.424 (0.007)	0.455 (0.006)	0.459 (0.010)
DSC ST	**0.889 (0.002)**	0.877 (0.002)	**0.884 (0.002)**	0.878 (0.003)
DSC bone	0.617 (0.004)	**0.625 (0.004)**	0.618 (0.003)	**0.631 (0.006)**
SSIM	0.813 (0.001)	**0.843 (0.001)**	**0.830 (0.001)**	0.824 (0.002)

The significantly better results are in bold.

enhance the model performance, if it contains useful data, or it will be ignored. Among the single modality networks, the performance of the MR water network was the best, and the other networks, MR fat and NASC PET, with similar overall performance, were the worst.

Table 4 assesses the effects of residual arm structures and metal-induced artifacts on the accuracy of pCT estimation. As expected for image slices containing metal artifacts, the discrepancy between the generated pCT and the reference CT is larger when using only NASC PET data as input. This occurs because, unlike in MRI, and especially in MR Dixon water, the effects of metal structures on the annihilation photons' scatter and attenuation are less apparent in PET. Furthermore, the absence of residual arm structures results in superior performan-

Table 5. Quantitative analysis of the influence of residual arm structure and metal artifacts (standard errors shown in parentheses)

	w arms w metal n=13	w/o arms w metal n=14	w arms w/o metal n=4	w/o arms w/o metal n=4
NASC PET + MR input				
MSE	13521 (202)	11534 (176)	15449 (312)	**10793 (299)**
MSE masked	51902 (13038)	43129 (2203)	46595 (1034)	**36437 (1012)**
MAE	39.7 (0.4)	33.1 (0.4)	42.5 (0.6)	**32.0 (0.6)**
MAE masked	180 (66)	123 (4)	129 (2)	**110 (2)**
PSNR	23.83 (0.06)	24.89 (0.10)	23.31 (0.19)	**25.30 (0.22)**
DSC air	0.584 (0.007)	0.547 (0.007)	0.574 (0.012)	0.568 (0.014)
DSC ST	0.921 (0.002)	0.911 (0.002)	0.906 (0.004)	0.918 (0.004)
DSC bone	0.716 (0.004)	0.718 (0.004)	0.654 (0.009)	**0.769 (0.006)**
SSIM	0.851 (0.002)	0.881 (0.001)	0.838 (0.002)	**0.887 (0.002)**
NASC PET input				
MSE	23910 (876)	17670 (281)	20284 (457)	**15192 (485)**
MSE masked	71028 (13996)	79482 (7002)	103600 (38111)	**52400 (1648)**
MAE	54.0 (0.8)	42.9 (0.4)	50.2 (0.7)	**40.1 (0.9)**
MAE masked	204 (59)	179 (9)	325 (165)	**142 (4)**
PSNR	21.82 (0.08)	22.91 (0.09)	22.03 (0.15)	**23.71 (0.23)**
DSC air	0.442 (0.007)	0.368 (0.008)	0.473 (0.012)	0.460 (0.015)
DSC ST	0.886 (0.004)	0.878 (0.004)	0.888 (0.005)	0.892 (0.005)
DSC bone	0.602 (0.004)	0.602 (0.005)	0.573 (0.010)	**0.686 (0.008)**
SSIM	0.808 (0.002)	0.847 (0.002)	0.812 (0.002)	**0.861 (0.003)**
MR water input				
MSE	17617 (202)	17194 (255)	22581 (435)	**15214 (407)**
MSE masked	51576 (978)	66573 (3246)	72166 (1844)	54079 (1924)
MAE	46.4 (0.4)	41.4 (0.4)	53.5 (0.7)	**38.9 (0.8)**
MAE masked	152 (17)	164 (9)	171 (4)	139 (4)
PSNR	22.57 (0.07)	23.08 (0.10)	21.50 (0.14)	**23.62 (0.21)**
DSC air	0.540 (0.007)	0.508 (0.008)	0.514 (0.013)	0.534 (0.014)
DSC ST	**0.908 (0.001)**	0.886 (0.003)	0.881 (0.003)	0.894 (0.005)
DSC bone	0.669 (0.004)	0.660 (0.004)	0.590 (0.010)	**0.705 (0.007)**
SSIM	0.828 (0.002)	0.856 (0.002)	0.807 (0.003)	**0.866 (0.003)**
MR fat input				
MSE	20525 (244)	19792 (279)	23546 (399)	19049 (514)
MSE masked	74605 (14369)	74042 (1845)	72864 (1541)	70602 (5750)
MAE	50.3 (0.5)	45.3 (0.5)	55.5 (0.7)	44.6 (0.9)
MAE masked	221 (71)	172 (4)	172 (3)	170 (16)
PSNR	21.88 (0.06)	22.38 (0.09)	21.36 (0.17)	**22.85 (0.28)**
DSC air	0.490 (0.007)	0.418 (0.008)	0.469 (0.013)	0.446 (0.015)
DSC ST	0.895 (0.002)	0.874 (0.003)	0.870 (0.005)	0.886 (0.003)
DSC bone	0.624 (0.004)	0.613 (0.005)	0.596 (0.009)	**0.669 (0.006)**
SSIM	0.817 (0.002)	0.842 (0.002)	0.803 (0.003)	0.848 (0.003)

The significantly better results are in bold.

ce across most evaluation metrics. Interestingly, for the MR-based models, the presence of metal artifacts results in pCT estimates that more closely resemble the CT. We speculate that this may be because metal artifacts are also visible in the MR images, as illustrated in **Figure 2**. However, since these artifacts in the CT images are themselves erroneous, this resemblance is not necessarily beneficial and should not be interpreted as indicating better AC accuracy.

Table 5 evaluates the combined effects of residual arm structures and metal-induced artifacts on the accuracy of pCT estimation. As expected, the most accurate results are obtained for patients who did not exhibit either of these confounding factors.

Table 6 evaluates the effect of photon attenuation caused by the arms on the accuracy of pCT estimation for models using NASC PET as input. Interestingly, the results show

Table 6. Quantitative analysis of the effect of arms photon attenuation on pCT estimation accuracy for models that use NASC PET as input (standard errors shown in parentheses)

	NASC PET		NASC PET + MR	
	without arms	with arms	without arms	with arms
MSE	**13504 (520)**	20437 (379)	**6956 (277)**	12998 (120)
MSE masked	121745 (48855)	73756 (7465)	49703 (15599)	45946 (5225)
MAE	**34.5 (0.9)**	48.4 (0.4)	**22.8 (0.6)**	37.4 (0.2)
MAE masked	229 (50)	201 (32)	128 (21)	145 (26)
PSNR	**24.10 (0.26)**	22.39 (0.06)	**27.05 (0.29)**	24.19 (0.06)
DSC air	0.282 (0.026)	**0.427 (0.005)**	0.552 (0.027)	0.568 (0.004)
DSC ST	0.830 (0.016)	**0.887 (0.002)**	0.891 (0.013)	**0.916 (0.001)**
DSC bone	**0.699 (0.014)**	0.602 (0.003)	**0.800 (0.011)**	0.710 (0.003)
SSIM	**0.895 (0.003)**	0.826 (0.001)	**0.927 (0.002)**	0.862 (0.001)

The significantly better results are in bold.

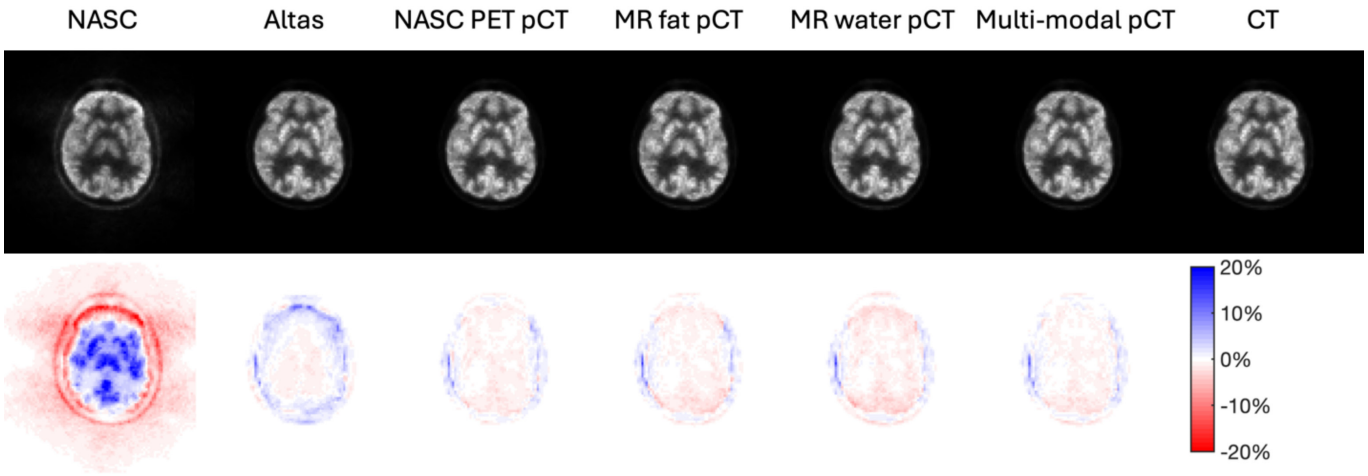

Figure 5. First row: left to right - NASC PET, AC PET using the atlas method implemented in the Signa system, AC PET using the generated pCT generated from NASC PET image, AC PET using the generated pCT generated from fat MR image, AC PET using the generated pCT generated from water MR image, AC PET using the generated pCT generated from multi-modal NASC PET + MR data, the ground truth attenuation and scatter corrected PET using the CT. Second row: the difference between the ground truth CT-based AC PET and the first row in percentage of the ground truth dynamic range. The color scale showing the % differences was compressed to be in the range of ± 20% to emphasize the difference between the different AC solution performances. As a result, the NASC difference image looks saturated (since its % difference from the ground truth was > 20%).

that, although most of the training data consisted of patients scanned with their arms positioned alongside their heads, and thus their PET images were affected by arm attenuation, the patients scanned without arm attenuation achieved more accurate pCT estimates. However, this subgroup includes only two patients: one without residual arm structures or metal-induced artifacts, and one without residual arm structures but with metal-induced artifacts. Therefore, it is difficult to determine whether the observed improvement in pCT accuracy is primarily due to the absence of arm-related photon attenuation, the lack of residual arm structures, the presence or absence of metal-induced artifacts, or a combination of these factors.

The PET results presented in **Figure 5** demonstrate the importance of using AC in brain PET. The magnitude of the difference between the NASC PET image and the ground truth PET image is significantly higher than the difference between the AC approaches and the ground truth PET image. The notably lower PSNR and SSIM in **Table 7** further support this claim.

When comparing the different AC approaches, all our networks produce superior results than the clinical atlas-based method used in the GE Signa system, according to **Figure 5** and **Table 7**. Since 2019, the Signa system has also included an alternative AC method that is based on the MRI ZTE sequence data [35]. Unfortunately, our datasets did not contain this sequence, so we could not compare our solution to this approach. As expected, the highest quality PET was achieved when using the multimodal network, having significantly better PSNR.

This paper highlights the advantages of using multimodal over single-modality DL model input. We selected the pix2pix network architecture for its simplicity, suitability for limited datasets, short training time, and proven suc-

Table 7. Quantitative comparison between the different methods of AC PET and NAC PET relative to the ground truth CTAC PET (standard errors shown in parentheses)

Metric	Avg PSNR	Mdn PSNR	Avg SSIM	Mdn SSIM
NASC PET + MR	50.0 (0.4)	49.0 (0.4)	0.9987 (0.0001)	0.9991 (0.0001)
MR Water	48.6 (0.4)	48.6 (0.4)	0.9963 (0.0011)	0.9988 (0.0011)
MR fat	48.3 (0.4)	47.7 (0.4)	0.9982 (0.0001)	0.9989 (0.0001)
NASC PET	47.7 (0.5)	47.6 (0.5)	0.9976 (0.0002)	0.9985 (0.0002)
Atlas	44.3 (0.3)	44.6 (0.3)	0.9913 (0.0024)	0.9974 (0.0024)
NASC	31.8 (0.2)	31.9 (0.2)	0.8637 (0.0040)	0.8753 (0.0040)

The significantly better results correspond to the multi-modal input.

cess in various applications. We then demonstrated our point using a limited dataset of only 35 patients, as this was the dataset available to us. However, we are confident that a more advanced model and a larger dataset would yield the same conclusion. We utilized the 2-point Dixon sequence since it was available to us, but we are certain that other sequences containing relevant data would also enhance pCT estimation.

We want to reiterate that the NASC PET data used in this study were acquired with a PET/CT scanner and therefore reflect the attenuation properties of that system. Consequently, applying the proposed solution to PET/MR data would require retraining the models with NASC PET data acquired directly from the PET/MR scanner.

Conclusion and future work

In this work, we demonstrated the superiority of using multimodal input - specifically, a combination of uncorrected PET, MR Water, and MR Fat - over single-modality input for more accurate pCT estimations. This indicates that if a 2-point Dixon MRI sequence is performed, we can achieve superior pCT at the minimal cost of additional PET reconstruction time (to generate NASC PET images). Additionally, we showed that the AC PET images using the pCT generated by the models proposed in this paper are closer to the ground truth than those produced using the GE Signa commercial system's atlas method for PET AC.

As mentioned in the discussion section, this paper demonstrates a proof of concept that can be further optimized. It will be interesting to investigate how more advanced network architectures and additional MRI sequences affect the pCT estimation, and we plan to continue exploring and optimizing the suggested solution in the future.

Acknowledgements

We would like to thank Dr. Mojtaba Jafaritadi for helpful discussions. We would also like to thank Dr. Mehdi Khalighi and Kim Halbert for providing the dataset. This work was partially supported by NIH-NIBIB grant R01-EB03103801A1. Jonathan Fisher was sponsored in part by Bruker BioSpin. Emily Anaya was supported by the Stanford NeuroTech Fellowship and Stanford Graduate Fellowship. These datasets used are based upon work supported by the National Science Foundation under Grant No. 1828993.

Disclosure of conflict of interest

None.

Address correspondence to: Dr. Craig S Levin, Molecular Imaging Instrumentation Laboratory, Department of Radiology, James H Clark Center, 318 Campus Drive, E1.1, Lab: E150, Stanford University, Stanford, CA 94305, USA. Tel: +1-650-736-7211; E-mail: cslevin@stanford.edu

References

[1] Gallamini A, Zwarthoed C and Borra A. Positron emission tomography (PET) in oncology. Cancers 2014; 6: 1821-1889.

[2] Barc K and Kuźma-Kozakiewicz M. Positron emission tomography neuroimaging in neurodegenerative diseases: Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis. Neurol Neurochir Pol 2019; 53: 99-112.

[3] Kinahan PE, Hasegawa BH and Beyer T. X-ray-based attenuation correction for positron emission tomography/computed tomography scanners. Semin Nucl Med 2003; 33: 166-79.

[4] Lee JS. A review of deep-learning-based approaches for attenuation correction in positron emission tomography. IEEE Trans Radiat Plasma Med Sci 2021; 5: 160-184.

[5] Wagenknecht G, Kaiser HJ, Mottaghy FM and Herzog H. MRI for attenuation correction in PET: methods and challenges. MAGMA 2013; 26: 99-113.

[6] Siegel S and Dahlbom M. Implementation and evaluation of a calculated attenuation correction for PET. IEEE Trans Nucl Sci 1992; 39: 1117-1121.

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

[8] Wiesinger F, Sacolick LI, Menini A, Kaushik SS, Ahn S, Veit-Haibach P, Delso G and Shanbhag DD. Zero TE MR bone imaging in the head. Magn Reson Med 2016; 75: 107-114.

[9] Kops ER and Herzog H. Alternative methods for attenuation correction for PET images in MR-PET scanners. IEEE NSS CR 2007; 6: 4327-4330.

[10] Kops ER and Herzog H. Template based attenuation correction for PET in MR-PET scanners. IEEE NSS CR 2008; 3786-3789.
[11] Schreibmann E, Nye JA, Schuster DM, Martin DR, Votaw J and Fox T. MR-based attenuation correction for hybrid PET-MR brain imaging systems using deformable image registration. Med Phys 2010; 37: 2101-2109.
[12] Nuyts J, Dupont P, Stroobants S, Benninck R, Mortelmans L and Suetens P. Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms. IEEE Trans Med Imaging 1999; 18: 393-403.
[13] Rezaei A, Defrise M, Bal G, Michel C, Conti M, Watson C and Nuyts J. Simultaneous reconstruction of activity and attenuation in time-of-flight PET. IEEE Trans Med Imaging 2012; 31: 2224-2233.
[14] Mihlin A and Levin CS. An expectation maximization method for joint estimation of emission activity distribution and photon attenuation map in PET. IEEE Trans Med Imaging 2017; 36: 214-224.
[15] Spuhler KD, Gardus J 3rd, Gao Y, DeLorenzo C, Parsey R and Huang C. Synthesis of patient-specific transmission data for PET attenuation correction for PET/MRI neuroimaging using a convolutional neural network. J Nucl Med 2019; 60: 555-560.
[16] Arabi H, Zeng G, Zheng G and 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: 2746-2759.
[17] Ladefoged CN, Marner L, Hindsholm A, Law I, Højgaard L and Andersen FL. Deep learning based attenuation correction of PET/MRI in pediatric brain tumor patients: evaluation in a clinical setting. Front Neurosci 2019; 12: 1005.
[18] Gong K, Yang J, Larson PEZ, Behr SC, Hope TA, Seo Y and Li Q. MR-based attenuation correction for brain PET using 3D cycle-consistent adversarial network. IEEE Trans Radiat Plasma Med Sci 2021; 5: 185-192.
[19] Lei Y, Wang T, Liu Y, Higgins K, Tian S, Liu T, Mao H, Shim H, Curran W, Shu HK and Yang X. MRI-based synthetic CT generation using deep convolutional neural network. Medical Imaging 2019: Image Processing 2019; 10949: 716-721.
[20] Shafai-Erfani G, Wang T, Lei Y, Tian S, Patel P, Jani AB, Curran WJ, Liu T and Yang X. Dose evaluation of MRI-based synthetic CT generated using a machine learning method for prostate cancer radiotherapy. Med Dosim 2019; 44: e64-e70.
[21] Liu F, Jang H, Kijowski R, Bradshaw T and McMillan AB. Deep learning MR imaging-based attenuation correction for PET/MR imaging. Radiology 2018; 286: 676-684.
[22] Dong X, Wang T, Lei Y, Higgins K, Liu T, Curran WJ, Mao H, Nye JA and Yang X. Yang, Synthetic CT generation from non-attenuation corrected PET images for whole-body PET imaging. Phys Med Biol 2019; 64: 215016.
[23] Wang T, Lei Y, Fu Y, Curran WJ, Liu T, Nye JA and Yang X. Machine learning in quantitative PET: a review of attenuation correction and low-count image reconstruction methods. Phys Med 2020; 76: 294-306.
[24] Shiri I, Ghafarian P, Geramifar P, Leung KH, Ghelichoghli M, Oveisi M, Rahmim A and Ay MR. Direct attenuation correction of brain PET images using only emission data via a deep convolutional encoder-decoder (Deep-DAC). Eur Radiol 2019; 29: 6867-6879.
[25] Dong X, Lei Y, Wang T, Higgins K, Liu T, Curran WJ, Mao H, Nye JA and Yang X. Deep learning-based attenuation correction in the absence of structural information for whole-body positron emission tomography imaging. Phys Med Biol 2020; 65: 05501.
[26] Jahangir R, Kamali-Asl A, Arabi H and Zaidi H. Strategies for deep learning-based attenuation and scatter correction of brain 18F-FDG PET images in the image domain. Med Phys 2024; 51: 870-880.
[27] Anaya E and Levin C. Automatic generation of MR-based attenuation map using conditional generative adversarial network for attenuation correction in PET/MR. IEEE NSS/MIC 2020; 1-3.
[28] Lowekamp BC, Chen DT, Ibanez L and Blezek D. The design of SimpleITK. Front Neuroinform 2013; 7: 45.
[29] Unser M, Aldroubi A and Gerfen C. Multiresolution image registration procedure using spline pyramids. Proc SPIE 1993; 2034: 160-170.
[30] Isola P, Zhu JY, Zhou T and Efros AA. Image-to-image translation with conditional adversarial networks. in 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 5967-5976, Honolulu, HI, 2017, IEEE.
[31] Zhu JY, Park T, Isola P and Efros AA. pytorch-CycleGAN-and-pix2pix. GitHub 2019; https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix.
[32] Bradshaw TJ, Huemann Z, Hu J and Rahmim A. A guide to cross-validation for artificial intelligence in medical imaging. Radiol Artif Intell 2023; 5: e220232.
[33] Liu F, Jang H, Kijowski R, Zhao G, Bradshaw T and McMillan AB. A deep learning approach for 18F-FDG PET attenuation correction. EJNMMI Phys 2018; 5: 24.
[34] Sekine T, Buck A, Delso G, ter Voert EE, Huellner M, Veit-Haibach P and Warnock G. Evaluation of atlas-based attenuation correction for integrated PET/MR in human brain: application of a head atlas and comparison to true CT-based attenuation correction. J Nucl Med 2016; 57: 215-220.
[35] Delso G, Wiesinger F, Sacolick LI, Kaushik SS, Shanbhag DD, Hüllner M and Veit-Haibach P. Clinical evaluation of zero-echo-time MR imaging for the segmentation of the skull. J Nucl Med 2015; 56: 417-422.