Revealing Anatomical Structures in PET to Generate CT for Attenuation Correction

Yongsheng Pan[1], Feihong Liu[1,2], Caiwen Jiang[1], Jiawei Huang[1], Yong Xia[3], and Dinggang Shen[1,4(✉)]

[1] School of Biomedical Engineering, ShanghaiTech University, Shanghai, China
{panysh,dgshen}@shanghaitech.edu.cn
[2] School of Information Science and Technology, Northwest University, Xi'an, China
[3] School of Computer Science, Northwestern Polytechnical University, Xi'an, China
yxia@nwpu.edu.cn
[4] Shanghai United Imaging Intelligence Co., Ltd., Shanghai, China

Abstract. Positron emission tomography (PET) is a molecular imaging technique relying on a step, namely attenuation correction (AC), to correct radionuclide distribution based on pre-determined attenuation coefficients. Conventional AC techniques require additionally-acquired computed tomography (CT) or magnetic resonance (MR) images to calculate attenuation coefficients, which increases imaging expenses, time costs, or radiation hazards to patients, especially for whole-body scanners. In this paper, considering technological advances in acquiring more anatomical information in raw PET images, we propose to conduct attenuation correction to PET by itself. To achieve this, we design a deep learning based framework, namely anatomical skeleton-enhanced generation (ASEG), to generate pseudo CT images from non-attenuation corrected PET images for attenuation correction. Specifically, ASEG contains two sequential modules, i.e., a skeleton prediction module and a tissue rendering module. The former module first delineates anatomical skeleton and the latter module then renders tissue details. Both modules are trained collaboratively with specific *anatomical-consistency constraint* to guarantee tissue generation fidelity. Experiments on four public PET/CT datasets demonstrate that our ASEG outperforms existing methods by achieving better consistency of anatomical structures in generated CT images, which are further employed to conduct PET attenuation correction with better similarity to real ones. This work verifies the feasibility of generating pseudo CT from raw PET for attenuation correction without acquising additional images. The associated implementation is available at https://github.com/YongshengPan/ASEG-for-PET2CT.

Keywords: PET · Attenuation correction · CT · Image generation

1   Introduction

Positron emission tomography (PET) is a general nuclear imaging technique, which has been widely used to characterize tissue metabolism, protein deposition, etc. [9]. According to the PET imaging principle, radioactive tracers injected into the body involve in the metabolism and produce γ decay signals externally. However, due to photoelectric absorption and Compton scattering, the decay signals are attenuated when passing through human tissues to external receivers, resulting in incorrect tracer distribution reasoning (see non-attenuation corrected PET (NAC-PET) in Fig. 1(a)). To obtain correct tracer distribution (see AC-PET in Fig. 1(a)), attenuation correction (AC) on the received signals is required.
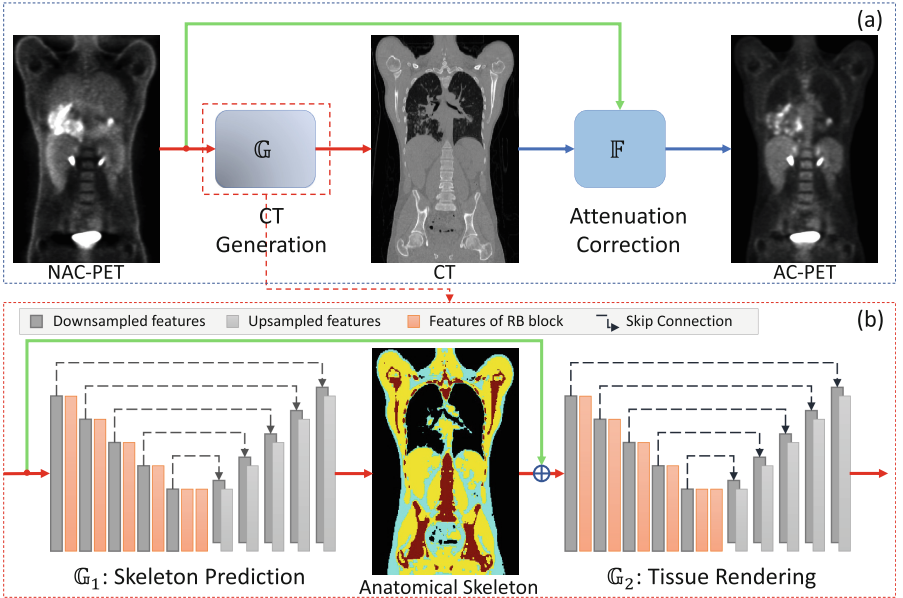

Fig. 1. ASEG framework. (a) The general process for reconstructing AC-PET by generating CT from NAC-PET. Rather than directly generating CT from NAC-PET, (b) our ASEG first delineates the anatomical skeleton and then renders the tissue details.

Traditional AC accompanies additional costs caused by the simultaneously obtained MR or CT images which are commonly useless for diagnosis. The additional costs are especially significant for advanced total-body PET/CT scanners [12,13], which have effective sensitivity and low radiation dose during PET scanning but accumulative radiation dose during CT scanning. In another word, CT becomes a non-negligible source of radiation hazards. To reduce the costs, including expense, time, and radiation hazards, some studies proposed to conduct AC by exploiting each PET image itself. Researchers have been motivated to generate pseudo CT images from NAC-PET images [2,7], or more directly,

to generate AC-PET images from NAC-PET images [5,11]. Since pseudo CT is convenient to be integrated into conventional AC processes, generating pseudo CT images is feasible in clinics for AC.

The pseudo CT images should satisfy two-fold requests. Firstly, the pseudo CT images should be visually similar in anatomical structures to corresponding actual CT images. Secondly, PET images corrected by pseudo CT images should be consistent with that corrected by actual CT images. However, current techniques of image generation tend to produce statistical average values and patterns, which easily erase significant tissues (e.g., bones and lungs). As a result, for those tissues with relatively similar metabolism but large variances in attenuation coefficient, these methods could cause large errors as they are blind to the correct tissue distributions. Therefore, special techniques should be investigated to guarantee the fidelity of anatomical structures in these generated pseudo CT images.

In this paper, we propose a deep learning framework, named anatomical skeleton enhanced generation (ASEG), to generate pseudo CT from NAC-PET for attenuation correction. ASEG focuses more on the fidelity of tissue distribution, i.e., anatomical skeleton, in pseudo CT images. As shown in Fig. 1(b), this framework contains two sequential modules structure prediction module $\mathbb{G}_1$ and tissue rendering module $\mathbb{G}_2$}. $\mathbb{G}_1$ devotes to delineating the anatomical skeleton from a NAC-PET image, thus producing a prior tissue distribution map to $\mathbb{G}_2$, while $\mathbb{G}_2$ devotes to rendering the tissue details according to both the skeleton and NAC-PET image. We regard $\mathbb{G}_1$ as a segmentation network that is trained under the combination of cross-entropy loss and Dice loss and outputs the anatomical skeleton. For training the generative module $\mathbb{G}_2$, we further propose the *anatomical-consistency constraint* to guarantee the fidelity of tissue distribution besides general constraints in previous studies. Experiments on four publicly collected PET/CT datasets demonstrate that our ASEG outperforms existing methods by preserving better anatomical structures in generated pseudo CT images and achieving better visual similarity in corrected PET images.

2   Method

We propose the anatomical skeleton enhanced generation (ASEG, as illustrated in Fig. (1) framework that regards the CT generation as two sequential tasks, i.e., skeleton prediction and tissue rendering, instead of simply mapping pseudo CT from NAC-PET. ASEG composes of two sequential generative modules $\{\mathbb{G}_1, \mathbb{G}_2\}$ to deal with them, respectively. $\mathbb{G}_1$ devotes itself to decoupling the anatomical skeleton from NAC-PET to provide rough prior information of attenuation coefficients to $\mathbb{G}_2$, particularly for lungs and bones that have the most influential variances. $\mathbb{G}_2$ then devotes to rendering the tissue details in the CT pattern exploiting both the skeleton and NAC-PET images. In short, the skeleton decoupled by $\mathbb{G}_1$ is a prior guidance to $\mathbb{G}_2$, and in turn, $\mathbb{G}_2$ can serve as a target supervision for $\mathbb{G}_1$. These two modules are trained with different constraints according to the corresponding tasks. Specially, the general Dice loss and cross-entropy loss [16]

are employed to guarantee $\mathbb{G}_1$ for the fidelity of tissue distributions while general *mean absolute error* and *feature matching* losses are utilized to guarantee $\mathbb{G}_2$ for potential coarse-to-fine semantic constraint. To improve the fidelity of anatomical structures, we further propose the *anatomical consistency loss* to encourage $\mathbb{G}_2$ to generate CT images that are consistent in tissue distributions with actual CT images in particular.

Network Architecture. As illustrated in Fig. 1(b), our ASEG has two generative modules for skeleton prediction and tissue rendering, respectively, where $\mathbb{G}_1$ and $\mathbb{G}_2$ share the same network structure but $\mathbb{G}_2$ is accompanied by an adversarial network $\mathbb{D}$ (not drawn, same structure in [8]). Each generative network consists of an input convolutional layer, four encoding blocks, two residual blocks (RBs) [8], four decoding blocks, and an output convolutional layer. Each encoding block contains a RB and a convolutional layer with strides of $2 \times 2 \times 2$ for downsampling while each decoding block contains an upsampling operation of $2 \times 2 \times 2$ and a convolutional layer. The kernel size for the input and output convolutional layers is $7 \times 7 \times 7$ while for others is $3 \times 3 \times 3$. Skip connections are further used locally in RBs and globally between corresponding layers to empower information transmission. Meanwhile, the adversarial network $\mathbb{D}$ consists of five $4 \times 4 \times 4$ convolutional layers with strides of $2 \times 2 \times 2$ for the first four layers and $1 \times 1 \times 1$ for the last layer.

Model Formulation. Let X_{nac} and X_{ac} denote the NAC-PET and AC-PET images, and Y be the actual CT image used for AC. Since CT image is highly crucial in conventional AC algorithms, they generally have a relationship as

$$X_{ac} = \mathbb{F}(X_{nac}, Y), \tag{1}$$

under an AC algorithm $\mathbb{F}$. To avoid scanning an additional CT image, we attempt to predict Y from X_{nac} as an alternative in AC algorithm. Namely, a mapping $\mathbb{G}$ is required to build the relationship between Y and X_{nac}, i.e., $\hat{Y} = \mathbb{G}(X_{nac})$. Then, X_{ac} can be acquired by

$$X_{ac} \approx \hat{X}_{ac} = \mathbb{F}(X_{nac}, \hat{Y}) = \mathbb{F}(X_{nac}, \mathbb{G}(X_{nac})). \tag{2}$$

This results in a pioneering AC algorithm that requires only a commonly reusable mapping function $\mathbb{G}$ for all PET images rather than a corresponding CT image Y for each PET image.

As verified in some previous studies [1,2,7], $\mathbb{G}$ can be assigned by some image generation techniques, e.g. GANs and CNNs. However, since these general techniques tend to produce statistical average values, directly applying them may lead to serious brightness deviation, for those tissues with large intensity ranges. To overcome this drawback, we propose ASEG as a specialized AC technique, which decouple the CT generation process $\mathbb{G}$ in two sequential parts, i.e., $\mathbb{G}_1$ for skeleton prediction and $\mathbb{G}_2$ for tissue rendering, as formulated as

$$\hat{Y} = \mathbb{G}(X_{nac}) = \mathbb{G}_2(Y_{as}, X_{nac}) \approx \mathbb{G}_2(\hat{Y}_{as}, X_{nac}) = \mathbb{G}_2(\mathbb{G}_1(X_{nac}), X_{nac}). \tag{3}$$

Herein, $\mathbb{G}_1$ devotes to delineating anatomical skeleton Y_{as} from X_{nac}, thus providing a prior tissue distribution to $\mathbb{G}_2$ while $\mathbb{G}_2$ devotes to rendering the tissue details from X_{nac} and $\hat{Y}_{as} = \mathbb{G}_1(X_{nac})$.

To avoid annotating the ground truth, Y_{as} can be derived from the actual CT image by a segmentation algorithm (denoted as $\mathbb{S} : Y_{as} = \mathbb{S}(Y)$). As different tissues have obvious differences in intensity ranges, we define $\mathbb{S}$ as a simple thresholding-based algorithm. Herein, we first smooth each non-normalized CT image with a small recursive Gaussian filter to suppress the impulse noise, and then threshold this CT image to four binary masks according to the Hounsfield scale of tissue density [6], including the air-lung mask (intensity ranges from -950HU to -125HU), the fluids-fat mask (ranges from -125HU to 10HU), the soft-tissue mask (ranges from 10HU to 100HU), and the bone mask (ranges from 100HU to 3000HU), as demonstrated by anatomical skeleton in Fig. 1(b). This binarization trick highlights the difference among different tissues, and thus is easier perceived.

General Constraints. As mentioned above, two generative modules $\{\mathbb{G}_1, \mathbb{G}_2\}$ work for two tasks, namely the skeleton prediction and tissue rendering, respectively. Thus, they are trained with different target-oriented constraints. In the training scheme, the loss function for $\mathbb{G}_1$ is the combination of Dice loss $\mathfrak{L}_{dice}$ and cross-entropy loss $\mathfrak{L}_{ce}$ [16], denoted as

$$\mathfrak{L}_1(\hat{Y}_{as}, Y_{as}) = \mathfrak{L}_{dice}(\hat{Y}_{as}, Y_{as}) + \mathfrak{L}_{ce}(\hat{Y}_{as}, Y_{as}). \tag{4}$$

Meanwhile, the loss function for $\mathbb{G}_2$ combines the mean absolute error (MAE) $\mathfrak{L}_{mae}$, perceptual feature matching loss $\mathfrak{L}_{fm}$ [14], and anatomical-consistency loss $\mathfrak{L}_{ac}$, denoted as

$$\mathfrak{L}_2(\hat{Y}, Y) = \mathfrak{L}_{mae}(\hat{Y}, Y) + \mathfrak{L}_{fm}(\hat{Y}; \mathbb{D}) + \mathfrak{L}_{ac}(\hat{Y}, Y). \tag{5}$$

where the anatomical consistency loss $\mathfrak{L}_{st}$ is explained below.

Anatomical consistency. It is generally known that CT images can provide anatomical observation because different tissues have a distinctive appearance in Hounsfield scale (linear related to attenuation coefficients). Therefore, it is crucial to ensure the consistency of tissue distribution in the pseudo CT images, tracking which we propose to use the tissue distribution consistency to guide the network learning. Based on the segmentation algorithm $\mathbb{S}$, both the actual and generated CTs $\{Y, \hat{Y}\}$ can be segmented to anatomical structure/tissue distribution masks $\{\mathbb{S}(Y), \mathbb{S}(\hat{Y})\}$, and their consistency can then be measured by Dice coefficient. Accordingly, the anatomical-consistency loss $\mathfrak{L}_{ac}$ is a Dice loss as

$$\mathfrak{L}_{ac}(Y, \hat{Y}) = \mathfrak{L}_{dice}(\mathbb{S}(\hat{Y}), \mathbb{S}(Y)). \tag{6}$$

During the inference phase, only the NAC-PET image of each input subject is required, where the pseudo CT image is derived by $\hat{Y} \approx \mathbb{G}_2(\mathbb{G}_1(X_{nac}), X_{nac})$.

3   Experiments

3.1   Materials

The data used in our experiments are collected from The Cancer Image Archive (TCIA) [4] (https://www.cancerimagingarchive.net/collections/), where a series of public datasets with different types of lesions, patients, and scanners are open-access. Among them, 401, 108, 46, and 20 samples are extracted from the Head and Neck Scamorous Cell Carcinoma (HNSCC), Non-Small Cell Lung Cancer (NSCLC), The Cancer Genome Atlas (TCGA) - Head-Neck Squamous Cell Carcinoma (TCGA-HNSC), and TCGA - Lung Adenocarcinoma (TCGA-LUAD), respectively. We use these samples in HNSCC for training and in other three datasets for evaluation.

Each sample contains co-registered (acquired with PET-CT scans) CT, PET, and NAC-PET whole-body scans. In our experiments, we re-sampled all of them to a voxel spacing of $2 \times 2 \times 2$ and re-scaled the intensities of NAC-PET/AC-PET images to a range of $[0, 1]$, of CT images by multiplying 0.001. The input and output of our ASEG framework are cropped patches with the size of $192 \times 192 \times 128$ voxels. To achieve full-FoV output, the consecutive outputs of each sample are composed into a single volume where the overlapped regions are averaged.

3.2   Comparison with Other Methods

We compared our ASEG with three state-of-the-art methods, including (i) a U-Net based method [3] that directly learns a mapping from NAC-PET to CT image with MAE loss (denoted as U-Net), (ii) a conventional GAN-based method [1,2] that uses the U-Net as the backbone and employ the style-content loss and adversarial loss as an extra constraint (denoted as CGAN), and (iii) an auxiliary GAN-based method [10] that uses the CT-based segmentation (i.e., the simple thresholding $\mathbb{S}$) as an auxiliary task for CT generation (denoted as AGAN). For a fair comparison, we implemented these methods by ourselves in a TensorFlow platform with an NVIDIA 3090 GPU. All methods share the same backbone structure as $\mathbb{G}_*$ in Fig. 1(b) and follow the same experimental settings. Particularly, the adversarial loss of methods (ii) and (iii) are replaced by the perceptual feature matching loss. These two methods could be considered as variants of our method without using predicted prior anatomic skeleton.

Quantitative Analysis of CT. As the most import application of CT that is to display the anatomical information, we propose to measure the anatomical consistency between the pseudo CT images and actual CT images, where the Dice coefficients on multiple anatomical regions that extracted from the pseudo/actual CT images are calculated. To avoid excessive self-referencing in evaluating anatomical consistency, instead of employing the simple thresholding segmentation (i.e., $\mathbb{S}$), we resort to the open-access TotalSegmentator [15] to finely segment the actual and pseudo CT images to multiple anatomical structures, and compose them to nine independent tissues for simplifying result

Table 1. Comparison of pseudo CT images generated by different methods.

(a) Anatomical consistency (%)

Method	SSIM	Dice$_l$	Dice$_h$	Dice$_k$	Dice$_b$	Dice$_d$	Dice$_r$	Dice$_v$	Dice$_{il}$
UNet	78.40	58.34	78.89	66.91	52.23	43.00	6.83	50.19	41.81
CGAN	**80.42**	80.97	79.25	67.38	52.86	43.91	18.95	52.40	41.19
AGAN	74.71	83.40	74.41	60.26	46.71	39.25	22.39	64.58	49.52
ASEG	76.31	**85.26**	**80.63**	**75.89**	**63.78**	**52.32**	**31.60**	**69.05**	**52.79**

(b) Effectiveness in PET attenuation correction (%)

CT Source	MAE (%)	PSNR (dB)	NCC (%)	SSIM (%)
No CT	1.55±0.56	34.83±2.58	97.13±2.04	94.08±3.44
UNet	1.47±0.50	34.76±2.45	97.20±1.75	95.18±2.49
CGAN	1.52±0.49	34.51±2.33	97.05±1.77	94.92±2.49
AGAN	1.45±0.51	35.09±2.38	97.11±1.80	95.22±2.61
ASEG	**1.36**±0.49	**35.89**±2.48	**97.66**±1.72	**95.67**±2.49
Actual CT	1.20±0.48	37.28±3.37	98.25±1.56	96.10±2.51

report, e.g., lung (Dice$_l$), heart (Dice$_h$), liver (Dice$_{li}$), kidneys (Dice$_k$), blood vessels (Dice$_k$), digestive system (Dice$_d$), ribs (Dice$_r$), vertebras (Dice$_v$), and iliac bones (Dice$_{ib}$). Additionally, the Structure Similarity Index Measure (SSIM) values are also reported to measure the global intensity similarity.

Results of various methods are provided in Table 1(a), where the following conclusions can be drawn. Firstly, U-Net and CGAN generate CT images with slightly better global intensity similarity but worse anatomical consistency in some tissues than AGAN and ASEG. This indicates that the general constraints (MAE and perceptual feature matching) cannot preserve the tissue distribution since they tend to produce statistical average values or patterns, particularly in these regions with large intensity variants. Secondly, AGAN achieves the worst intensity similarity and anatomical consistency for some organs. Such inconsistent metrics suggest that the global intensity similarity may have a competing relationship with anatomical consistency in the learning procedure, thus it is not advisable to balance them in a single network. Thirdly, CGAN achieves better anatomical consistency than U-Net, but worse than ASEG. It implies that the perceptual feature matching loss can also identify the variants between different tissues implicitly but cannot compare to our strategy to explicitly enhance the anatomical skeleton. Fourthly, our proposed ASEG achieves the best anatomical consistency for all tissues, indicating it is reasonable to enhance tissue variations. In brief, the above results supports the strategy to decouple the skeleton prediction as a preceding task is effective for CT generation.

Effectiveness in Attenuation Correction. As the pseudo CT images generated from NAC-PET are expected to be used in AC, it is necessary to further evaluate the effectiveness of pseudo CT images in PET AC. Because we cannot access the original scatters [5], inspired by [11], we propose to resort CGAN to simulate the AC process, denoted as ACGAN and trained on HNSCC dataset. The input of ACGAN is a concatenation of NAC-PET and actual CT, while the output is actual AC-PET. To evaluate the pseudo CT images, we simply use them to take place of the actual CT. Four metrics, including the Peak Signal to Noise Ratio (PSNR), Mean Absolute Error (MAE), Normalized Cross Correlation (NCC), and SSIM, are used to measure ACGAN with pseudo CT images on test datasets (NSCLC, TCGA-HNSC, and TCGA-LUDA). The results are reported in Table 1(b), where the fourth column list the ACGAN results with actual CT images. Meanwhile, we also report the results of direct mapping NAC-PET to AC-PET without CT images in the third column ("No CT"), which is trained from scratch and independent from ACGAN.

It can be observed from Table 1(b) that: (1) ACGAN with actual CT images can predict images very close to the actual AC-PET images, thus is qualified to simulate the AC process; (2) With actual or pseudo CT images, ACGAN can predict images closer to the actual AC-PET images than without CT, demonstrating the necessity of CT images in process of PET AC; (3) These pseudo CTs cannot compare to actual CTs, reflecting that there exist some relative information that can hardly be mined from NAC-PET; (4) The pseudo CTs generated by ASEG achieve the best in three metrics (MAE, PSNR, NCC) and second in the other metric (SSIM), demonstrating the advance of our ASEG. Figure 2 displayed the detailed diversity of the AC-PET corrected by different pseudo
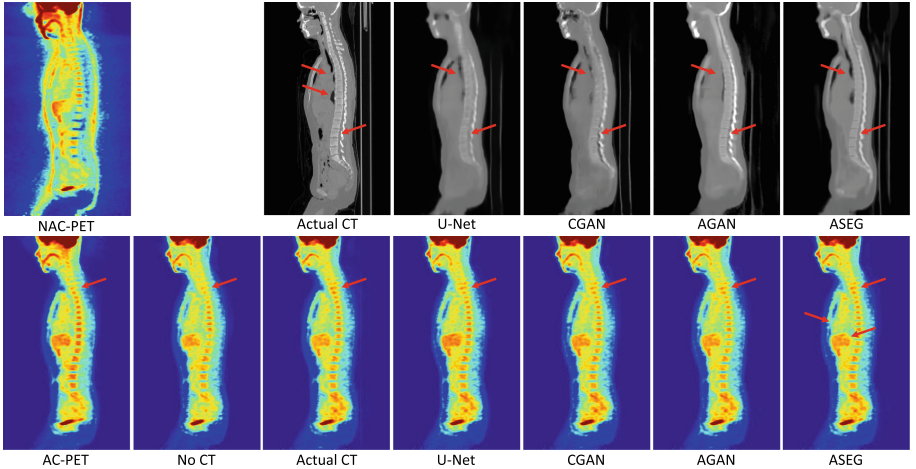

Fig. 2. Visualization of different pseudo CT images (top) and their AC effect (bottom). From left to right are actual NAC-PET/AC-PET and PET corrected without CT (no CT) or with actual or pseudo CT generated by U-Net, CGAN, AGAN, and our ASEG.

CTs. It can be found that the structures of AC-PET are highly dependent on CT, particularly the lung regions. However, errors in corners and shapes are relatively large (see these locations marked by red arrows), which indicates there are still some space in designing more advanced mapping methods. Nonetheless, compared to other pseudo CTs, these generated by ASEG result in more realistic AC-PET with fewer errors, demonstrating the AC usability of ASEG.

4   Conclusion

In this paper, we proposed the anatomical skeleton-enhance generation (ASEG) to generate pseudo CT images for PET attenuation correction (AC), with the goal of avoiding acquiring extra CT or MR images. ASEG divided the CT generation into the skeleton prediction and tissue rendering, two sequential tasks, addressed by two designed generative modules. The first module delineates the anatomical skeleton to explicitly enhance the tissue distribution which are vital for AC, while the second module renders the tissue details based on the anatomical skeleton and NAC-PET. Under the collaboration of two modules and specific *anatomical-consistency constraint*, our ASEG can generate more reasonable pseudo CT from NAC-PET. Experiments on a collection of public datasets demonstrate that our ASEG outperforms existing methods by achieving advanced performance in anatomical consistency. Our study support that ASEG could be a promising and lower-cost alternative of CT acquirement for AC. Our future work will extend our study to multiple PET tracers.

References

1. Armanious, K., et al.: Independent attenuation correction of whole body [18F] FDG-PET using a deep learning approach with generative adversarial networks. EJNMMI Res. **10**(1), 1–9 (2020)
2. Dong, X., et al.: Synthetic CT generation from non-attenuation corrected PET images for whole-body PET imaging. Phys. Med. Biol. **64**(21), 215016 (2019)
3. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3D U-Net: learning dense volumetric segmentation from sparse annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9901, pp. 424–432. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46723-8_49
4. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., et al.: The cancer imaging archive (TCIA): maintaining and operating a public information repository. J. Digit. Imaging **26**(6), 1045–1057 (2013)

5. Guo, R., Xue, S., Hu, J., Sari, H., Mingels, C., et al.: Using domain knowledge for robust and generalizable deep learning-based CT-free PET attenuation and scatter correction. Nat. Commun. **13**, 5882 (2022)
6. Häggström, M.: Hounsfield units. https://radlines.org/Hounsfield_unit
7. Liu, F., Jang, H., Kijowski, R., Bradshaw, T., McMillan, A.B.: Deep learning MR imaging-based attenuation correction for PET/MR imaging. Radiology **286**(2), 676–684 (2018)
8. Pan, Y., Liu, M., Xia, Y., Shen, D.: Disease-image-specific learning for diagnosis-oriented neuroimage synthesis with incomplete multi-modality data. IEEE Trans. Pattern Anal. Mach. Intell. **44**, 6839–6853 (2021)
9. Rabinovici, G.D., Gatsonis, C., Apgar, C., Chaudhary, K., Gareen, I., et al.: Association of amyloid positron emission tomography with subsequent change in clinical management among medicare beneficiaries with mild cognitive impairment or dementia. JAMA **321**(13), 1286–1294 (2019)
10. Rodríguez Colmeiro, R., Verrastro, C., Minsky, D., Grosges, T.: Towards a whole body [18F] FDG positron emission tomography attenuation correction map synthesizing using deep neural networks. J. Comput. Sci. Technol. **21**, 29–41 (2021)
11. Shiri, I., et al.: Direct attenuation correction of brain PET images using only emission data via a deep convolutional encoder-decoder (Deep-DAC). Eur. Radiol. **29**(12), 6867–6879 (2019). https://doi.org/10.1007/s00330-019-06229-1
12. Spencer, B.A., Berg, E., Schmall, J.P., Omidvari, N., Leung, E.K., et al.: Performance evaluation of the uEXPLORER total-body PET/CT scanner based on NEMA NU 2-2018 with additional tests to characterize PET scanners with a long axial field of view. J. Nucl. Med. **62**(6), 861–870 (2021)
13. Tan, H., et al.: Total-body PET/CT: current applications and future perspectives. Am. J. Roentgenol. **215**(2), 325–337 (2020)
14. Wang, T.C., Liu, M.Y., Zhu, J.Y., Tao, A., Kautz, J., Catanzaro, B.: High-resolution image synthesis and semantic manipulation with conditional GANs. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 8798–8807 (2018)
15. Wasserthal, J., Meyer, M., Breit, H., Cyriac, J., Yang, S., Segeroth, M.: TotalSegmentator: robust segmentation of 104 anatomical structures in CT images. arXiv preprint arXiv:2208.05868 (2022)
16. Zhao, R., et al.: Rethinking dice loss for medical image segmentation. In: 2020 IEEE International Conference on Data Mining (ICDM), pp. 851–860. IEEE (2020)