

# Whole Body Positron Emission Tomography Attenuation Correction Map Synthesizing using 3D Deep Generative Adversarial Networks

Ramiro Rodriguez Colmeiro (✉ [rodriguez.colmeiro@gmail.com](mailto:rodriguez.colmeiro@gmail.com))

CONICET San Luis <https://orcid.org/0000-0001-9279-5717>

**Claudio Verrastro**

Comision Nacional de Energia Atomica

**Daniel Minsky**

Consejo Nacional de Investigaciones Cientificas y Tecnicas

**Thomas Grosjes**

Universite de Technologie de Troyes

---

## Original research

**Keywords:** Attenuation Correction, Deep Learning, Generative Models, Positron Emission Tomography

**Posted Date:** July 28th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-46953/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

## RESEARCH

# Whole Body Positron Emission Tomography Attenuation Correction Map Synthesizing using 3D Deep Generative Adversarial Networks

Ramiro G Rodríguez Colmeiro<sup>1,2,4\*</sup>, Claudio A Verrastro<sup>2,3</sup>, Daniel Minsky<sup>3,4</sup> and Thomas Grosge<sup>1</sup>

\*Correspondence:

[ramiro.german.rodriguez.colmeiro@utt.fr](mailto:ramiro.german.rodriguez.colmeiro@utt.fr)

<sup>1</sup>GAMMA3 (UTT-INRIA),

Université de Technologie de

Troyes, 12 Rue Marie Curie, 10004

Troyes cedex, France

Full list of author information is

available at the end of the article

## Abstract

**Background:** The correction of attenuation effects in Positron Emission Tomography (PET) imaging is fundamental to obtain a correct radiotracer distribution. However direct measurement of this attenuation map is not error-free and normally results in additional ionization radiation dose to the patient. Here, we propose to obtain the whole body attenuation map using a 3D U-Net generative adversarial network. The network is trained to learn the mapping from non attenuation corrected <sup>18</sup>F-fluorodeoxyglucose PET images to a synthetic Computerized Tomography (sCT) and also to label the input voxel tissue. The sCT image is further refined using an adversarial training scheme to recover higher frequency details and lost structures using context information. This work is trained and tested on public available datasets, containing several PET images from different scanners with different radiotracer administration and reconstruction modalities. The network is trained with 108 samples and validated on 10 samples.

**Results:** The sCT generation was tested on 133 samples from 8 distinct datasets. The resulting mean absolute error of the network is  $103 \pm 18$  HU and a peak signal to noise ratio of  $18.6 \pm 1.5$  dB. The generated images show good correlation with the unknown structural information.

**Conclusions:** The proposed deep learning topology is capable of generating whole body attenuation maps from uncorrected PET image data. Moreover, the method accuracy holds in the presence of data from multiple sources and modalities and is trained on publicly available datasets.

**Keywords:** Attenuation Correction; Deep Learning; Generative Models; Positron Emission Tomography

## Background

The correct estimation of attenuation correction maps of positron emission tomography (PET) images is fundamental to their correct reconstruction, but direct measurement of this map means additional ionization radiation dose to the patient. A safer approach to obtain this information is to use image analysis methods. These methods create an attenuation structure from other image modality, such as Magnetic Resonance Imaging (MRI) studies or the Non Attenuation Corrected PET (NAC-PET) image. This image translation is specially difficult in whole-body NAC-PET images, since the information it presents is incomplete. In this scenario,

where the translation process also needs to fill information blanks, the generative adversarial networks (GANs) are specially powerful.

The application of GANs in image to image translation tasks has been successfully exploited in many medical imaging domains, including PET attenuation map synthesizing. However, most methods of attenuation map generation analyze the MRI to CT translation using convolutional networks [1] and GANs with paired [2] and unpaired data [3], requiring a co-registered MRI image which contains anatomical information that is not present in NAC-PET images. The PET (and NAC-PET) to Computed Tomography (CT) image translation remains as one of the less explored domains. The studies in this particular domain focus on PET-CT image translation of corrected images on head scans. Liu [4] proposes using a 2D U-Net architecture to translate NAC-PET head scans to CT, showing promising results in head region scans. Armanious [2] proposes a general GAN application composed of a cascaded 2D U-Net generator and a discriminator used to evaluate the perceptual loss and style of the generated image. They show the capability of the topology to translate PET scans to CT, using only axial slices and, again, only for head region scans. Both methods provide no information on their capability for whole-body image translation which is a harder problem to solve, given that the possible modes in the attenuation structures is larger. Solving this problem is essential to the application of this technique in the praxis.

We propose to use a fully 3D GAN topology with mixed loss in order to generate high quality whole body CT images from NAC-PET images. Given that, the dimensionality of the 3D volumes is comparable to the generation of high resolution 2D images. Therefore we perform a two step training, starting with supervised training of labels and then we add an adversarial loss block to enhance the image resolution. This results in faster convergence and lower computational cost. Our model is trained on a public available data-set from the Cancer Image Archive [5], the data-set contains series of registered CT, PET and NAC-PET scans of head and neck squamous cell carcinoma (HNSCC) [6]. We use 8 different datasets for the testing process, containing 5 different types of carcinomas and scanners from multiple manufacturers.

## Methods

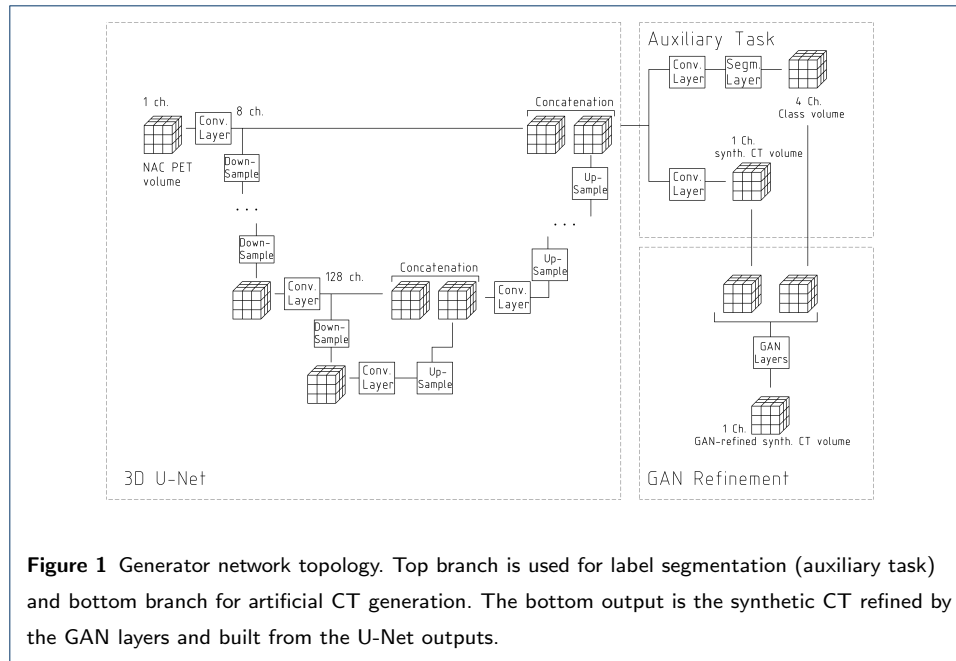
In this section the topology of the network is presented, followed by the implementation of the loss function and implementation details.

### Topology Description

The network topology is composed of a 3D U-Net generator and a convolutional critic (or discriminator). An additional segmentation branch is used to regularize the training. Nevertheless, the adversarial loss gradient flow is limited to the last part of the network.

#### *Generator:*

The initial section of the generator is a 3D U-Net topology [7], after which the network forks into two branches. The model representation can be seen in Fig. 1.



The first branch is used for segmentation, it is composed of three convolutional layers and ends in a softmax layer. The second branch is responsible for the synthetic CT (sCT) generation, it is composed of a convolutional layer with an hyperbolic tangent activation and the GAN layers. The outputs of the U-Net are merged and processed by the GAN layers. These are a collection of 5 convolutional layers with 8 filters each, used during the adversarial training. All convolutional operations use a filter of size  $3 \times 3 \times 3$  except the output layer which use a  $1 \times 1 \times 1$  filter. The network posses 5 resolution levels, each of them composed of two convolutional layers with filter shape of  $3 \times 3 \times 3$  and Rectified Linear Unit (ReLU) activation. Each resolution level posses an skip connection between the down-sample and up-sample path. Instead of convolutional resampling the resolution changes are performed using trilinear up- or down-sampling. After each convolutional layer we apply a voxel normalization along feature maps, dividing each voxel value by  $\sqrt{\frac{1}{N} \sum_{i=0}^{N-1} (vx_{x,y,z}^i)^2 + e}$ , where  $N$  is the number of channels in the feature map,  $vx_{x,y,z}^i$  is the voxel value of the  $i^{th}$  voxel in the position  $(x, y, z)$  and  $e = 1.0 \times 10^{-8}$ . We also apply at each convolutional layer, a scaling factor to the filter kernel based on He's [8] scaled initialization of weights.

#### Critic:

The critic or discriminator network is a fully convolutional network with ReLU activation in all layers, only the last layer has no activation. The input of this network is a two channel volume composed of the NAC-PET volume and the real or sCT image. The output of the network is a value proportional to quality value of the generated image. The network is conformed by 4 resolution levels with two convolutional layers per level. Each convolution has a filter size of  $3 \times 3 \times 3$  and ReLU activation. No batch or pixel normalization is applied. The last two layers of the critic are a flatten operation followed by a single dense layer with linear output.

### Training Scheme

The training of the network is divided into two stages: first the generator network is trained in a supervised manner using a composed loss. The segmentation branch of the network applies a 3D-DICE as shown in (1),

$$L_D = \frac{1}{N_c} \sum_{c=0}^{N_c} \frac{2 \sum_i^{N_v} p_{i,c} g_{i,c}}{\sum_i^{N_v} p_{i,c}^2 + \sum_i^{N_v} g_{i,c}^2}, \quad (1)$$

where  $N_c$  is the number of objective classes,  $N_v$  the number of voxels in the volume,  $g_{i,c}$  are the voxels of the ground truth and  $p_{i,c}$  the values of the softmaxed output of the network. The dice loss ranges from 0 to 1. It produces its maximum value when all voxels of the ground truth ( $g_{i,c}$ ) have the same value as the softmaxed output voxels ( $p_{i,c}$ ). Since the output is softmaxed the denominator of (1) is always larger than its numerator except when  $g_{i,c}$  and  $p_{i,c}$  are identical. In the case of a multi-class problem ( $N_c > 1$ ) the final value is divided by the number of classes.

The CT synthetization branch, up to the GAN layers, is trained using as loss function the euclidean distance between the sCT and the objective CT image,  $L_e = \|s, r\|^2$ , where  $s$  is the sCT and  $r$  is the real CT volume. The loss for the supervised training stage is shown in (2),

$$L_{gen}^{sup.} = (1 - L_D) + K_e L_e, \quad (2)$$

where  $K_e$  is a coupling constant. After the initial training, the adversarial training starts. The adversarial training uses the Wasserstein-GAN (W-GAN) strategy [9], resulting in a generator loss as shown in (3),

$$L_W = -f_c(f_g(x)), \quad (3)$$

where  $f_c()$  is the critic network function,  $f_g()$  is the generator network function and  $x$  is the input NAC-PET image. During the adversarial training the GAN layers become active and are trained using the W-GAN loss. The gradient of the GAN does not flow into the 3D U-Net layers. The critic is trained using a coupled pairs of NAC-PET and CT images, real or fake. It is trained using the wasserstein loss shown in (4),

$$L_{crit.}^{GAN} = f_c(r) - f_c(f_g(x)) + \lambda G_p(s, r), \quad (4)$$

where  $G_p$  is the gradient penalty [10] and  $\lambda = 10.0$ . The critic is trained 5 steps for each generator step. At the initial step of the GAN training stage the critic is optimally trained before starting the GAN training loop. The generators are trained using an Adaptive Moment Estimation (ADAM) optimizer with parameters  $\beta_1 = 0.0$ ,  $\beta_2 = 0.99$  and  $\epsilon = 1.0 \times 10^8$  and learning rate  $lr = 0.0001$ . The discriminator uses a RMSprop optimizer with learning rate  $lr = 0.0005$ .

## Results

In this section, we first introduce the train dataset and the preprocessing operations, then we present the test datasets and finally, the test metrics and the results are given.

### Train Dataset Description

The HNSCC dataset is composed of a series of registered CT, PET and NAC-PET scans of head and neck squamous cell carcinoma. The dataset is first stripped of all samples that do not contain PET, NAC-PET and CT matched samples. Then the matched samples are tested for overlapping and cropped to contain only axial slices with information from all the image types. After the dataset is clean, we normalize the image size to a  $128 \times 128 \times 256$  voxels FoV, with a voxel size of  $5.46 \times 5.46 \times 5.08$  mm<sup>3</sup>. The final dataset contains 118 images from 71 different patients, from which 7 patients (10 images) were separated as validation dataset. Before feeding each sample to the network, the volume is randomly sliced into a  $128 \times 128 \times 32$  volume and the input NAC-PET voxels values are randomly shifted by a 10% and re normalized.

#### *Objective CT normalization:*

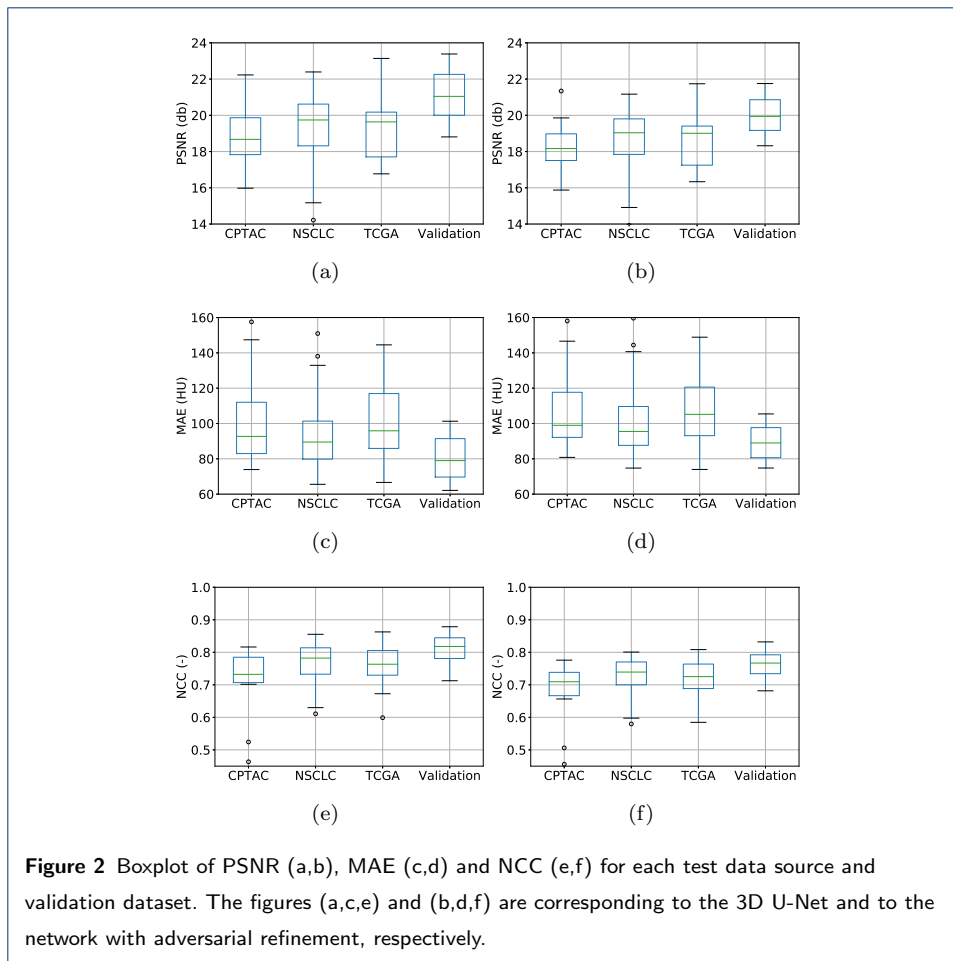
The objective CT must be free of the couch structures, this task is performed using a method based on the voxel variance along the axial axis [11]. Then the image dynamic range is clipped between  $-125$  and  $1300$  Hounsfield Units (HU) and normalized between 0 and 1, to maximize the distance between soft and bone tissue.

#### *Label Generation:*

Four labels classes are extracted from the couch stripped CT image using a voxel value threshold and open and closing filters. The *Air-Lung* mask values ranges from  $-1000$  HU to  $-125$  HU, the *Fluids-Fat* mask ranges from  $-125$  HU to  $10$  HU, the *Parenchyma* mask ranges from  $10$  HU to  $90$  HU and the *Bone* mask ranges from  $90$  HU to  $1300$  HU.

### Test Datasets Description

The test datasets are series of public dataset, also from TCIA, including different types of lesions, patients and scanner technologies. These datasets were cleaned from non matching samples and re normalized to fit voxel size expected by the network, resulting in 133 test samples: 73 from the Non-Small Cell Lung Cancer (NSCLC) [12] dataset, 25 from The Cancer Genome Atlas - Head-Neck Squamous Cell Carcinoma (TCGA-HNSC) [13], 20 from The Cancer Genome Atlas Lung Adenocarcinoma (TCGA-LUAD) [14], 1 from The Cancer Genome Atlas - Thyroid Cancer (TCGA-THCA) [15], 4 from Clinical Proteomic Tumor Analysis Consortium - Lung Adenocarcinoma (CPTAC-LUAD) [16], 1 from Clinical Proteomic Tumor Analysis Consortium - Pancreatic Ductal Adenocarcinoma (CPTAC-PDA) [17], 3 from Clinical Proteomic Tumor Analysis Consortium - Uterine Corpus Endometrial Carcinoma (CPTAC-UCEC) [18] and 6 from Clinical Proteomic Tumor Analysis Consortium - Lung Squamous Cell Carcinoma (CPTAC-LSCC)[19].



During the testing procedure the NAC-PET is fed in consecutive slices of  $128 \times 128 \times 32$  voxels. The intensity is re-scaled to  $[0, 1]$  for each slice. The resulting sCT slices are composed into a single volume using a weighted average operation.

### Experimental Results

The network was tested on the quality of the generated CT images using three different metrics, Peak Signal to Noise Ratio (PSNR), Mean Absolute Error (MAE) and Normalized Cross Correlation (NCC). The metrics are presented in the Fig. 2, each of the box plots corresponds to the metrics of the 3D U-Net and GAN on the different test sets. The datasets are presented by source since the number of samples of the individual datasets is too small in some cases.

An ablation test was done to assess the influence of the different network properties. Five different networks were trained for the same amount of epochs and have the same number of parameters. The baseline network has no supervised loss (shown in E1. 2), the adversarial gradient is allowed through the U-Net and is pre-trained. The code *Sup.* indicates supervised loss during GAN training, *(No) Pre-Train* indicates whether or not the network has a supervised pre-train step, *No GAN.Loss* indicates no adversarial training and *RestrictedGrad.* indicates that the adversarial gradient

is not allowed into the U-Net. The results of the study using the PSNR, MAE and MCC metrics of all datasets is summarized in table 1.

Topology	Dataset	PSNR [dB]	MAE [HU]	NCC [-]
Baseline	Validation	$18.3 \pm 1.36$	$121 \pm 20.0$	$0.64 \pm 0.076$
	Testing	$17.3 \pm 1.46$	$133 \pm 21.9$	$0.60 \pm 0.074$
Sup. Pre-Train	Validation	$19.6 \pm 1.25$	$92.2 \pm 14.1$	$0.75 \pm 0.058$
	Testing	$18.3 \pm 1.41$	$105 \pm 19.8$	$0.70 \pm 0.066$
Sup. No Pre-Train	Validation	$17.5 \pm 1.76$	$147 \pm 38.3$	$0.54 \pm 0.14$
	Testing	$17.0 \pm 1.37$	$143 \pm 24.9$	$0.57 \pm 0.082$
Sup. Pre-Train No GAN.Loss.	Validation	$21.0 \pm 1.49$	$80.3 \pm 13.4$	$0.80 \pm 0.050$
	Testing	$19.3 \pm 1.68$	$96.8 \pm 20.4$	$0.76 \pm 0.064$
<b>Sup. Pre-Train RestrictedGrad.</b>	Validation	<b><math>19.9 \pm 1.35</math></b>	<b><math>88.9 \pm 10.5</math></b>	<b><math>0.76 \pm 0.046</math></b>
	Testing	<b><math>18.6 \pm 1.49</math></b>	<b><math>103 \pm 18.5</math></b>	<b><math>0.72 \pm 0.059</math></b>

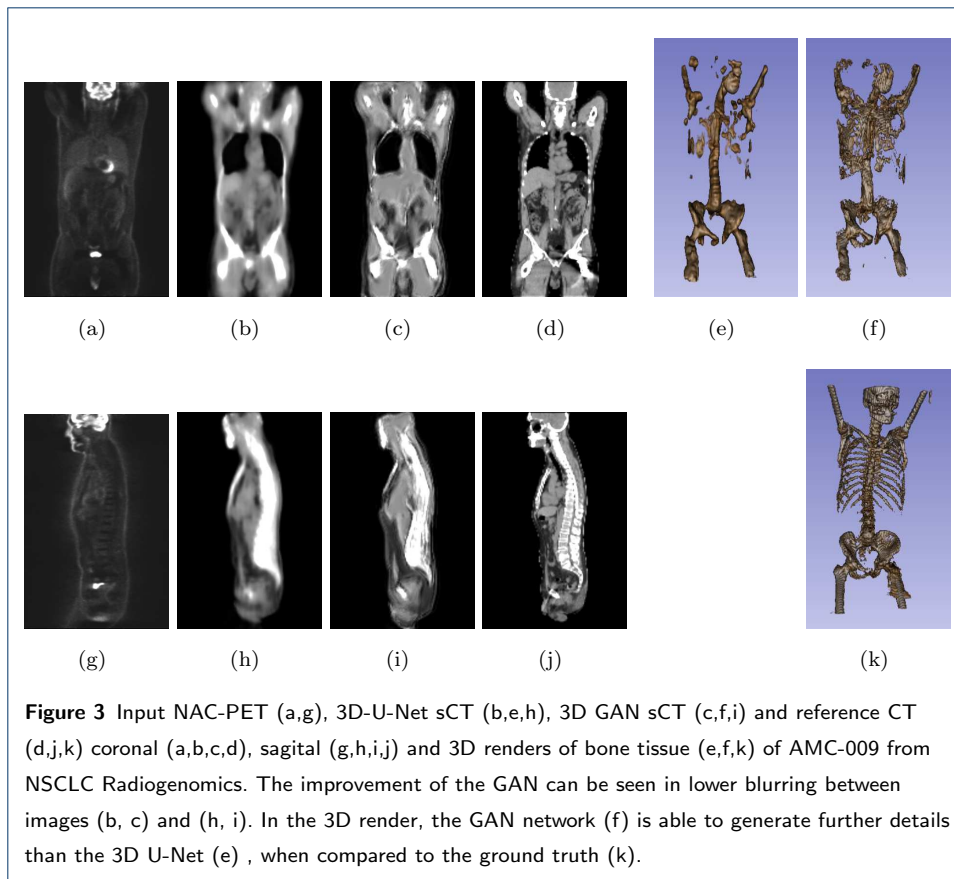
**Table 1** Validation (10 samples) and testing (133 samples) scores for topologies with additional features. Supervised loss (*Sup.*), pre-train step (*Pre-Train*), no adversarial training (*No GAN.Loss.*) and adversarial gradient restricted to last layers (*RestrictedGrad.*). The presented network is in bold text.

Three samples from the test datasets can be seen in Fig. 3, Fig. 4, and Fig. 5, two with the patients arms elevated over the head (arms up) and other with the patients arms positioned at the side (arms down). These images correspond to the NSCLC Radiogenomics, CPTAC-PDA and TCGA-HNSC datasets respectively. In the sub-figures (b,h) are the sCT images generated using only the supervised loss shown in Eq. 2 and in sub-figures (c,i), are the sCT images generated using the adversarial loss shown in Eq. 3.

## Discussion

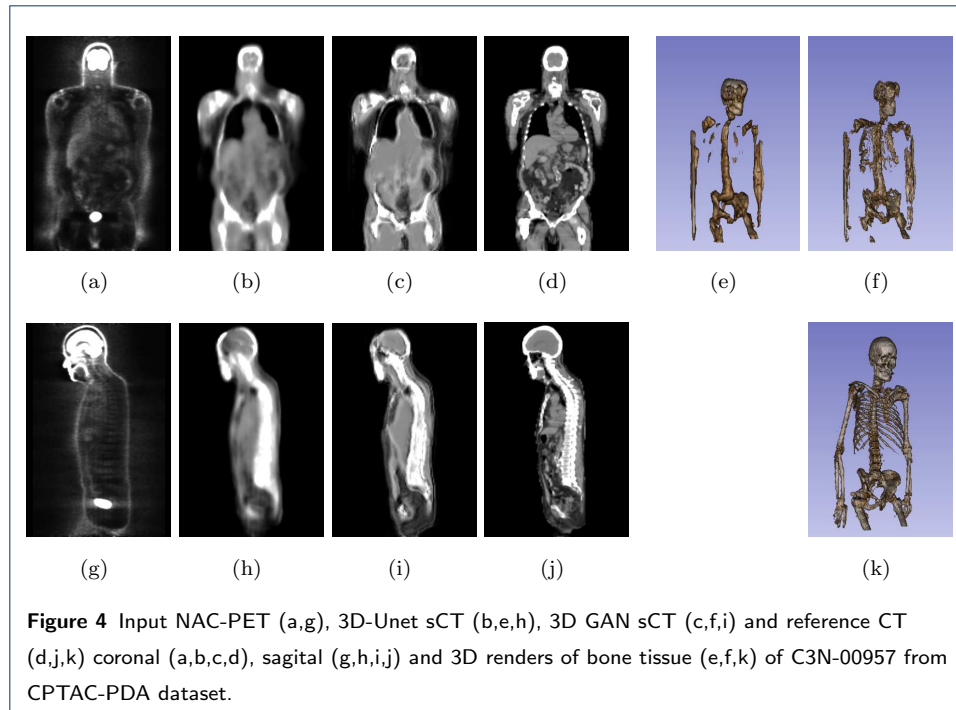
Our technique shows to be resistant to multiple reconstruction techniques and scanners technologies, as shown by the test metrics in Fig. 2. The basic 3D U-Net topology generates synthetic attenuation correction images with a PSNR of  $19.3 \pm 1.68$  dB, a MAE of  $96.7 \pm 20.4$  HU and NCC of  $0.76 \pm 0.064$ . The addition of the GAN layers achieve a PSNR of  $18.6 \pm 1.45$  dB, a MAE of  $103.2 \pm 18.53$  HU and NCC of  $0.72 \pm 0.059$ . These scores are obtained on test samples from different scanners, patients and lesions, showing that our technique can be used on multiple sources. The adversarial loss enables the network to learn higher frequency details, as shown in Fig. 3. It can be seen that the supervised network is unable to generate fine bone structures (see Fig. 3(f)) whereas the GAN trained network shows a more complete structure, including part of the rib cage and the sacrum and coccyx. Nevertheless these improvements are not reflected in the performance of the metrics, resulting in lower but not significantly important changes when compared to the base 3D U-Net. This illustrates the limits of simplistic metrics to reflect anatomic improvements in the image generation process. It is expected that more specific losses based on learnt representations could reflect this improvement [20].





The U-Net and the adversarial networks fail to generalize the upper section of the body where less training data were available. This is also reflected in the arms, that appear in different positions inside the dataset. This could be mitigated by anatomically matching the training data and train region specific networks. Further improvement in this direction will be to train the network using a full-size intermediate space to map each anatomical section, such as the intermediate representations presented in [21] and [22]. The current sCT generation can also be used as prior on attenuation reconstruction techniques such as the maximum-likelihood reconstruction of attenuation and activity (MLAA) [23, 24] single scatter modeling [25]. These techniques can potentially eliminate artefacts from the generated attenuation maps, such as the CT contrast in the stomach observed in Fig. 5(d) that is not present in Fig 5(b) nor in Fig 5(b), however its effect is clearly seen in Fig. 5(a).

Recently, cycle-GAN architectures were proposed for NAC-PET to CT translation [26] on controlled dose administration and reconstruction procedures for the Discovery 690 PET/CT scanner (General Electric) with time of flight capabilities. They achieve a MAE of  $108.4 \pm 19.1$  HU in the reconstruction of the CT image. In a later work they provide proof of the capacity of sCT images to be used as attenuation correction maps [27]. Their results are within our accuracy but no direct comparison was possible since their data is private. We consider that is important to possess a common dataset to test these methods, for this reason the code and dataset used in this work is released.



## Conclusion

We presented a deep learning approach to the task of attenuation map generation from uncorrected PET image data. The method performs with accuracy comparable to other methods showing their fitting for PET image correction.

### Abbreviations

PET: Positron Emission Tomography ; MRI: Magnetic Resonance Imaging ; NAC-PET: Non Attenuation Corrected PET ; GANs: generative adversarial networks ; CT: Computed Tomography ; HNSCC: head and neck squamous cell carcinoma ; sCT: synthetic CT ; ReLU: Rectified Linear Unit ; W-GAN: Wasserstein-GAN ; ADAM: Adaptive Moment Estimation ; HU: Hounsfield Units ; NSCLC: Non-Small Cell Lung Cancer ;TCGA HNSC: The Cancer Genome Atlas - Head-Neck Squamous Cell Carcinoma ; TCGA-LUAD: The Cancer Genome Atlas Lung Adenocarcinoma ; CPTAC-LUAD: Clinical Proteomic Tumor Analysis Consortium - Lung Adenocarcinoma ; CPTAC-PDA: Clinical Proteomic Tumor Analysis Consortium - Pancreatic Ductal Adenocarcinoma ; CPTAC-UCEC: Clinical Proteomic Tumor Analysis Consortium - Uterine Corpus Endometrial Carcinoma ; CPTAC-LSCC: Clinical Proteomic Tumor Analysis Consortium - Lung Squamous Cell Carcinoma ; PSNR: Peak Signal to Noise Ratio ; MAE: Mean Absolute Error ; NCC: Normalized Cross Correlation ; MLAA: maximum-likelihood reconstruction of attenuation and activity

### Ethics approval and consent to participate

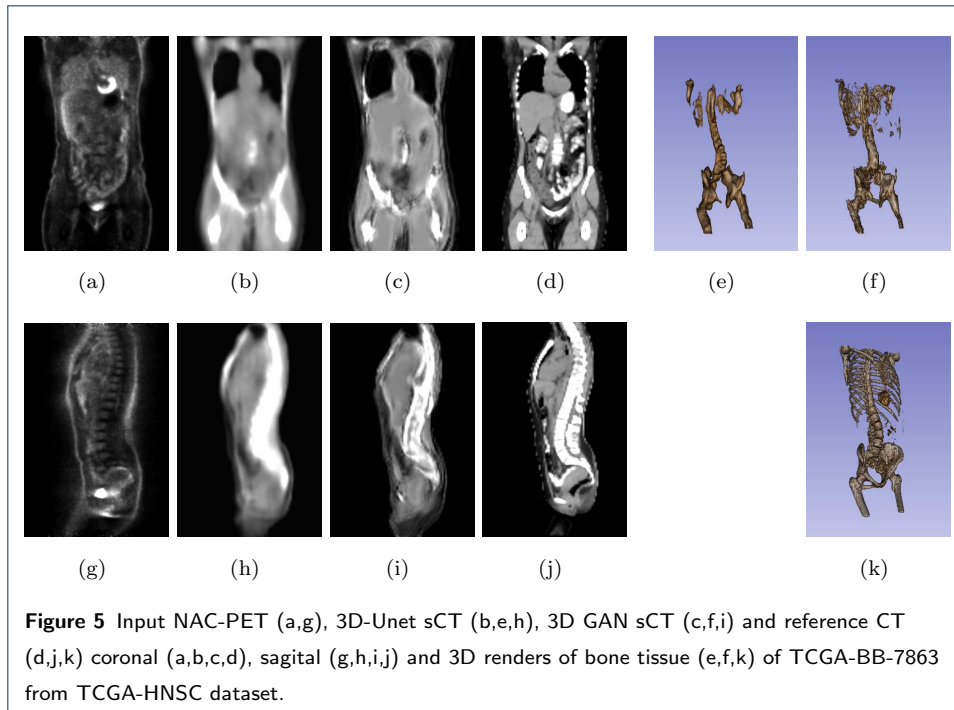
Not applicable.

### Consent for publication

Not applicable.

### Availability of data and material

The source code of the presented work, including the data pre-processing code and the testing code are available at — (will be made public). The datasets generated and/or analysed during the current study are available in the TCIA repository <https://www.cancerimagingarchive.net/>.



#### Competing interests

The authors declare that they have no competing interests.

#### Funding

This work was supported by the Universidad Tecnológica Nacional, the Université de Technologie de Troyes, the Comisión Nacional de Energía Atómica and the National Scientific and Technical Research Council (CONICET).

#### Author's contributions

RRC, CAV, DM, and TG are the guarantors of integrity of the entire study. CAV, DM, and TG to the data interpretation, manuscript drafting and revision for important intellectual content, literature research, and manuscript editing. All authors approved the final version of the manuscript.

#### Acknowledgements

Not applicable.

#### Author details

<sup>1</sup>GAMMA3 (UTT-INRIA), Université de Technologie de Troyes, 12 Rue Marie Curie, 10004 Troyes cedex, France.

<sup>2</sup>Universidad Tecnológica Nacional, Sarmiento 440, C1041AAJ Buenos Aires, Argentina. <sup>3</sup>Comisión Nacional de Energía Atómica, Av. del Libertador 8250, C1429BNP, Buenos Aires, Argentina. <sup>4</sup>CONICET, Godoy Cruz 2290, C1425FQB, Buenos Aires, Argentina.

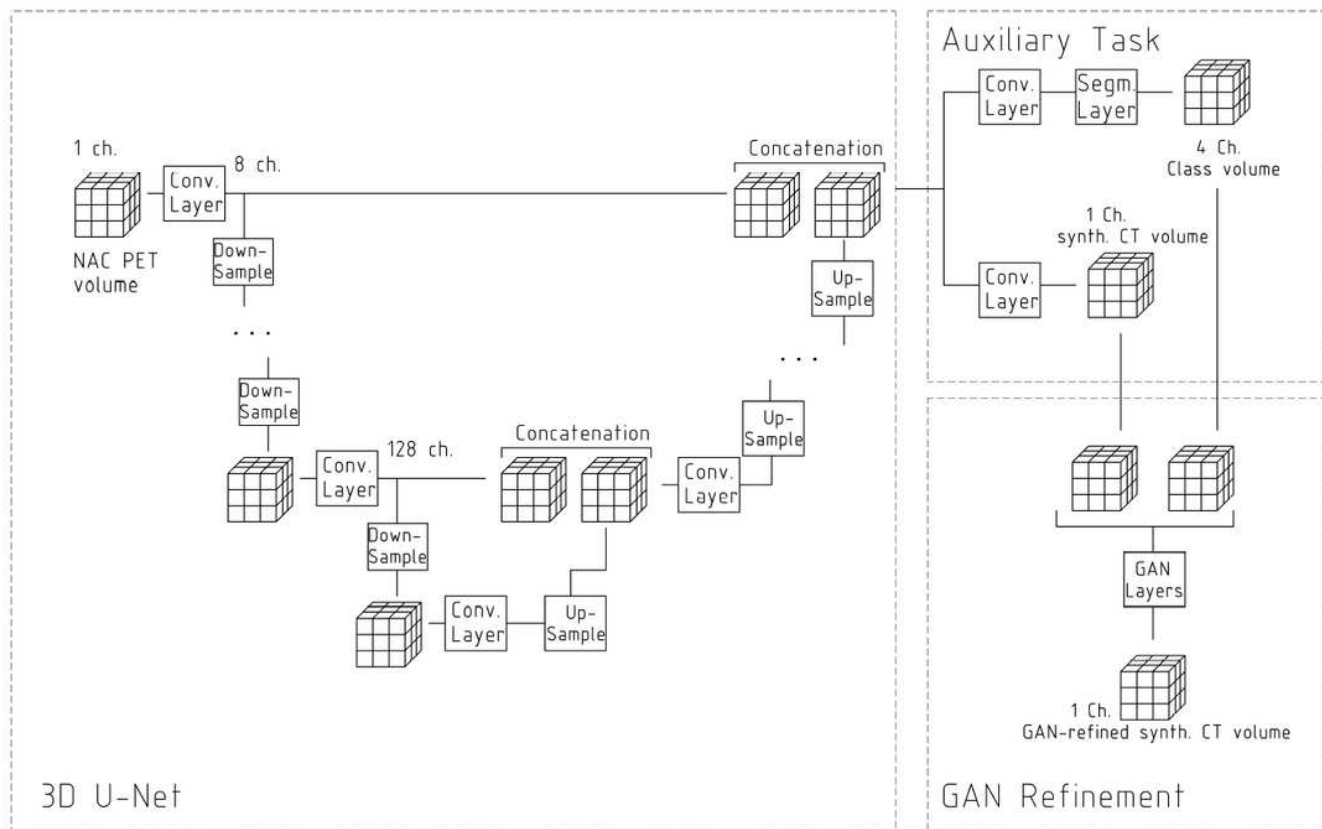
#### References

- Nie, D., Cao, X., Gao, Y., Wang, L., Shen, D.: Estimating ct image from mri data using 3d fully convolutional networks. In: *Deep Learning and Data Labeling for Medical Applications*, pp. 170–178. Springer, ??? (2016)
- Armanious, K., Jiang, C., Fischer, M., Küstner, T., Hepp, T., Nikolaou, K., Gatidis, S., Yang, B.: Medgan: Medical image translation using gans. *Computerized Medical Imaging and Graphics*, 101684 (2019)
- Wolterink, J.M., Dinkla, A.M., Savenije, M.H., Seevinck, P.R., van den Berg, C.A., Išgum, I.: Deep mr to ct synthesis using unpaired data. In: *International Workshop on Simulation and Synthesis in Medical Imaging*, pp. 14–23 (2017). Springer

4. Liu, F., Jang, H., Kijowski, R., Zhao, G., Bradshaw, T., McMillan, A.B.: A deep learning approach for 18 f-fdg pet attenuation correction. *EJNMMI physics* **5**(1), 24 (2018)
5. Clark, K., Vendt, B., Smith, K., Freymann, J., Kirby, J., Koppel, P., Moore, S., Phillips, S., Maffitt, D., Pringle, M., *et al.*: The cancer imaging archive (tcia): maintaining and operating a public information repository. *Journal of digital imaging* **26**(6), 1045–1057 (2013)
6. Tatiana, B., De Ornelas Couto, M., B. Mihaylov, I.: Head-and-neck squamous cell carcinoma patients with CT taken during pre-treatment, mid-treatment, and post-treatment Dataset. (2018).  
doi:[10.7937/K9/TCIA.2018.13upr2xf](https://doi.org/10.7937/K9/TCIA.2018.13upr2xf)
7. Milletari, F., Navab, N., Ahmadi, S.-A.: V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: *3D Vision (3DV), 2016 Fourth International Conference On*, pp. 565–571 (2016). IEEE
8. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1026–1034 (2015)
9. Arjovsky, M., Chintala, S., Bottou, L.: Wasserstein generative adversarial networks. In: Precup, D., Teh, Y.W. (eds.) *Proceedings of the 34th International Conference on Machine Learning, Proceedings of Machine Learning Research*, vol. 70, pp. 214–223. PMLR, International Convention Centre, Sydney, Australia (2017).  
<http://proceedings.mlr.press/v70/arjovsky17a.html>
10. Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A.C.: Improved training of wasserstein gans. In: *Advances in Neural Information Processing Systems*, pp. 5767–5777 (2017)
11. Bandi, P., Zsoter, N., Seres, L., Toth, Z., Papp, L.: Automated patient couch removal algorithm on ct images. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 7783–7786 (2011). IEEE
12. Bakr, S., Gevaert, O., EcheGARAY, S., Ayers, K., Zhou, M., Shafiq, M., Zheng, H., Zhang, W., Leung, A., Kadoch, M., *et al.*: Data for nscl radiogenomics collection. *The Cancer Imaging Archive* (2017).  
doi:[10.7937/K9/TCIA.2017.7hs46erv](https://doi.org/10.7937/K9/TCIA.2017.7hs46erv)
13. Zuley, M.L., Jarosz, R., Kirk, S., Lee, Y., Colen, R., Garcia, K., Aredes, N.: Radiology data from the cancer genome atlas head-neck squamous cell carcinoma [tcga-hnsc] collection. *Cancer Imaging Archive*. doi (2016).  
doi:[10.7937/K9/TCIA.2016.LXKQ47MS](https://doi.org/10.7937/K9/TCIA.2016.LXKQ47MS)
14. Albertina, B., Watson, M., Holback, C., Jarosz, R., Kirk, S., Lee, Y., Lemmerman, J.: Radiology data from the cancer genome atlas lung adenocarcinoma [tcga-luad] collection. *Cancer Imaging Arch* (2016).  
doi:[10.7937/K9/TCIA.2016.JGNIHEP5](https://doi.org/10.7937/K9/TCIA.2016.JGNIHEP5)
15. Kirk, S., Lee, Y., Roche, C., *et al.*: Radiology data from the cancer genome atlas thyroid cancer [tcga-thca] collection. *Cancer Imaging Archive*. doi (2016). doi:[10.7937/K9/TCIA.2016.9ZFRVF1B](https://doi.org/10.7937/K9/TCIA.2016.9ZFRVF1B)
16. Meldo, A., Utkin, L., Lukashin, A., Muliukha, V., Zaborovsky, V.: Database acquisition for the lung cancer computer aided diagnostic systems. In: *2019 25th Conference of Open Innovations Association (FRUCT)*, pp. 220–227 (2019). IEEE
17. Menegotto, A.B., Becker, C.D.L., Cazella, S.C.: Computer-aided hepatocarcinoma diagnosis using multimodal deep learning. In: *International Symposium on Ambient Intelligence*, pp. 3–10 (2019). Springer
18. National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC), N.C.I.C.P.T.A.C.C.: Radiology Data from the Clinical Proteomic Tumor Analysis Consortium Uterine Corpus Endometrial Carcinoma (CPTAC-UCEC) Collection. *The Cancer Imaging Archive* (2019). doi:[10.7937/K9/TCIA.2018.3R3JUISW](https://doi.org/10.7937/K9/TCIA.2018.3R3JUISW).  
<https://wiki.cancerimagingarchive.net/display/Public/CPTAC-UCEC>
19. National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC), N.C.I.C.P.T.A.C.C.: Radiology Data from the Clinical Proteomic Tumor Analysis Consortium Lung Squamous Cell Carcinoma [CPTAC-LSCC] Collection. *The Cancer Imaging Archive* (2018). doi:[10.7937/k9/tcia.2018.6emub5I2](https://doi.org/10.7937/k9/tcia.2018.6emub5I2).  
<https://wiki.cancerimagingarchive.net/display/Public/CPTAC-LSCC>
20. Mansilla, L., Milone, D.H., Ferrante, E.: Learning deformable registration of medical images with anatomical constraints. *Neural Networks* **124**, 269–279 (2020)

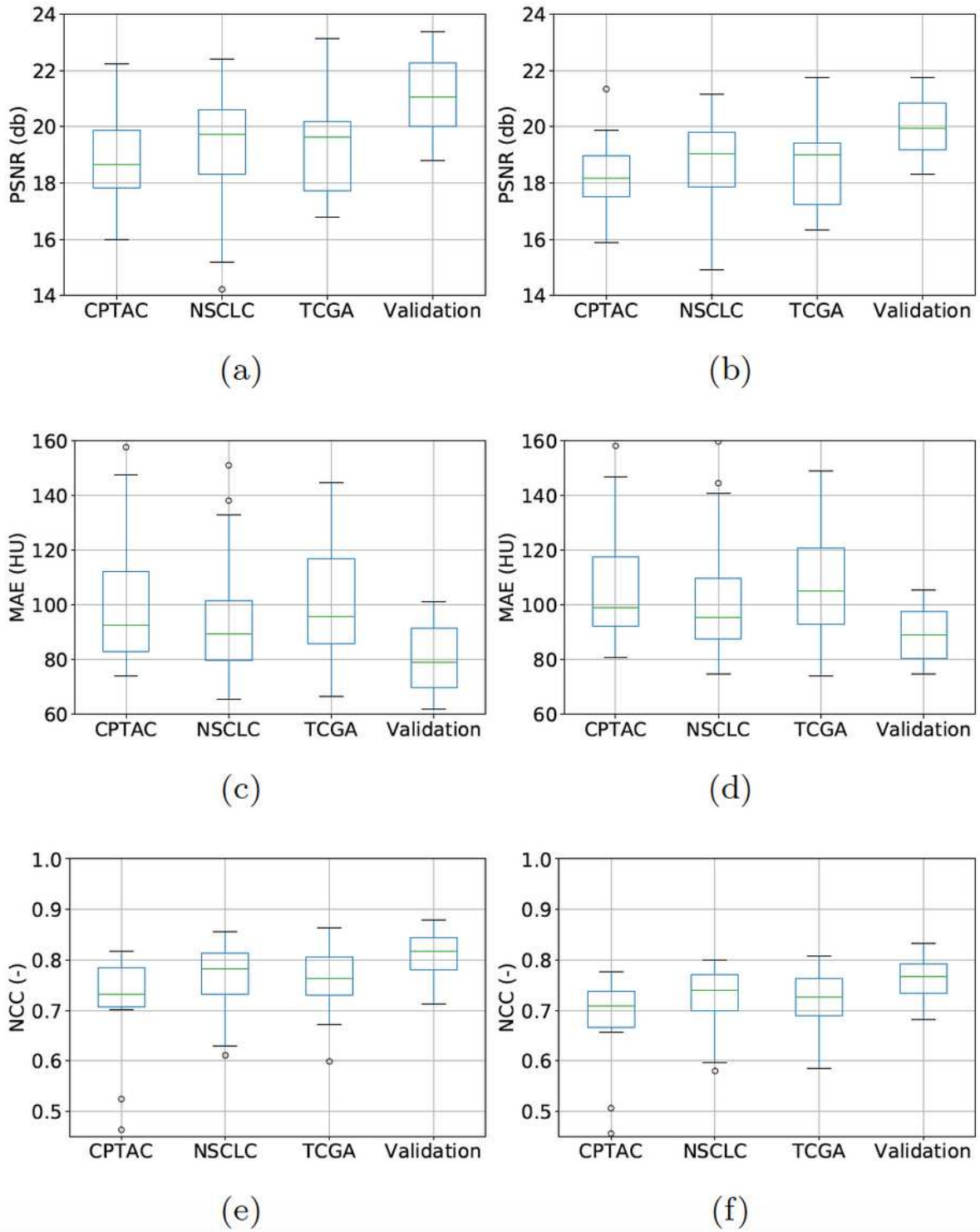
21. Zhu, J.-Y., Zhang, Z., Zhang, C., Wu, J., Torralba, A., Tenenbaum, J., Freeman, B.: Visual object networks: image generation with disentangled 3d representations. In: *Advances in Neural Information Processing Systems*, pp. 118–129 (2018)
22. Nguyen-Phuoc, T., Li, C., Theis, L., Richardt, C., Yang, Y.-L.: Hologan: Unsupervised learning of 3d representations from natural images. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 7588–7597 (2019)
23. Nuyts, J., Dupont, P., Stroobants, S., Beninck, R., Mortelmans, L., Suetens, P.: Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms. *IEEE transactions on medical imaging* **18**(5), 393–403 (1999)
24. Mihlin, A., Levin, C.S.: Gpu formulated mlem joint estimation of emission activity and photon attenuation in positron emission tomography. In: *2014 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pp. 1–2 (2014). IEEE
25. Brusaferrri, L., Bousse, A., Efthimiou, N., Emond, E., Atkinson, D., Ourselin, S., Hutton, B.F., Arridge, S., Thielemans, K.: Potential benefits of incorporating energy information when estimating attenuation from pet data. In: *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*, pp. 1–4 (2017). IEEE
26. Dong, X., Wang, T., Lei, Y., Higgins, K., Liu, T., Curran, W.J., Mao, H., Nye, J.A., Yang, X.: Synthetic ct generation from non-attenuation corrected pet images for whole-body pet imaging. *Physics in Medicine & Biology* **64**(21), 215016 (2019)
27. Dong, X., Lei, Y., Wang, T., Higgins, K., Liu, T., Curran, W.J., Mao, H., Nye, J.A., Yang, X.: Deep learning-based attenuation correction in the absence of structural information for whole-body pet imaging. *Physics in Medicine & Biology* (2019)

# Figures



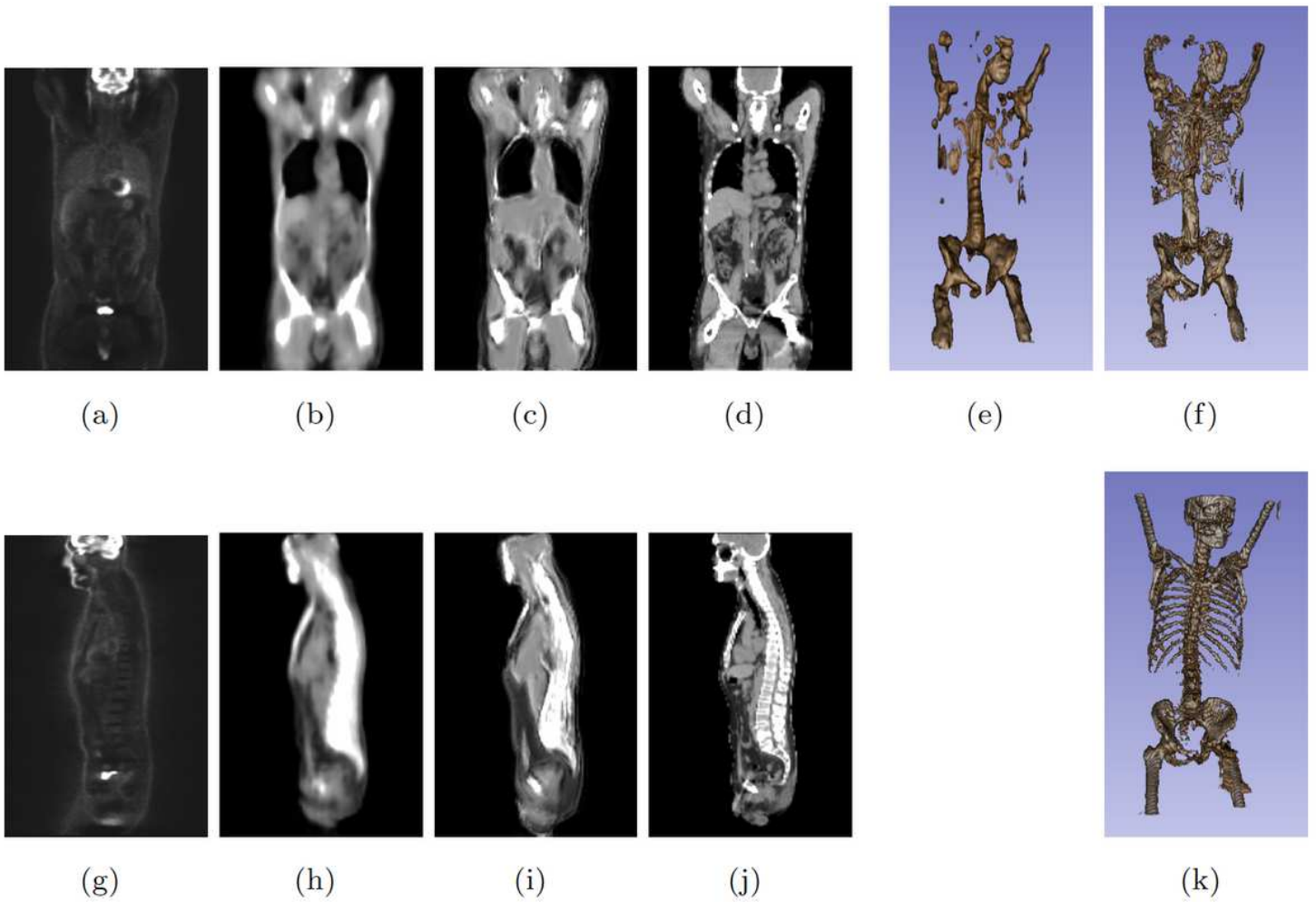
**Figure 1**

Generator network topology. Top branch is used for label segmentation (auxiliary task) and bottom branch for artificial CT generation. The bottom output is the synthetic CT refined by the GAN layers and built from the U-Net outputs.



**Figure 2**

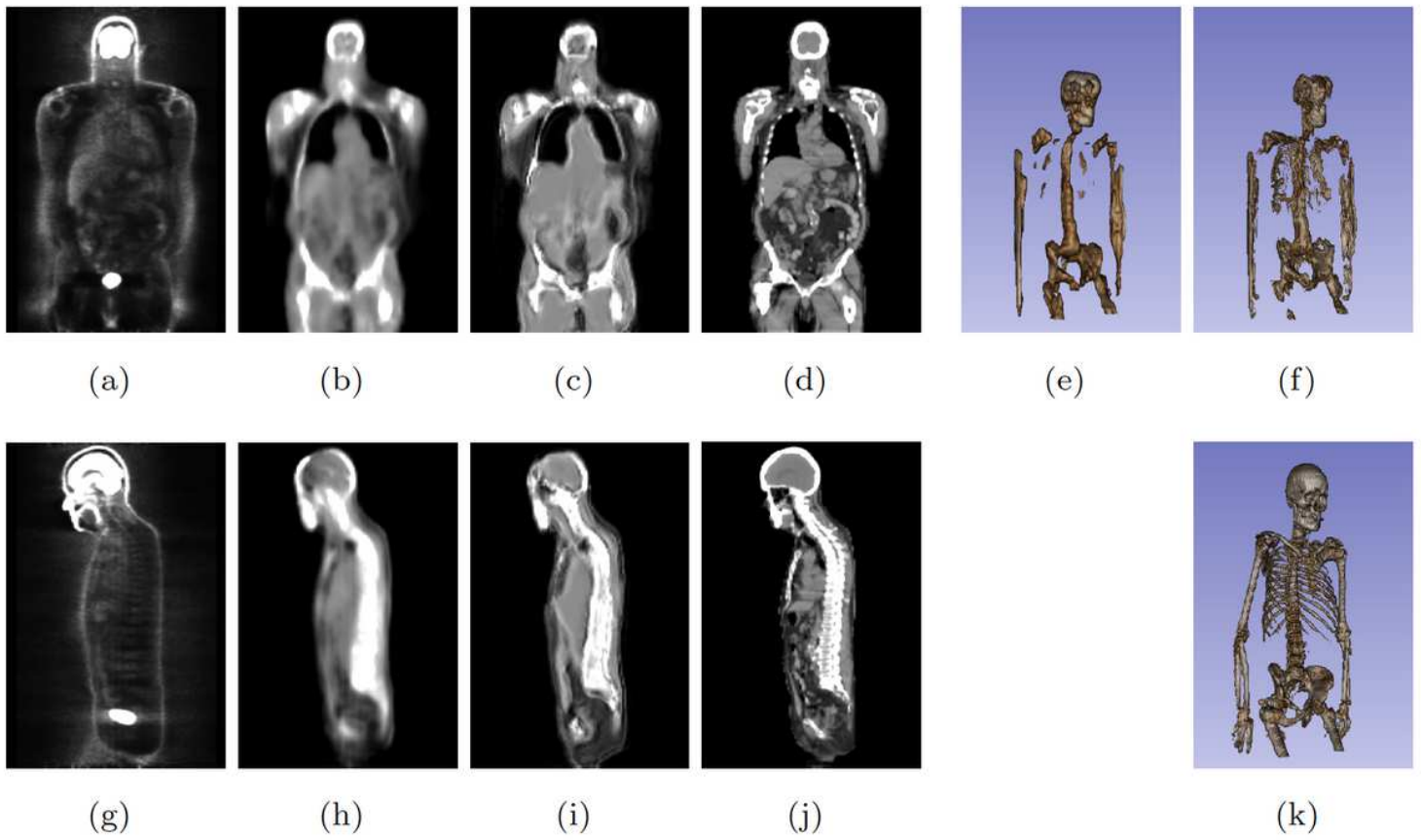
Boxplot of PSNR (a,b), MAE (c,d) and NCC (e,f) for each test data source and validation dataset. The figures (a,c,e) and (b,d,f) are corresponding to the 3D U-Net and to the network with adversarial refinement, respectively.



**Figure 3**

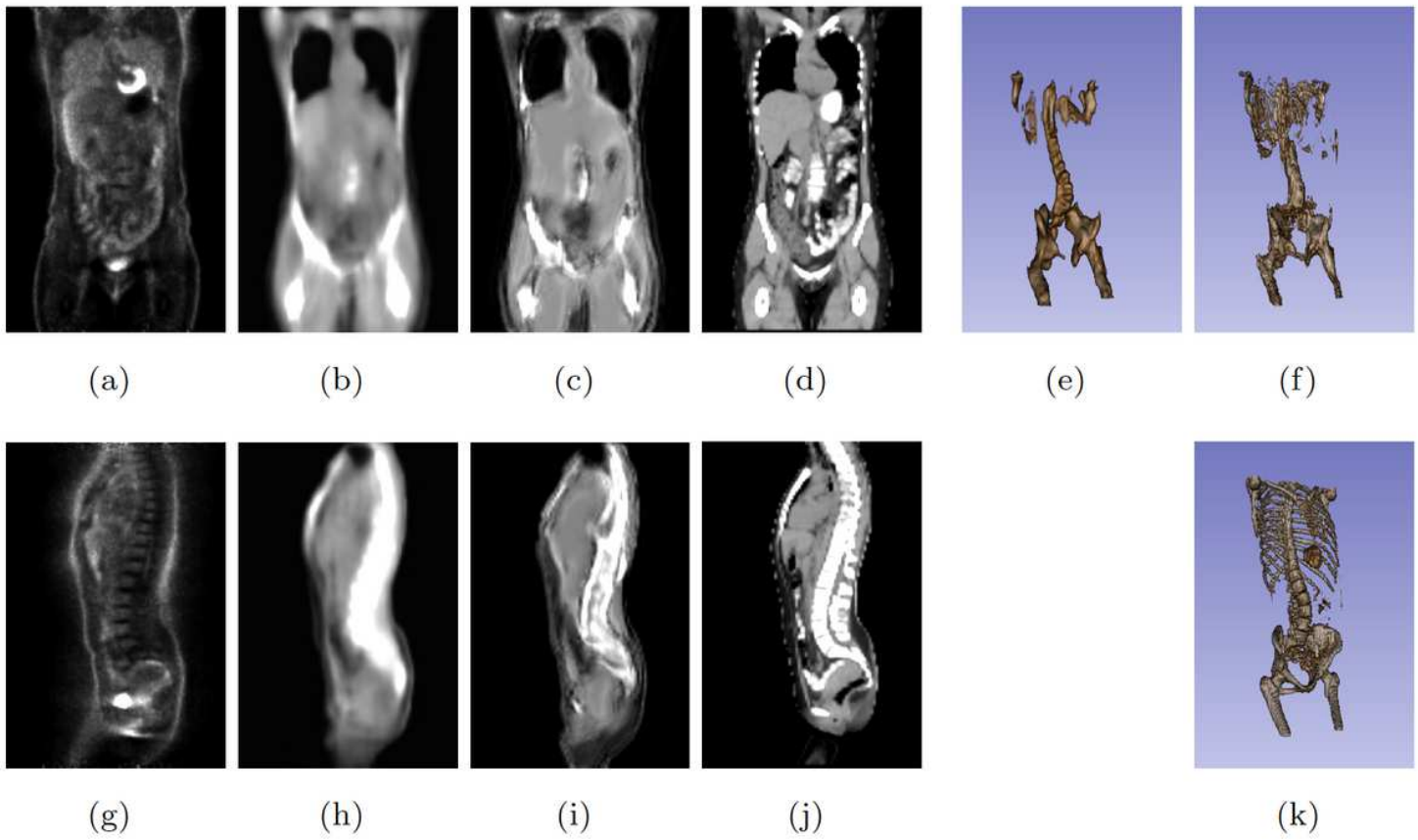
Input NAC-PET (a,g), 3D-U-Net sCT (b,e,h), 3D GAN sCT (c,f,i) and reference CT (d,j,k) coronal (a,b,c,d), sagittal (g,h,i,j) and 3D renders of bone tissue (e,f,k) of AMC-009 from NSCLC Radiogenomics. The improvement of the GAN can be seen in lower blurring between images (b, c) and (h, i). In the 3D render, the GAN network (f) is able to generate further details than the 3D U-Net (e), when compared to the ground truth (k).





**Figure 4**

Input NAC-PET (a,g), 3D-Unet sCT (b,e,h), 3D GAN sCT (c,f,i) and reference CT (d,j,k) coronal (a,b,c,d), sagittal (g,h,i,j) and 3D renders of bone tissue (e,f,k) of C3N-00957 from CPTAC-PDA dataset.



**Figure 5**

Input NAC-PET (a,g), 3D-Unet sCT (b,e,h), 3D GAN sCT (c,f,i) and reference CT (d,j,k) coronal (a,b,c,d), sagittal (g,h,i,j) and 3D renders of bone tissue (e,f,k) of TCGA-BB-7863 from TCGA-HNSC dataset.