

Different Deep Learning Training Strategies for Attenuation and Scatter Correction in PET

Florence M. Muller^{1,2}, Margaret E. Daube-Witherspoon¹, Michael J. Parma¹, Christian Vanhove², Stefaan Vandenberghe², Peter B. Noël¹, Joel S. Karp¹

I. INTRODUCTION

For an accurate quantification of the tracer distribution, attenuation correction (AC) and scatter correction (SC) are essential in PET. As the emitted 511 keV photons travel through the body (before being detected), they are likely to interact with the underlying tissue, resulting in both undetected coincidences (attenuated photons) and detected scattered coincidences. Most AC methods rely on CT scans to obtain AC factors which are derived from a μ -map. Single scatter simulation (SSS) is widely used to estimate the scatter fraction for each particular line-of-response. SSS requires the μ -map and emission data to calculate the scatter distribution using the Klein-Nishina formula.

Most commercial PET systems have an integrated CT component to not only provide an anatomical reference but also to derive a μ -map for AC and SC (ASC). However, there are some concerns: 1) the radiation burden resulting from PET-CT scans limits applications in pediatric and longitudinal imaging, 2) CT-based correction methods are challenging for stand-alone PET systems, and 3) image artefacts frequently originate from a spatial mismatch between PET and CT (due to respiratory as well as rigid body motion between PET and CT). For SC, in addition to being sensitive to errors in the μ -map, SSS does not model multiple-scattering, requires sinogram tail-fitting to scale the estimate, and relies on the (iterative) reconstruction of emission images that can be time-consuming [1-2].

While there are conventional methods for CT-less AC (with transmission sources, MR images, or algorithms for joint reconstruction of activity and attenuation maps with time-of-flight) and energy-based scatter estimation methods, deep learning (DL) shows promise for achieving more accurate and faster PET ASC [2]. The most common DL approach for CT-less AC predicts a CT-like image from a non-corrected (NC) PET that can then be used in the conventional reconstruction pipeline. DL-based SC methods either classify which events originated from scatter based on position and energy or predict the scatter profile from NC PET data. Other neural networks attempt to directly output a fully corrected (AC + SC) PET image from a NC PET input.

However, one challenge is to develop a single DL solution for both AC and SC that works for long axial field-of-view (AFOV) images, while being tracer-independent. This work investigates different image-based DL frameworks for AC and SC, starting from a NC PET image as input. We propose two training strategies (Fig 1) to fine-tune such neural networks for long AFOV PET and multiple tracer studies.

(i) We present a two-stage neural network that is separately trained for the tasks of AC and SC (Fig 1A), under the premise that AC is multiplicative and object-dependent, while SC is subtractive and uptake (distribution) dependent. This will be compared to a DL framework that predicts a transmission (Tr) image, followed by conventional reconstruction with the SSS scatter estimate, and a network that directly predicts an ASC PET, both from a NC PET input.

(ii) We propose a co-learning strategy using NC PET and radiographic projection images, so-called scout scans used to position patients, as inputs in a multi-branch neural network (Fig 1B). Building upon previous work on a DL model for 3D pseudo-CT reconstruction from 2D projections for specific organ regions [3], we extend this approach to long AFOV data. We investigate the potential of utilizing scout scans, rich in anatomical information, to constrain network training for PET AC and SC. The clinical advantage lies in operating current PET-CTs with only scout scans at significantly lower X-ray doses (by one order of magnitude).

II. MATERIALS AND METHODS

A. Datasets

35 research scans acquired on the 142 cm AFOV PennPET Explorer were collected: 19 ¹⁸F-fluorodeoxyglucose (¹⁸F-FDG), 7 ¹⁸F-labeled nitric oxide synthase (¹⁸F-NOS) and 9 ¹¹C-carfentanil (¹¹C-CFN) datasets. Each study included a NC PET image, a CT-rescaled Tr image, a PET image reconstructed only with CT-based AC (no SC), and a PET image reconstructed with CT-based AC and 3 iterations of TOF-enhanced SSS for SC. All NC PET images and all PET images with CT-based AC and ASC were reconstructed using list-mode TOF OSEM. 21 from those 35 datasets studies also included dual-view radiographic projection images (coronal and sagittal scouts). Note that for training purposes, datasets with minimal motion mismatch between PET and CT were selected. We plan to collect further data, including ¹⁸F-labeled fluoranthrace studies.

B. Deep Learning Frameworks

This study compares different DL training strategies for AC and SC in PET, and more specifically considers different input-target image pairs. The DL frameworks (Fig 1) use coronal and sagittal slices of NC PET images as input, either independently or combined with the corresponding coronal or sagittal scout image. The target training images vary based on the specific DL model,

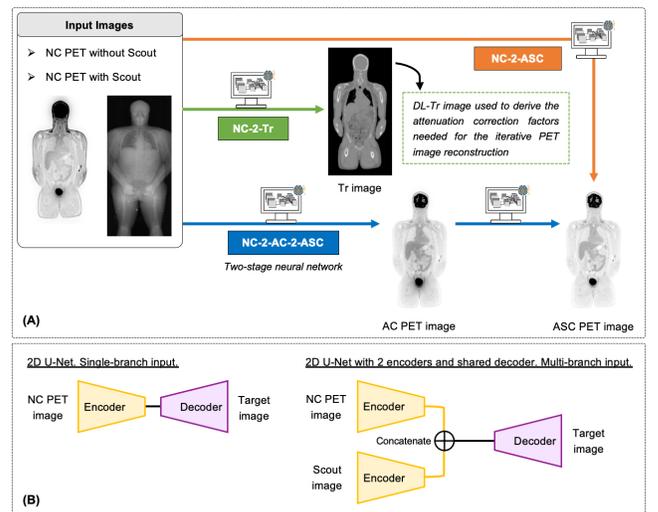


Fig 1. (A) Schematic of the investigated DL frameworks, with different input-target image pairs. The NC-2-Tr model needs further reconstruction steps and SSS, as opposed to the other DL models that directly predict the final PET image. (B) The DL models are either trained with only NC PET images (left), or with NC PET and scout images in a multi-branch neural network (right).

[1] Department of Radiology, University of Pennsylvania, Philadelphia, United States. [2] Faculty of Engineering and Architecture, Ghent University, Ghent, Belgium.

and include CT-derived Tr images, CT-based AC PET images, or CT-based ASC images. The **NC-2-Tr model** predicts a Tr image that is then utilized to derive the μ -map needed for image reconstruction (using list-mode TOF OSEM). The **NC-2-ASC model** directly predicts an ASC PET image from a NC PET image, tasking this single network to jointly predict AC and SC. The **NC-2-AC-2-ASC model** consists of a first network that predicts a PET image with only AC from a NC PET image (= attenuation prediction), followed by a second network that predicts a PET image with ASC from a PET image with only AC (= scatter prediction). So, the output of the first network serves as input to the second network.

C. Neural Network Architectures and Training settings

All DL models adopt a U-Net type structure. For the encoder, the feature maps are increased from 64 to 1024 through 2D convolutional layers (with batch normalization and parametric rectified linear unit activation) and max-pooling is used for down-sampling. The decoder uses 2D transposed convolutions for up-sampling. For the multi-branch Scout/NC(PET) model (Fig 1B), two separate encoders are implemented to extract modality-specific features and latent features are concatenated before the decoder. Both encoders and the shared decoder have the same network topology as the single-branch input models.

For training, an element-wise loss function, using the mean squared error (L2) or mean absolute error (L1), was combined with a feature-wise perceptual loss function derived from a pre-trained VGG16 network [4]. When predicting Tr images, L2 loss was used to improve sensitivity in reducing larger discrepancies

on bone regions of high pixel values. For the models predicting AC or ASC PET images, L1 loss was applied.

III. RESULTS AND DISCUSSION

Fig 2 compares the CT-based ASC PET images to those obtained from the different DL frameworks. For the ^{11}C -CFN example in Fig 2A, enhanced lung activity is seen for the NC-2-ASC and NC-2-AC-2-ASC models using only NC PET data as input. Quantitative comparisons in the liver and lung suggest that the co-learning strategy with a multi-branch Scout/NC(PET) input model significantly reduces the SUV_{mean} bias. The DL models that directly predict a PET image (in less than 1 min) eliminate the extra reconstruction steps with the SSS scatter estimate (which take a few hours) resulting in much faster AC and SC, but lead to higher biases (in the order of ± 1 SUV unit) than the NC-2-Tr model, as exemplified by the ^{18}F -FDG example in Fig 2B. The comparatively higher biases seen for the head region suggest that the DL training strategies are not yet optimized to capture the more refined brain structures. Recall that training used coronal and sagittal image slices, where the relative size of the brain region compared to the whole-body is small, which may lead to insufficient attention during training.

IV. CONCLUSIONS

This work studies different DL training strategies for AC and SC in PET, eliminating artefacts seen with CT-based AC and accelerating PET image reconstruction thanks to DL-based SC. The resultant reduction in radiation burden opens up applications in pediatric imaging, longitudinal studies, and the use of stand-alone (dedicated) PET systems.

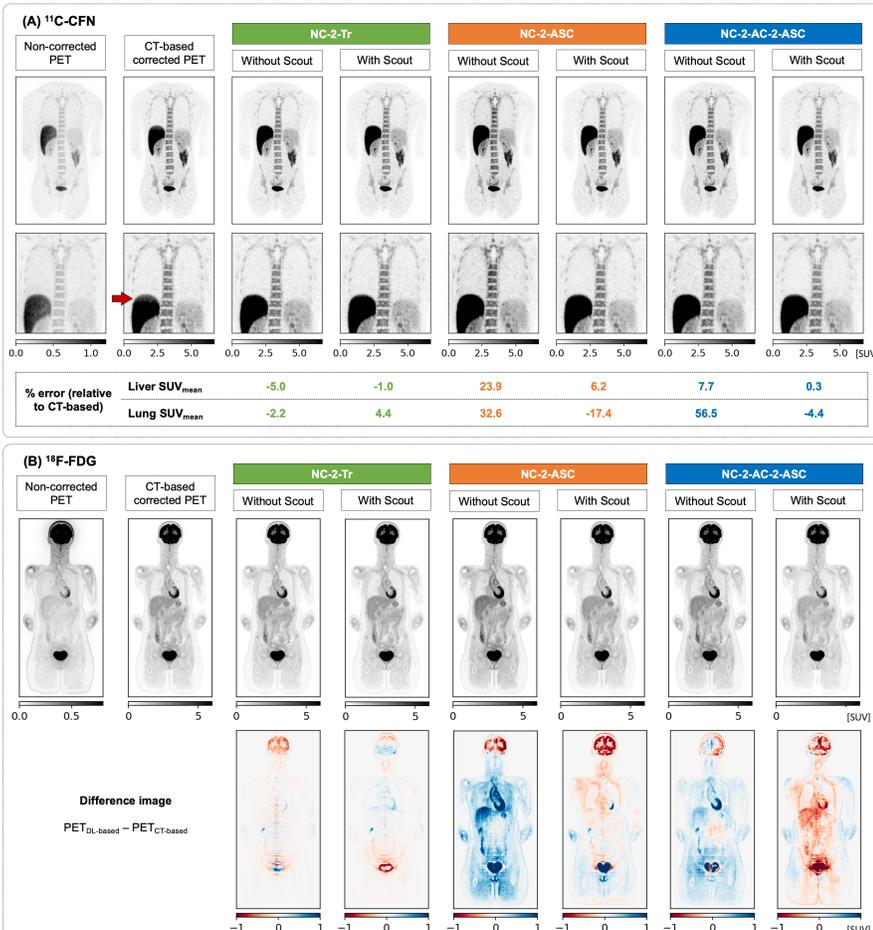


Fig 2. For ^{11}C -CFN (A) and ^{18}F -FDG (B), visual comparison between the NC PET, reference CT-based ASC PET and DL-based images obtained from the NC-2-Tr, NC-2-ASC and NC-2-AC-2-ASC models when trained only with NC PET images or in combination with scout images. In (A), the red arrow points to the ‘banana’ artefact at the liver dome in the CT-based ASC PET (due to respiratory motion causing a mismatch between PET and CT) that is eliminated in the DL-based images. In (A), liver and lung SUV_{mean} measures are reported as percent error relative to the CT-based PET image. In (B), difference images ($\text{PET}_{\text{DL-based}} - \text{PET}_{\text{CT-based}}$) are shown (note scale change) with some systematic differences on the order of ± 1 SUV unit.

Acknowledgments

FMM acknowledges the support of Fulbright, B.A.E.F and the Research Foundation Flanders (FWO, File number: 11P0E24N). Data courtesy: J. Dubroff, C. Wiers, A. Pantel, H. Wachtel, E. McDonald (UPenn).

References

- [1] Alessio AM, et al. (2004) PET/CT scanner instrumentation, challenges, and solutions. *Radiol Clin.* 42(6):1017-32.
- [2] McMillan AB, Bradshaw TJ. (2021) AI-based data corrections for attenuation and scatter in PET and SPECT. *PET Clin.* 16(4):543-52.
- [3] Shapira N, et al. (2023) Convolutional Encoder-Decoder Networks for Volumetric Computed Tomography Surveys from Single- and Dual-View Topograms. *Res Sq [Preprint]*.
- [4] Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. *arXiv [Preprint]*.