A novel convolutional neural network for predicting full dose from low dose PET scans

Amirhossein Sanaat, Hossein Arabi, Habib Zaidi *IEEE Fellow*

Abstract— The use of radiolabeled tracers in PET imaging raises concerns owing to potential risks from radiation exposure. Therefore, to reduce this potential risk in diagnostic PET imaging, efforts have been made to decrease the amount of radiotracer administered to the patient. However, decreasing the injected activity reduces the signal-to-noise Ratio (SNR) and deteriorates image quality, thus adversely impacting clinical diagnosis. Previously proposed techniques are complicated and slow, yet they yield satisfactory results at significantly low dose. In this work, we propose a deep learning algorithm to reconstruct full-dose (FD) from low-dose (LD) PET images using a fully convolutional encoder-decoder deep neural network model. The goal is to train a model to learn to reconstruct from images with only 5% of the counts to produce images corresponding to 100% of the dose. Brain PET/CT images of 140 patients acquired on the Siemens Biograph mCT with a standard injected activity of ^{18}F-FDG (205 ± 10 MBq). Images were acquired for about 20 min. The sinograms of each scan were used to produce a low-dose sinogram by randomly selecting only 1/20th of the counts. To avoid over fitting, data augmentation was used. A modified 3D U-Net, was developed to predict standard-dose sinogram (PSS) from their corresponding LD sinogram. Detailed quantitative and qualitative comparison demonstrated the proposed method can generate artefact-free diagnostic quality images that preserve internal structures without noise amplification. The structural similarity index (SSIM) and peak signal to noise ratio (PSNR) were used as quantitative metrics for assessment. For instance, the PSNR and SSIM in selected slices were 37.30±0.71 and 0.97±0.02, respectively. The proposed algorithm operates in the projection space and is capable of producing diagnostic quality images with only 5% of the standard injected activity.

I. Introduction

DIAGNOSTIC PET imaging is commonly used in a number of clinical indications, particularly in oncology for the assessment of malignant disease, staging and restaging, treatment monitoring and therapy planning. Patients are usually injected with a full dose (FD) of radiotracer prior to scanning. However, there is always a desire to reduce the injected dose to decrease the radiation exposure to patients.

However, reducing the injected activity below a certain threshold would require longer acquisition times, thus impacting throughput and increasing imaging costs. Special logistics might also need to be developed when fast-decaying tracers are used [1]. However, lower injected activities will undesirably affect PET image quality, resulting in lower signal-to-noise ratio (SNR). As such, higher doses are generally administered in the clinic. A limited number of studies were carried out to address this problem, specifically focusing on denoising low dose (LD) medical images using machine learning techniques [2-4]. The connection between LD and FD images is learned by the machine. Previous work used a regular convolutional neural network (CNN) to estimate FD PET images from images taken at 1/4th of a FD [2]. Over-smoothing is one of the side effects of CNNs, which can be improved by using a different loss function than the mean squared error (MSE) during training [5] or convolutional generative adversarial networks (GANs) [3]. The kernel method is another approach that showed clear improvement in the retention of details and reduction of noise for reduced dose reconstructions [5].

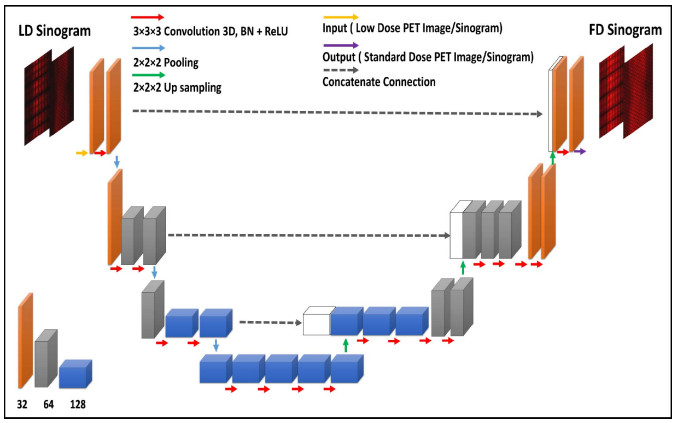

Fig 1. A schematic flowchart of the encoder-decoder CNN used in this work. PET sinograms are the input and output of the network.

While the above referenced previous works were relatively effective in decreasing the noise in low dose medical images, the resulting images were of poor quality at the edges and did not preserve well the internal structures.

In this work, we propose a residual 3D U-Net to estimate FD PET sinograms from 5% of the FD PET sinograms that preserve more details by specifically accounting for them in the loss function during training. To the best of our knowledge, this is the first time that a deep learning algorithm operating in the projection domain is proposed to predict FD sinograms from LD PET sinograms.

Manuscript received January 02, 2020. This work was supported by the Swiss National Science Foundation under grant SNFN 320030-176052 and the Swiss Cancer Research Foundation under Grant KFS-3855-02-2016.

A. Sanaat, H. Arabi and H. Zaidi are with the Division of Nuclear Medicine & Molecular Imaging, Geneva University Hospital, CH-1211, Geneva, Switzerland (e-mail: Amirhossein.sanaat@etu.unige.ch; hossein.arabi@unige.ch; habib.zaidi@hcuge.ch)

978-1-7281-4164-0/19/$31.00 ©2019 IEEE

II. Materials & Methods

A. Clinical brain PET/CT studies

One hundred forty ^{18}F-FDG brain PET/CT clinical studies were acquired for about 20 min on the Siemens Biograph mCT following injection of a standard dose of 205 ± 10 MBq. PET data were acquired in listmode format to enable the generation of synthesized LD data by randomly selecting 5% of the events, spread uniformly over the entire acquisition period. The original and generated sinograms were then reconstructed using Siemens e7 tools. To avoid over-fitting, three types of data augmentation methods were adopted during the training process to simulate a larger dataset.

Table 1. Demographic information of patients included in this work.

	Training	Test	Validation
Number	100	20	20
Male/Female	45/55	11/9	8/12
Age (Mean ± SD)	73±8	68±18	73±4.5
Weight(Mean ± SD)	70±13	67±12	71±11
Indication/Diagnosis	Cognitive symptoms of possible neurodegenerative a etiology		

B. Deep neural network implementation

Figure 1 displays the architecture of the purposed 3D U-Net consisting of an encoder-decoder modules. In the encoder module, two 3D convolutions were used for each layer [6] followed by a 3D max pooling with a stride size of 2 and a rectified linear unit (ReLu) as activation function. In the decoder part, 3D up-sampling was used for each layer with stride and activation function similar to the encoder part. The size of all convolutional kernels were set to 3 × 3 × 3 voxels in each convolutional layer. To address the gradient vanishing problem, shortcut connections between the outputs of each layer in the encoder network and the corresponding layer in the decoder network was used. For minimizing the loss function, Adam optimizer with learning rate of 0.001 was implemented. The algorithm was employed on a graphics processing unit with NVIDIA 2080Ti GPU and 8 GB of memory. The proposed network was implemented on the Tensor Flow platform. We first applied a low-pass filter to the input low-dose sinograms to aid the training process through noise reduction without losing important details. Then by using a loss function that combines specific features, namely the gradient and total variation, with the MSE and an adversarial network to ensure the estimated full dose images preserve essential details. The OSEM algorithm (5 iterations, 21 subsets) was used to reconstruct images from the predicted sinograms.

C. Performance evaluation

We evaluated qualitatively the generated images and used standard quantitative metrics for the assessment of image quality, including the root mean square error (RMSE), peak signal-to-noise ratio (PSNR), and structural similarity index (SSIM) between the LD and FD images serving as reference. The scatter and linear regression between the tracer uptakes of the LD vs. FD was also calculated.

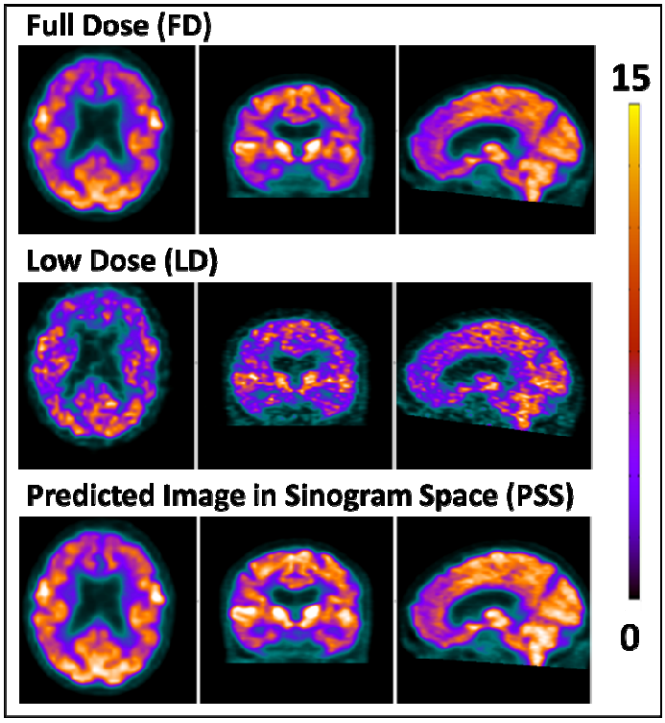

Fig. 2 Representative brain PET images of a reconstruction showing the reference full dose (top row), image with 5% of the full dose (middle row), and the estimated full dose image using the proposed algorithm (middle row).

III. Results & Discussion

Figure 2 shows representative slices of the generated images comparing low dose images generated using PSS and reference FD images after reconstruction, the enhancement in image quality is striking. The images predicted by the PSS deep learning technique are almost similar to those produced by FD scanning.

The quantitative results in terms of RMSE, SSIM, and PSNR are summarized in Fig 3. The high values of SSIM and PSNR and the low RMSE for the PSS images prove that the image quality produced by the model is comparable to reference FD images. Fig 4 illustrates linear regression plots depicting the correlation between the standardized uptake value (SUV) of the LD vs FD and PSS vs FD.

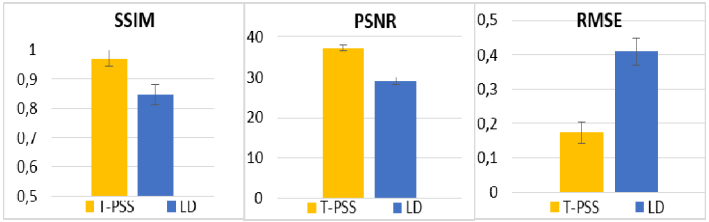

Fig. 3 Image quality metrics compare images from low-dose PET (LD) and the standard dose PET images in sinogram space (PSS) for test dataset. For the three metrics, comparison is to the reference full-dose (FD) PET images.

The scatter and linear regression plots presented high correlation between PSS and FD (RMSE = 0.28, R^2 = 0.99). A lower correlation coefficient (R^2 = 0.97) along with underestimation of tracer uptake was observed in LD PET

images reflected by the relatively higher RMSE (0.42) and the slope of regression line.

Unlike previous methods, our designed deep learning approach is fed by sinograms instead of reconstructed images. As such, the training was performed in the projection domain prior to image reconstruction. By using sinograms instead of reconstructed images, the model does not focus on learning

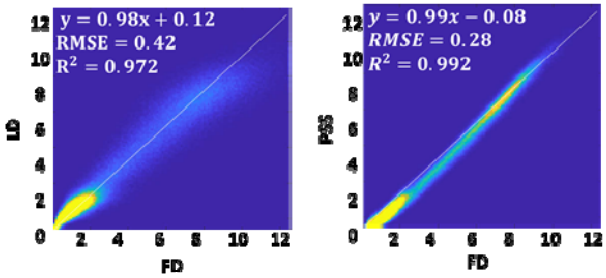

Fig. 4 Joint histograms analysis of the low dose PET images (left) and predicted FD images (right) versus full dose PET images (reference).

plenty of features present in the images tainted by the reconstruction. Instead, it focuses on essential features in the sinogram, which project meaningful information after reconstruction.

REFERENCES

[1] J. Xu, E. Gong, J. Pauly, and G. Zaharchuk, "200x low-dose pet reconstruction using deep learning," *arXiv preprint arXiv:1712.04119,* 2017.

[2] L. Xiang *et al.*, "Deep auto-context convolutional neural networks for standard-dose PET image estimation from low-dose PET/MRI," *Neurocomputing,* vol. 267, pp. 406-416, 2017.

[3] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose CT," *IEEE transactions on medical imaging,* vol. 36, no. 12, pp. 2536-2545, 2017.

[4] K. T. Chen *et al.*, "Ultra-low-dose (18)F-Florbetaben amyloid PET imaging using deep learning with multi-contrast MRI inputs.," *Radiology,* vol. 290, no. 3, pp. 649-656, 2019.

[5] Q. Yang, P. Yan, M. K. Kalra, and G. Wang, "CT image denoising with perceptive deep neural networks," *arXiv preprint arXiv:1702.07019,* 2017.

[6] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3D U-Net: learning dense volumetric segmentation from sparse annotation," in *International conference on medical image computing and computer-assisted intervention*, 2016, pp. 424-432: Springer.