Whole-body PET Image Synthesis from Low-Dose Images Using Cycle-consistent Generative Adversarial Networks

Amirhossein Sanaat, Isaac Shiri, Hossein Arabi, Ismini Mainta, René Nkoulou, and Habib Zaidi, *Fellow, IEEE*

2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC) | 978-1-7281-7693-2/20/$31.00 ©2020 IEEE | DOI: 10.1109/NSS/MIC42677.2020.9507947

***Abstract*–This work sets out to investigate the performance of full-dose (FD) PET prediction from fast or low-dose (LD) whole-body (WB) PET scans using convolutional neural networks. One hundred patients who underwent WB PET/CT examinations were retrospectively used to develop LD to FD PET conversion models. The patients underwent two separate WB examinations lasting ~27 min (regular scan) and ~3 min (fast or LD scan) acquisition times. The fast (3 min) WB PET examinations are equivalent to 1/8th of the full-dose PET scan. A residual neural network (ResNet) and a modified cycle-consistent generative adversarial network (CycleGAN) architecture were employed to model the LD to FD PET conversion. The quality of synthetic PET images produced by ResNet and CycleGAN models, referred to as RNET and CGAN, respectively, were evaluated by two nuclear medicine physicians using a five-point scoring. Quantitative metrics, including structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), mean square error (MSE) and standardized uptake value (SUV) bias were calculated within the left/right lung, brain, liver, and one hundred hot spots (malignant lesions) in PET images predicted using CGAN and RNET models. The physicians assigned scores of 3.88 and 4.92 (out of 5) to the CGAN-predicted FD PET images for the neck & trunk and brain regions, respectively. Considering the PSNR and SSIM metrics, the CGAN model exhibited superior performance with PSNR=39.08±3.56 and SSIM=0.97±0.02 compared to the RNET model with PSNR=34.91±1.50 and SSIM=0.93±0.04. Overall, the CGAN model outperformed the RNET model exhibiting lower SUV bias and higher image quality in the predicted FD PET images.**

I. Introduction

DIAGNOSTIC quality and quantitatively accurate whole-body PET images could be synthesized using deep learning algorithms from images acquired with a noticeably reduced radiotracer injected dose and/or scan duration [1-5]. There is a tendency for dose reduction or fast image acquisition in WB PET examination. However, increased noise levels in the resulting images would hinder the clinical value of the low-dose PET images. The conventional post-reconstruction denoising approaches enable effective noise reduction in the PET images [6]. However, these approaches may cause significant signal loss, quantitative bias, and image artifact, which limit their application in low-dose PET imaging . In this regard, deep learning-based noise suppression approaches and/or FD PET estimation from the LD images have exhibited superior performance to the conventional denoising approaches [7-9]. In this work, we set out to compare two state-of-the-art deep learning architectures, namely the residual network (ResNet) and cycle-consistent generative adversarial network (CycleGAN) models, for the task of FD ^{18}F-FDG WB PET prediction from the LD counterparts. To this end, a clinical database of WB PET/CT examinations with conventional/standard and fast acquisition protocols was created. One hundred patients underwent two separate PET scans with normal and short scan durations mimicking FD and LD image acquisition with a ratio of 1/8th. The predicted WB FD PET images by the CGAN and RNET models were quantitatively (using established image analysis metrics) and qualitatively (by the nuclear medicine specialists) assessed.

Manuscript was submitted December 20, 2020. This work was supported by the Swiss National Science foundation under grant SNFN 320030_176052 and the Private Foundation of Geneva University Hospitals under Grant RC-06-01.

A. Sanaat, I. Shiri, H. Arabi, I. Mainta, R. Nkoulou and H. Zaidi are with the Division of Nuclear Medicine & Molecular Imaging, Geneva University Hospital, Geneva, Switzerland (e-mail: Amirhossein.sanaat@unige.ch, Isaac.shirilord@unige.ch, hossein.arabi@unige.ch, rene.nkoulou@hcuge.ch, IsminiCharis.Mainta@hcuge.ch, habib.zaidi@hcuge.ch).

R. Nkoulou and H. Zaidi are with the Geneva University Neurocenter, Geneva University, Geneva, Switzerland.

II. Materials and Methods

One hundred patients who underwent WB ^{18}F-FDG PET/CT scans on a Biograph mCT PET/CT scanner. In a single imaging session, a WB PET LD and/or fast-scan was acquired 60 min post-injection of 240±50 MBq of ^{18}F-FDG. Continuous bed motion protocol was selected for PET acquisitions with a speed of 5 mm/sec. Shortly afterward, a standard/conventional PET scan (regarded as FD) was acquired with a bed speed of 0.7 mm/sec. Both fast and standard WB PET images were reconstructed with ordinary Poisson ordered subsets-expectation maximization (OP-OSEM) algorithm using 2 iterations and 21 subsets. Two state-of-the-art convolutional neural network architectures, namely CycleGAN and ResNet [10], were employed to create LD to FD PET conversion models. The training, testing, and validation of the developed models were performed on 60, 15, and 10 patients, respectively. 15 subjects were excluded due to the presence of gross image artifacts.

Two experienced nuclear medicine specialists evaluated the PET images synthesized by the two deep learning models. The PET images were randomly enumerated and anonymized prior to the evaluation.

978-1-7281-7693-2/20/$31.00 ©2020 IEEE

Peak signal-to-noise ratio (PSNR) mean squared error (MSE), and structural similarity index metrics (SSIM) were calculated to evaluate the quantitative accuracy of the predicted FD from LD PET images.

III. Result

Regarding the quality of the predicted PET images in Fig. 1, the CGAN model led to superior recovery of signals and underlying ^{18}F-FDG uptake patterns within the different anatomical regions compared to the RNET model. The inter- and intra-rates' agreement was assessed via Weighted Cohen's Kappa and Krippendorff alpha tests, wherein a high Krippendorff alpha score (>8 for the entire comparisons) was obtained for ^{18}F-FDG uptake patterns and the overall image quality. For the failed/accepted task, the Kappa for the entire cases was more than 0.7, except for the LD images with Kappa of 0.52.

The LD images received poor mean quality score of 2.6 and the largest failed cases (56%) within the neck and trunk regions. However, good quality score (4.2) and zero failed cases were given to these images within the brain region (Fig. 2). On the other hand, the synthesized PET images by the CGAN model received a good mean image quality score of 3.88 and a high percentage of the accepted cases (86%) within the neck and trunk regions. For the brain region, the CGAN model achieved an excellent mean image quality score of 4.92 and no failed cases, outperforming the RNET model with mediocre results.

PET images predicted by CGAN showed the highest PSNR, SSIM, MSE, better noise properties, and higher quantitative accuracy with statistically significant differences with respect to RNET (Table 1).

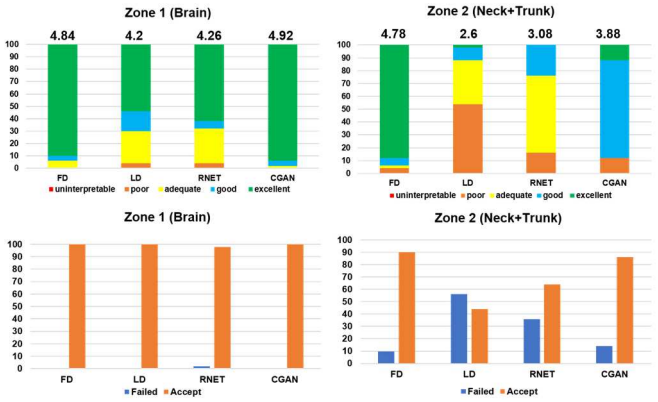

Fig. 2. First row: assessment of the clinical quality of the LD, FD, RNET and CGAN PET images by the two nuclear medicine specialists. The corresponding mean scores are indicated on the top of each bar. In this scoring scheme, #1 represents uninterpretable, #2 poor, #3 adequate, #4 good, #5 excellent. Second row: the bar plots present the percentage of accepted/failed cases in terms of lesion detectability.

The joint histogram analysis of the FD PET images predicted by the RNET and CGAN models are presented in Fig. 3. The voxel-wise scatter plots of the activity concentration as well as the linear regression analysis between predicted FD and reference standard PET images demonstrated the superior performance of the CGAN model with RMSE = 0.18 (SUV) and R^2 = 0.98 compared to RNET model with RMSE = 0.32 (SUV) and R^2 = 0.92. A mean RMSE of 0.51 (SUV) was observed in the original LD PET images.

IV. Discussion

The CGAN model exhibited promising performance to predict FD PET images from the corresponding LD ones with excellent signal recovery and low SUV bias. The CGAN model imposes certain constraints to the generator component

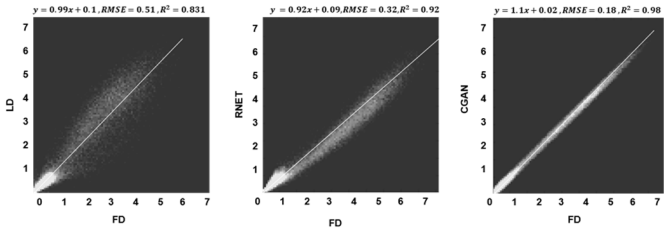

Fig. 3. Voxel-wise SUV correlation and joint histogram analysis of the LD (left), and synthesized PETs by the RNET model (middle), and CGAN model (right) versus the FD/standard PET images.

of the model through establishing an inverse transformation circle from the discriminator component. This architecture benefits from an effective optimization process to offer/find optimal solutions for the regression problem. This study demonstrated the effectiveness of the CycleGAN structure (compared to the powerful ResNET model) for the task of noise suppression in PET images.

The clinical assessment of the PET image quality carried out by the two nuclear medicine specialists, revealed the acceptable quality of the synthesized PET images by the CGAN model with clinically tolerable errors. MSEs of 0.15 ±0.09, 0.12 ± 0.10, and 0.03 ± 0.04 SUV were observed in the original LD and the synthesized PET images by the RNET and CGAN models (p-value<0.05), respectively. Regarding the SSIM metric, the CGAN model showed improved image quality (0.98±0.08) compared to the RNET model (0.94±0.10) and the original LD PET images (0.89±0.11). It should be mentioned that the quantitative metrics measured on the

TABLE I
COMPARISON OF THE QUANTITATIVE METRICS OBTAINED FROM EVALUATION OF THE LD PET IMAGES, AND THE PREDICTED FD PET IMAGES BY THE RNET AND CGAN MODELS ON THE EXTERNAL TEST AND VALIDATION DATASETS.

Validation dataset	MSE	SSIM	PSNR
CGAN	0.03 ± 0.04	0.98±0.08	41.08±3.90
RNET	0.12 ± 0.10	0.94±0.10	35.41±5.56
LD	0.15 ± 0.09	0.89±0.11	31.21±3.08
Test dataset	**MSE**	**SSIM**	**PSNR**
CGAN	0.03 ± 0.07	0.97±0.02	39.08±3.56
RNET	0.13 ± 0.10	0.93±0.04	34.91±1.50
LD	0.17 ± 0.04	0.9±0.03	29.21±2.43

synthesized PET images should be relatively considered/compared with respect to those obtained from the original LD images. This study demonstrated that high-quality WB ^{18}F-FDG PET images could be predicted from the fast/low-dose PETscans, provided a powerful deep learning architecture, such as the CycleGAN, is properly

optimized/trained to be able to recover the underlying radiotracer uptake pattern.

REFERENCES

[1] A. Sanaat, H. Arabi, I. Mainta, V. Garibotto, and H. Zaidi, "Projection-space implementation of deep learning-guided low-dose brain PET imaging improves performance over implementation in image-space," *J Nucl Med,* Jan 10 2020.

[2] Y. Lei *et al.*, "Whole-body PET estimation from low count statistics using cycle-consistent generative adversarial networks," *Physics in Medicine & Biology,* vol. 64, no. 21, p. 215017, 2019.

[3] A. Sanaat, I. Shiri, H. Arabi, I. Mainta, R. Nkoulou, and H. Zaidi, "Deep Learning-assisted Ultra-Fast/Low-Dose Whole-Body PET/CT Imaging," *Eur J Nucl Med Mol Imaging.,* In press 2020.

[4] A. Sanaat and H. Zaidi, "Depth of Interaction Estimation in a Preclinical PET Scanner Equipped with Monolithic Crystals Coupled to SiPMs Using a Deep Neural Network," *Applied Sciences,* vol. 10, no. 14, p. 4753, 2020.

[5] H. Arabi, A. AkhavanAllaf, A. Sanaat, I. Shiri, and H. Zaidi, "The promise of artificial intelligence and deep learning in PET and SPECT imaging," *Phys Med,* vol. 83, pp. 122-137, Mar 22 2021.

[6] H. Arabi and H. Zaidi, "Non-local mean denoising using multiple PET reconstructions," (in eng), *Ann Nucl Med,* Nov 26 2020.

[7] L. Zhou, J. D. Schaefferkoetter, I. W. K. Tham, G. Huang, and J. Yan, "Supervised learning with cyclegan for low-dose FDG PET image denoising," (in eng), *Med Image Anal,* vol. 65, p. 101770, Oct 2020.

[8] A. Sanaat, H. Arabi, M. Reza Ay, and H. Zaidi, "Novel preclinical PET geometrical concept using a monolithic scintillator crystal offering concurrent enhancement in spatial resolution and detection sensitivity: a simulation study," *Phys Med Biol,* vol. 65, no. 4, p. 045013, Feb 13 2020.

[9] A. Sanaat, A. Ashrafi-Belgabad, and H. Zaidi, "Polaroid-PET: a PET scanner with detectors fitted with Polaroid for filtering unpolarized optical photons-a Monte Carlo simulation study," *Phys Med Biol,* vol. 65, no. 23, p. 235044, Dec 2 2020.

[10] W. Li, G. Wang, L. Fidon, S. Ourselin, M. J. Cardoso, and T. Vercauteren, "On the compactness, efficiency, and representation of 3D convolutional networks: brain parcellation as a pretext task," in *International conference on information processing in medical imaging*, 2017, pp. 348-360: Springer.