Does Prior Knowledge in the Form of Multiple Low-dose PET Images (at Different Dose Levels) Improve Standard-dose PET Prediction?

Behnoush Sanaei, Reza Faghihi, Hossein Arabi, and Habib Zaidi, Fellow, IEEE

Abstract—To acquire a high-quality PET image, a standard dose of radioactive tracer is injected into the patient which may pose a high risk of radiation exposure damage. On the other hand, reducing the injected dose increases the statistical noise within the PET images. To improve the image quality of low-dose PET (L-PET) images, deep learning methods have been introduced to denoise the L-PET images, wherein the relationship between the L-PET images and the standard-dose PET (S-PET) images is learned by the model to predict the S-PET images from their low-dose counterparts. The existing deep learning-based approaches solely focus on a single level of L-PET imaging to predict the S-PET images. In this work, we investigate the benefits of exploiting multiple PET images at lower dose levels (in addition to the target low-dose level) as prior knowledge to predict the S-PET images. To this end, a high-resolution residual deep learning network was employed to develop two S-PET prediction models. First, the network was trained using a single input channel for 8% L-PET images. In the second model, multiple L-PET images (6% and 4%, in addition to 8% L-PET) were considered as inputs to the network. The performance of the two models was evaluated using peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), root mean square error (RMSE) of standard uptake value (SUV) within the entire head region. Moreover, mean SUV bias (SUV\textsubscript{bias}) was calculated for the malignant lesions. The quantitative analysis of 20 patients in the external validation dataset demonstrated the superior performance of the multi-input model. The RMSE within the entire head region reduced from 0.12±0.04 in 8% L-PET, to 0.09±0.03 and 0.06±0.02 for the single- and multi-input models, respectively. Moreover, the SUV\textsubscript{bias} reduced from -4.18±1.14% in the single-input model to -1.44±0.56% in the multi-input model. This study demonstrated the benefits of using multiple L-PET images to estimate the S-PET images.

I. INTRODUCTION

POSITRON emission tomography (PET) images are used in many clinical applications such as cancer diagnosis, tumor detection, and evaluating lesion malignancy, disease stage, and treatment monitoring [1, 2]. Although a high dose of radioactive tracer is necessary to obtain high-quality PET images, a standard dose of radioactive tracer is injected into the patient prior to scanning [3-5]. The PET system detects pairs of gamma rays emitted indirectly from the radioactive quality PET images, this inevitably poses a high risk of radiation exposure damage. On the other hand, reducing the injection dose has an inverse effect on the image quality by increasing the noise in the PET images that limits their clinical value [6]. To address this issue, deep learning methods have been introduced to denoise the low-dose PET images in which the relationship between the low-dose images (at a certain dose-level) and the standard-dose PET (S-PET) images is learned by the model to predict the standard-dose images from their low-dose counterparts [7-9]. These methods mainly focused on a single specific dose level of L-PET imaging to predict the S-PET images. In other words, the input of the deep learning models is only from a single low-dose PET imaging [10, 11]. To the best of our knowledge, no study so far has utilized multiple levels of low-dose PET imaging as input to the deep learning models as prior knowledge. For instance, if the goal is 8% low-dose PET imaging, lower-dose PET images (for instance 4%, 6% low-dose PET images which could be easily reconstructed from the 8% raw data) could be exploited as prior knowledge.

In this work, we investigate the benefits of utilizing additional information/prior knowledge from multiple L-PET images in a deep learning model to predict S-PET images.

II. MATERIAL AND METHODS

Data acquisition

This study was conducted on 120 \(^{18}\text{F}\)-FDG PET/CT brain scans from patients with head and neck malignant lesions (48 males and 52 females, 71 ± 9 yrs, mean age ± standard deviation (SD)) (100 subjects for training and 20 subjects for evaluation). The PET scans were performed for an acquisition time of 20 min after the injection of 210 ± 8 MBq of \(^{18}\text{F}\)-FDG. The PET/CT acquisitions were performed on a Biograph-6 scanner (Siemens Healthcare, Erlangen, Germany) about 40 minutes after the injection. The PET raw data, recorded in list-mode format, was randomly sampled to contain 8%, 6%, and 4% of the total counts in the standard-dose PET imaging (the 6% and 4% L-PET images were reconstructed from the 8% raw data). In this regard, we reconstructed the low-dose versions of the standard PET images based on the abovementioned percentages utilizing ordered subsets-expectation maximization (OSEM) algorithm (4 iterations, 18 subsets).
**Deep neural network structure**

NiftyNet, an open-source convolutional neural networks (CNNs) platform built upon Tensorflow module in Python environment was adopted to predict S-PET images from L-PET images [12-14]. The full-dose PET estimation was implemented using a high-resolution Resnet model. This network consists of twenty $3 \times 3 \times 3$ convolution layers. The first seven layers with 16 kernels are designed to extract low-level features of the images such as corners and edges. In order to capture medium and high-level features, the 12 remaining layers with 32 and 64 kernels are employed [14, 15]. Two independent deep learning-based models were developed to investigate the benefits of prior knowledge in the form of multiple low-dose PET images. First, the network was trained using a single input of 8% L-PET images (conventional framework). For the second model, in addition to the 8% L-PET images, 6% and 4% L-PET images were fed into the deep learning model in a multi-input configuration. In this procedure, the 6% and 4% low-dose PET images were reconstructed from the 8% L-PET image and are considered as prior/addition knowledge to estimate the full-dose PET images. The training and evaluation of the models were performed using 100 and 20 clinical brain studies, respectively.

**Evaluation strategy**

We evaluated the quality of the predicted images using four standard quantitative metrics, including peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), root mean square error (RMSE) of standard uptake value (SUV) in both the entire head region and malignant lesions. Moreover, the bias of mean SUV ($SUV_{mean}$) between the L-PET image and predicted S-PET images versus the ground truth S-PET images was calculated for the malignant lesions. These quantitative metrics reference were employed to compare the benefits of using the prior knowledge in the form of multiple low-dose PET images.

**III. RESULTS AND DISCUSSION**

Sagittal views of standard-dose PET images predicted by the single-input and the multi-input networks, L-PET image (8% of S-PET image), and the ground truth S-PET image are shown in Fig. 1.

**Table I. Quantitative metrics calculated for the L-PET image, S-PET image predicted using single-input, and multi-input deep learning models.**

|            | L-PET (8%) | Single-input (L-PET image 8%) | Multi-input (L-PET images 8%, 6%, 4%) |
|------------|------------|-------------------------------|--------------------------------------|
| PSNR±SD    | 37.37±1.26 | 40.11±9.33                    | 43.32±10.01                          |
| SSIM±SD    | 0.95±0.01  | 0.98±0.22                     | 0.99±0.22                            |
| RMSE±SD (Head) | 0.12±0.04 | 0.09±0.03                     | 0.06±0.02                            |
| RMSE±SD (Lesion) | 0.49±0.09 | 0.56±0.25                     | 0.36±0.16                            |
| $SUV_{mean}$ bias±SD | 0.04±0.42 | -4.18±1.14                    | -1.44±0.56                           |

Table 1 summarizes the results of the quantitative analysis in terms of PSNR, SSIM, RMSE for the entire head region. Moreover, the mean SUV bias between the 8% L-PET images, S-PET images predicted from single-input (8%), and multi-input (8%, 6%, and 4%) models versus the ground truth S-PET images is presented for the malignant lesions.

The analysis of these metrics demonstrated the superiority of the multi-input network over the single-input network. Utilizing the additional/prior knowledge in the multi-input network reduced the RMSE within the entire head region, RMSE within the lesions, and lesion $SUV_{mean}$ bias by 33.33%, 35.71%, and 65.55%, respectively. Furthermore, the multi-input network (multiple low-dose PET images) achieved higher values of PSNR and SSIM.

**Fig. 2.** Boxplots of RMSE and $SUV_{mean}$ bias within the malignant lesions for the different models.
IV. CONCLUSION

The quantitative investigation of the predicted S-PET images from the corresponding L-PET images indicated the superior performance of the network trained with multiple low-dose PET images. This study demonstrated the benefits of using multiple low-dose images (at different dose levels) as prior knowledge to estimate the full-dose PET images.

REFERENCES

[1] H. Arabi and H. Zaidi, "Applications of artificial intelligence and deep learning in molecular imaging and radiotherapy," European Journal of Hybrid Imaging, vol. 4, no. 1, pp. 1-23, 2020.

[2] H. Arabi, A. AkhavanAllaf, A. Sanaat, I. Shiri, and H. Zaidi, "The promise of artificial intelligence and deep learning in PET and SPECT imaging," (in eng), Phys Med, vol. 83, pp. 122-137, Mar 2021, doi: 10.1016/j.ejmp.2021.03.008.

[3] B. Sanaei, R. Faghhihi, and H. Arabi, "Quantitative investigation of low-dose PET imaging and post-reconstruction smoothing," arXiv preprint arXiv:2103.10541, 2021.

[4] A. Sanaat, H. Arabi, I. Maina, V. Garibotto, and H. Zaidi, "Projection-space implementation of deep learning-guided low-dose brain PET imaging improves performance over implementation in image-space," Journal of Nuclear Medicine, p. jnumed.119.239327, 2018.

[5] A. Sanaat, I. Shiri, H. Arabi, I. Maina, R. Nkoulo, and H. Zaidi, "Deep learning-assisted ultra-fast/low-dose whole-body PET/CT imaging," (in eng), Eur J Nucl Med Mol Imaging, vol. 48, no. 8, pp. 2405-2415, Jul 2021, doi: 10.1007/s00259-020-05167-1.

[6] H. Arabi and H. Zaidi, "Improvement of image quality in PET using post-reconstruction hybrid spatial-frequency domain filtering," Physics in Medicine & Biology, vol. 63, no. 21, p. 215010, 2018.

[7] H. Arabi and H. Zaidi, "Non-local mean denoising using multiple PET reconstructions," (in eng), Ann Nucl Med, vol. 35, no. 2, pp. 176-186, Feb 2021, doi: 10.1007/s12149-020-01550-y.

[8] A. Sanaat, I. Shiri, S. Ferdowsi, H. Arabi, and H. Zaidi, "Robust-Deep: A Method for Increasing Imaging Datasets to Improve Deep Learning Models’ Robustness and Quantitative/Qualitative Metrics," ed: Soc Nuclear Med, 2021.

[9] H. Arabi and H. Zaidi, "Spatially guided nonlocal mean approach for denoising of PET images," (in eng), Med Phys, vol. 47, no. 4, pp. 1656-1669, Apr 2020, doi: 10.1002/mp.14024.

[10] N. Aghakhan Olia et al., "Deep learning-based denoising of low-dose SPECT myocardial perfusion images: quantitative assessment and clinical performance," (in eng), Eur J Nucl Med Mol Imaging, Nov 15 2021, doi: 10.1007/s00259-021-05614-7.

[11] I. Shiri et al., "Standard SPECT myocardial perfusion estimation from half-time acquisitions using deep convolutional residual neural networks," (in eng), J Nucl Cardiol, Apr 28 2020, doi: 10.1007/s12350-020-02119-y.

[12] A. Bahrami, A. Karimian, and H. Arabi, "Comparison of different deep learning architectures for synthetic CT generation from MR images," (in eng), Phys Med, vol. 90, pp. 99-107, Oct 2021, doi: 10.1016/j.ejmp.2021.09.006.

[13] A. Bahrami, A. Karimian, E. Fatemizadeh, H. Arabi, and H. Zaidi, "A new deep convolutional neural network design with efficient learning capability: Application to CT image synthesis from MRI," (in eng), Med Phys, vol. 47, no. 10, pp. 5158-5171, Oct 2020, doi: 10.1002/mp.14418.

[14] 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: Springer, pp. 348-360.

[15] F. Gholamiankhah, S. Mostafapour, and H. Arabi, "Deep learning-based synthetic CT generation from MR images: comparison of generative adversarial and residual neural networks," arXiv preprint arXiv:2103.01609, 2021.