Fast Dynamic Brain PET Imaging Using a Generative Adversarial Network

Amirhossein Sanaat, Ehsan Mirsadeghi, Behrooz Razeghi, Nathalie Ginovart, 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/MIC42677.2020.9507894

***Abstract*– This work aims to present and evaluate a novel recurrent deep learning model for reduction of the acquisition time in dynamic brain PET imaging without forfeiting clinical information. The clinical dataset included 46 dynamic ^{18}F-DOPA brain PET/CT images used to evaluate a model for generation of complete dynamic PET images from 27% of the total acquisition time. The dataset was split into 35, 6, and 5 for training, validation, and test, respectively. Each dynamic PET scan lasts 90 minutes acquired in list-mode format used to reconstruct 26 dynamic frames). A video prediction deep learning algorithm consisted of two generative adversarial networks and one variational autoencoder was developed and optimized to depict the tracer variation trend from the initial 13 frames (0 to 25 min) and synthesize the last 13 frames (25 to 90 min), respectively. The generated image was analyzed quantitatively by calculating standard metrics, such as the peak signal-to-noise ratio (PSNR), structural similarity index metric (SSIM), and time-activity curve (TAC). The PSNR and SSIM varied from 43.24 ± 0.4 to 38.82 ± 0.74 and from 0.98±0.03 to 0.81±0.09 for synthesized frames (14 to 26), respectively. The TAC trend showed that our model is able to predict images with similar tracer distribution compared to reference images. We demonstrated that the proposed method can generate the last 65 min time frames from the initial 25 min frames in dynamic PET imaging, thus reducing the total scanning time.**

I. Introduction

POSITRON emission tomography (PET) has been widely adopted in the clinic owing to its role in monitoring various diseases including neurological disorders at the molecular and cellular level[1-3]. Conventional image reconstruction of dynamic PET data based on independent handling of single frames has limited performance, particularly in the initial dynamic frames owing to the low statistics. Recently in the clinic, many applications have been considered for dynamic PET scanning, but the long scanning time restricted the usage of that in clinic and research approaches, particularly for old and child patients. Several studies attempted to decrease the overall PET acquisition time in brain PET studies using deep learning algorithms [4-7]. In this work, we used a recurrent frame generation to reduce the acquisition time by estimating half of the late frames of dynamic ^{18}F-Dopa PET images from its first half frames. Based on our literature review, this is the first work that uses video prediction deep learning models for reduction of the acquisition time from 90 min to 25 min, which makes the process of scanning more comfortable and increases the scanner throughput.

Manuscript was submitted Dec 17, 2020. This work was supported by the Swiss National Science foundation under grant SNFN 320030_176052.

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

E. Mirsadeghi is with Amirkabir University of Technology, Electrical Engineering Department, Tehran, Iran (e-mail: e.mirsadeghi@aut.ac.ir).

B. Razeghi is with Department of Computer Sciences, Geneva University, Geneva, Switzerland (e-mail: Behrooz.Razeghi@unige.ch).

N. Ginovart is with Department of Psychiatry and Department of Basic Neurosciences, Geneva University, Geneva, Switzerland (e-mail: Nathalie.Ginovart@unige.ch).

H. Zaidi is with Geneva University Neurocenter, Geneva University, Geneva, Switzerland.

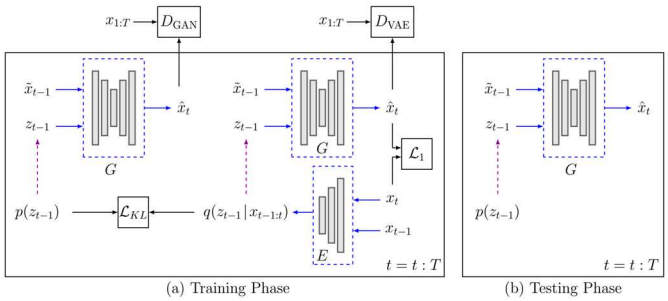

Fig. 1. Testing usage of SAVP model: New frames generated by sampling from latent codes z at every time step from a prior distribution p(z). The generator G synthesize new frame from a combination of previous frame and a latent code z. Indeed, for next time step, predicted frames are used into the generator. X_{t-1} indicate a ground truth frame (for the initial frames) or the last prediction.

II. Materials and Methods

In this work, a database consisted of 46 ^{18}F-DOPA dynamic brain PET/CT subjects consisting of 12 healthy subjects, 25 patients who suffer from cannabis use disorder (CUD), and 9 subjects with Internet gaming disorder (IGD) was used. A statistic brain PET scan was performed for each patient that takes around 90-min after tracer injection of 190 ± 10 MBq of ^{18}F-DOPA. The PET list-mode data under-sampled into projections and reconstructed into 26 various dynamic frames (2 × 30 s; 4 × 60 s; 3 × 120 s; 3 × 180 s; 14 × 300 s). An ordinary Poisson-ordered subsets expectation-maximization algorithm with 5 iterations and 21 subsets accompanied by time-of-flight information was performed for image reconstruction. Fig.1 shows the Stochastic Adversarial Video Prediction (SAVP) model proposed to generate raw pixels of future dynamic frames, given a sequence of initial frames.

To evaluate how much the predicted second half of the frames sequence (frames 14 to 26) are similar to the corresponding frames in the reference dataset two quantitative metrics, namely the peak signal-to-noise ratio (PSNR) and the structural similarity index metrics (SSIM) were calculated for each frame. The Patlak graphical analysis was performed in

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

PMOD software for extracting the maps of the influx rate constant (Ki in min^{-1}) and distribution volume (V) for ^{18}F-DOPA PET parametric images. Furthermore, to assess the performance of the model to follow tracer variation during the time, the time-activity curve (TAC) attained over the reference region corresponding to the cerebellum for predicted and reference images for left and right Caudate, Putamen, and Cellarium. TAC graph helps us to figure out the tracer variation trapped in the tissue during the time. TAC curve use for extracting kinetic parameter in Patlak analysis and have great influence of Ki and V.

III. RESULTS

Fig. 2 shows the PSNR, and SSIM calculated between predicted and reference frames from 14 to 26. The earlier predicted frames (14–20) show better image quality, lower noise, and higher quantitative accuracy than the last frames (21-26). The graphs show the value of these two metrics gets worse by going to the last frames, but even in the last frames, the magnitude is reasonable.

Fig. 3 depicts the TAC curve for all generated and reference frames for seven brain regions. The dashed line illustrates the endpoint of the scan at 25 min instead of 90 min. the frames on the left side have been fed to the model and the right side will be predicted by the model.

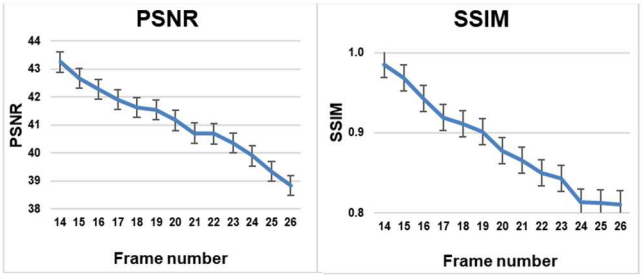

Fig. 2. The structural similarity index metrics (SSIM), and peak signal to noise ratio (PSNR) calculated between synthesized and reference images as an image quality metric.

IV. DISCUSSION

In this study, a VAE-GAN model was implemented to synthesize the last half-frames contain 13 frames with an overall duration of 65 min of dynamic brain PET images from the earliest half-frames include 13 frames with 25 min duration. The main goal was to reduce the scanning time (around 73 %) in dynamic ^{18}F-DOPA brain PET images and preserve the diagnostic quality as same as full-time scanning. Our video prediction model was developed to predict the future movement in a movie through learning previous actions and their variation trend. Here we used the same approach and fed the model with time frames to obtain the trend of the biodistribution variation during a full scan and after training, by feeding the model with only half of the frames the model will predict the future biodistribution based on the previous trend. This study is the first attempt at reducing the acquisition time of long dynamic PET imaging protocols using deep video prediction algorithms. Our results demonstrated that the predicted dynamic PET frames from the first half frames have

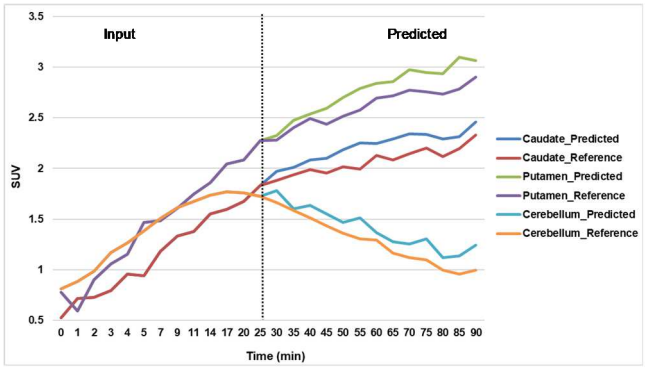

Fig. 3. Time activity curves for predicted and reference data for the caudate, putamen and cerebellum for a 26-year-old normal female patient. The black dotted line shows and separate the earliest frames that fed to model and last frames that predicted by model.

a similar diagnostic quality and biodistribution patterns, as well as compartmental modeling and distribution volume values, relative to reference frames particularly for the primary, predicted dynamic frames.

The PSNR that present the level of noise and signal calculated on predicted frames and it gradually gets worse from ~43 for frame number 14 to ~38 for frame number 26 means that in last frames the noise increase and signal reduce. The reason behind this hampering in quantitative and qualitative accuracy with time is in the Data Processing Inequality (DPI) concept. Base on DPI, data that pass through a channel, the information cannot be improved. Consequently, the last frames damage more than the earliest frames with a subsequent decrease of the PSNR. Furthermore, the similarity between the generated frames and reference dynamic frames calculated by the SSIM metric shows a variation from 0.98 for frame number 14 to 0.81 for the last frame. The comparison of the TACs between predicted and reference images shows an optimal fitting curve for the radiotracer distribution. The compartmental modeling of generated images resulted in a K_i and V with less than 7% and 14% bias in most brain regions.

In this work, we demonstrated that a novel video prediction model based on variational autoencoder and GAN model is able to use for fast dynamic brain PET imaging and time-dependent data. Our model was fed by only the initial 25 min frames and trained to generate the last 65 min time frames. The model's performance was very promising as revealed by the good image quality and low quantification bias.

REFERENCES

[1] 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.

[2] 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.

[3] A. Sanaat, M. S. Zafarghandi, and M. R. Ay, "Design and performance evaluation of high resolution small animal PET scanner based on monolithic crystal: a simulation study," *Journal of Instrumentation,* vol. 14, no. 01, pp. P01005-P01005, 2019/01/03 2019.

[4] 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,* vol. 61, no. 9, pp. 1388-1396, Sep. 2020.

[5] 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," *Appl Sci.,* vol. 10, no. 14, p. 4753, 2020.

[6] 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,* Jan 25 2021.

[7] 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.