Towards Deep-Learning Partial Volume Correction for SPECT

Théo Kaprélian, Ane Etxebeste, David Sarrut

Towards Deep-Learning Partial Volume Correction for SPECT

Théo Kaprelian[1], Ane Etxebeste[1], and David Sarrut[1]

[1]Université de Lyon, CREATIS; CNRS UMR5220; Inserm U1044; INSA-Lyon; Université Lyon 1; Centre Léon Bérard, France.

Abstract Partial Volume Effect impacts the spatial resolution of SPECT images. We investigated the feasibility of a deep learning-based Partial Volume Correction method (PVCNet) that compensates for the effect of collimator blurring on 2D projections, before reconstruction. A large dataset containing 600,000 pairs of synthetic projections was generated and used to train two consecutive UNets (one for denoising, one for PVC). Scatter and attenuation were not yet considered in the database. Our proposed PVCNet method achieves 12.8% NMAE reduction compared to conventional Resolution Modeling on the IEC phantom but Recovery Coefficients were not always better for smallest spheres.

1 Introduction

Single-Photon Emission Computed Tomography (SPECT) images are impacted by several physical effects that need to be compensated in order to achieve a reasonable image quality [1]: photon attenuation, photon scatter and Partial Volume Effect (PVE). Several works have been able to effectively reduce the impact of the three effects [2–4], but PVE still remains the main limiting factor, leading to inaccurate quantification of the radioactive tracer uptake [5]. PVE is defined as the apparent underestimation of activity in an object of interest due to limited spatial resolution. It is particularly an issue for objects whose size is smaller than the system's resolution volume [6], approximately defined as twice the Full-Width at Half Maximum (FWHM) of the Point Spread Function (PSF) obtained by imaging a point source located at the center of the Field Of View (FOV). Several elements contribute to degrade the spatial resolution: detector and electronics response, collimator septal penetration, collimator scatter and collimator geometric response. The main one is the geometrical response of the collimator that depends on the source-to-collimator distance and the characteristic of the collimator such as hole diameter, septal thickness, length and material. Typical values for the FWHM at 10 cm from the collimator front surface range between 5-15 mm according to the collimator type.

A widely used Partial Volume Correction (PVC) is the Resolution Modeling (RM) method [7] applied in the system's matrix used for forward and back-projections of the Ordered Subset Expectation Maximisation (OSEM) reconstruction algorithm. This method has good noise reduction properties and theoretically converges to the true activity distribution but the resolution gain achieved in practice is limited because of the loss of high frequency information leading to Gibbs artefacts as the number of iteration increases [4]. For this reason, a regularization can be applied, e.g. [8], resulting in a smoother image and requiring additional parameter tuning. Other PVC techniques [4] include image deconvolution [9] or region-based correction [10]. A drawback of deconvolution methods is that they tend to amplify noise. Region-based corrections rely on a segmentation mask of Regions of Interest (ROI) which may not be easy to define.

Recently, Deep Learning methods have shown promising results in various tasks in nuclear medicine [11]. In SPECT, recent works showed that some neural networks architectures were able to perform scatter correction [12], image reconstruction [13] or projection interpolation [14]. However, to our knowledge, only very few works investigated deep learning-based PVC, e.g. in [15, 16] the net was trained with small datasets and with ground truth images obtained with conventional PVC.

In this work, we propose a deep learning framework trained to compensate the effect of the PSF due to the collimator on the 2D projections, before 3D reconstruction. A large training dataset is first generated by simulation and contains 600,000 pairs of corresponding projections with and without PVE+noise. Our PVC networks are two consecutive UNets henceforth denoted as PVCNet.

2 Materials and Methods

2.1 Database

We generated a large dataset of simulated pairs of corresponding *input* and *target* 2D SPECT projections. Input projections are the realistic ones with PVE and noise, while target ones are artefact-free.

We first created 3D sources of ^{99m}Tc made of a large elliptic cylinder background with variable axis size (90-260 mm) with several hot sources (between 1 and 8) of ellipsoidal shapes (8-128 mm axis) randomly oriented and located within the background. Hot source to background activity ratios between 1/1000 and 1/8 were considered. A total of 5 000 3D voxelized (256^3 voxels of 2 mm size) activity sources were randomly generated. Then, 2D projections were obtained by forward-projecting each one of the 5 000 sources with ray-tracing using RTK [17] in two different ways: once without resolution modeling (P_{noPVE} projection) and once with resolution modeling (P_{PVE} projection). Resolution modeling was performed during forward-projection operator by applying depth-dependant Gaussian convolutional kernel [7] whose parameters were derived from the dimensions of the Siemens-Intevo LEHR collimator following the analytical analysis provided by [6]. We obtain $\mathrm{FWHM}(d) = 0.048 + 1.11d$

where d is the distance from the source to the collimator front face. For P_{noPVE}, the same operator was applied but with $FWHM(d) = 0$, so that the simulated projection is the one that would have been obtained without collimator blurring. P_{noPVE} will thus serve as target projections. Each projection contains 256^2 pixels with a size of 2.3976^2 mm. Sources were simulated in air to avoid attenuation and scattering, and the 140 keV photo-peak window of ^{99m}Tc was considered. Poisson noise was applied to P_{PVE} to roughly mimic the data detection process. The resulting projection was denoted $P_{PVE,noisy}$. For each of the 5000 sources, we applied the same projection process for 120 evenly distributed angles between 0° and 360°, resulting in 120 triplets ($P^i_{PVE,noisy}$, P^i_{PVE}, P^i_{noPVE}), for $i = 1,...,120$. The source distributions were randomly scaled so that the total number of counts in each projection was comprised between 5,000 and 500,000 such as in realistic clinical applications.

2.2 Networks and training

The database described in the previous section was employed to train simultaneously two neural networks: a *Denoiser* and a *PVC* network. The *Denoiser* network was trained to take $P_{PVE,noisy}$ as inputs and to output projections close to the corresponding unnoisy P_{PVE} projections. Input and output of the *Denoiser* have the same number of channels. The *PVC* network then takes as input the output of the *Denoiser* and is trained with P_{noPVE} as target to perform PVC with one projection angle as output. The idea behind this is that since the database generation is completely analytical, we have access to useful intermediate information that can be used to divide training into these two supervised tasks.
Both networks were UNets with 3 encoding/decoding residual blocks with skip-connections. The first layer was a Conv2d expanding the number of channels to 32. Then, each encoding (resp. decoding) block was composed by a sequence of Conv2d (resp. TransposeConv2d)-InstNorm-LeakyRelu-Conv2d-LeakyRelu-InstNorm-MaxPooling(resp. Conv2d). Both networks end by a final convolution layer that outputs the needed number of channels. All kernels were (3,3) convolutions.
The input of the *Denoiser* network was extended to consider several $P_{PVE,noisy}$ projections corresponding to different projections angles of the same source. Considering that the projection to be corrected is at angle i, the *Denoiser* takes as input projections of angles : $(i°, i-3°, i+3°, i+90°, i+180°, i+270°)$, *i.e.* the projection to be corrected, two adjacent angles, two orthogonal and opposite ones. Moreover, we further enrich the input of the *Denoiser* to take an additional channel previously obtained by using the full sinogram (120 angles) $P_{PVE,noisy}$ to reconstruct a coarse volume with one iteration of OSEM and RM and then forwardprojecting this volume (without RM) on the same 120 angles to obtain one additional channel per angle. The input/output of *Denoiser* then have 7 channels whereas the *PVC* network has 7 channels as input and outputs only one channel (i.e. the estimated projection $\hat{P}^i_{noPVE}$). The idea here was to exploit additional information contained in the data to help solving this ill-posed inverse problem (different noise realisation, source depth, RM) and to ensure a continuity in the corrected projections. Parameters of both networks were optimized to minimize a L1 loss functions. Networks were trained during 100 epochs with Adam optimizer, with 4 GPUs, a batch size of 256 per GPU and a learning rate of 10^{-4} halved every 20 epochs.

2.3 Evaluation data and metrics

Performance of the proposed method was evaluated with three experiments. First, we considered an analytical version of the standard NEMA IEC phantom composed of six spheres with increasing diameters 10-37 mm, with 1/40 background ratio and projections obtained with RTK like for the training database. Several reconstructed images were compared: the images reconstructed from P_{noPVE} projections (noPVE-noPVC), from $P_{PVE,noisy}$ projections and RM (PVE-RM), from projections corrected by the networks (PVE-PVCNet) and from projections without any PVC (PVE-noPVC). Then, we evaluated the performance of our proposed method on real acquisition of the NEMA IEC phantom, obtained with a 1/10 background-to-source ratio and the Siemens-Intevo SPECT/CT system. PVCNet was applied to the primary energy window and scatter window independently before scatter correction with the DEW was applied (k=1.1). Finally, we tested our proposed network on real patient data.
All reconstructions were performed with OSEM with 8 subsets and 5 iterations (except for PVE-RM for which 20 iterations were needed), scatter and attenuation correction. Resulting image resolutions were compared by computing the hot sphere contrast Recovery Coefficient (RC) [18] for each sphere and correction method. We also computed the Normalized Root Mean Square Error (NRMSE), Normalized Mean Absolute Error (NMAE), Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) of each image.

3 Results

Database generation took 3 hours using one hundred parallel CPUs. Training took 50 hours for 100 epochs using 4 GPUs. The proposed PVCNet method was compared to the ground-truth image (noPVE-noPVC), to the widely used Resolution Modeling (RM) method and to the un-corrected one (PVE-noPVC). From now on, FWHM refers to the collimator resolution value at a distance of 28 cm, which corresponds to the employed isocenter-collimator distance to generate the simulated projections and for the acquisition. Visualisation and RC results for the analytical IEC phantom are shown in Figure 1 and show promising results in terms of activity recovery, Gibbs artefact reduction and error reduction. NMAE was divided by two and PVCNet was significantly better in

term of all global metrics shown in Figure 1c. However, small spheres (size<FWHM) were completely lost by the network.
Reconstructed images and RC curves for the real IEC acquisition are shown in Figure 2. Figure 2b underlines that we only achieved better correction than RM on the largest sphere (of size 37 mm) in terms of RC. Similarly to the previous experiment, the network struggles to correct PVE on small spheres and some distortion artefacts are visible on the corrected spheres. On the other side, Figure 1a shows that for sphere with size>FWHM, the homogeneity was better retrieved with PVE-PVCNet than PVE-RM and regarding the other studied criteria, PVE-PVCNet outperfomed PVE-RM (Figure 2c).
Real patient reconstructed images are shown in Figure 3 for visual assessment only, as no reference was available.

4 Discussion and conclusion

For the first time, this work investigated the feasibility of training a network from simulated projections to compensate the effect of the PSF and to denoise projections, before reconstruction. We showed that building such a database is feasible, and we designed an adapted deep learning architecture to correct both noise and PVE. On simple test cases, PVCNet reduced Partial Volume Effect compared to standard RM method, while requiring less iterations and no regularization. However, on real data acquisition, while NRMSE, NMAE, PSNR and SSIM were better than the values obtained with RM, RC was not. Small spheres were not well recovered. We now envision to improve the realism of the training database by using tumor-like source shapes, heterogeneous activities and projections generated by (fast) Monte Carlo simulations [19, 20]. Finally, considering more projection angles as input could be useful to increase source-depth information.

Acknowledgement

This research was funded, in part, by MOCAMED (ANR-20-CE45-0025), LYriCAN (INCa-INSERM-DGOS-12563), LABEX PRIMES (ANR-11-LABX-0063, ANR- 11-IDEX-0007), POPEYE (ANR-19-PERM-0007-04). We gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research. This work was granted access to the HPC resources of IDRIS under the allocation 2019-101203 made by GENCI (Jean Zay computing center).

References

[1] E. C. Frey, J. L. Humm, and M. Ljungberg. "Accuracy and Precision of Radioactivity Quantification in Nuclear Medicine Images". *Seminars in Nuclear Medicine*. Theranostics 42.3 (May 1, 2012), pp. 208–218. DOI: 10.1053/j.semnuclmed.2011.11.003.

[2] G. T. Gullberg, R. H. Huesman, J. A. Malko, et al. "An attenuated projector-backprojector for iterative SPECT reconstruction". *Physics in Medicine & Biology* 30.8 (Aug. 1985), p. 799. DOI: 10.1088/0031-9155/30/8/004.

[3] R. J. Jaszczak, K. L. Greer, C. E. Floyd, et al. "Improved SPECT Quantification Using Compensation for Scattered Photons". *Journal of Nuclear Medicine* 25.8 (Aug. 1, 1984). Publisher: Society of Nuclear Medicine Section: Basic Sciences, pp. 893–900.

[4] K. Erlandsson, I. Buvat, P. H. Pretorius, et al. "A review of partial volume correction techniques for emission tomography and their applications in neurology, cardiology and oncology". *Physics in Medicine and Biology* 57.21 (Oct. 2012). Publisher: IOP Publishing, R119–R159. DOI: 10.1088/0031-9155/57/21/R119.

[5] D. Bailey, H. Marquis, and K. Willowson. "Partial Volume Effect in SPECT & PET Imaging and Impact on Radionuclide Dosimetry Estimates". *Asia Oceania Journal of Nuclear Medicine and Biology* (Online First May 2022). DOI: 10.22038/aojnmb.2022.63827.1448.

[6] S. R. Cherry, J. A. Sorenson, and M. E. Phelps. "chapter 14 - The Gamma Camera: Performance Characteristics". *Physics in Nuclear Medicine (Fourth Edition)*. Ed. by S. R. Cherry, J. A. Sorenson, and M. E. Phelps. Philadelphia: W.B. Saunders, Jan. 1, 2012, pp. 209–231. DOI: 10.1016/B978-1-4160-5198-5.00014-9.

[7] G. L. Zeng, G. T. Gullberg, C. Bai, et al. "Iterative Reconstruction of Fluorine-18 SPECT Using Geometric Point Response Correction". *Journal of Nuclear Medicine* 39.1 (Jan. 1, 1998). Publisher: Society of Nuclear Medicine Section: General Nuclear Medicine, pp. 124–130.

[8] A. R. De Pierro. "A Modified Expectation Maximization Algorithm for Penalized Likelihood Estimation in Emission Tomography". *IEEE Transactions on Medical Imaging* 14.1 (Mar. 1995), pp. 132–137. DOI: 10.1109/42.370409.

[9] J. Tohka and A. Reilhac. "Deconvolution-based partial volume correction in Raclopride-PET and Monte Carlo comparison to MR-based method". *NeuroImage* 39.4 (2008), pp. 1570–1584. DOI: 10.1016/j.neuroimage.2007.10.038.

[10] O. G. Rousset, Y. Ma, and A. C. Evans. "Correction for Partial Volume Effects in PET: Principle and Validation". *Journal of Nuclear Medicine* 39.5 (May 1, 1998). Publisher: Society of Nuclear Medicine Section: General Nuclear Medicine, pp. 904–911.

[11] H. Arabi, A. AkhavanAllaf, A. Sanaat, et al. "The promise of artificial intelligence and deep learning in PET and SPECT imaging". *Physica Medica* 83 (Mar. 1, 2021), pp. 122–137. DOI: 10.1016/j.ejmp.2021.03.008.

[12] H. Xiang, H. Lim, J. A. Fessler, et al. "A deep neural network for fast and accurate scatter estimation in quantitative SPECT/CT under challenging scatter conditions". *European Journal of Nuclear Medicine and Molecular Imaging* 47.13 (Dec. 1, 2020), pp. 2956–2967. DOI: 10.1007/s00259-020-04840-9.

[13] G. Wang, J. C. Ye, and B. De Man. "Deep learning for tomographic image reconstruction". *Nature Machine Intelligence* 2.12 (Dec. 2020). Number: 12 Publisher: Nature Publishing Group, pp. 737–748. DOI: 10.1038/s42256-020-00273-z.

[14] T. Rydén, M. Van Essen, I. Marin, et al. "Deep-Learning Generation of Synthetic Intermediate Projections Improves ^{177} Lu SPECT Images Reconstructed with Sparsely Acquired Projections". *Journal of Nuclear Medicine* 62.4 (Apr. 2021), pp. 528–535. DOI: 10.2967/jnumed.120.245548.

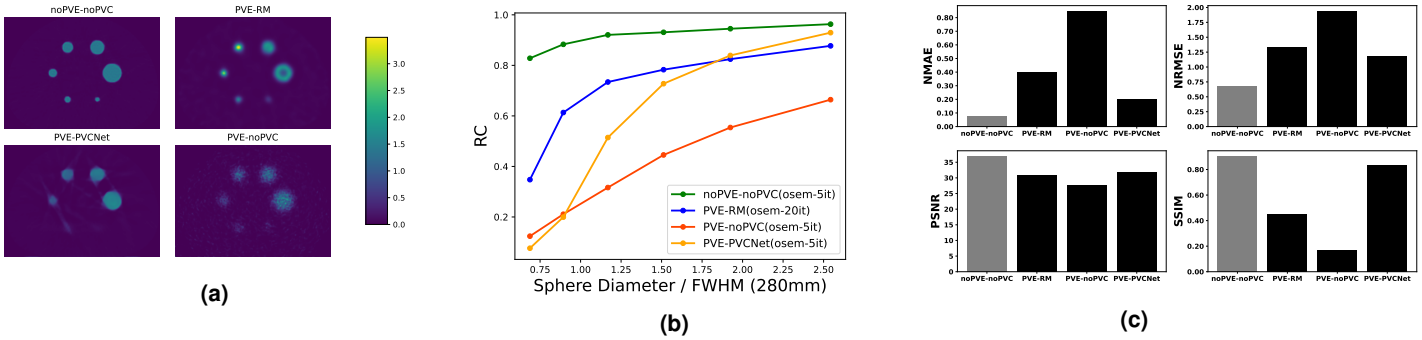

Figure 1: (a) Slice of the four reconstructed images (noPVE-noPVC, PVE-RM, PVE-noPVC, PVE-PVCNet), (b) the corresponding Recovery Coefficient (RC) for each reconstructed sphere with respect to the ratio sphere diameter / FWHM (c) comparison of NRMSE, NMAE, PSNR and SSIM. The reference image was the initial voxelised 3D source used for forward projection.

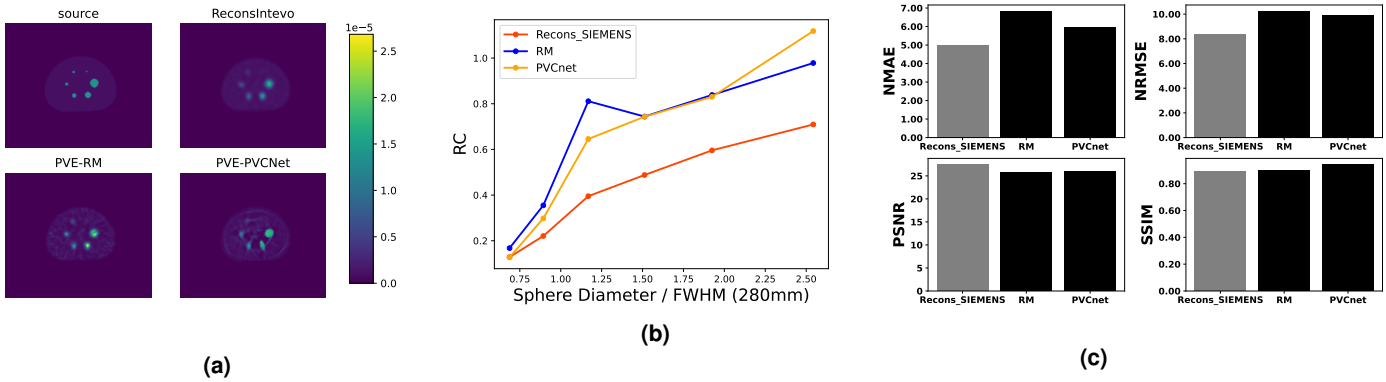

Figure 2: (a) Slice of the four images (source, Reconstruction by the INTEVO SPECT system, PVE-RM, PVE-PVCNet) divided by the total number of counts in each image (b) the corresponding Recovery Coefficient (RC) for each reconstructed sphere with respect to the ratio sphere diameter / FWHM (c) NRMSE, NMAE, PSNR and SSIM with a manually contoured reference image knowing the injected activity concentrations in spheres/background

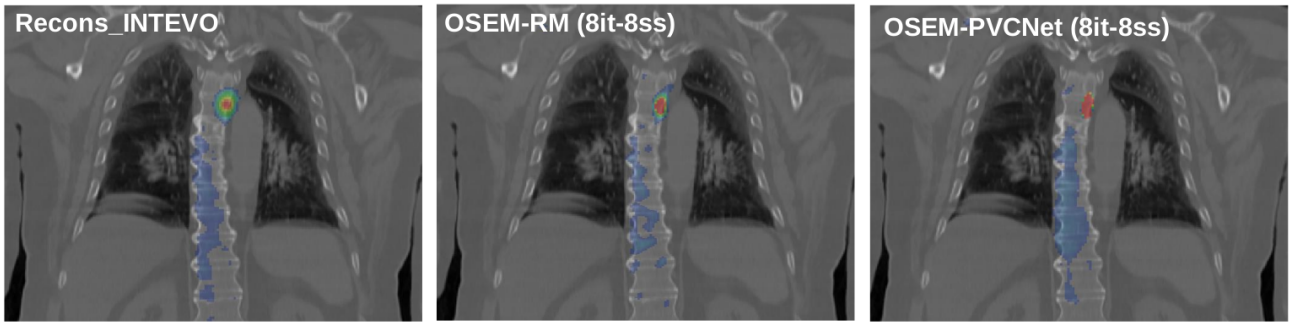

Figure 3: Visual comparison of patient SPECT/CT images with different correction methods. All three images were reconstructed with OSEM algorithm (20 iterations, 4 subsets), attenuation and scatter correction.

[15] K. Matsubara, M. Ibaraki, and T. Kinoshita. *DeepPVC: Prediction of A Partial Volume-Corrected Map for Brain Positron Emission Tomography Studies Via a Deep Convolutional Neural Network*. ISSN: 2693-5015 Type: article. 2022. DOI: 10.21203/rs.3.rs-1190960/v1.

[16] H. Xie, Z. Liu, L. Shi, et al. "Segmentation-free PVC for Cardiac SPECT using a Densely-connected Multi-dimensional Dynamic Network". *IEEE Transactions on Medical Imaging* (2022), pp. 1–1. DOI: 10.1109/TMI.2022.3226604.

[17] S. Rit, M. V. Oliva, S. Brousmiche, et al. "The Reconstruction Toolkit (RTK), an open-source cone-beam CT reconstruction toolkit based on the Insight Toolkit (ITK)". *Journal of Physics: Conference Series* 489 (Mar. 2014). Publisher: IOP Publishing, p. 012079. DOI: 10.1088/1742-6596/489/1/012079.

[18] M. E. Daube-Witherspoon, J. S. Karp, M. E. Casey, et al. "PET Performance Measurements Using the NEMA NU 2-2001 Standard". *Journal of Nuclear Medicine* 43.10 (Oct. 1, 2002). Publisher: Society of Nuclear Medicine Section: Special Contribution, pp. 1398–1409.

[19] D. Sarrut, A. Etxebeste, N. Krah, et al. "Modeling Complex Particles Phase Space with GAN for Monte Carlo SPECT Simulations: A Proof of Concept". *Physics in Medicine and Biology* (2021). DOI: 10.1088/1361-6560/abde9a.

[20] A. Saporta, A. Etxebeste, T. Kaprelian, et al. "Modeling Families of Particle Distributions with Conditional GAN for Monte Carlo SPECT Simulations". *Physics in Medicine & Biology* (2022). DOI: 10.1088/1361-6560/aca068.