Deep Image Prior PET Reconstruction using a SIRF-Based Objective

Imraj RD. Singh†, *Student Member, IEEE*, Riccardo Barbano†, Robert Twyman, *Student Member, IEEE*
Željko Kereta, Bangti Jin, Simon Arridge, and Kris Thielemans, *Senior Member, IEEE*

***Abstract*—Widespread adoption of deep learning in medical imaging has been hampered, in part, due to a lack of integration with clinically applicable software. In this work, we establish a direct connection between an established PET reconstruction suite, SIRF, and PyTorch. This allows for advanced reconstruction methodologies to be deployed on clinical data with an unsupervised deep learning approach: the Deep Image Prior. Results show consistent quality metrics for DIP in comparison to OSMAP.**

I. Introduction

AS deep learning becomes ubiquitous in medical image reconstruction research, there still exists a lack of compatibility between established reconstruction suites and deep learning toolkits. Few exceptions exist and those that do are primarily applied to computed tomography (e.g., Operator Discretization Library [1] and Tomosipo [2]) while positron emission tomography (PET) is mostly neglected. In this work, we develop a PyTorch wrapper for CCP SyneRBI's Synergistic Image Reconstruction Framework (SIRF) [3]. SIRF provides access to Software for Tomographic Image Reconstruction (STIR), an advanced PET and SPECT reconstruction library [4], as well as Gadgetron for MRI reconstruction. The integration is publicly available on GitHub and we showcase its potential by reconstructing a clinically-realistic PET XCAT torso phantom via the deep image prior (DIP) framework [5]. DIP is a well-established unsupervised learning approach requiring only a single set of measured data. The results show performance comparable to ordered subset methodologies.

A. Penalised Maximum Likelihood

Penalised maximum likelihood methods for PET solve the following optimisation problem

$$\text{argmin}_{\mathbf{x}\geq 0}\Phi(\mathbf{x}) = -L(\mathbf{x}) + \beta R(\mathbf{x}), \tag{1}$$

where $L(\mathbf{x})$ is the Poisson log-likelihood describing the goodness of fit to the measurements, $R(\mathbf{x})$ is the penalty, and $\beta > 0$ balances the two terms. In this work, we consider two penalties. The first is the Quadratic Prior (QP), defined by $R(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j\in\mathcal{N}_i} w_{ij}(x_i - x_j)^2$, where $\mathcal{N}_i$ is a 3×3 neighbourhood of the i-th image voxel and w_{ij} are the neighbourhood weights. The other is the Relative Difference Prior (RDP), defined by $R(\mathbf{x}) = \sum_{i=1}^{N}\sum_{j\in\mathcal{N}_i} w_{ij}\frac{(x_i-x_j)^2}{x_i+x_j+\gamma|x_i-x_j|}$, where the edge preservation parameter is set as $\gamma = 2$.

B. Wrapping SIRF-Objective as `torch.nn.Module`

SIRF is integrated into PyTorch by wrapping the SIRF `ObjectiveFunction` class as a PyTorch module via an *ad hoc* class named `ObjectiveFunctionModule`. Our custom autograd function subclasses `torch.autograd.Function` and uses the value and derivative of (1), which is computed via SIRF (with a C++ backend provided by STIR).

C. Deep Image Prior

DIP represents the image $\mathbf{x}$ through learnable parameters $\boldsymbol{\theta}\in\mathbb{R}^p$ of a CNN $\mathbf{f}(\mathbf{z};\boldsymbol{\theta})$ with a fixed random input $\mathbf{z}$. The optimisation problem (1) is then recast as

$$\boldsymbol{\theta}^\star \in \text{argmin}_{\boldsymbol{\theta}\in\mathbb{R}^p}\Phi(\mathbf{f}(\mathbf{z};\boldsymbol{\theta})), \tag{2}$$

and we denote the recovered image as $\mathbf{x}^\star = \mathbf{f}(\mathbf{z};\boldsymbol{\theta}^\star)$. In line with [6], we use a U-Net [7] (without skip connections) with 128 channels at each scale. To enforce the non-negativity constraint, ReLU is applied on the output $\mathbf{x}^\star$. The unsupervised nature of DIP is attractive for imaging tasks as the scarcity of high-quality training data has hindered the deployability of supervised deep learning solutions to clinically-realistic settings. Additionally, test time optimisation of DIP ensures desired data-consistency. Previous applications of DIP to PET have been implemented indirectly via optimising the augmented Lagrangian of the objective function with ADMM [8], where early stopping was required to deliver best reconstruction quality. In this work, the use of an explicit penalty alleviates the dependency on early stopping for regularisation.

II. Experiments and Results

A. Dataset

A numerical GATE simulation [9] of a 1 ring GE Discovery 690 scanner [10], with uniform crystal efficiencies, is performed via the STIR-GATE-Connection (SGC) [11]. A photon emission simulation is performed using back-to-back 511 keV photon emissions from a voxelised XCAT torso phantom [12] with activity concentrations representative of an

Manuscript recieved December 9, 2022. This work is supported by the EPSRC-funded UCL Centre for Doctoral Training in Intelligent, Integrated Imaging in Healthcare (i4Health) (EP/S021930/1), the Department of Health's NIHR funded Biomedical Research Centre at UCL Hospitals, GE Healthcare and the Alan Turing Institute (EPSRC EP/N510129/1). Software used in this project is partially maintained by CCP SyneRBI (EPSRC EP/T026693/1).
I. Singh, K. Thielemans and R. Twyman are with the Institute of Nuclear Medicine, UCL, UK. S. Arridge, R. Barbano, B. Jin, Ž. Kereta and I. Singh are with the Department of Computer Science, UCL, UK. Corresponding authors are I. Singh (e-mail: imraj.singh.20@ucl.ac.uk) and R. Barbano (e-mail: riccardo.barbano.19@ucl.ac.uk). † denotes equal contribution.

978-1-6654-8872-3/22/$31.00 ©2022 IEEE

^{18}F-FDG study. A 1 cm diameter, 1 cm long, cylindrical hot lesion, with 2.6:1 lesion to lung contrast, is inserted into the lung of the XCAT emission. Cardiac and respiratory motion, and radioactive decay, are not modelled. The resulting list mode data are binned into sinograms with a 0.93:1 true-to-background ratio, where background is random plus scattered events. Random, scatter and normalisation are modelled using SGC tools and are included in the system model.

B. Quality Metrics

Standard metrics related to quantification and detectability of a lung lesion are used to assess image quality. Contrast recovery coefficient (CRC) values are calculated between the lung lesion and lung regions of interest (ROIs) by $(\frac{\bar{a}}{\bar{b}} - 1)/(\frac{a_\mathrm{t}}{b_\mathrm{t}} - 1)$, where $\bar{a}$ and $\bar{b}$ are average emissions over lung lesion and lung ROIs, respectively. The subscript t denotes ground truth emission values. Standard deviation (STD) is calculated for an eroded lung ROI (excluding the lesion) by $\mathrm{STD} = \left(N_\mathrm{roi}^{-1} \sum (b_i - b_{i,\mathrm{t}})^2\right)^{1/2}$, where i indexes a voxel from the lung ROI and N_roi is the number of voxels in the ROI.

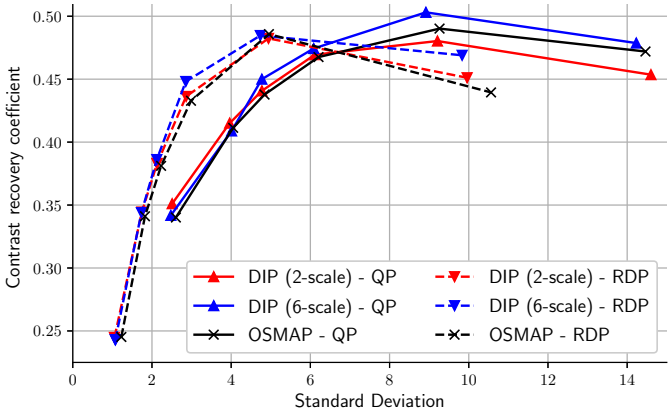

Fig. 1. For the baseline, we use 200 epochs of the One Step Late MAP (OSMAP) modification of OSEM [13] with 4 ordered subsets. For DIP both 6-scale and 2-scale U-Net architectures are tested and 50,000 epochs are used. OSMAP and DIP results on CRC of lung/lung lesion ROIs and STD of lung ROI. Both the QP and RDP are tested with various β values: QP $\beta \in \{20, 10, 7.5, 5, 2.5, 1\} \cdot \mathrm{e}{-4}$ and RDP $\beta \in \{20, 10, 7.5, 5, 2.5, 1\} \cdot \mathrm{e}{-2}$.

III. Discussion and Conclusion

The SIRF objective function enables the use of penalty functions (e.g., QP and RDP) and more complex system models allowing for an investigation of penalised DIP reconstructions on realistic datasets.

From Fig. 1 it follows that the over-parameterised 6-scale DIP network out-preforms the 2-scale network with larger CRCs and smaller STDs, improvements are consistent yet marginal. Similarly there are increases in performance for exemplary DIP reconstructions as compared with OSMAP with the same β values (see Fig. 2 for the reconstructions).

In future work, an improved comparison with [8] will be conducted along with investigations into subset methods, applications to 3D data, more noise realisations, and more diverse network architectures.

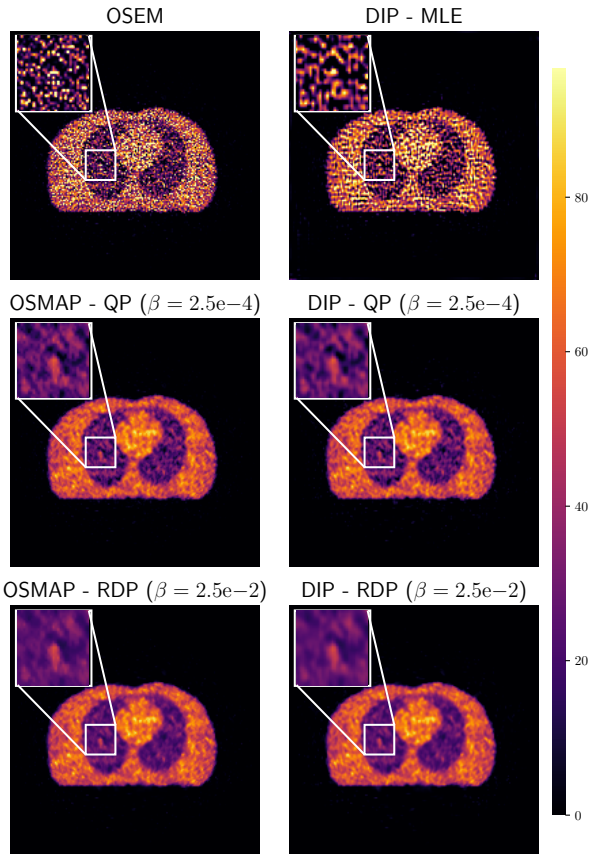

Fig. 2. OSEM, OSMAP and DIP (6-scale) reconstruction images. DIP - Maximum Likelihood Estimate (MLE) was obtained after 50,000 epochs, and OSEM after 200 epochs with 4 subsets. The β values for penalised reconstructions were determined from greatest CRC values for RDP and QP from Fig. 1. The inset shows the reconstruction of the lung lesion ROI.

References

[1] J. Adler *et al.*, "Odlgroup/odl: Odl 0.7. 0," 2018.

[2] A. Hendriksen *et al.*, "Tomosipo: Fast, flexible, and convenient 3D tomography for complex scanning geometries in Python," *Optics Express*, 2021.

[3] E. Ovtchinnikov *et al.*, "SIRF: synergistic image reconstruction framework," *Computer Physics Communications*, 2020.

[4] K. Thielemans *et al.*, "STIR: software for tomographic image reconstruction release 2," *Physics in Medicine and Biology*, 2012.

[5] D. Ulyanov *et al.*, "Deep image prior," in *CVPR*, 2018.

[6] D. O. Baguer *et al.*, "Computed tomography reconstruction using deep image prior and learned reconstruction methods," *Inv. Problems*, 2020.

[7] O. Ronneberger *et al.*, "U-net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015.

[8] K. Gong *et al.*, "PET image reconstruction using deep image prior," *IEEE Trans. Med. Imaging*, 2018.

[9] S. Jan *et al.*, "GATE V6: A major enhancement of the GATE simulation platform enabling modelling of CT and radiotherapy," *Physics in Medicine and Biology*, 2011.

[10] V. Bettinardi *et al.*, "Physical performance of the new hybrid PET/CT Discovery-690," *Medical Physics*, 2011.

[11] R. Twyman *et al.*, "A demonstration of STIR-GATE connection," *IEEE NSS/MIC*, 2021.

[12] W. P. Segars *et al.*, "4D XCAT phantom for multimodality imaging research," *Medical Physics*, 2010.

[13] H. Hudson and R. Larkin, "Accelerated image reconstruction using ordered subsets of projection data," *IEEE Trans. Med. Imaging*, 1994.