LEARNED PRIMAL-DUAL RECONSTRUCTION FOR MEDICAL IMAGING

Executive summary

dw661

Word Count: 878

Background & Motivation

Computed Tomoraphy (CT) is one of the most commonly used medical imaging techniques. It involves using emission of X-rays to infer the internal structures of patients, which can then be used to detect abnormalities and thus lead to more accurate diagnosis. Normally, to ensure high quality of reconstructions, intensities of emitted X-rays need to be high, otherwise noisy artefacts will emerge in the reconstructions. However, X-rays are known to have cancer inducing effects, making extensive exposure to them undesirable. Therefore, current active research areas revolve around obtaining good quality reconstructions whilst maintaining amount of dosage low. With the surge of ready clinical data and modern machine learning methods, Learned Primal Dual [1] attempts to address the aforementioned challenges and resulted in remarkable success.

Theoretical Foundations

Traditionally, in CT imaging, there are two main non machine learning based reconstruction methods, namely filtered backprojection (FBP) and Total Variation (TV) [9]. The former suffers substantially in presence of noise, as it involves a filtering step in the Fourier domain (multiplication by $|\omega|$), which magnifies noise components with high frequency, thus making the algorithm unstable. Hence, alternative methods for reconstruction uses variational regularisation approaches [3], which often involves minimising convex objective functions. Amongst many variants of objective functions, TV is one of the best performing ones, due to its strong edge preserving abilities.

However, it could be challenging to optimise the TV objective function, as it's non-smooth, which hinders traditional optimisation algorithms such as Gradient Descent as smoothness is a perquisite most times. To this end, proximal methods are developed to overcome this shortcoming. The Chambolle-Pock algorithm [4], was specifically proposed to optimise convex objective functions which are composed of a data fidelity term and a regularising term; such objective functions often arise in Computed Tomography problems.

$$\min_{\mu} \underbrace{\|A\mu - p\|_2^2}_{\text{Data fidelity term}} + \underbrace{R(\mu)}_{\text{Regularisation term}} \tag{1}$$

The novelty of Learned Primal Dual, lies in that it replaces the update steps from Chambolle-Pock algorithm with learned neural networks; this additional induced relaxation allows more flexibility, thus making finding optimal updates easier whilst still maintaining the algorithm's overall structure. During the unrolled iterations, physical information are incorporated through repeatedly taking forward projections and back projections, which eliminates solutions that are infidel to the proposed physical geometry. Moreover, learned methods like such (replacing vanilla optimisation algorithm updates with learned neural networks) had already resulted in success in previous works [2], which provides further empirical motivation for employing a similar approach here.

Methodology

In the models, Convolution Neural Networks are the chosen replacements for the proximal operators, as they are known for their strength in imaging problems as well as its low parametrisation. Since the operators are 'learned', they are inferred through training on clinical data in a supervised manner. Specifically, the models are trained on images from the LoDoPab-CT dataset [8], which hosts over 40000 readily processed lungs images.

For benchmarking, the learned models' performances are compared with classical reconstruction methods (as well as other variants) across three different metrics: MSE, SSIM and PSNR. During training, metrics are calculated across all validation set images to determine best epochs. All mathematical transformations are done using `tomosipo` [8], leveraging its faster GPU computation speeds and its compatibility with `Pytorch`.

Results

The results are split into two sections. We first reproduced the baseline results quoted by the paper, under similar physical geometry (exact environment can't be reproduced due to difference in datasets, but the overall trend of observed results is very similar to the original experiment barring small differences). Then, the performances of the standard LPD model is compared with TV-LPD, our custom model, and FBPConvNet [7], another state-of-the-art model used for CT reconstructions, under more adversarial conditions (more noise and limited projection geometry). In the tables below, we quote the results obtained from the experiments.

- Baseline Results:

Method	Image Metric		
	MSE	PSNR	SSIM
FBP	1.957×10^{-3}	23.36	0.3469
TV	3.575×10^{-4}	31.81	0.7473
Learned PDHG	$2.807 \times 10^{-4} \pm 3.820 \times 10^{-4}$	33.68 ± 3.85	0.793 ± 0.151
Learned Primal	$2.591 \times 10^{-4} \pm 3.702 \times 10^{-4}$	34.30 ± 4.11	0.808 ± 0.153
LPD	$1.939 \times 10^{-4} \pm 3.409 \times 10^{-4}$	$\mathbf{37.29 \pm 5.55}$	0.865 ± 0.140
FBPConvNet	$\mathbf{1.206 \times 10^{-4} \pm 2.581 \times 10^{-4}}$	$\mathbf{39.66 \pm 5.39}$	$\mathbf{0.923 \pm 0.083}$

Table 1: Comparison of different image metrics under different methods of reconstruction in standard conditions

- **Extension Results**:

Condition	Image Metric		
	MSE	PSNR	SSIM
LPD Standard	$1.94 \times 10^{-4} \pm 3.41 \times 10^{-4}$	37.29 ± 5.55	0.865 ± 0.140
TV-LPD Standard	$1.88 \times 10^{-3} \pm 3.40 \times 10^{-4}$	37.68 ± 5.73	0.872 ± 0.138
FBPConvNet Standard	$1.21 \times 10^{-4} \pm 2.58 \times 10^{-4}$	39.66 ± 5.39	0.923 ± 0.083
LPD Limited	$1.35 \times 10^{-3} \pm 6.42 \times 10^{-4}$	25.28 ± 2.11	0.643 ± 0.142
TV-LPD Limited	$1.21 \times 10^{-3} \pm 5.80 \times 10^{-4}$	25.74 ± 2.11	0.655 ± 0.144
FBPConvNet Limited	$2.36 \times 10^{-3} \pm 2.11 \times 10^{-3}$	23.34 ± 2.60	0.536 ± 0.110
LPD Sparse	$3.00 \times 10^{-4} \pm 4.30 \times 10^{-4}$	33.70 ± 4.10	0.795 ± 0.161
TV-LPD Sparse	$2.78 \times 10^{-4} \pm 4.24 \times 10^{-4}$	34.39 ± 4.41	0.808 ± 0.161
FBPConvNet Sparse	$4.02 \times 10^{-4} \pm 4.44 \times 10^{-4}$	31.45 ± 3.17	0.729 ± 0.150

Table 2: Comparison of different image metrics under different conditions and methods in adversarial conditions; Limited correspond to 60 projection angles over $[0, 180]$ degrees and 1000 as initial intensity; 'Sparse' correspond 60 projection angles over $[0, 60]$ degrees and 1000 as initial intensity

Strengths, Weaknesses and Future Work

From the results, it can be seen that Learned Primal Dual substantially outperformed classical methods by sizable margins. As for comparison with other State-of-the-Art models, its performance is almost on par with FBPConvNet's under standard conditions and better under noisier settings (where emitted X-ray intensities and number of projection angles are low). This demonstrates LPD's robustness to adversarial conditions, thus making it more suitable to use under low dosage.

However, despite its success, there are still few major caveats present. Firstly, the trained models aren't generalisable. Slight changes in physical geometry would often lead to enlarged deviations in performance, which could be undesirable, as the physical setup in clinical settings often aren't exact. Moreover, the model requires lots of training resources to ensure good performance; the extent of resources required might not always be achievable in clinical settings.

Nevertheless, Learned Primal Dual is still extremely influential. Not only does it bridge deep learning with theories from convex analysis, it also provides a general framework for solving other inverse problems that aren't necessarily related to CT imaging, as its algorithm can be naturally extended to any problems that requires the minimisation of a non-smooth convex function that's compoased of a sum of two terms. There had already been many extensions of LPD in similar domains, such as PET reconstructions [6] and MRI imaging [5].

References

[1] J. Adler and O. Oktem, *Learned primal-dual reconstruction*, IEEE Transactions on Medical Imaging, 37 (2018), p. 1322–1332.

[2] M. Andrychowicz, M. Denil, S. Gomez, M. W. Hoffman, D. Pfau, T. Schaul, B. Shillingford, and N. de Freitas, *Learning to learn by gradient descent by gradient descent*, 2016.

[3] M. Benning and M. Burger, *Modern regularization methods for inverse problems*, Acta Numerica, 27 (2018), p. 1–111.

[4] A. Chambolle and T. Pock, *A first-order primal-dual algorithm for convex problems with applications to imaging*, Journal of Mathematical Imaging and Vision, 40 (2010), p. 120–145.

[5] J. Cheng, H. Wang, L. Ying, and D. Liang, *Model learning: Primal dual networks for fast mr imaging*, 2019.

[6] A. Guazzo and M. Colarieti-Tosti, *Learned primal dual reconstruction for pet*, Journal of Imaging, 7 (2021).

[7] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, *Deep convolutional neural network for inverse problems in imaging*, IEEE Transactions on Image Processing, 26 (2017), pp. 4509–4522.

[8] J. Leuschner, M. Schmidt, D. O. Baguer, and P. Maass, *LoDoPaB-CT, a benchmark dataset for low-dose computed tomography reconstruction*, Sci. Data, 8 (2021), p. 109.

[9] L. I. Rudin, S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D: Nonlinear Phenomena, 60 (1992), pp. 259–268.