Primal-Dual UNet for Sparse View Cone Beam Computed Tomography Volume Reconstruction

Philipp Ernst[1,2]   PHILIPP.ERNST@OVGU.DE

Soumick Chatterjee[1,2]   SOUMICK.CHATTERJEE@OVGU.DE

Georg Rose[2,3]   GEORG.ROSE@OVGU.DE

Andreas Nürnberger[1]   ANDREAS.NUERNBERGER@OVGU.DE

[1] *Faculty of Computer Science, University of Magdeburg, Germany*

[2] *Research Campus STIMULATE, University of Magdeburg, Germany*

[3] *Institute for Medical Engineering, University of Magdeburg, Germany*

Abstract

In this paper, the Primal-Dual UNet for sparse view CT reconstruction is modified to be applicable to cone beam projections and perform reconstructions of entire volumes instead of slices. Experiments show that the PSNR of the proposed method is increased by 10dB compared to the direct FDK reconstruction and almost 3dB compared to the modified original Primal-Dual Network when using only 23 projections. The presented network is not optimized wrt. memory consumption or hyperparameters but merely serves as a proof of concept and is limited to low resolution projections and volumes.

Keywords: Sparse view CT, cone beam CT, deep learning, convolutional neural networks

1. Introduction

During CT-guided medical interventions, surgeons and patients are exposed to harmful X-radiation. Keeping the dose low is essential but also results in high noise or streaking artifacts in the reconstructions. Ernst et al. (2022) proposed the Primal-Dual UNet, based on Primal-Dual Network (Adler and Öktem, 2018), for sparse view parallel and fan beam CT reconstruction. In medical interventions, however, surgeons make use of cone beam CT for imaging. Therefore, the main contributions of this work are: (i) modifying the network to process cone beam projections and (ii) reconstruct entire volumes instead of axial slices.

2. Methods

The network architecture used in this work is a modified Primal-Dual UNet (Ernst et al., 2022). The two-dimensional convolutions of the dual space blocks were replaced with their three-dimensional counterparts. The two-dimensional UNet in the primal space was replaced with a three-dimensional UNet by replacing convolutions, batch normalizations, average poolings and linear upsamplings with their three-dimensional counterparts. Instead of the parallel or fan beam projection layer, a cone beam geometry (detector: 310×240px, 1.232mm pixel size; SID=160mm; SDD=400mm) on a circular trajectory was used. The FBP reconstruction layer was replaced with its FDK counterpart.

For comparability, the data normalization, the L_1 loss function, the Adam optimizer (lr=1e-3, β_1=0.9, β_2=0.999) and the number of epochs (151) were kept the same. The effective batch size was set to 16. Training data was simulated by downsampling LungCT-Diagnosis (Grove et al., 2015) volumes (42/9/10 for training/validation/test) to cubes with side lengths of 128px = 128mm due to memory limitations. Random flips, rotations and scalings of the volumes were used as augmentation during training. Sparse views were simulated by retaining every 8th or 16th of 360 equiangular projections (called *Sparse 8* or *Sparse 16*, respectively).

3. Results

Table 1: Mean and standard deviation over all axial test slices for *Sparse 16*.

Model/Method	SSIM [%]	PSNR [dB]	RMSE [HU]
FDK	43.54±8.27	17.92±2.64	388.57±108.15
FDKConvNet	67.37±8.99	24.72±1.93	177.30±60.94
Primal-Dual Network	69.87±7.66	25.19±2.08	169.22±64.35
Primal-Dual UNet	78.76±7.50	27.93±2.33	128.89±57.78

Table 1 shows the results of the different models evaluated on the test set. All models outperform the direct sparse view FDK reconstruction by a large margin, while the Primal-Dual models further increase the quality compared to FDKConvNet (Jin et al., 2017). The proposed Primal-Dual UNet results in the lowest errors. Wilcoxon signed-rank tests reveal that the proposed model significantly outperforms any other model/method pair-wise (p-value < 0.5%).
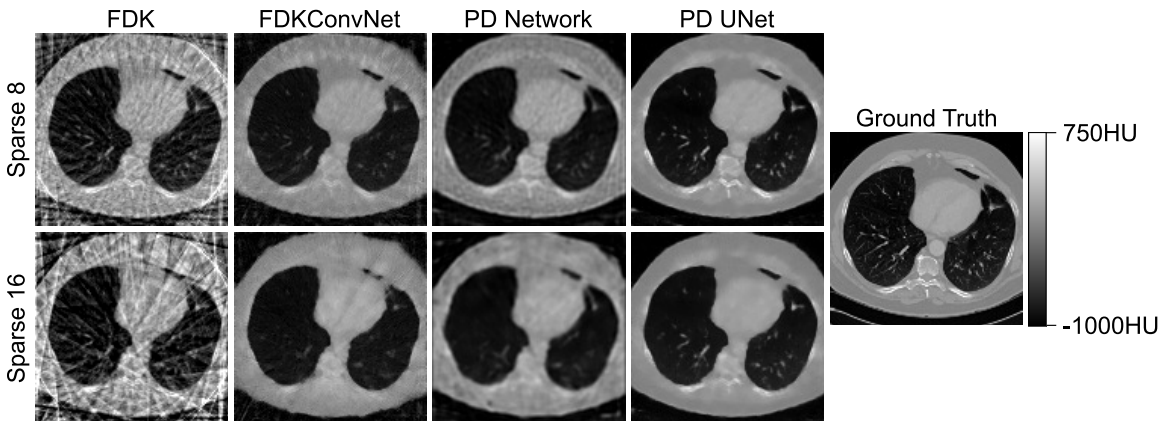

Figure 1: Exemplary axial slice from different models/methods.

Figure 1 shows an exemplary axial slice from the different models for *Sparse 8* (top row) and *Sparse 16* (bottom row). FDKConvNet does not seem to have learned anatomical

structures and merely attempts to suppress streaking artifacts. Primal-Dual Network produces results that look blurrier with more low frequency noise than FDKConvNet's outputs but anatomical structures, e.g. costal cartilage, are preserved better. The reconstructions of Primal-Dual UNet are superior compared to Primal-Dual Network. Tissues with high attenuation coefficients are clearly distinguishable from soft tissues and edges are well preserved, e.g. vertebrae, even for the higher sparsity factor *Sparse 16*.

4. Discussion and Conclusion

The proposed Primal-Dual UNet for cone beam reconstruction not only outperforms other methods – Primal-Dual Network in particular – in quality but also in memory requirements and is more than twice as fast during both training and inference while retaining data consistency wrt. the cone beam projections, as opposed to FDKConvNet. Moreover, the training of the proposed network is much more stable compared to Primal-Dual Network. However, the main limitation is still the memory consumption: with enabled mixed precision, the inference takes $\sim$9GB of GPU RAM for even these unrealistically low resolution volumes and projections and a batch size of 1. Training consumes even more space: a *Sparse 4* version of Primal-Dual Network did not even fit into the 48GB of an Nvidia RTX A6000.

Since usually, not the entire volume needs to be reconstructed during an intervention, future work will focus on reducing the memory requirements by only reconstructing volumes of interest. Moreover, this preliminary work is based on simulations and has to be evaluated for real cone beam CT data. The Pytorch implementation is available on Github[1].

References

Jonas Adler and Ozan Öktem. Learned primal-dual reconstruction. *IEEE Transactions on Medical Imaging*, 37(6):1322–1332, 2018. doi: 10.1109/TMI.2018.2799231.

Philipp Ernst, Soumick Chatterjee, Georg Rose, Oliver Speck, and Andreas Nürnberger. Sinogram upsampling using primal-dual unet for undersampled ct and radial mri reconstruction. In *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, 2022.

Olya Grove, Anders E. Berglund, Matthew B. Schabath, et al. Data from: Quantitative computed tomographic descriptors associate tumor shape complexity and intratumor heterogeneity with prognosis in lung adenocarcinoma, 2015. URL https://wiki.cancerimagingarchive.net/x/8IUiAQ.

Kyong Hwan Jin, Michael T. McCann, Emmanuel Froustey, and Michael Unser. Deep convolutional neural network for inverse problems in imaging. *IEEE Transactions on Image Processing*, 26(9):4509–4522, 2017. doi: 10.1109/TIP.2017.2713099.

---

1. https://github.com/phernst/pd-unet-conebeam