BPConvNet for compressed sensing recovery in bioimaging

Kyong Hwan Jin[†], Michael T. McCann[†‡], Michael Unser[†]

[†] Biomedical Imaging Group, EPFL, Lausanne, Switzerland
[‡] Center for Biomedical Imaging, Signal Processing Core, EPFL, Lausanne, Switzerland
Email: {kyong.jin,michael.mccann,michael.unser}@epfl.ch

I. Introduction

Iterative reconstruction methods have become the standard approach to solving inverse problems in imaging including denoising [1], [2], [3], deconvolution [4], and interpolation [5]. With the appearance of compressed sensing [6], our theoretical understanding of these approaches evolved further with remarkable outcomes [7], [8]. These advances have been particularly influential in the field of biomedical imaging, e.g., in magnetic resonance imaging (MRI) [9] and X-ray computed tomography (CT) [10]. A more recent trend is deep learning [11], which has arisen as a promising framework providing state-of-the-art performance for image classification [12] and segmentation [13], regression-type neural networks [14], [15], [16]. In this paper, we explore the relationship between CNNs (Convolutional neural netwrok, ConvNet) and iterative optimization methods for one specific class of inverse problems: those where the normal operator associated with the forward model is a convolution. Based on this connection, we propose a method for solving these inverse problems by combining a fast, approximate solver with a CNN. We demonstrate the approach on low-view CT reconstruction and accelerated MRI using residual learning [16] and multilevel learning [13].
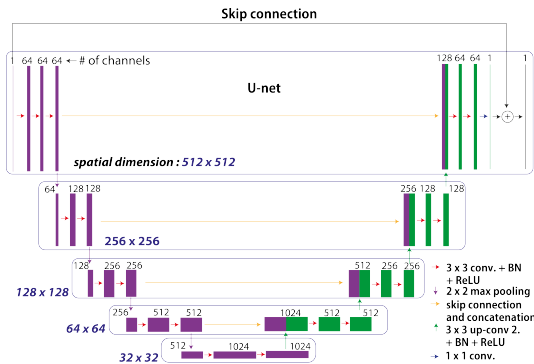
II. Method and Results

Fig. 1. Structure of the proposed convolutional network.

Method (BPConvNet : BackProjection CONVolutional Network) : Our goal here is not to follow iterative methods (e.g. by building a network that corresponds to an unrolled version of some iterative method), but rather to explore a state-of-the-art CNN architecture. We base our CNN on the U-net [13], which was originally designed for segmentation. There are several properties of this architecture that recommend it for our purposes.

Multilevel decomposition: The U-net explores a dyadic scale decomposition based on max pooling, so that the effective filter size in the middle layers is larger than that of the early and late layers. This is critical for our application because the filters corresponding to H^*H (and its inverse) may have non-compact support, e.g. in CT. Thus, a CNN with a small, fixed filter size may not be able to effectively invert H^*H. This decomposition also has a nice analog to the use of multiresolution wavelets in iterative approaches.

Multichannel filtering: U-net employs multichannel filters, such that there are multiple feature maps at each layer. This is the standard approach in CNNs [12] to increase the expressive power of the network [17]. The multiple channels also have an analog in iterative methods: In the ISTA (iterative shrinkage thresholding algorithm) formulation [18], we can think of the wavelet coefficient vector as being partitioned into different channels, with each channel corresponding to one wavelet subband [19], [20]. The CNN architecture greatly generalizes this by allowing filters to make arbitrary combinations of filters.

Residual learning: As a refinement of the original U-net, we add a skip connection [16] between input and output, which means that the network actually learns the difference between input and output. This approach mitigates the vanishing gradient problem [21] during training. This yields a noticeable increase in performance compared to the same network without the skip connection.

Implementation details: We made two additional modification to U-net. First, we use zero-padding so that the image size does not decrease after each convolution. Second, we replaced the last layer with a convolutional layer which reduces the 64 channels to a single output image. For the multichannel data, the number of channels in the last layer will be changed.

Results The experiments provide strong evidence for the feasibility of the BPConvNet for sparse-view CT reconstruction and accelerated MRI. In CT, down sampling arose regularly in angular dimension, and in MRI, we chose a variable density down sampling mask with the factor of 6. In real datasets, the SNR improvement of the BPConvNet came from its ability to preserve fine details in the images. This points to one advantage of the proposed method over iterative methods: the iterative methods must explicitly impose regularization, while the BPConvNet effectively learns a regularizer from the data. The computation time for the BPConvNet was about 200 ms for the FBP and 200∼300 ms in GPU for the CNN for a 512×512 image. In accelerated MRI, BPConvNet spent 100∼150ms in GPU for the CNN for a $320 \times 320 \times 8$ volume (multichannel images [22]). The numbers of training dataset are 475 for both modalities.

III. Conclusion

In this paper, we proposed the *(F)BPConvNet*. The structure of the proposed CNN is based on U-net, with the addition of residual learning. This approach was motivated by the convolutional structure of several biomedical inverse problems, including CT, MRI, etc. Specifically, we showed conditions on a linear operator that ensure that its normal operator is a convolution.

Acknowledgment

Ground truth | **FBP** SNR 13.43 | **TV** SNR 24.89 | **FBPConvNet** SNR 28.53
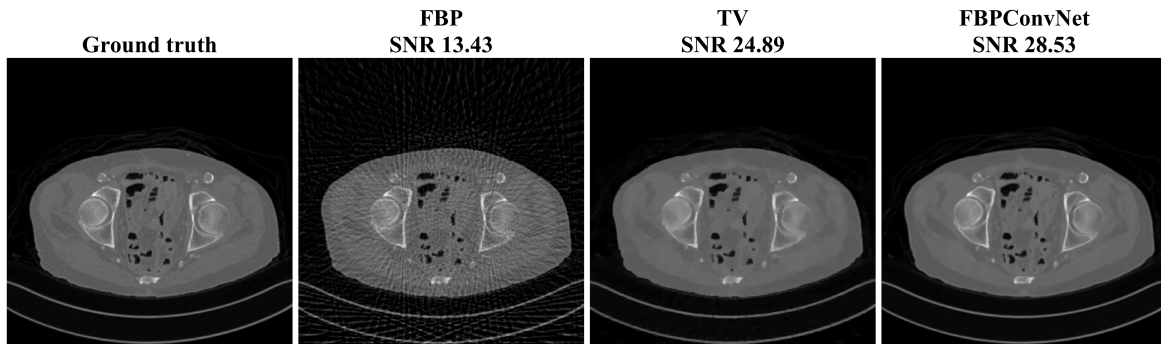
Fig. 2. Ground truth image and reconstructed images of experimental dataset from 50 views using FBP in CT, TV regularized convex optimization [10] ('TV'), and the FBPConvNet.

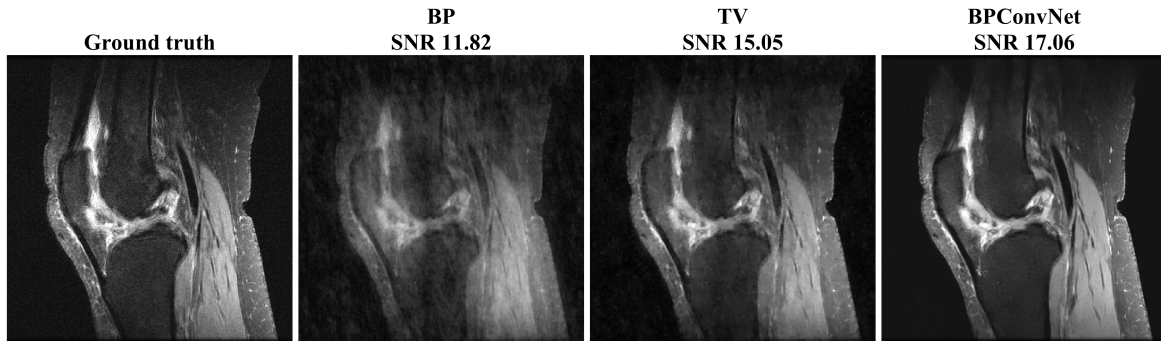**Ground truth** | **BP** SNR 11.82 | **TV** SNR 15.05 | **BPConvNet** SNR 17.06

Fig. 3. Ground truth image and reconstructed images of experimental dataset [22] from 6-fold down sampling in MRI using zero inserted inverse Fourier transform ('BP'), TV regularized convex optimization [23] (channel by channel processing) ('TV'), and the BPConvNet.

REFERENCES

[1] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, pp. 259–268, 1992.

[2] A. Chambolle, "An algorithm for total variation minimization and applications," *J. Math. imaging and vision*, vol. 20, no. 1-2, pp. 89–97, 2004.

[3] F. Luisier, T. Blu, and M. Unser, "A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding," *IEEE Transactions on image processing*, vol. 16, no. 3, pp. 593–606, 2007.

[4] T. F. Chan and C.-K. Wong, "Total variation blind deconvolution," *IEEE transactions on Image Processing*, vol. 7, no. 3, pp. 370–375, 1998.

[5] P. Thévenaz, T. Blu, and M. Unser, "Interpolation revisited [medical images application]," *IEEE Transactions on medical imaging*, vol. 19, no. 7, pp. 739–758, 2000.

[6] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.

[7] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med.*, vol. 58, no. 6, pp. 1182–1195, 2007.

[8] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Physics in medicine and biology*, vol. 53, no. 17, p. 4777, 2008.

[9] H. Jung, K. Sung, K. S. Nayak, E. Y. Kim, and J. C. Ye, "k-t FOCUSS: A general compressed sensing framework for high resolution dynamic MRI," *Magnetic Resonance in Medicine*, vol. 61, no. 1, pp. 103–116, Jan. 2009.

[10] M. T. McCann, M. Nilchian, M. Stampanoni, and M. Unser, "Fast 3D reconstruction method for differential phase contrast X-ray CT," *Opt. Express*, vol. 24, no. 13, pp. 14 564–14 581, Jun 2016.

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[13] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015, pp. 234–241.

[14] L. Xu, J. S. Ren, C. Liu, and J. Jia, "Deep convolutional neural network for image deconvolution," in *Advances in Neural Information Processing Systems*, 2014, pp. 1790–1798.

[15] C. Dong, C. C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 38, no. 2, pp. 295–307, 2016.

[16] J. Kim, J. K. Lee, and K. M. Lee, "Accurate image super-resolution using very deep convolutional networks," in *Proc. of IEEE Converence on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[17] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *European Conference on Computer Vision*. Springer, 2014, pp. 818–833.

[18] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *arXiv preprint arXiv:1611.03679*, 2016.

[19] A. L. Da Cunha, J. Zhou, and M. N. Do, "The nonsubsampled contourlet transform: theory, design, and applications," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3089–3101, 2006.

[20] S. Mallat, *A wavelet tour of signal processing*. Academic press, 1999.

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *arXiv preprint arXiv:1512.03385*, 2015.

[22] [Online]. Available: http://mridata.org/fullysampled/knees

[23] T. Goldstein and S. Osher, "The split bregman method for L1-regularized problems," *SIAM J. Imaging Sci.*, vol. 2, no. 2, pp. 323–343, 2009.