2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)
April 4-7, 2018, Washington, D.C., USA

3D BPCONVNET TO RECONSTRUCT PARALLEL MRI

Kyong Hwan Jin and Michael Unser

Biomedical Imaging Group, École polytechnique fédérale de Lausanne (EPFL), Switzerland

ABSTRACT

In recent years, compressed sensing techniques have been applied to the reconstruction of parallel magnetic resonance (MR) images. Particularly for 3D MR signal, it is crucial to acquire fewer samples to reduce the distortions caused by long-time acquisitions (e.g., motion, organ dynamics). Motivated by the recent success of ConvNet in 2D image reconstruction, we propose to extend the approach to 3D volume reconstruction and parallel MR imaging. The structure of the proposed network follows FBPConvNet with additional coil compression by SSoS and wavelet transform. A parallelism using two GPUs is also applied to overcome the memory shortage. The proposed method is able to reconstruct a $(320 \times 320 \times 256 \times 8)$ volume in less than 10s with 2 GPUs, while the iterative algorithm ℓ_1-ESPIRiT takes over 5 min in CPU.

Index Terms— Convolutional neural network, FBPConvNet, parallel MRI, 3D MRI reconstruction

1. INTRODUCTION

The full-scale acquisition of magnetic resonance imaging (MRI) suffers from artifacts caused by patients' movements, breathing, and other internal dynamics. To bring advantage of coil redundancy, parallel MRI (pMRI) was proposed to reduce the duration of acquisition, then attenuating the motion artifacts [1, 2]. A further reduction is achieved by combining it with compressed sensing, where a random subsampling of the k-space is deployed to alleviate the aliasing that a regular subsampling would induce [3]. In particular, researchers have investigated iterative reconstructions with regularization such as Tikhonov (ℓ_2) or sparsity (ℓ_1) norms [4, 5]. In another approach, low-rank matrix completion has been applied for Hankel matrices with missing samples in k-space [6].

Each physical coil is associated to a limited spatial field of view. This association is recorded in a map called a sensitivity encoding. Some reconstruction methods require their explicit knowledge [1], possibly obtained through a dedicated calibration step. Another strategy is to rely on autocalibration signals (ACS) to estimate the reconstruction kernel [2]. Both solutions suffer from inconsistencies due to the motions experienced during the additional time needed to deal with the acquisition of the sensitivity maps.

By contrast, the approach [6] requires no coil-sensitivity estimates; however, it is not applicable to 3D because of the prohibitive size of the annihilating Hankel matrices. In [5], data-driven kernels are estimated from ACS regions to remove the aliasing that results from k-space subsampling, but they require high-accuracy hyper-parameters and intensive computations.

Recently, deep neural networks have been applied to inverse problems for pMRI. They require no tuning of the hyper-parameters and are usually computationally faster than compressed-sensing algorithms. Moreover, they tend to achieve better visual reconstruction quality [7]. Unfortunately, they are often memory-hungry and cannot guarantee that the reconstruction is consistent with the data. While hybrid methods were proposed in [8, 9] to solve these issues, all previous works have been restricted to 2D images.

In this paper, we describe a first attempt at the calibrationless reconstruction of 3D pMRI volumes based on convolutional networks. This work is an extension of our 2D convolutional network [10]. Here, we propose a parallelization that reduces the memory requirements of 3D BPConvNet and spreads the computations over several (here, two) GPUs. We use a square-root of sum-of-squares operation (SSoS) volume as input and train our network without coil-sensitivity maps, which has the added benefit of shortening the physical acquisition time, thus resulting in fewer motion-related artifacts. Since our multilayered convolutional neural network involves a multiresolution decomposition, it naturally results in an increased size of the receptive fields.

2. METHODS

2.1. BPConvNet for Multidimensional pMRI

In [10], FBPConvNet employed U-Net [11] with external skip-connection to reconstruct CT images for a small set of views. FBPConvNet allows us to reconstruct signals by adapting the architecture of U-Net [11] with skip-connection [12]. The scheme also exploits the analogy between convolutional networks and iterative solvers for the class of inverse

This project has received funding from the European Unions Horizon 2020 Framework Programme for Research and Innovation under Grant Agreement 665667 (call 2015). Authors would also like to thank the support of NVIDIA Corporation.

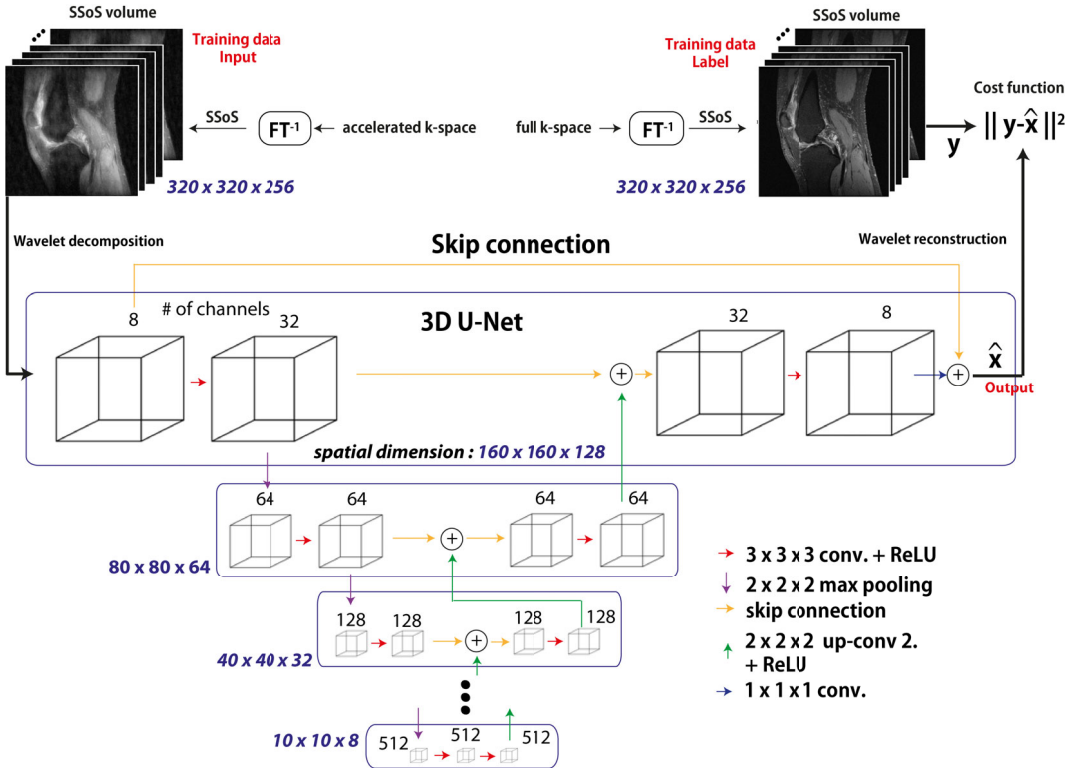

Fig. 1. Architecture of 3D BPConvNet extended from FBPConvNet [10]. In 3D BPConvNet, all convolutional layers are comprised of 3D based convolutions. Then, FT^{-1} is three-dimensional inverse Fourier transform, SSoS means the square-root of sum-of-squares operation, and ReLU is a rectified linear unit, conv. is a convolution, and up-conv. is a convolution transpose.

problems where the normal operator of the forward model is a convolutional operation. This implies that (F)BPConvNet is also applicable to 2D or 3D pMRI because MRI has the required structure[1]. Indeed, it can be shown that, for any k-space sampling pattern in 3D pMRI, a normal operator from (F)BPConvNet as $M = H^H H$ is always a convolutional operation, where H is the forward model of pMRI and M is the kernel obtained from the inverse Fourier transform of the sampling mask.

Because the sampling trajectory is defined on a complete discrete k-space, the associated kernel has values on an entire spatial domain. Accordingly, the ideal size of the receptive field for a convolutional network would have to match the size of the target 3D volume. This full ROI scheme would then generate huge intermediate feature maps, as shown in Figure 1, and would cause serious memory shortage. To overcome this, we apply a model parallelism that we describe next.

As shown in Figure 1, we first compress the multicoil information into a single-channel volume by an SSoS operation in order to eliminate complex values of input and label data. The SSoS volume ($S \in \mathbb{R}^N, N = XYZ$) is given as

$$S(f(\mathbf{x})) = \sqrt{|f_1(\mathbf{x})|^2 + |f_2(\mathbf{x})|^2 + \cdots + |f_C(\mathbf{x})|^2},$$

where $f_i \in \mathbb{C}^N$ is the ith coil 3D pMRI signal and $\mathbf{x} = (x, y, z)$. The effect of the SSoS operation over subsampled parallel MR images can be modeled by

$$S(f_d(\mathbf{x})) = \sqrt{|M(\mathbf{x}) * f_1(\mathbf{x})|^2 + \cdots + |M(\mathbf{x}) * f_C(\mathbf{x})|^2},$$

where f_d is a complex-valued image obtained from subsampling in k-sapce and $M(\mathbf{x})$ is the kernel associated with the sampling mask. This implies that the reconstruction performance after coil compression by SSoS is not much different than the reconstruction performance without SSoS. This SSoS process yields calibration-less reconstruction in 3D BPConvNet.

Additionally, inspired by the work on denoising in low-dose CT [13], we propose to use a wavelet transform before feeding both input and label volumes to the network. The authors of [13] claim that the directional wavelet transform exploits the intra- and inter-band correlations resulting in an effective suppression of CT-specific noises. Moreover, this preserves shift-invariance due to the filtering structure of the wavelet transform.

[1]pMRI is categorized into BPConvNet because there is no filtering step before the direct inverse or adjoint operation. BP (backprojection) simply corresponds to a zero-filled inverse Fourier transform.

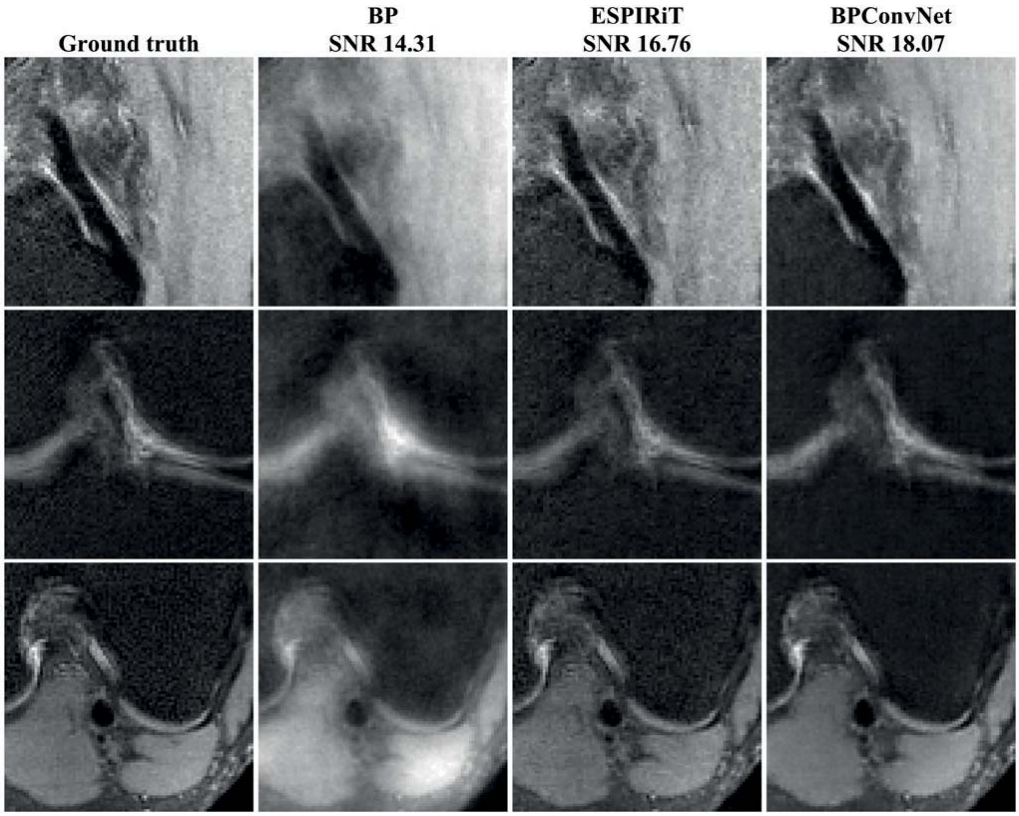

Fig. 2. Reconstructed region of interest from a knee dataset (http://mridata.org/) [14]. BP: Backprojection (zero-filled inverse Fourier transform). ESPIRiT: [5]. BPConvNet: proposed method.

2.2. Implementation Details

The architecture of the proposed network is based on a 3D convolutional network inspired from BPConvNet [10]. Three-dimensional convolutions take the place of 2D convolution layers. Similarly, ReLU, max pooling, and interpolation layers are replaced by their 3D counterparts. In addition, we apply zero-padding after each convolution to keep the original volume size. Besides the domain of applications, the primary difference with 3D U-Net [15] is external skip-connection as in [10]. Also, we replace the concatenation layers with summation layers ('plus' notation in Figure 1) in order to save on the memory allocations of the intermediate feature maps.

To favor scalability, we use a model parallelism with two GPUs. We partition the whole network in several pieces and assign fragmented tasks to different devices. Memory-intensive calculations such as the back-propagation of the intermediate feature maps of different layers are assigned to different GPUs. In our case, the first GPU handles the first half of the network in Figure 1 while the second GPU computes the other half. As pre-processing, we used the Haar wavelet transform (`waverec3` and `wavedec3`) to exploit channel diversity.

We used two TITAN GTX graphic cards (Maxwell and Pascal architectures) with i7-4770k CPU. The codes were written in Python 2.7 on the Tensorflow library[2]. We used the 'ADAM' learning algorithm in the Tensorflow library with learning rate 10^{-3} and batchsize 1. The training phase took three days.

3. RESULTS

Among the 20-patient dataset from [14], we assigned 17 patients to training and 3 to validation. Each dataset consists in an 8-channel HD knee coil volume (GE Healthcare, Milwaukee, WI, USA) [14]. The scans were obtained by 3D FSE CUBE sequences. The size of the data is $(320\times320\times256\times8)$ $(X, Y, Z, Coil)$. To avoid overfitting issues, data augmentation (flipping around each axis) was applied. We simulated

[2]Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dan Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu and Xiaoqiang Zheng, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems," 2015, Software available from tensorflow.org

Table 1. Comparison results from averages of three validation sets for the knee data. For a fair comparison of the CPU times, we always perform in CPU the inverse fast Fourier transforms (BP) and the wavelet transforms (WT). ℓ_1-ESPIRiT used only CPU, Patch CNN used 1 GPU and BPConvNet used 2 GPUs.

x6	BP	ℓ_1-ESPIRiT	Patch CNN	BPConvNet
SNR (dB)	13.68	17.63	16.67	17.96
Time (s)	1.91	321.28	27.12	8.91

a subsampling by 6 by applying a binary mask where elements follow a centered Gaussian distribution in k-space.For ESPIRiT [5], we provide the central ACS (fully sampled region) corresponding to a $(21 \times 21 \times 21)$ cube around the zero frequency.

We evaluated the performance of l_1-ESPIRiT [5] and patch-based CNN. Patch-based CNN ('Patch CNN' in Table 1) used the same architecture of 3-level 3D BPConvNet without wavelet decomposition/reconstruction in order to confirm the effect of whole ROI processing. The size of a patch volume is $(64 \times 64 \times 64)$. In Figure 2, BP images correspond to a zero-filled inverse Fourier transform and show severe blurring effects over the entire ROI. ESPIRiT and BPConvNet clearly suffer from less severe artifacts in all imaging planes (XY, XZ, YZ). BPConvNet tends to produce less noisy reconstructions than ESPIRiT.

4. CONCLUSION

To the best of our knowledge, this is the first neural-network reconstruction of 3D parallel-MRI. It provides full-ROI scheme without requiring coil-sensitivity estimates. Our network is competitive with the state-of-the-art ESPIRiT and is significantly faster (a sub-10-seconds reconstruction time). By comparison, the iterative algorithm l_1-ESPIRiT obtains a similar quality but takes over 5 minutes when implemented in CPU. Even when implemented with 4 GPUs and a 20-core CPU, it still requires nearly 1 minute for problems of similar size [16].

5. REFERENCES

[1] Klaas P Pruessmann, Markus Weiger, Markus B Scheidegger, and Peter Boesiger, "SENSE: Sensitivity encoding for fast MRI," *Magn. Reson. Med.*, vol. 42, no. 5, pp. 952–962, 1999.

[2] Mark A Griswold, Peter M Jakob, Robin M Heidemann, Mathias Nittka, Vladimir Jellus, Jianmin Wang, Berthold Kiefer, and Axel Haase, "Generalized autocalibrating partially parallel acquisitions (GRAPPA)," *Magn. Reson. Med.*, vol. 47, no. 6, pp. 1202–1210, 2002.

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

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

[5] Martin Uecker, Peng Lai, Mark J Murphy, Patrick Virtue, Michael Elad, John M Pauly, Shreyas S Vasanawala, and Michael Lustig, "ESPIRiT—An eigenvalue approach to autocalibrating parallel MRI: Where SENSE meets GRAPPA," *Magn. Reson. Med.*, vol. 71, no. 3, pp. 990–1001, 2014.

[6] Kyong Hwan Jin, Dongwook Lee, and Jong Chul Ye, "A general framework for compressed sensing and parallel MRI using annihilating filter based low-rank Hankel matrix," *IEEE Trans. Comput. Imag.*, vol. 2, no. 4, pp. 480–495, 2016.

[7] Dongwook Lee, Jaejun Yoo, and Jong Chul Ye, "Deep residual learning for compressed sensing MRI," in *Proc. of IEEE Int. Symp. Biomedical Imaging*, 2017, pp. 15–18.

[8] Shanshan Wang, Zhenghang Su, Leslie Ying, Xi Peng, Shun Zhu, Feng Liang, Dagan Feng, and Dong Liang, "Accelerating magnetic resonance imaging via deep learning," in *Proc. of IEEE Int. Symp. Biomedical Imaging*, 2016, pp. 514–517.

[9] Yan Ynag, Jian Sun, Huibin Li, and Zongben Xu, "Deep ADMM-net for compressive sensing MRI," in *Advances in Neural Information Processing Systems*, 2016, pp. 10–18.

[10] Kyong Hwan Jin, Michael T McCann, Emmanuel Froustey, and Michael Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, 2017.

[11] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Int. Conf. Medical Image Computing and Computer-Assisted Intervention*, 2015, pp. 234–241.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *Proc. of IEEE Computer Soc. Conf. on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[13] Eunhee Kang, Junhong Min, and Jong Chul Ye, "A deep convolutional neural network using directional wavelets for low-dose x-ray CT reconstruction," *Med. Phys.*, vol. 44, no. 10, 2017.

[14] Kevin Epperson, Anne M. Sawyer, Michael Lustig, Marcus Alley, Martin Uecker, Patrick V. Virtue, Peng Lai, and Shreyas Vasanawala, "Creation of fully sampled MR data repository for compressed sensing of the knee," in *SMRT Conf., Salt Lake City, UT*, 2013.

[15] Özgün Çiçek, Ahmed Abdulkadir, Soeren S Lienkamp, Thomas Brox, and Olaf Ronneberger, "3D U-Net: Learning dense volumetric segmentation from sparse annotation," in *Int. Conf. Medical Image Computing and Computer-Assisted Intervention*, 2016, pp. 424–432.

[16] Martin Uecker, Frank Ong, Jonathan I Tamir, Dara Bahri, Patrick Virtue, Joseph Y Cheng, Tao Zhang, and Michael Lustig, "Berkeley advanced reconstruction toolbox," in *Proc. Intl. Soc. Mag. Reson. Med*, 2015, vol. 23, p. 2486.