DEEP LEARNING FOR MRI RECONSTRUCTION USING A NOVEL PROJECTION BASED CASCADED NETWORK

Deniz Kocanaogullari *Ender M. Eksioglu*

Istanbul Technical University
Electronics and Communication Engineering Department
Istanbul, Turkey
{kocanaogullarid, eksioglue}@itu.edu.tr

ABSTRACT

After their triumph in various classification, recognition and segmentation problems, deep learning and convolutional networks are now making great strides in different inverse problems of imaging. Magnetic resonance image (MRI) reconstruction is an important imaging inverse problem, where deep learning methodologies are starting to make impact. In this work we will develop a new Convolutional Neural Network (CNN) based variant for MRI reconstruction. The developed algorithm is based on the recently proposed deep cascaded CNN (DC-CNN) structure. In the original DC-CNN network, the regular data consistency (DC) layer acts as a periodic enforcer of data fidelity. Here, we introduce a novel DC layer structure which also calculates the projection of the intermediary image estimate onto the unobserved subspace of the Fourier domain. These intermediary innovation images are saved and reutilized in the final stage of the overall structure via skip connections. This enhanced cascaded deep network results in improved reconstruction performance when compared to not only the original DC-CNN structure but also another recent deep network approach, where similar number of parameters get utilized in the competing deep methods.

***Index Terms*—** Image reconstruction, magnetic resonance image, MRI, deep learning, CNN, cascaded network

1. INTRODUCTION

Deep learning methodologies using multilayered neural network structures have been at the forefront of recent attempts at modeling data. The revolution started with object recognition and classification type of algorithms, and it has spread to various other applications [1, 2]. A rather recent stronghold for the application of deep learning (DL) and convolutional neural networks (CNNs) are inverse problems in imaging where the actual desired outputs are images [3, 4]. The inverse problems where CNNs have been applied include but are not limited to image denoising [5], image deblurring and deconvolution, superresolution and image reconstruction [6], including both computerized tomography (CT) and magnetic resonance image (MRI) reconstruction. Recently, the important problem of MRI reconstruction has been addressed by DL practitioners using a plethora of approaches. The utilized approaches include both k-domain methods and image domain methods. In the k-domain methods, the deep networks are trained on the actual Fourier domain data [7]. In the image domain methods, the zero-filled reconstruction image is taken to be the input of the network, and the ground-truth image is enforced as the desired output. The image-domain strategy using a variety of different networks has been the prevailing approach for DL based MRI reconstruction [8–10]. The graphical description of the image-domain strategy for deep learning based MRI reconstruction is given in Fig. 1.

After the initial usage of vanilla CNNs [8], more advanced networks such as encoder-decoder networks, U-Nets [10] and generative adversarial networks (GANs) [11] have been employed in MRI reconstruction. A recent and outstanding method has been the deep cascaded CNN (DC-CNN) network as introduced in [9]. This network utilizes a series of separate CNNs interspersed with data consistency (DC) layers. The DC layer in a MRI reconstruction setting reenforces the actual acquired k-space data onto an intermediary reconstructed image.

In this work, we introduce a new projection based updated data consistency (UDC) layer, which has two outputs as opposed to the single rectified output of the regular DC layer. The UDC layer has a secondary output which stores the residual image which presents the innovation at that particular stage projected onto the Fourier subspace outside the observation domain. The sequence of innovation carrying residual images are transported to the input layer of a final CNN stage via skip connections. The final reconstruction network works on the residual images and the intermediary rectified image

This work was supported by BAP Research Fund of the Istanbul Technical University. Project Number: MGA-2018-41028.

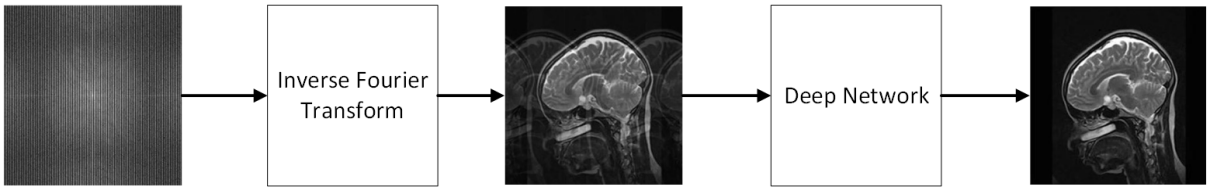

Fig. 1: The deep learning based MRI reconstruction framework utilized for the newly developed approach.

to create the ultimate reconstruction. MRI reconstruction experiments using a rather large database of MR images suggest that the introduced novel DL strategy maintains the performance superiority of CNN based methods over advanced model based reconstruction strategies, and it also displays improved performance when compared to the original DC-CNN MRI approach. The next section gives an overview of the problem under consideration and introduces the novel DL structure to solve this problem. In the simulations section we give detailed performance comparisons with competing DL based and state-of-the-art model (sparsity) based algorithms.

2. PROJECTED DEEP CASCADE CNN FOR MRI RECONSTRUCTION

2.1. MR Reconstruction Prior

The data acquisition forward model for MRI reconstruction can be given as follows [12].

$$\boldsymbol{y} = \boldsymbol{\mathcal{F}}_{\Omega}\hat{\mathbf{x}} + \boldsymbol{\eta} \tag{1}$$

Here, $\hat{\mathbf{x}} \in \mathbb{C}^N$ denotes the underlying original image in a vectorized form. The operator $\boldsymbol{\mathcal{F}}_{\Omega} : \mathbb{C}^N \rightarrow \mathbb{C}^M$, with $M < N$, designates the undersampled Fourier transform operator, with $\Omega \in \{1, 2, \ldots, N\}^M$ being the set of indices for the subsampled Fourier data stored in $\boldsymbol{y}$. Ω specifies the positions of the samples included in the acquired Fourier (k-space) data [13]. Hence, Ω also dictates the downsampling ratio M/N. $\boldsymbol{\eta} \in \mathbb{C}^M$ designates the additive noise. Here, we have outlined the general case of complex valued MRI data. However, in our study of the proposed MRI reconstruction system and in the simulations we will be limiting our studies to real valued data without additive observation noise. MRI reconstruction boils down to the ill-conditioned inverse problem of estimating $\hat{\mathbf{x}}$ from $\boldsymbol{y}$.

A standard way of reconstructing $\hat{\mathbf{x}}$ is by straightforward backprojection $\mathbf{x}_{\text{zf}} = \boldsymbol{\mathcal{F}}_{\Omega}^{H}\boldsymbol{y}$, also called as zero-filling (ZF) reconstruction. Until recently model based reconstruction methods using iterative solutions to regularized variational formulations offered the best performance results for MRI reconstruction. Total variation (TV), wavelet transforms or self-similarity have been widely utilized as source of sparsity models for compressed sensing (CS) MRI reconstruction algorithms [12–15]. CS-MRI has already found its way into patents and actual current MRI hardware [16]. However, CS-MRI does have its own pitfalls. Firstly, the iterative solutions to the CS variational problems are computationally expensive and time-consuming. Secondly, the utilized models are not universal for the overall domain of the MR images, and the performance of a particular image model might depend on the type of the reconstructed MRI image. As another disadvantage, the iterative algorithms incorporate a plethora of coefficients which necessitate fine-tuning for different reconstruction scenarios.

2.2. Deep Learning Meets MRI Reconstruction

Inverse problems in imaging form a recent stronghold for DL approaches. DL based solutions have been proposed for a variety of inverse imaging problems including the fundamental problem of image denoising, image superresolution, image inpainting, deconvolution, CS reconstruction [4], and last but not least MRI reconstruction. One main class of DL based MR reconstruction methods are based on deep networks which transform a ZF based initial image estimate to the final reconstructed image. Networks of different types such as regular CNNs [8], U-Nets [10] or more recently GANs [11] have been utilized in this image domain approach. There have also been attempts to learn networks which start the reconstruction directly from the k-domain data, though these attempts are fewer in number [7]. One recent and effective DL approach for MR reconstruction has been the deep cascaded CNN (DC-CNN) network developed in [9]. The DC-CNN includes a succession of individual CNN stages interlinked together by so-called DC layers. The output of each individual CNN stage is an intermediary reconstructed image, and this output is given as input to the subsequent DC layer. The DC layer acts as a block which reimposes the original k-space data onto the intermediary reconstructed image. In its noise-free formulation as implemented in [9], the DC layer simply replaces the Fourier coefficients of the input image with the actual measured k-space data. The input-output relationship of the DC layer in the noise-free setting can be formulated as follows.

$$\mathbf{x}_{\text{out}} = \boldsymbol{\mathcal{F}}^{-1}\left\{\overline{\mathbf{M}} \circ (\boldsymbol{\mathcal{F}}\mathbf{x}_{\text{in}}) + \widehat{\boldsymbol{y}}\right\} \tag{2}$$

Here $\boldsymbol{\mathcal{F}}$ and $\boldsymbol{\mathcal{F}}^{-1}$ designate the nonsampled, full 2D DFT forward and inverse operators, respectively. $\mathbf{M}$ is a matrix of proper size which has unit values at the k-domain sample positions corresponding to Ω and zeroes elsewhere. $\overline{\mathbf{M}}$ is the

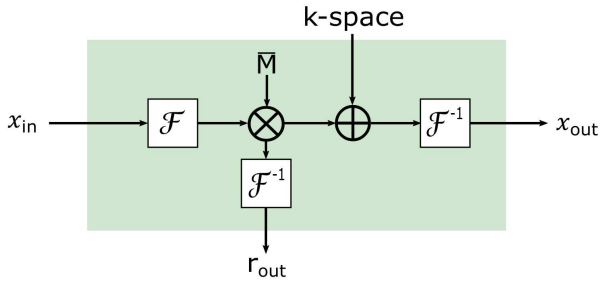

Fig. 2: The novel updated data consistency (UDC) layer as utilized in the proposed algorithm.

complement matrix of $\mathbf{M}$, and "$\circ$" denotes a pointwise multiplication. $\widehat{\boldsymbol{y}}$ is the actual k-domain observation as calculated in (1). Hence, the regular DC layer as utilized in [9] and elsewhere has two inputs (input image $\mathbf{x}_{\text{in}}$ and k-domain data $\widehat{\boldsymbol{y}}$) and a single output (the rectified image $\mathbf{x}_{\text{out}}$).

2.3. An Updated DC Layer Leading to a New DL Network: Projective Deep Cascade CNN

The DC-CNN network [9] uses a successive string of CNNs and DC layers. The final reconstructed estimate is the output of the finishing CNN. We observe that each of these intermediary networks are generating an estimate for the unobserved, missing k-domain information of the original image. However, this innovation part of the intermediary reconstructed image which corresponds to the unobserved Fourier information is not utilized directly. In this work we propose an updated DC layer (UDC). This novel UDC layer not only produces a rectified output image $\mathbf{x}_{\text{out}}$, but also outputs a residual or innovation image $\boldsymbol{r}_{\text{out}}$. This second output is calculated as follows using the notation of (2).

$$\boldsymbol{r}_{\text{out}} = \boldsymbol{\mathcal{F}}^{-1}\left\{\overline{\mathbf{M}} \circ (\boldsymbol{\mathcal{F}}\mathbf{x}_{\text{in}})\right\} \tag{3}$$

The block diagram for the newly developed UDC layer is depicted in Fig. 2. The new secondary output $\boldsymbol{r}_{\text{out}}$ is a projection of the intermediary estimate $\mathbf{x}_{\text{in}}$ onto the Fourier subspace which corresponds to the nonobserved coefficients not included in Ω. We extract these projected images $\boldsymbol{r}_i$ from each UDC stage. These projection images are delivered via skip connections to a concatenation layer which is at the input of a final reconstruction CNN. These $\boldsymbol{r}_i$ projection images carry the innovation introduced by each CNN stage. The final CNN reprocesses the information in these projection images and the intermediary output of the first phase to form the final output image. We should state here that the novel UDC layer does not entail additional computational complexity when compared to the standard DC layer. Fig. 4 illustrates the overall structure of the newly developed deep learning structure, which we entitle as Projective Deep Cascaded CNN (PDC-CNN) for MRI reconstruction. In this novel structure,

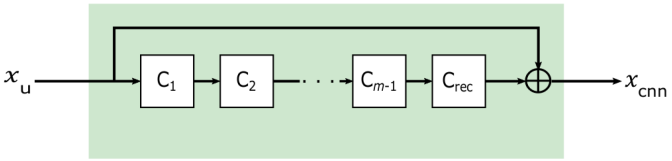

Fig. 3: The inner structure of the CNN stages used in the cascaded network.

n is the overall number of cascaded CNNs. Each CNN has the same structure with m convolutional layers, as depicted in Fig. 3. In our simulations we have made sure to implement both DC-CNN and PDC-CNN using the same n and m values to facilitate a fair comparison.

3. SIMULATION RESULTS

3.1. Dataset

We have realized the training and testing utilizing the IXI T1 dataset which contains close to 600 lateral brain scans[1]. The image set we used is from the Guy's Hospital scan set (1.5T with a repetition time of 9.813ms, echo time of 4.603ms). Each scan array in this set contains a total of 150 slices, each slice being a real valued image matrix of size 256×256.

3.2. Downsampling

We have used the same Cartesian downsampling strategy as explained and employed in [10]. The utilized mask applies a 4-step Cartesian downsampling in the Fourier domain by sampling one in every four rows of k-space horizontally. An additional 4% of the Fourier data from center, which corresponds to very low frequencies gets also sampled. The resulting undersampling ratio is 29%.

3.3. Model Analysis

We have used a total of $n = 10$ CNN blocks in the cascaded structures of both DC-CNN and the novel PDC-CNN. Each CNN block has $m = 6$ convolutional layers with the last layer being a reconstruction layer. All convolutional layers have 64 filters with a kernel size of 3. For DC-CNN and PDC-CNN we have employed networks of similar depths and similar total number of parameters. Both structures in the given setting have around a total of 1.49 million parameters, and the complexity difference between the two structures is negligible. We have also realized the U-Net deep network structure of [10] for MRI reconstruction. Although the U-Net structure is fundamentally different from the cascaded structure, we have made sure to have comparable complexity for the U-Net structure. The U-Net structure we realized has around 1.33M parameters.

[1] https://brain-development.org/ixi-dataset/

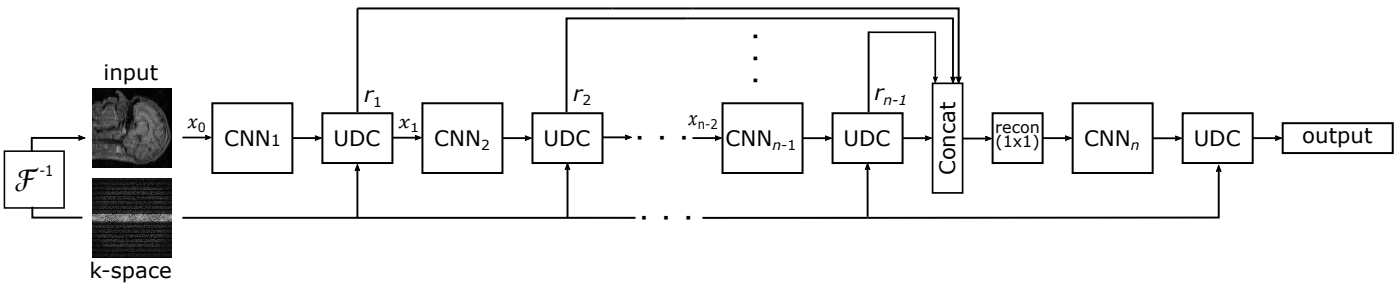

Fig. 4: The structure for the newly proposed Projective Deep Cascaded CNN (PDC-CNN) framework for MRI reconstruction.

3.4. Evaluation Methodology

In the simulations, 35 centrally stationed slices are taken from each scan. Hence, 980 images from 28 scans are used for training, 140 images from 4 scans are used for validation and a total of 210 images from 6 other scans are used for testing. We work on real valued ground-truth MRI images normalized to the range $[0, 1]$. The ground-truth images are downsampled using the 29% downsampling ratio Cartesian mask as described above. The ZF recontructed images are employed as inputs of to the DL models. In the training phase, the ground-truth MRI images are used as desired outputs. Given the real valued nature of these images, in the reconstruction algorithms we have converted complex valued image estimates to real valued ones via absolute value and real part operations. Mean squared error (MSE) is utilized as the cost function in training. We have trained the DL models for 1000 epochs with a mini-batch size of 8 images. Despite the rather high memory demand of the realized DL architectures, we were still able to use a mini-batch size of 8 leading to faster training times, while ensuring convergence. The network weights are initialized using Glorot uniform initialization [17]. Adam optimizer [18] is used with parameters $\alpha = 0.001$, $\beta_1 = 0.9$ and $\beta_2 = 0.999$. We have not explicitly set any weight decays. We have realized all three DL architecture models using Keras [19] and TensorFlow [20]. The simulations for the DL methods have been implemented on an NVIDIA Titan V GPU. The Titan V utilized in this work was donated by the NVIDIA Corporation. The training times for all three DL architectures are on the same range, all three taking around 24 hours to train with the above specifications. We should reiterate that this training procedure is an offline, one-time procedure which is not repeated during the testing phase.

3.5. Results

From the DL side, we have implemented the simulations for DC-CNN, U-Net and the newly proposed PDC-CNN. We have also realized MR reconstruction simulations using ZF, and sparsity based iterative methods TV CS-MRI [12] and PANO [13]. The CS-MRI and PANO have been implemented using the public software of PANO algorithm[2]. The average PSNR and structural similarity (SSIM) index performance results for all six algorithms are listed in Table 1. The DL based models have surpassed the iterative methods by a rather large margin. For the PSNR, our proposed PDC-CNN model brings performance improvement around 0.36 dB and 1.37 dB when compared to DC-CNN [9] and U-Net [10], respectively. The utilized sampling mask and one sample ground truth MR image are shown in Fig. 5. For this particular ground truth image, the reconstructed images and the corresponding error maps are produced in Fig. 6 and Fig. 7, respectively. Especially the error map images confirm the PSNR and SSIM results visually. In Table 1, we also list the average runtimes per image for the reconstruction procedures. Despite their lengthy, offline training procedure, the actual forward pass reconstruction for DL based methods is quite faster when compared to the iterative model based algorithms.

Model	PSNR (dB)	SSIM	Avg. time (sec)
Zero-filled	25.53	0.589	0.013
CS-MRI [12]	26.86	0.723	1.80
PANO [13]	27.95	0.781	70.80
U-Net [10]	33.73	0.908	0.019
DC-CNN [9]	34.74	0.921	0.028
PDC-CNN	**35.10**	**0.934**	0.038

Table 1: Performance and reconstruction time per image comparison for the competing methods.

4. CONCLUSION

Deep learning is successfully making inroads into inverse problems in imaging, with MRI reconstruction being one of the highly impacted applications. Recent examples for DL based approaches have utilized the data consistency layer as a rectification apparatus to enforce data fidelity in the reconstructed image. In this work, we have proposed an updated data consistency layer which not only rectifies the input, but

[2]https://sites.google.com/site/xiaoboxmu/publication

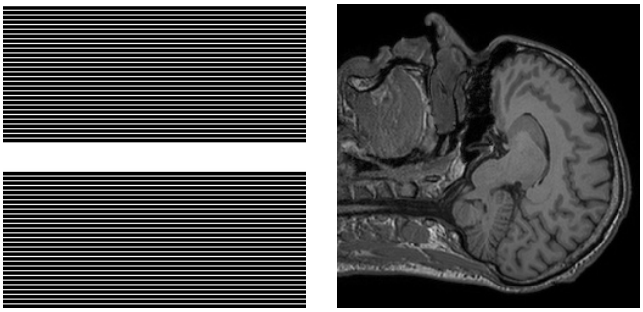

Fig. 5: The utilized sampling mask and the ground truth for a particular sample image from the test set.

also produces a secondary output which calculates by projection the innovation as produced by the preceding network stage. We propose to connect these innovation or residual images to the input of a final reconstruction network together with an intermediary reconstructed image. Simulation tests over a large set of MR images indicate that this novel structure improves the learning capacity and image reconstruction capability when compered to previously proposed deep networks of similar complexity.

5. REFERENCES

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

[2] O. Russakovsky *et al.*, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, Dec. 2015.

[3] 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*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.

[4] A. Lucas, M. Iliadis, R. Molina, and A. K. Katsaggelos, "Using deep neural networks for inverse problems in imaging: Beyond analytical methods," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 20–36, Jan. 2018.

[5] D. Yang and J. Sun, "BM3D-Net: A convolutional neural network for transform-domain collaborative filtering," *IEEE Signal Processing Letters*, vol. 25, no. 1, pp. 55–59, Jan. 2018.

[6] M. T. McCann, K. H. Jin, and M. Unser, "Convolutional neural networks for inverse problems in imaging: A review," *IEEE Signal Processing Magazine*, vol. 34, no. 6, pp. 85–95, Nov. 2017.

[7] T. Eo, Y. Jun, T. Kim, J. Jang, H.-J. Lee, and D. Hwang, "KIKI-net: cross-domain convolutional neural networks for reconstructing undersampled magnetic resonance images," *Magnetic Resonance in Medicine*, vol. 80, no. 5, pp. 2188–2201, Apr. 2018.

[8] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, and D. Liang, "Accelerating magnetic resonance imaging via deep learning," in *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*. IEEE, Apr. 2016.

[9] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 491–503, Feb. 2018.

[10] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, "Deep learning for undersampled MRI reconstruction," *Physics in Medicine & Biology*, vol. 63, no. 13, p. 135007, Jun. 2018.

[11] G. Yang, S. Yu, H. Dong, G. Slabaugh, P. L. Dragotti, X. Ye, F. Liu, S. Arridge, J. Keegan, Y. Guo, and D. Firmin, "DAGAN: Deep de-aliasing generative adversarial networks for fast compressed sensing MRI reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1310–1321, Jun. 2018.

[12] M. Lustig, D. Donoho, and J. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magnetic Resonance in Medicine*, vol. 58, no. 6, pp. 1182–1195, 2007.

[13] X. Qu, Y. Hou, F. Lam, D. Guo, J. Zhong, and Z. Chen, "Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator," *Medical Image Analysis*, vol. 18, no. 6, pp. 843–856, 2014.

[14] E. M. Eksioglu and A. K. Tanc, "Denoising AMP for MRI reconstruction: BM3D-AMP-MRI," *SIAM Journal on Imaging Sciences*, vol. 11, no. 3, pp. 2090–2109, Jan. 2018.

[15] E. M. Eksioglu, "Decoupled algorithm for MRI reconstruction using nonlocal block matching model: BM3D-MRI," *Journal of Mathematical Imaging and Vision*, vol. 56, no. 3, pp. 430–440, 2016.

[16] J. C. Ye, "Compressed sensing MRI: a review from signal processing perspective," *BMC Biomedical Engineering*, vol. 1, no. 1, p. 8, Mar. 2019.

[17] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proceedings of the International Conference on Artificial Intelligence and Statistics (AISTATS10). Society for Artificial Intelligence and Statistics*, 2010.

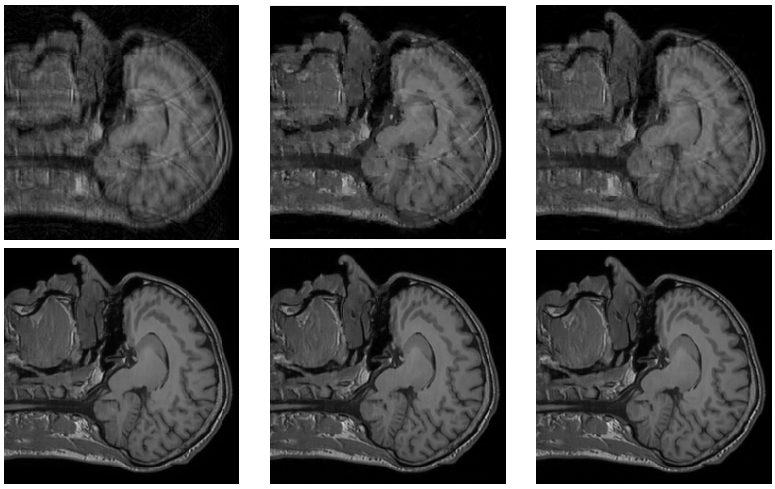

Fig. 6: Sample reconstructed images. First row: ZF, CS-MRI [12], PANO [13]. Second row: U-Net [10], DC-CNN [9], proposed PDC-CNN.

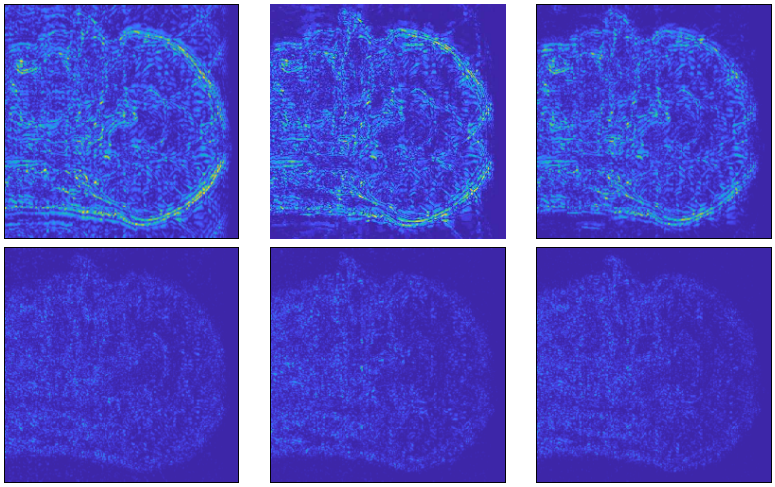

Fig. 7: Sample error images. First row: ZF, CS-MRI [12], PANO [13]. Second row: U-Net [10], DC-CNN [9], proposed PDC-CNN.

[18] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, 2014.

[19] F. Chollet *et al.*, "Keras," https://keras.io, 2015.

[20] M. Abadi *et al.*, "Tensorflow: A system for large-scale machine learning," in *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016, pp. 265–283.