FAST AUTOMATIC PARAMETER SELECTION FOR MRI RECONSTRUCTION

Tanjin Taher Toma and Daniel S. Weller

Electrical and Computer Engineering, University of Virginia, Charlottesville, USA

ABSTRACT

This paper proposes an automatic parameter selection framework for optimizing the performance of parameter-dependent regularized reconstruction algorithms. The proposed approach exploits a convolutional neural network for direct estimation of the regularization parameters from the acquired imaging data. This method can provide very reliable parameter estimates in a computationally efficient way. The effectiveness of the proposed approach is verified on transform-learning-based magnetic resonance image reconstructions of two different publicly available datasets. This experiment qualitatively and quantitatively measures improvement in image reconstruction quality using the proposed parameter selection strategy versus both existing parameter selection solutions and a fully deep-learning reconstruction with limited training data. Based on the experimental results, the proposed method improves average reconstructed image peak signal-to-noise ratio by a dB or more versus all competing methods in both brain and knee datasets, over a range of subsampling factors and input noise levels.

Index Terms— parameter selection, image reconstruction, magnetic resonance imaging, convolutional neural network

1. INTRODUCTION

Inverse problems in magnetic resonance image (MRI) reconstruction are regularized often to enhance the resulting image quality in various aspects. For example, such reconstruction algorithms can remove aliasing artifacts and noise in MRI data. Many of these regularized model-based solutions have one or more tuning parameters, which make the model suitable for a wide range of variations in the input data. But, manually adjusting such parameters is cumbersome, limits reproducibility, and becomes more challenging the more parameters there are to tune. Automatic parameter selection can overcome such limitations. We propose a novel data-driven approach for automatic parameter selection that exploits a convolutional neural network in lieu of hand-picked features. We demonstrate the effectiveness of our proposed approach over existing parameter selection algorithms for model-based MRI reconstruction.

Automatic parameter estimation techniques include empirical methods, such as generalized cross validation, the discrepancy principle, L-curve methods, and meta-heuristic techniques. Nonempirical approaches include Stein's unbiased risk estimator, other no-reference image quality measures, and adaptive selection techniques. Although these methods are widely used for automatic parameter selection, many are computationally intensive. For instance, quality-guided approaches reconstruct the same image with many different candidate parameter values in order to select among them. The accuracy of quality/risk estimates also influences the success of these methods for parameter selection. Further, tuning multiple parameters exacerbates the computational cost of such methods. Recently proposed data-driven learning-based methods to address parameter selection in other restoration and reconstruction problems include a regression-based framework [1] for denoising, deblurring, and demosaicing, as well as a deep reinforcement learning network [2] for X-ray computed tomography. Since regression directly estimates the parameters, in contrast to iterative adjustments with reinforcement learning or other methods, regression appears more appropriate for real-time parameter estimation. But, this method requires hand-selecting suitable features for learning based on domain knowledge.

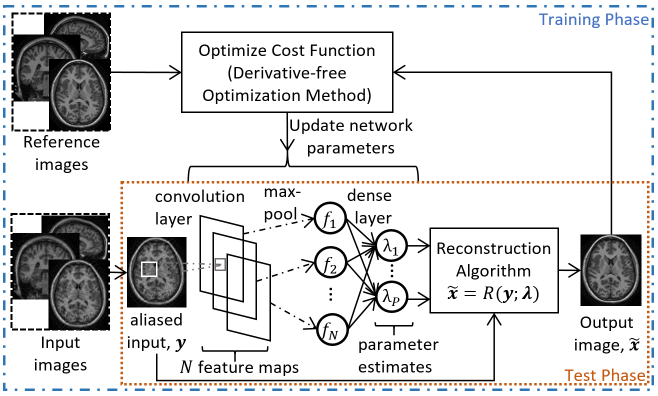


Fig. 1: Training and testing phases of the proposed convolutional neural network for parameter estimation

The proposed method overcomes various limitations of these existing solutions. As in regression [1], our method directly outputs parameter values, and is thus computationally very efficient. In addition, instead of relying on hand-picked features, the proposed method learns features via a convolutional layer. Further, the proposed network architecture is trainable even with limited data, while fully learning-based solutions to the high-dimensional reconstruction problem [3,4] benefit from more substantial training data, which may not always be available, to train deeper networks. Overall, the proposed method allows users (e.g., healthcare professionals or instrument operators) to produce better quality reconstructions exploiting model-based algorithms through fast and reliable estimation of regularization parameters.

2. THEORY

We begin with a noisy, aliased, blurred, or otherwise corrupted initial image $\boldsymbol{y}_m$ obtained from undersampled k-space. Then, we produce a new image $\tilde{\boldsymbol{x}}_m$ using a reconstruction algorithm $R(\boldsymbol{y}_m; \boldsymbol{\lambda}_m)$, with tuning parameters $\boldsymbol{\lambda}_m = [\lambda_{1,m}, \lambda_{2,m}, \ldots, \lambda_{P,m}]^T$. Parameter selection aims to estimate $\boldsymbol{\lambda}_m$ using information from input $\boldsymbol{y}_m$, to optimize the image quality of $\tilde{\boldsymbol{x}}_m$. This section describes a data-driven convolutional neural network-based method to estimate $\boldsymbol{\lambda}_m$.

Fig. 1 illustrates our overall approach with an image reconstruction example. The convolutional network layer convolves the magnitude of the 2D aliased input image $\boldsymbol{y}_m$ with N real-valued $r \times r$

kernels, $\boldsymbol{h}_1, \ldots, \boldsymbol{h}_N$, to generate N feature maps,

$$\boldsymbol{F}_{n,m} = ReLU(|\boldsymbol{y}_m| \circledast \boldsymbol{h}_n + \boldsymbol{b}_n), \quad \forall n, m. \tag{1}$$

where bias matrix $\boldsymbol{b}_n = b_n \boldsymbol{U}$, with $\boldsymbol{U}$ as the unit matrix of the same size as the convolved image, and the rectified linear unit ($ReLU$) performs nonlinear activation. Next, max-pooling produces single feature units from the feature maps $\boldsymbol{F}_{n,m}$:

$$f_{n,m} = \max(\boldsymbol{F}_{n,m}), \quad \forall n, m. \tag{2}$$

These pooled feature neurons, along with a bias neuron, form the feature vector $\boldsymbol{f}_m = [1\ f_{1,m} ... f_{N,m}]^T$. A dense layer connects these to the output neurons, generating the reconstruction parameter estimates $\boldsymbol{\lambda}_m^{\boldsymbol{\Theta}}$ (call the network parameters $\boldsymbol{\Theta}$),

$$\boldsymbol{\lambda}_m^{\boldsymbol{\Theta}} = (\boldsymbol{\lambda}^{max} - \boldsymbol{\lambda}^{min}) \circ S(\boldsymbol{W} \times \boldsymbol{f}_m) + \boldsymbol{\lambda}^{min}, \quad \forall m. \tag{3}$$

In (3), $\boldsymbol{W} = [\boldsymbol{w}_0\ \boldsymbol{w}_1 ... \boldsymbol{w}_N]$ is the weight matrix of the dense layer, where $\boldsymbol{w}_i = [w_{i1}\ w_{i2} ... w_{iP}]^T$; S is the sigmoid/logistic activation function bounded between user-defined minimum and maximum values $\boldsymbol{\lambda}^{min} = \left[\lambda_1^{min}\ \lambda_2^{min} ... \lambda_P^{min}\right]^T$ and $\boldsymbol{\lambda}^{max} = [\lambda_1^{max}\ \lambda_2^{max} ... \lambda_P^{max}]^T$ respectively; and $\circ$ is the Hadamard product. The vector of network model parameters $\boldsymbol{\Theta}$ collects all biases b_n, kernels $\boldsymbol{h}_n$, and dense layer coefficients $\boldsymbol{W}$. This bounded formulation keeps the parameter estimates within a reasonable range.

Supervised learning for parameter selection is slightly different from traditional supervised learning (e.g., image classification). Firstly, ground truth values of $\boldsymbol{\lambda}_m$ are not directly accessible for a given reconstruction. Instead, relatively clean image examples, such as fully sampled reconstructions, are the target images $\boldsymbol{x}_m$. Subsampling them in k-space (frequency domain) and adding noise, followed by an inverse Fourier transform, produces the aliased input images $\boldsymbol{y}_m$ for training. Then, the learning cost function uses a full-reference image quality measure such as mean squared error (MSE), peak signal-to-noise ratio (PSNR), or SSIM [5]. Secondly, the reconstruction operator R can be non-differentiable with respect to the processing parameters $\boldsymbol{\lambda}_m^{\boldsymbol{\Theta}}$, so estimating the network model parameters $\boldsymbol{\Theta}$ should use derivative-free optimization. For a training set of input-target pairs $\{\boldsymbol{y}_m, \boldsymbol{x}_m\}_{m=1}^M$, maximizing the PSNR of the reconstruction solves

$$\arg\max_{\boldsymbol{\Theta}} \frac{1}{M} \sum_{m=1}^{M} PSNR(\boldsymbol{x}_m, R(\boldsymbol{y}_m; \boldsymbol{\lambda}_m^{\boldsymbol{\Theta}})). \tag{4}$$

We optimize $\boldsymbol{\Theta}$ in (4) using the Nelder-Mead (NM) simplex method [6]. Once network parameters $\boldsymbol{\Theta}$ are learned, the network estimates the regularization parameters, tuning the reconstruction in one shot, in the on-line/testing phase. Unlike hand-picked features used with logistic regression [1], the general feature set $\boldsymbol{f}_m$ is extracted through the convolutional layer of the proposed network.

3. METHODS

We demonstrate our proposed parameter tuning approach using transform learning (TL) for MRI reconstruction [7, 8]. This chosen model-based reconstruction algorithm is a well-established, regularized iterative algorithm that simultaneously learns a sparsifying transform and reconstructs an image from highly undersampled k-space data. We tune the 'error threshold' parameter C that constrains reconstructed image patches in each TL iteration (see (P2c) in [8]). Common practice monotonically decreases C in successive iterations, since the initial input data are more degraded than the reconstructed image in subsequent iterations. Our objective is to estimate the initial C^u and final C^l values for C. The value of C in intermediate iterations decreases linearly in the interval $[C^l, C^u]$. Our proposed network, shown in Fig. 1, estimates these parameters $\{C^u, C^l\}$ for the magnitude of any aliased input image $\boldsymbol{y}_m$ (a sum-of-squares image can be used for parallel imaging reconstructions). We chose TL MRI partly because it permits us to demonstrate tuning multiple parameters. The convolutional layer has $N = 3$ kernels, each of size 3×3, and the dense layer produces $P = 2$ ($\lambda_1 = C^u$, $\lambda_2 = C^l$) processing parameters. The user-defined maximum and minimum values for these parameters are set in (3) as $\lambda_1^{max} = 2.5$, $\lambda_1^{min} = 0.25$, $\lambda_2^{max} = 0.20$, $\lambda_2^{min} = 0.01$.

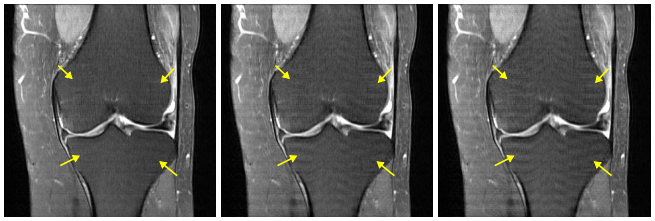
3.1. Dataset & Validation Experiments

We validated the proposed method on two different MRI datasets: the NYU fastMRI Initiative database[1] [9] and the OASIS brain database [10]. From the fastMRI database, we selected 36 fully-sampled knee MRI images including axial, coronal and sagittal cross-sections from randomly chosen subjects and generated separate training and testing sets. For each image in both training and test sets, 15 different realizations were created at 5 different subsampling rates $\{2.0, 2.5, 3.0, 3.5, 4.0\}$, each at 3 different input noise levels, SNR=$\{25, 30, 35$ dB$\}$ using additive complex Gaussian noise in k-space. Overall, we created 360 training images and 180 test images from the fastMRI dataset. Similarly for the OASIS database, we generated separate training and testing sets for 36 fully-sampled randomly selected brain cross-sectional images, each producing 10 different realizations at the same 5 subsampling rates, each at 2 different input noise levels, SNR=$\{25,\ 30$ dB$\}$. In total, we created 240 training images and 120 test images from the OASIS dataset.

We compare the reconstruction PSNR (dB) and SSIM using our parameter selection method against logistic regression (LR) [1], and image quality-based Metric-Q [11] and Deep-IQA [12]. Since the LR method extracts hand-picked features from the corrupted input image, we include two different versions of the LR method with different choices. In one, we use signal processing-based features (LR+SPF): the median high frequency quadrant coefficient of the discrete cosine transform and the area-under-the-curve of the spatial autocorrelation function [13], both extracted from $\boldsymbol{y}_m$. In the other, we use local statistical features (LR+LSF): the 5×5 and 9×9 local grey level average, standard deviation, and their ratio used for deblurring in [1]. We train both LR and our network separately for each dataset. Beyond LR, we compare against Metric-Q, commonly used for quality estimation, which is based on statistical measures of image content. Also, Deep-IQA is a recent data-driven deep learning approach for image quality assessment, pre-trained on the TID [14] natural image database. These approaches compute no-reference image quality measurements instead of direct parameter estimates. NM optimization finds suitable parameter values by maximizing these quality measurements. We also study how learning-based parameter selection compares against a fully-deep-learning reconstruction, the Deep Unet [3]. In addition to presenting visual and quantitative comparisons, we compute the statistical significance of the comparative results using repeated-measures one-factor ANOVA and post-hoc paired-sample two-sided t-tests with Bonferroni correction.

[1]NYU fastMRI investigators provided data but did not participate in analysis or writing of this article. A listing of NYU fastMRI investigators, subject to updates, can be found at fastmri.med.nyu.edu. The primary goal of fastMRI is to test whether machine learning can aid in the reconstruction of medical images.

(a) Reference image | (b) Zero-filled PSNR: 32.23 dB, SSIM: 0.8369 | (c) Proposed PSNR: 36.74 dB, SSIM: 0.9074



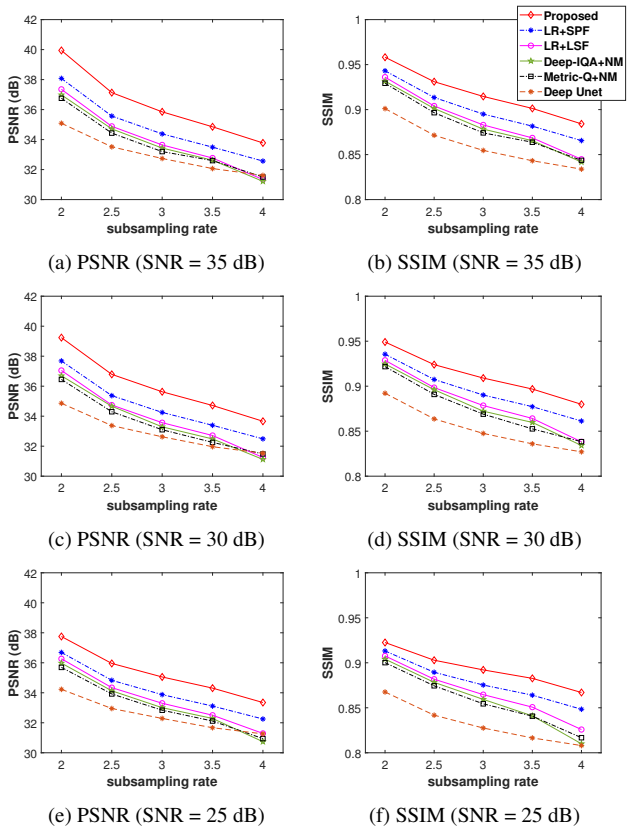
(d) LR+SPF PSNR: 35.04 dB, SSIM: 0.8906 | (e) LR+LSF PSNR: 34.03 dB, SSIM: 0.8769 | (f) Deep-IQA+NM PSNR: 33.49 dB, SSIM: 0.8679

Fig. 2: TL MRI reconstructions with different parameter selection methods for fastMRI [9] example image.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In Fig. 2, we portray reconstructions of one example knee image from the fastMRI dataset. The reference/clean image has been subsampled in k-space by a factor of 3 using a Gaussian variable-density mask, with complex Gaussian noise added at SNR=30 dB. The resulting zero-filled image is the input to reconstruction algorithm. Reconstructed images using parameter values estimated by the proposed and several competing methods are presented in the figure. We observe that the reconstruction using the proposed parameter estimation method has fewer artifacts and superior image quality compared to the other results. These quality differences observed in the reconstructed images are reflected both in PSNR (dB) and SSIM values. The LR method varies based on the type of features used in the model. The quality metric-based parameter estimates are far less suitable values, so the resulting reconstructions retain severe artifacts, shown for Deep-IQA method in the figure. One reason may be the pre-training places the Deep-IQA method at a disadvantage versus the other learning-based approaches. In Fig. 4, we demonstrate the reconstruction of a OASIS brain image subsampled by factor of 2.5 and at input SNR=25 dB. For this particular test image, two reconstruction parameters $\{\lambda_1, \lambda_2\}$ estimated by the proposed method, LR+SPF and Metric-Q+NM are $\{0.2500, 0.0315\}$,$\{0.9485, 0.1002\}$ and $\{2.1017, 0.1953\}$, respectively. From the results, we again observe that the TL MRI reconstruction using parameter estimates from the proposed method has superior image quality, while other comparative results retain blur and aliasing artifacts.

The comparative results among different methods at various subsampling factors and input SNR levels are shown in Figures 3 and 5 on the two datasets. We observe that the proposed method consistently outperforms other methods irrespective of input subsampling factor and SNR level. Further, the difference in average reconstruction results between the proposed method and its closest counterpart, LR+SPF, is significant in all cases (Bonferroni-corrected $p < 0.01$). Also, we observe that the average quality with the LR+LSF method remains below that of the LR+SPF method, confirming the draw-

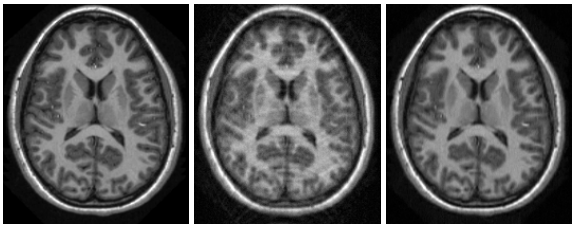


(a) PSNR (SNR = 35 dB) (b) SSIM (SNR = 35 dB)

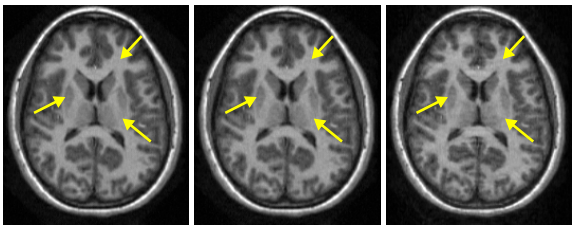
(c) PSNR (SNR = 30 dB) (d) SSIM (SNR = 30 dB)

(e) PSNR (SNR = 25 dB) (f) SSIM (SNR = 25 dB)

Fig. 3: Comparisons of average reconstruction quality in terms of PSNR and SSIM for various input subsampling rates and input noise levels on the fastMRI [9] dataset.

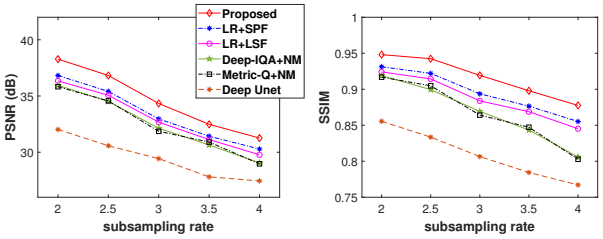


(a) Reference image | (b) Zero-filled PSNR: 28.86 dB SSIM: 0.7857 | (c) Proposed PSNR: 35.62 dB SSIM: 0.9039

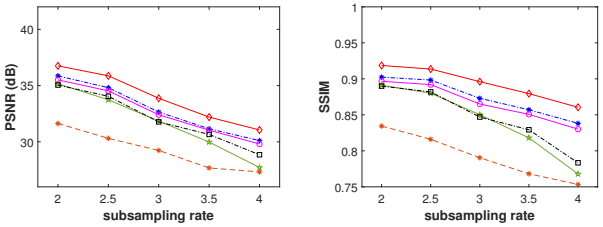


(d) LR+SPF PSNR: 34.60 dB SSIM: 0.8908 | (e) Metric-Q+NM PSNR: 33.96 dB SSIM: 0.8798 | (f) Deep Unet PSNR: 31.26 dB SSIM: 0.8342

Fig. 4: TL MRI reconstructions with different parameter selection methods and a fully deep-learning reconstruction for OASIS [10] example image.



(a) PSNR (SNR = 30 dB) (b) SSIM (SNR = 30 dB)



(c) PSNR (SNR = 25 dB) (d) SSIM (SNR = 25 dB)

Fig. 5: Comparisons of average reconstruction quality in terms of PSNR and SSIM for various input subsampling rates and input noise levels on the OASIS [10] dataset.

Table 1: Time for TL MRI Parameter Selection

Methods	Average Computation Time (seconds)
Proposed	0.0112
LR + SPF	0.0083
LR + LSF	0.2233
Metric-Q + NM	473.376
Deep-IQA + NM	790.80

back of using the LR approach with less effective hand-picked feature representations. Besides, it is evident from both Figures 3 and 5 that average reconstruction qualities using Deep Unet are unsatisfactory due to limited training data. We also note the ANOVA and t-test p-values versus our method are significant ($p < 0.05$) in all cases.

Moreover, we compare between different methods in terms of running time for parameter selection. Table 1 mentions average time (in seconds) to estimate the two TL MRI parameters given a test image from the fastMRI dataset. Also, the average reconstruction time by the TL MRI algorithm and Deep Unet are 24.773 and 0.429 seconds, respectively. Our experiments were performed using MATLAB (The Mathworks, Natick, MA) on a machine with Intel Core i7-4770 CPU and Ubuntu 16.04 LTS operating system, and the Deep-IQA and Deep Unet approaches also used an NVIDIA Titan X Pascal GPU. We observe the proposed and LR-based methods are both much faster computationally than the quality metric-based approaches. This difference is likely due to the first methods directly producing parameter estimates, as the quality-metric-based methods have to run many reconstruction instances to find a suitable parameter. The much higher computation time for Metric-Q and Deep-IQA limit comparison with these methods to relatively small testing sets.

5. CONCLUSION

We proposed a convolutional neural network architecture for automatic estimation of image reconstruction regularization parameters and demonstrated this approach for TL MRI reconstruction. Our data-driven approach is computationally very efficient and suitable for tuning multiple parameters. We demonstrated for image reconstruction on two datasets that the proposed method provides simpler and more reliable parameter tuning, in terms of both visual and quantitative image quality, than existing techniques. We also demonstrated that for limited training data, the combination of model-based reconstruction and our automatic tuning method can outperform a fully-deep-learning reconstruction.

6. ACKNOWLEDGEMENTS

This work was supported by National Science Foundation Grant No. 1759802 and by the National Institutes of Health (R56 EB028254). The authors thank NVIDIA Corporation for the Titan X Pascal GPU used in this research. The authors have no conflicts of interest.

7. REFERENCES

[1] J. Dong, I. Frosio, and J. Kautz, "Learning adaptive parameter tuning for image processing," *Electronic Imaging*, vol. 2018, no. 13, pp. 1–8, 2018.

[2] C. Shen, Y. Gonzalez *et al.*, "Intelligent parameter tuning in optimization-based iterative CT reconstruction via deep reinforcement learning," *IEEE Transactions on Medical Imaging*, vol. 37, no. 6, pp. 1430–1439, 2018.

[3] C. M. Hyun, H. P. Kim *et al.*, "Deep learning for undersampled MRI reconstruction," *Physics in Medicine & Biology*, vol. 63, no. 13, p. 135007, 2018.

[4] J. Schlemper, J. Caballero *et al.*, "A deep cascade of convolutional neural networks for dynamic MR image reconstruction," *IEEE Transactions on Medical Imaging*, vol. 37, no. 2, pp. 491–503, 2017.

[5] Z. Wang, A. C. Bovik *et al.*, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[6] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.

[7] S. Ravishankar and Y. Bresler, "Learning sparsifying transforms," *IEEE Transactions on Signal Processing*, vol. 61, no. 5, pp. 1072–1086, 2012.

[8] ——, "Sparsifying transform learning for compressed sensing MRI," in *2013 IEEE 10th International Symposium on Biomedical Imaging*. IEEE, 2013, pp. 17–20.

[9] J. Zbontar, F. Knoll *et al.*, "fastMRI: An open dataset and benchmarks for accelerated MRI," 2018, arXiv:1811.08839.

[10] A. F. Fotenos, A. Snyder *et al.*, "Normative estimates of cross-sectional and longitudinal brain volume decline in aging and AD," *Neurology*, vol. 64, no. 6, pp. 1032–1039, 2005.

[11] X. Zhu and P. Milanfar, "Automatic parameter selection for denoising algorithms using a no-reference measure of image content," *IEEE Transactions on Image Processing*, vol. 19, no. 12, pp. 3116–3132, 2010.

[12] S. Bosse, D. Maniry *et al.*, "Deep neural networks for no-reference and full-reference image quality assessment," *IEEE Transactions on Image Processing*, vol. 27, no. 1, pp. 206–219, 2017.

[13] C. Chatfield, *The Analysis of Time Series: An Introduction*. Chapman and Hall/CRC, 2016.

[14] N. Ponomarenko, O. Ieremeiev *et al.*, "Color image database TID2013: Peculiarities and preliminary results," in *European Workshop on Visual Information Processing (EUVIP)*. IEEE, 2013, pp. 106–111.