G-022

K-space Enhancement Based Sparse Directional Diffusion for Compressed Sensing MRI

Yichao Zuo† Sei-ichiro Kamata†

Abstract

Recently, Compressed Sensing (CS) has effectively accelerated Magnetic Resonance Imaging (MRI), with deep learning-based CS-MRI methods significantly outperforming traditional algorithms. Diffusion models using Stochastic Differential Equations (SDE) have shown great promise in CS-MRI reconstruction. However, existing SDE methods often overlook specific CS-MRI sampling rules, such as fully sampling the low-frequency space to retain essential image details. To address this, we propose a k-space enhanced diffusion model (KE-SDE) that maintains low-frequency sampling consistency during the diffusion process. Additionally, we introduce a Gaussian dilution-based sparse sampling operator that reduces the sampling rate without compromising reconstruction quality, thereby improving acceleration. Experiments on the public fastMRI dataset demonstrate our method's effectiveness, surpassing existing mainstream supervised deep learning algorithms and other generative models.

Keywords: Diffusion Models, MRI, Compressed Sensing, Image Reconstruction

1. Introduction

Magnetic Resonance Imaging (MRI) is a widely used medical imaging technique that provides detailed images of anatomical structures and physiological processes in the human body [1]. Despite recent advancements, the imaging speed remains relatively slow due to the sequential sampling process in k-space, which is a significant limitation for motion-sensitive and time-critical tasks [2]. In 2008, Compressed Sensing (CS) was applied to MRI, successfully enabling accelerated imaging with sampling rates below the Nyquist rate [3]. Since then, researchers have explored integrating CS with MRI, yielding promising results [4]-[6].

CS theory is based on two main conditions: the signal must be sparse in some transform domain, and the observation matrix and the sparse representation basis must be incoherent [7]. Based on this theory, iterative optimization algorithms can be designed to achieve MRI reconstruction from undersampled k-space data. Thus, the primary tasks of CS-MRI are: 1) finding an appropriate observation matrix, and 2) designing effective reconstruction algorithms. Recently, deep learning-based reconstruction algorithms have shown great potential, focusing on eliminating undersampling artifacts [8]-[10]. Deep learning addresses reconstruction primarily in two ways [11]: end-to-end supervised techniques, and distribution-based unsupervised learning techniques. End-to-end methods show good performance in specific tasks but have weak generalization capabilities [12] [13].

This study focuses on the second approach, using unsupervised learning techniques based on distribution learning. Generative models, particularly Compressed Sensing Generative Models (CSGM), are well-suited for these tasks and have been successfully applied to many inverse problems [14][15]. Recently, diffusion model-based CSGM has demonstrated superior performance to traditional end-to-end algorithms in some tasks [16]-[18].

Song et al. integrated two diffusion models—Denoising Diffusion Probabilistic Models (DDPM) and Score Matching with Langevin Dynamics (SMLD)—to propose a score-based continuous Stochastic Differential Equation (SDE) generative model [19]. DDPM and SMLD correspond to Variance Exploding (VE) SDE and Variance Preserving (VP) SDE, respectively. In SDE, the gradient of the data distribution (score function) is estimated using a neural network to solve the SDE backward. Score-based SDE has shown great potential in MRI reconstruction [20][21].

This research proposes a k-space enhanced method (KE-SDE), which fully samples the low-frequency space and applies the diffusion process only in the high-frequency space. This approach ensures lower errors in the low-frequency space, reduces the number of parameters in the feedforward process, and decreases the model size. Additionally, we introduce a Gaussian dilution-based sparse sampling mask to reduce the sampling rate while enhancing reconstruction results. These methods are detailed in the following sections.

2. Background

2.1 Framework

The purpose of CS-MRI is to reconstruct a full-space MR image without artifacts from undersampled K-space measurements, which can be formulated as:

$$y = Ux + \epsilon, \quad (1)$$

Where y is the undersampled measurements in the K-space, x is the target image which needs to be reconstructed, U is the encoding matrix and $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$.

† Yichao Zuo and Sei-ichiro Kamata are with the Image Media Laboratory, Graduate School of Information, Production and Systems, Waseda University, Kitakyushu 8080135, Japan.

Copyright © 2024 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

For 2D image, $\mathrm{x} \in \mathbb{C}^{\mathrm{n}}$, $\mathrm{y} \in \mathbb{C}^{m}$,and $\mathrm{U} \in \mathbb{C}^{\mathrm{m}\times\mathrm{n}}$. When reconstructing the MR images, $\mathrm{U} = \mathrm{M}_{\mathrm{sample}} \cdot \mathrm{F} \cdot \mathrm{csm}$, where $\mathrm{M}_{\mathrm{sample}}$ is the Under-sampling operator, F is the Fourier operator and csm is the coil sensitivity.

Eq. (1) can be constructed as an optimization problem:

$$\mathrm{x}^{*} = \arg\min_{\mathrm{x}} \frac{1}{2}\|\mathrm{Ux} - \mathrm{y}\|_2^2 + \mathcal{R}(\mathrm{x}), \tag{2}$$

Where $\mathcal{R}(\mathrm{x})$ is the prior constraint in the image domain of MR images.

2.2 Score-based SDE

Denoising Diffusion Probabilistic Models (DDPM) and Score Matching with Langevin Dynamics (SMLD) both follow a unified framework. Initially, Gaussian noise perturbs the data, aligning the original data distribution with a standard normal distribution. Then, data is sampled from the standard normal distribution, and the original data is generated using Langevin MCMC sampling. Score-based Stochastic Differential Equations (SDE) extend this approach to an infinite noise scale. The continuous diffusion process $\{\mathrm{x}(t)\}_{t=0}^{T}$ over the time variable $t \in [0, T]$can be described by the following SDE:

$$\mathrm{dx} = f(\mathrm{x}, t)\mathrm{d}t + g(t)\mathrm{dw}, \tag{3}$$

where the function f is drift coefficient of $\mathrm{x}(t)$; the function g is the diffusion coefficient of $\mathrm{x}(t)$; and w is the standard Wiener process.

The reverse process can accurately simulate the forward process in reverse time and is also a diffusion process. The reverse-time SDE is as follows:

$$\mathrm{dx} = [f(\mathrm{x}, t) - g(t)^2 \nabla_{\mathrm{x}} \log p_t(\mathrm{x})]\mathrm{d}t + g(t)\mathrm{d}\bar{\mathrm{w}}, \tag{4}$$

where $\bar{\mathrm{w}}$ is the standard Wiener process and $\mathrm{d}t$ is an infinitesimal negative time step. $\nabla_{\mathrm{x}} \log p_t(\mathrm{x})$ is the score function. As long as this function is known for all time t, the reverse diffusion process can be derived from this Eq. (4). Therefore, the neural network $s_\theta(x(t), t)$ is designed to evaluate the score function, with the optimization defined as:

$$\theta^{*} = \underset{\theta}{\mathrm{argmin}}\mathbb{E}_t\{\lambda(t)\mathbb{E}_{\mathrm{x}(0)}\mathbb{E}_{\mathrm{x}(t)|\mathrm{x}(0)}[|| s_\theta(\mathrm{x}(t), t) - \nabla_{\mathrm{x}} \log\, p_{0t}(\mathrm{x}(t)|\mathrm{x}(0))]||_2^2\}, \tag{5}$$

where $\lambda(t)$ is a positive weighting function, $t \in [0, T]$, $\mathrm{x}(0) \sim \mathcal{p}_{data}$, $\mathrm{x}(T) \sim \mathcal{p}_{noise}$, and $\mathcal{p}_{0t}(\mathrm{x}(t)|\mathrm{x}(0))$ is the perturbation kernel.

3. **Proposed method**

3.1 K-space Enhancement SDE

To utilize the characteristics of compressed sensing sampling in practical applications, namely full sampling in the low-frequency space and undersampling in the high-frequency space, a mask that constrains the training process is needed. This is referred to as the k-space mask M_{k}. In SDE, continuous noise of known intensity is added to assist in training the neural network. Applying the k-space mask to the added noise can effectively direct the network's training. At any time $t \in [0, T]$ during the training process, M_{k} can constrain the added noise as follows:

$$\begin{aligned}\mathrm{x}_{\mathrm{t}} = {} & \mathrm{F}^{-1}\big(\mathrm{M}_{\mathrm{k}} \cdot \mathrm{F}(\mathrm{x}_{t-1})\big) \\ & + \sqrt{1 - \beta_t}\mathrm{F}^{-1}\big((1 - \mathrm{M}_{\mathrm{k}}) \cdot \mathrm{F}(\mathrm{x}_{t-1})\big) \\ & + \sqrt{\beta_t}\mathrm{F}^{-1}\big((1 - \mathrm{M}_{\mathrm{k}}) \cdot \mathrm{F}(z_{t-1})\big), \\ & t = 1, \cdots, \mathrm{N},\end{aligned} \tag{7}$$

where x_t is the image with noise added at time step t; $F(\cdot)$ represents the fast Fourier transform; β_t is the parameter controlling the noise intensity at time step t; $z_t \sim \mathcal{N}(0, I)$; and M_{k} acts as a sampling filter, which is a 0-1 matrix of the same size as the input image. The parameter n can be used to control the size of the low-frequency space. In the extreme case where $n = 0$, the forward process is consistent with VP-SDE.

When $\Delta t = \frac{1}{N} \to 0$, Eq. (7) can be converted into the following SDE:

$$\begin{aligned}\mathrm{dx} = {} & -\frac{1}{2}\beta(t)\mathrm{F}^{-1}\big((1 - \mathrm{M}_{\mathrm{k}}) \cdot \mathrm{F}(\mathrm{x})\big)\mathrm{d}t \\ & + \sqrt{\beta(t)}\mathrm{F}^{-1}\big((1 - \mathrm{M}_{\mathrm{k}}) \cdot \mathrm{F}\big)\mathrm{dw},\end{aligned} \tag{8}$$

where w is the standard Wiener process. The optimization objective simultaneously transforms into:

$$\begin{aligned}\theta^{*} = {} & \underset{\theta}{argmin}\mathbb{E}_t\{\lambda(t)\mathbb{E}_{\mathrm{x}(0)}\mathbb{E}_{\mathrm{x}(t)|\mathrm{x}(0)}[|| \\ & s_\theta(\mathrm{F}^{-1}\Big((1 - \mathrm{M}_{\mathrm{k}}) \cdot \mathrm{F}\big(\mathrm{x}(t)\big)\Big), t) \\ & - \nabla_{\mathrm{x}} \log\, p_{0t}(\mathrm{x}(t)|\mathrm{x}(0))]||_2^2\},\end{aligned} \tag{9}$$

where $\lambda(t)$ is a positive weighting function. The network only predicts the high frequency noise, which can improve training efficiency.

3.2 Reverse of KE-SDE

The reverse process can be calculated according to the forward process of time, this is one of the important advantages of SDE. According to Eq. (8) and research [21], the inversion formula of KE-SDE can be derived as follows:

$$\begin{aligned}\mathrm{dx} = {} & (-\frac{1}{2}\beta(t)\mathrm{F}^{-1}((1 - \mathrm{M}_{\mathrm{k}}) \cdot \mathrm{F}(\mathrm{x})) \\ & -\beta(t)\mathrm{F}^{-1}((1 - \mathrm{M}_{\mathrm{k}}) \cdot \mathrm{F}(\nabla_{\mathrm{x}} \log p_t(\mathrm{x}\,|\,\mathrm{y}))))\mathrm{d}t \\ & + \sqrt{\beta(t)}\mathrm{d}\bar{\mathrm{w}},\end{aligned} \tag{10}$$

According to Bayes` theorem, the score function $\nabla_x \log\, p_t(x|y)$ can be formulate as:

$$\begin{aligned}\nabla_{\mathrm{x}} \log p_t(\mathrm{x}(t)\,|\,\mathrm{y}) \approx {} & s_{\theta^{*}}(\mathrm{x}(t), t) \\ & + \frac{\mathrm{U}^{\mathrm{H}}\big(\mathrm{y} - \mathrm{Ux}(t)\big)}{\sigma_\epsilon^2},\end{aligned} \tag{11}$$

Where σ_ϵ^2 is a hyperparameter that needs annealing, and $s_{\theta^{*}}$ is the neural network that evaluates the score. $\mathrm{U} = \mathrm{M}_{\mathrm{sample}} \cdot \mathrm{F} \cdot \mathrm{csm}$, $\mathrm{M}_{\mathrm{sample}}$ represents the sampling operator, F is the Fourier operator, and csm refers to the coil sensitivity map. It is important to note that for multi-coil MR images, computations need to be performed separately for each coil, and then summed to achieve the complete image reconstruction.

Copyright © 2024 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

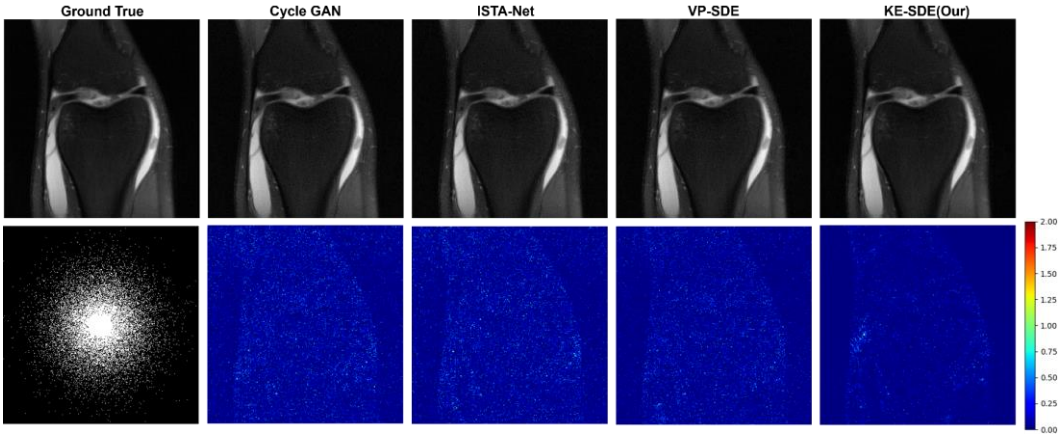

Fig 1: The reconstruction results of fastMRI multi-coil knee data at gaussian dilution sampling. The first row shows the ground truth and the reconstruction of ISTA-Net, cycleGAN, VP-, and KE-SDE (ours). The second row shows the error map of the reconstruction. The undersampling mask used for the test is shown in the lower left corner

3.3 Gaussian Dilution Sparse Sampling Operator

The sampling operator is crucial in CS-MRI reconstruction, as the sampling rate directly determines MRI acceleration and the relevance of the sampled content impacts model performance. Increasing the acceleration rate while minimizing the adverse effects of low sampling rates on reconstruction quality is a significant challenge. To address this, we propose a Gaussian dilution-based sparse sampling operator, as illustrated below:

$$\mathrm{M}_{\mathrm{sample}}(\mathrm{x},\mathrm{y}) = \mathrm{e}^{\left(-\frac{\mathrm{x}^2+\mathrm{y}^2}{2\alpha^2(\mathrm{H}^2+\mathrm{W}^2)}\right)} \sim \mathcal{N}\left(0, \alpha\left(\sqrt{\mathrm{H}^2+\mathrm{W}^2}\right)\right), \quad (11)$$

where (x, y) are the coordinates centered at the image origin, H and W are the image height and width, respectively, and α represents the dilution coefficient. The sampling density of this operator decreases from the center towards the edges, corresponding well to the frequency increase from the center outward in k-space. To ensure full sampling in the low-frequency space, the sampling operator includes a rectangular central frame with a side length l. This approach ensures that all information is captured in the low-frequency space while reducing the number of samples in the high-frequency space. Additionally, it serves as a condition to enhance the model's fitting performance. Experimental results have validated the effectiveness of this sampling strategy.

4. **Experiment**

4.1 Experiment data and Configuration

The experimental data was sourced from the fastMRI dataset [23]-[24]. During training, data from 97 individuals were randomly selected, discarding the first six slices from each to eliminate low-quality data, resulting in 2809 T1-weighted slices for training. Additionally, 35 T1-weighted slices were used as a validation set. For testing, 87 T1-weighted slices were used. Images were cropped to 320×320, and normalized by their standard deviation before being input into the scoring network.

Three metrics were used to evaluate reconstruction performance: Normalized Mean Squared Error (NMSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) [26]. Lower NMSE and higher PSNR and SSIM values indicate better reconstruction quality.

Supervised deep learning methods ISTA-Net [25], GAN-based methods CycleGAN [10], and score-based methods VP-SDE [19] were compared. VP-SDE served as an ablation model for KE-SDE. All algorithms were trained for 200 epochs. For score-based algorithms, the time step T was set to 1000, the exponential moving average (EMA) to 0.9999, and noise levels β_{max} and β_{min} to 20 and 0.1, respectively. A modified U-net was used for training as described in [19].

4.2 Experimental Results

For the sampling operator, we used the commonly employed 10-fold uniform sampling as a baseline to verify the effectiveness of the proposed method. The 10-fold uniform sampling refers to full sampling in the 24 center lines, while other areas are sampled every 10 lines, resulting in an actual acceleration rate of 5.9. For gaussian dilution sampling, the length of the side l of the frame is 24 and dilution coefficient α is 0.1, resulting in an actual acceleration rate of 8.0.

The results of Gaussian dilution sampling are shown in the Fig. 1. The first row displays the ground truth, ISTA-Net, CycleGAN, VP-SDE, and KE-SDE reconstruction results, respectively. The second row shows the error map of the reconstruction. The undersampling mask used for the test is shown in the lower left corner. The reconstructions produced by ISTA Net and CycleGAN all showed significant loss of detail. Although VP-SDE suffers less loss of detail, it introduces artifacts. Obviously,

Copyright © 2024 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.

Table Ⅰ: The average quantitative metrics on the fastMRI knee dataset under 12-fold uniform and gaussian dilution undersampling.

AF	Method	NMSE (%)	PSNR (db)	SSIM (%)
10-fold 16.9%	ISTA-Net	2.38	29.84	70.66
	CycleGAN	2.41	29.78	70.57
	VP-SDE	2.29	30.02	71.19
	KE-SDE	**1.93**	**30.75**	**77.16**
Gaussian Dilution 12.4%	ISTA-Net	1.41	31.77	73.38
	CycleGAN	1.43	31.75	73.25
	VP-SDE	1.31	32.27	75.62
	KE-SDE	**0.91**	**33.45**	**81.72**

KE-SDE has the lowest noise intensity and presents the best visual effects.

Additionally, Table I lists the related metrics for each model, comparing the results of different sampling operators. KE-SDE, shows the best result in every evaluation among all the models. Notably, the Gaussian dilution sampling operator demonstrates better reconstruction performance than the 10-fold uniform sampling even at lower sampling rates, proving its effectiveness.

5. Conclusion

In this study, the KE-SDE model based on K-space enhancement is proposed to solve the acceleration problem for CS-MRI, and the effectiveness of diffusion only in high frequency space is verified. In addition, the proposed Gaussian sparse sampling operator also successfully improves the acceleration rate and reconstruction results.

Acknowledgement

The authors declare no conflict of interest. Yichao Zuo: Conceptualization, Methodology, Research, Experiment, Validation. Sei-ichiro Kamata: Conceptualization, Research, Validation. I would like to express my great appreciation to supervisor Sei-ichiro Kamata for his consecutive suggestions.

References

[1] Liang, Z.P., Lauterbur, P.C., 2000. Principles of magnetic resonance imaging: a signal processing perspective. *SPIE Optical Engineering Press*

[2] Tsao, J., Kozerke, S., 2012. Mri temporal acceleration techniques. *Journal of Magnetic Resonance Imaging* 36, 543–560.

[3] Lustig, M., Donoho, D.L., Santos, J.M., Pauly, J.M., 2008. Compressed sensing mri. *IEEE signal processing magazine* 25, 72.

[4] Emmanuel J Candes. The restricted isometry property and its implications for compressed sensing. *Comptes rendus mathematique*, 346(9-10):589–592, 2008.

[5] J. P. Haldar, D. Hernando, and Z.-P. Liang, "Compressed-sensing mri with random encoding," *IEEE Transactions on Medical Imaging*, vol. 30, no. 4, pp. 893–903, 2010

[6] A. Majumdar, "Improving synthesis and analysis prior blind compressed sensing with low-rank constraints for dynamic mri reconstruction," *Magnetic Resonance Imaging*, vol. 33, no. 1, pp. 174–179, 2015.

[7] Donoho, D.L., 2006. Compressed sensing. *IEEE Transactions on Information Theory* 52, 1289–1306.

[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," i*n 2016 IEEE 13th International Symposium on Biomedical Imaging. IEEE*, 2016, pp. 514–517.

[9] Y. Han, L. Sunwoo, and J. C. Ye, "k-space deep learning for accelerated mri," *IEEE Transactions on Medical Imaging*, vol. 39, no. 2, pp. 377–386, 2019.

[10] G. Oh, B. Sim, H. Chung, L. Sunwoo, and J. C. Ye, "Unpaired deep learning for accelerated mri using optimal transport driven cyclegan," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 1285–1296, 2020.

[11] Gregory Ongie, Ajil Jalal, Christopher A Metzler, Richard G Baraniuk, Alexandros G Dimakis, and Rebecca Willett. *Deep learning techniques for inverse problems in imaging*. arXiv preprint arXiv:2005.06001, 2020.

[12] Hemant K Aggarwal, Merry P Mani, and Mathews Jacob. Modl: Model-based deep learning architecture for inverse problems. *IEEE transactions on medical imaging*, 38(2):394–405, 2018.

[13] Vegard Antun, Francesco Renna, Clarice Poon, Ben Adcock, and Anders C Hansen. On instabilities of deep learning in image reconstruction and the potential costs of ai. *Proceedings of the National Academy of Sciences*, 117(48):30088–30095, 2020.

[14] Ashish Bora, Ajil Jalal, Eric Price, and Alexandros G Dimakis. Compressed sensing using generative models. *In Proceedings of the 34th International Conference on Machine Learning* Volume 70, pages 537–546. JMLR. org, 2017.

[15] Paul Hand and Babhru Joshi. Global guarantees for blind demodulation with generative priors. *In Advances in Neural Information Processing Systems*, pages 11531–11541, 2019.

[16] P. Dhariwal and A. Nichol, "Diffusion models beat gans on image synthesis," *Advances in Neural Information Processing Systems*, vol. 34, pp. 8780–8794, 2021.

[17] A. Q. Nichol and P. Dhariwal, "Improved denoising diffusion probabilistic models," *in International Conference on Machine Learning*. PMLR, 2021, pp. 8162–8171.

[18] L. Yang, Z. Zhang, Y. Song, S. Hong, R. Xu, Y. Zhao, Y. Shao, W. Zhang, B. Cui, and M.-H. Yang, "Diffusion models: A comprehensive survey of methods and applications," arXiv preprint arXiv:2209.00796, 2022.

[19] Y. Song, J. Sohl-Dickstein, D. P. Kingma, A. Kumar, S. Ermon, and B. Poole, "Score-based generative modeling through stochastic differential equations," *in International Conference on Learning Representations*, 2021.

[20] A. Jalal, M. Arvinte, G. Daras, E. Price, A. G. Dimakis, and J. Tamir, "Robust compressed sensing mri with deep generative priors," *Advances in Neural Information Processing Systems*, vol. 34, pp. 14 938–14 954, 2021.

[21] H. Chung and J. C. Ye, "Score-based diffusion models for accelerated mri," *Medical Image Analysis*, p. 102479, 2022.

[22] B. D. Anderson, "Reverse-time diffusion equation models," *Stochastic Processes and their Applications*, vol. 12, no. 3, pp. 313–326, 1982

[23] J. Zbontar, F. Knoll, A. Sriram, T. Murrell, Z. Huang, M. J. Muckley, A. Defazio, R. Stern, P. Johnson, M. Bruno et al., "fastmri: An open dataset and benchmarks for accelerated mri," arXiv preprint arXiv:1811.08839, 2018.

[24] F. Knoll, J. Zbontar, A. Sriram, M. J. Muckley, M. Bruno, A. Defazio, M. Parente, K. J. Geras, J. Katsnelson, H. Chandarana et al., "fastmri: A publicly available raw k-space and dicom dataset of knee images for accelerated mr image reconstruction using machine learning," Radiology. *Artificial Intelligence*, vol. 2, no. 1, 2020.

[25] J. Zhang and B. Ghanem, "Ista-net: Interpretable optimization-inspired deep network for image compressive sensing," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 1828–1837

[26] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

Copyright © 2024 by
The Institute of Electronics, Information and Communication Engineers and
Information Processing Society of Japan All rights reserved.