UFC-Net: Unrolling Fixed-point Continuous Network for Deep Compressive Sensing

Supplementary Material

1. Traditional FPC Algorithm

Fixed-Point Continuation (FPC), a frequently employed signal reconstruction algorithm in compressive sensing (CS), incrementally enhances signal sparsity through a series of iterative steps to achieve signal recovery. In comparison to iterative shrinkage/thresholding algorithm (ISTA), FPC leverages operator splitting techniques combined with path-following strategies to address the ℓ_1-norm regularization problem, as follows:

$$\min_{\mathbf{x} \in \mathbb{R}^N} \|\mathbf{x}\|_{\ell_1} + \frac{\mu}{2} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|_{\ell_2}^2, \tag{1}$$

where μ serves as a penalty coefficient, striking a balance between the data fitting term and the ℓ_1-norm regularization term.

Let $g(\mathbf{x}) = \|\mathbf{x}\|_{\ell_1}$, $f(\mathbf{x}) = \frac{1}{2}\|\mathbf{A}\mathbf{x} - \mathbf{y}\|_{\ell_2}^2$. Hence, we can reformulate Eq. (1) as follows: $F(\mathbf{x}) = g(\mathbf{x}) + \mu f(\mathbf{x})$. Furthermore, let $T(\mathbf{x}) = \partial F(\mathbf{x})$. According to the principles of operator splitting theory, it is feasible to decompose the function F into the summation of two convex functions: $F = F_1 + F_2$. Likewise, the operator T can also be disassembled into the sum of two maximally monotone operators: $T = T_1 + T_2$. Through rigorous derivation (for comprehensive details, please refer to [4]), the iterative formula for obtaining the minimal solution of $F(\mathbf{x})$ can be articulated as follows:

$$\mathbf{x}^{k+1} = (I + \tau T_1)^{-1}(I - \tau T_2)\mathbf{x}^k. \tag{2}$$

For the minimization problem, T_1 is defined as $\nabla g(\mathbf{x}^k)$ and T_2 is defined as $\mu \nabla f(\mathbf{x}^k)$, while $(I + \tau T_1)^{-1}$ corresponds to a soft thresholding operation.

This subsequently leads to the specific iterative steps of the FPC algorithm as outlined below:

$$\begin{aligned} \mathbf{x}^{k+1} &= s_\nu \circ h(\mathbf{x}^k) \\ &= \text{sgn}(h(\mathbf{x}^k)) \odot \max\{|h(\mathbf{x}^k)| - \nu, 0\}, \end{aligned} \tag{3}$$

where

$$h(\mathbf{x}^k) = I(\mathbf{x}^k) - \tau \nabla f(\mathbf{x}^k), \tag{4}$$

where $h(\cdot)$ bears a resemblance to the gradient descent iterations of $f(\mathbf{x}^k)$ intuitively, $s_\nu(\cdot)$ embodies the soft thresholding function. τ represents the gradient descent step size, ν signifies the soft thresholding shrinkage operator, and there exists an equivalence relationship that $\mu\nu = \tau$.

Moreover, FPC introduces an continuation strategy for μ and incorporates a regularization shrinkage parameter β that escalates the regularization parameter μ during the iterative process, resulting in a sequence of growing regularization parameters $\{\mu_i\}$, instead of using a fixed soft thresholding shrinkage operator. When addressing a new problem associated with μ_{i+1}, the approximate solution of the current problem with μ_i is employed as the starting point, thereby enhancing the algorithm's convergence speed.

2. Implementation Details

We selectively procure a subset of 800 images from the extensively employed large-scale training dataset, COCO2017 [9]. In conjunction with data augmentation techniques, we extract 300 individual 32×32 pixel patches from each image, employing random horizontal and vertical flips, rotations, and scaling procedures. Moreover, Set11 [8] dataset is incorporated as a validation set for the evaluation of the generalization capabilities of our UFC-Net.

We harness the Adam optimization strategy [6] along with a cosine annealing scheme, encompassing a total of 100 iterations. The learning rate progressively diminishes from 5×10^{-4} to 5×10^{-5}, with a warm-up phase spanning three iterations. Image manipulations are executed within the YCbCr color space, with evaluations performed on the Y channel. To assess the quality of reconstruction images, we conduct sufficient experiments on extensive datasets, including large-scale datasets such as CIFAR10 and CIFAR100 [7], as well as various widely-recognized benchmark datasets, such as Set5 [1], Set11 [8], Set14 [17], McM18 [18], BSD68 [10], Urban100 [5], and General100 [2]. Furthermore, we employ the widely adopted evaluation metrics, namely peak signal noise ratio (PSNR) and structural similarity (SSIM), in conjunction with a visual examination of the reconstructed images, to comprehensively evaluate the quality of the reconstruction images.

Moreover, as far as CS-MRI related experiments are concerned, the training dataset for this experiment comprises 4865 images extracted from the FastMRI knee dataset [16], with an additional set of 995 images designated for testing purposes. The experiment spans a total of 200 epochs, with a batch size of 1 and a learning rate cosinely decayed from 5×10^{-5} to 1×10^{-6}. The inputs are magnetic resonance images of dimensions 320×320. Within each iteration stage in the reconstruction module, we first perform the gradient descent computations in the GDU on images with a scale of $(b, 1, 320, 320)$, where b means batch size. Subsequently, it is reshaped into $(b \cdot 100, 1, 32, 32)$ and fed into the CAM module for further processing. Finally, the output of CAM
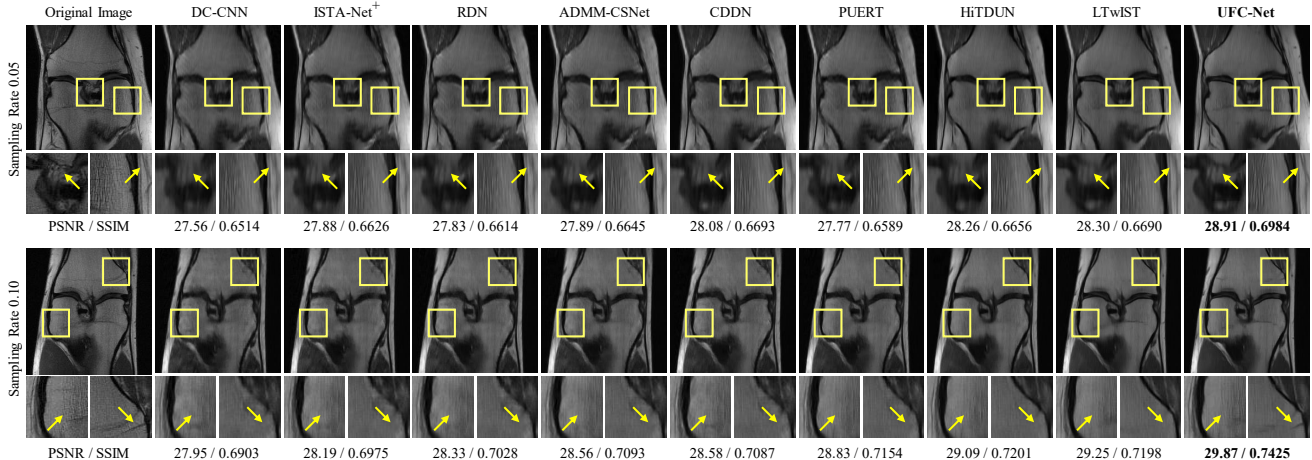
Figure 1. Comparisons of reconstructed MRI images from the FastMRI dataset [16] between UFC-Net and other eight competing algorithms. The first row pertains to comparative images at $\delta = 0.05$, while the second row corresponds to comparative images at $\delta = 0.10$. Regions of interest are magnified, with distinctive differences indicated by arrows for ease of observation. It is evident that our method reconstructs textures with enhanced clarity and incorporates a greater level of fine-grained details.
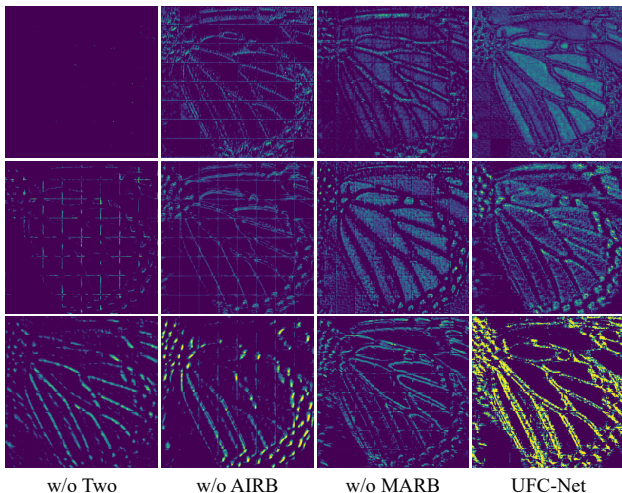

Figure 2. Visual analysis of feature maps of our UFC-Net and the three variants derived from it at different iteration stages when $\delta = 0.10$. The first, second and third rows correspond to the feature maps of the thrid, fourth and fifth stages in the reconstrction module respectively.

is restored to $(b, 1, 320, 320)$ as the input of the next stage. Furthermore, to expedite the training process in the CS-MRI experiments, the number of iterative stages is adjusted to 6.

3. Ablation Study

3.1. Visual Analysis of Feature Maps

Fig. 2 depicts the visual analysis of the feature maps derived from our UFC-Net and three variants w/o Two, w/o AIRB, and w/o MARB at different iteration stages when $\delta = 0.10$.

It is obviously that our MARB and AIRB are capable of capturing more image details and textures.

3.2. UFC-Net with Different Iterative Stages

To comprehensively investigate the performance of our UFC-Net, we derive four variations with distinct numbers of iterative stages, designated as UFC-Net-{8, 9, 10, 11}, with UFC-Net-10 being the default configuration, while the others adhere to the same training specifications as previously described.

Table 1. PNSR (dB) and SSIM comparisons of UFC-Net-{8, 9, 10, 11} on datasets Set11 [8] and BSD68 [10] at different sampling rates $\delta = \{0.01, 0.04, 0.10, 0.25\}$.

Datasets	δ	UFC-Net-8		UFC-Net-9		UFC-Net-10 (default)		UFC-Net-11	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Set11	0.01	21.22	0.5591	21.19	0.5589	**21.24**	0.5607	21.23	**0.5612**
	0.04	25.81	0.7904	25.87	0.7934	**25.92**	**0.7943**	25.81	0.7920
	0.10	29.96	0.8930	30.01	0.8938	**30.15**	**0.8960**	30.08	0.8948
	0.25	35.26	0.9559	35.30	0.9560	**35.42**	**0.9567**	35.34	0.9561
BSD68	0.01	22.33	0.5261	22.35	0.5264	**22.37**	**0.5268**	22.35	0.5264
	0.04	25.27	0.6868	25.29	**0.6880**	25.30	0.6877	25.26	0.6865
	0.10	27.89	0.8065	27.91	0.8071	**27.95**	**0.8086**	27.93	0.8071
	0.25	31.70	0.9087	31.71	0.9091	**31.74**	**0.9093**	31.72	0.9087

As presented in Tab. 1, we carry out evaluations on the Set11 and BSD68 datasets at four diverse sampling rates $\delta \in \{0.01, 0.04, 0.10, 0.25\}$. The optimal outcomes are predominantly concentrated in UFC-Net-10 (the default configuration), while the secondary best results are unevenly distributed between UFC-Net-9 and UFC-Net-11. Overall, with an increase of the number of iterative stages, the performance of our UFC-Net gradually improves initially, reaching a saturation at UFC-Net-10, after which it begins to decline, accompanied by a progressive increase

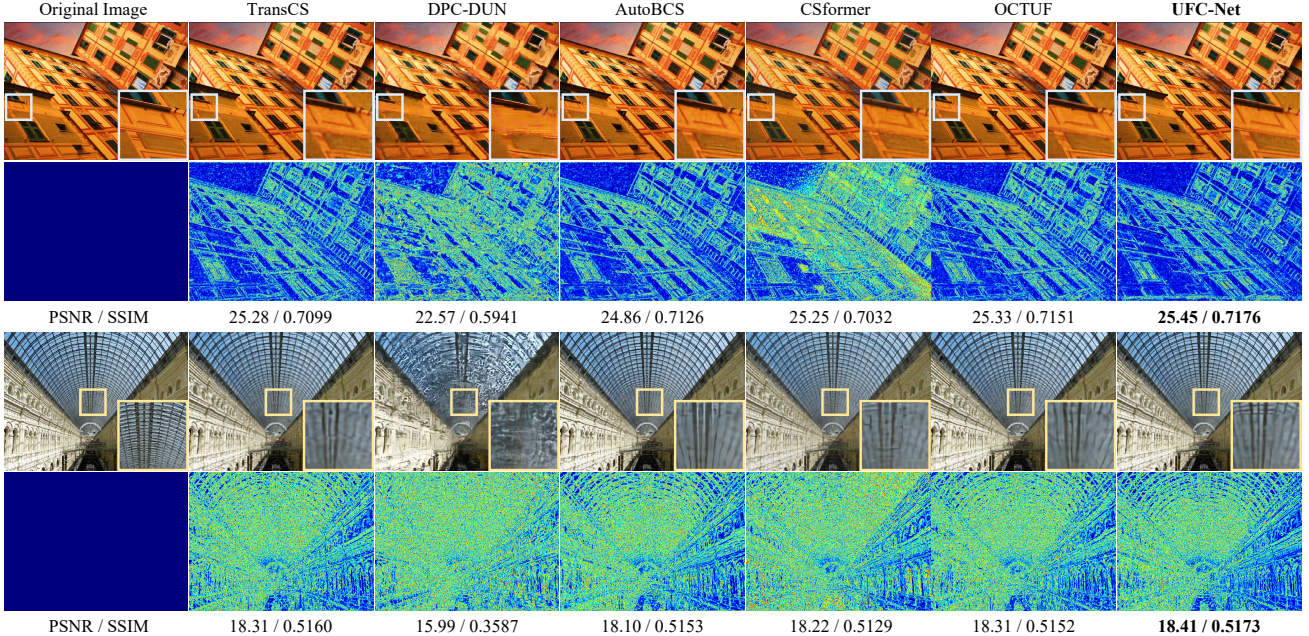|	Original Image	TransCS	DPC-DUN	AutoBCS	CSformer	OCTUF	**UFC-Net**
PSNR / SSIM		25.28 / 0.7099	22.57 / 0.5941	24.86 / 0.7126	25.25 / 0.7032	25.33 / 0.7151	**25.45 / 0.7176**
PSNR / SSIM		18.31 / 0.5160	15.99 / 0.3587	18.10 / 0.5153	18.22 / 0.5129	18.31 / 0.5152	**18.41 / 0.5173**

Figure 3. Visual comparisons of reconstruction images and error maps of other five competing methods and our proposed UFC-Net under noises when the sampling rate $\delta = 0.04$. The first and second rows correspond to the reconstructed images and error maps respectively when the Gaussian noise $\sigma = 0.004$. The third and fourth rows correspond to the reconstructed images and error maps respectively at the salt and pepper noise $\epsilon = 0.98$. Local areas are zoomed in for better comparisons.

in complexity. Therefore, it is a better choice to define the number of iterative stages in reconstruction module as 10.

3.3. UFC-Net under Different Noises

This part is primarily dedicated to evaluating the robustness of our proposed UFC-Net in the presence of varying levels of noise. On the CIFAR10 dataset, we examine three types of zero-mean Gaussian noise, characterized by distinct variance levels, denoted as $\sigma \in \{0.001, 0.002, 0.004\}$. Subsequently, we conduct comparative evaluations between our UFC-Net and TransCS [12], DGUNet$^+$ [11], OCTUF [14], focusing on their Gaussian noise robustness at sampling rates of 0.01 and 0.04. The outcomes are outlined in Tab. 2. It is evident that our UFC-Net consistently maintains peak PSNR and SSIM values, significantly surpassing the second-ranking OCTUF algorithm. For instance, UFC-Net outperforms OCTUF by a substantial margin, enhancing PSNR (percentage gains) by 0.65dB ($\sim$3.69%) and SSIM (percentage gains) by 0.0478 ($\sim$15.57%) when $\sigma = 0.002$ and $\delta = 0.01$.

Moreover, we study different levels of salt and pepper noise on the CIFAR100 dataset, expressed as $\epsilon \in \{0.94, 0.96, 0.98\}$. We conduct a comparative analysis in the Tab. 3 of DPC-DUN [13], CSformer [15], Auto-BCS [3] and our UFC-Net regarding the salt and pepper noise robustness at sampling rates of 0.01 and 0.04. From the numerical results in the Tab. 3, our proposed UFC-

Net consistently exhibits superior performance. Particularly, in comparison to the second-best algorithm, Auto-BCS, our UFC-Net demonstrates a significant improvement of 0.72dB ($\sim$4.06%) in PSNR and a notable increase of 0.0424 ($\sim$13.16%) in SSIM, when $\epsilon = 0.98$ and $\delta = 0.01$. These results show that our UFC-Net can reconstruct high-quality images even under the influence of different noises.

Furthermore, when $\delta = 0.04$, we show the reconstructed images and error maps of TransCS [12], DPC-DUN [13], AutoBCS [3], CSformer [15], OCTUF [14] and our proposed UFC-Net under Gaussian noise $\sigma = 0.004$, and the salt and pepper noise $\epsilon = 0.98$ respectively in Fig. 3. It can be seen that our UFC-Net has higher visual performance and lower reconstruction error.

Table 2. PNSR (dB) and SSIM comparisons of UFC-Net and state-of-the-art methods with various Gaussian noise levels on dataset CIFAR10 [7] at $\delta = \{0.01, 0.04\}$.

σ	δ	TransCS [12]		DGUNet$^+$ [11]		OCTUF [14]		UFC-Net	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
0.001	0.01	16.49	0.2946	15.23	0.2720	17.61	0.3073	**18.26**	**0.3550**
	0.04	19.29	0.5746	19.41	0.5601	21.35	0.6278	**21.66**	**0.6387**
0.002	0.01	16.48	0.2945	15.20	0.2717	17.60	0.3070	**18.25**	**0.3548**
	0.04	19.26	0.5728	19.41	0.5617	21.29	0.6253	**21.60**	**0.6362**
0.004	0.01	16.47	0.2939	15.13	0.2704	17.59	0.3067	**18.24**	**0.3545**
	0.04	19.19	0.5690	19.40	0.5626	21.18	0.6201	**21.50**	**0.6314**

Table 3. PNSR (dB) and SSIM comparisons of UFC-Net and state-of-the-art methods with various Salt and Pepper noise levels on dataset CIFAR100 [7] at $\delta = \{0.01, 0.04\}$.

ϵ	δ	DPC-DUN [13]		CSformer [15]		AutoBCS [3]		UFC-Net	
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
0.98	0.01	14.64	0.2013	16.53	0.3004	17.74	0.3222	18.46	0.3646
	0.04	16.72	0.3718	20.29	0.5667	21.25	0.6197	21.80	0.6325
0.96	0.01	14.32	0.1914	16.57	0.2999	17.74	0.3211	18.41	0.3612
	0.04	16.03	0.3348	20.07	0.5520	21.01	0.6064	21.44	0.6147
0.94	0.01	14.08	0.1835	16.59	0.2990	17.73	0.3194	18.35	0.3581
	0.04	15.45	0.3043	19.85	0.5387	20.77	0.5927	21.11	0.5983

Table 4. Comparisons of UFC-Net and its variants on dataset Set11 at the sampling rate $\delta = 0.10$.

Methods	woT-RB	woT-U	woMA-RB	woMA-U	woAI-RB	woAI-U	UFC-Net
PSNR	29.68	29.64	29.82	29.63	29.78	29.59	30.15
SSIM	0.8883	0.8878	0.8911	0.8890	0.8887	0.8881	0.8960
LPIPS	0.0786	0.0796	0.0774	0.0787	0.0789	0.0787	0.0755

Table 5. Comparisons of UFC-Net with different loss functions on different datasets at the sampling rate $\delta = 0.25$.

Loss Function	Set11		Set14		Urban100		General100	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
MSE Loss(default)	35.42	0.9567	33.81	0.9259	32.82	0.9423	37.75	0.9624
L1 Loss	35.31	0.9561	33.69	0.9258	32.72	0.9409	37.66	0.9622
Perceptual Loss	32.99	0.9423	31.33	0.9046	30.54	0.9242	35.20	0.9485

3.4. UFC-Net and Its Variants

We substitute MARB and AIRB in CAM with the standard residual block and UNet, resulting in six variations: woT-RB, woT-U, woMA-RB, woMA-U, woAI-RB, and woAI-U. Among them, woT-RB means using the standard residual block to replace the two blocks MARB and AIRB. Other variants are obtained through similar approaches. Tab. 4 illustrates the comparison results on the dataset Set11 at $\delta = 0.10$. Our UFC-Net consistently exhibits optimal performance and demonstrates significant improvement compared to the second-best variant, providing evidence for the superiority of our proposed MARB and AIRB.

3.5. UFC-Net with Different Loss Functions

We perform ablation studies on UFC-Net using various loss functions, including Mean Squared Error (MSE), L1 loss, and perceptual loss functions. As depicted in Tab. 5, UFC-Net exhibits superior performance when using MSE function, while showing suboptimal results under the L1 loss function.

3.6. UFC-Net under Perceptual Loss

As depicted in Tab. 6, we conduct a comparison of the LPIPS metric between proposed UFC-Net and different methods on datasets CIFAR10 and CIFAR100 at $\delta = 0.04$. The results consistently demonstrate that our UFC-Net outperforms other methods on different datasets.

Table 6. LPIPS comparisons of UFC-Net and other methods at $\delta = 0.04$.

Methods	TransCS	DGUNet+	DPC-DUN	LTwIST	AutoBCS	CSformer	OCTUF	UFC-Net
CIFAR10	0.4247	0.1695	0.1692	0.3910	0.3910	0.1469	0.1416	0.1351
CIFAR100	0.4182	0.1722	0.1763	0.3816	0.3839	0.1496	0.1440	0.1376

References

[1] Marco Bevilacqua, Aline Roumy, Christine Guillemot, and Marie Line Alberi-Morel. Low-complexity single-image super-resolution based on nonnegative neighbor embedding. In *British Machine Vision Conference (BMVC)*, pages 135.1–135.10, 2012.

[2] Chao Dong, Chen Change Loy, and Xiaoou Tang. Accelerating the super-resolution convolutional neural network. In *European Conference on Computer Vision (ECCV)*, pages 391–407, 2016.

[3] Hongping Gan, Yang Gao, Chunyi Liu, Haiwei Chen, Tao Zhang, and Feng Liu. AutoBCS: Block-based image compressive sensing with data-driven acquisition and noniterative reconstruction. *IEEE Transactions on Cybernetics*, 53 (4):2558–2571, 2023.

[4] Elaine T. Hale, Wotao Yin, and Yin Zhang. Fixed-Point Continuation for l_1-minimization: Methodology and convergence. *SIAM Journal on Optimization*, 19(3):1107–1130, 2008.

[5] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[6] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2015.

[7] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.

[8] Kuldeep Kulkarni, Suhas Lohit, Pavan Turaga, Ronan Kerviche, and Amit Ashok. ReconNet: Non-iterative reconstruction of images from compressively sensed measurements. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 449–458, 2016.

[9] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft COCO: Common objects in context. In *European Conference on Computer Vision (ECCV)*, pages 740–755, 2014.

[10] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *IEEE International Conference on Computer Vision (ICCV)*, pages 416–423, 2001.

[11] Chong Mou, Qian Wang, and Jian Zhang. Deep generalized unfolding networks for image restoration. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 17399–17410, 2022.

[12] Minghe Shen, Hongping Gan, Chao Ning, Yi Hua, and Tao Zhang. TransCS: A transformer-based hybrid architecture for image compressed sensing. *IEEE Transactions on Image Processing*, 31:6991–7005, 2022.

[13] Jiechong Song, Bin Chen, and Jian Zhang. Dynamic path-controllable deep unfolding network for compressive sensing. *IEEE Transactions on Image Processing*, 32:2202–2214, 2023.

[14] Jiechong Song, Chong Mou, Shiqi Wang, Siwei Ma, and Jian Zhang. Optimization-inspired cross-attention transformer for compressive sensing. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6174–6184, 2023.

[15] Dongjie Ye, Zhangkai Ni, Hanli Wang, Jian Zhang, Shiqi Wang, and Sam Kwong. CSformer: Bridging convolution and transformer for compressive sensing. *IEEE Transactions on Image Processing*, 32:2827–2842, 2023.

[16] Jure Zbontar, Florian Knoll, Anuroop Sriram, Tullie Murrell, Zhengnan Huang, Matthew J Muckley, Aaron Defazio, Ruben Stern, Patricia Johnson, Mary Bruno, et al. FastMRI: An open dataset and benchmarks for accelerated MRI. *arXiv preprint arXiv:1811.08839*, 2018.

[17] Roman Zeyde, Michael Elad, and Matan Protter. On single image scale-up using sparse-representations. In *International Conference Curves and Surfaces*, pages 711–730. Springer, 2012.

[18] Lei Zhang, Xiaolin Wu, Antoni Buades, and Xin Li. Color demosaicking by local directional interpolation and nonlocal adaptive thresholding. *Journal of Electronic Imaging*, 20(2): 023016, 2011.