CPP-Net: Embracing Multi-Scale Feature Fusion into Deep Unfolding CP-PPA Network for Compressive Sensing

Zhen Guo
Northwestern Polytechnical University
Xi'an 710072, China
guozhen2022@mail.nwpu.edu.cn

Hongping Gan*
Northwestern Polytechnical University
Xi'an 710072, China
ganhongping@nwpu.edu.cn

1. Overview

In this *Supplementary Material*, we first conduct the convergence analysis of our proposed deep unfolding CP-PPA framework in Sec. 2. Subsequently, we provide more details and comparisons in Sec. 3, including detailed experiments settings in Sec. 3.1, more comparisons results in Sec. 3.2, performance under salt-and-pepper noise in Sec. 3.3, more CS-MRI comparisons results in Sec. 3.4 and more visualization results in Sec. 3.5.

2. Convergence Analysis

In the previous section (Sec. 3.2 in the main text), we reformulate the Eq. (1) into a more tractable form denoted as Eq. (6). From this, we derive the following iterative steps: Eq. (8), Eq. (9), and Eq. (10). Furthermore, we obtain the deep unfolding form of the CP-PPA denoted as Eq. (15).

First, consider that $\mathcal{D}(\mathbf{x}) = \mathbf{C}_2\text{ReLU}(\mathbf{C}_1)$, where $\mathbf{C}_1$ and $\mathbf{C}_2$ are convolutional operators. Let $\mathbf{r}^{(k)} = \mathbf{x}^{(k-1)} - \eta^{(k)}\mathbf{\Phi}^\top\mathbf{y}^{(k-1)}$. By approximating based on Theorem 1 [27], we obtain:

$$\|\mathcal{D}(\mathbf{x}) - \mathcal{D}(\mathbf{r}^{(k)})\|_{\ell_2}^2 \approx \alpha\|\mathbf{x} - \mathbf{r}^{(k)}\|_{\ell_2}^2, \tag{20}$$

where α is a constant related to $\mathcal{D}(\cdot)$ and can be merged into the learnable parameter $\lambda^{(k)}$. Incorporating this relationship into Eq. (8), we derive:

$$\mathbf{x}^{(k)} = \arg\min_{\mathbf{x}} \lambda^{(k)}\eta^{(k)}\|\mathcal{D}(\mathbf{x})\|_{\ell_1} + \frac{1}{2}\|\mathcal{D}(\mathbf{x}) - \mathcal{D}(\mathbf{r}^{(k)})\|_{\ell_2}^2. \tag{21}$$

Thus, we obtain the closed form of $\mathcal{D}(\mathbf{x}^{(k)})$:

$$\mathcal{D}(\mathbf{x}^{(k)}) = \mathcal{S}_{\lambda^{(k)}\eta^{(k)}}(\mathcal{D}(\mathbf{r}^{(k)})). \tag{22}$$

Let $\widetilde{\mathcal{D}}(\cdot)$ be the left inverse of $\mathcal{D}(\cdot)$, we derive:

$$\mathbf{x}^{(k)} = \widetilde{\mathcal{D}}(\mathcal{S}_{\lambda^{(k)}\eta^{(k)}}(\mathcal{D}(\mathbf{r}^{(k)}))). \tag{23}$$

*Corresponding author.

Thus, the optimal solution to Eq. (8) is proven to be Eq. (23).

Lemma 1 Let $\chi \in \Re^n$ be a closed convex set, $\xi(\mathbf{x})$ and $\varphi(\mathbf{x})$ be convex functions, with $\varphi(\mathbf{x})$ being differentiable. Assume that the solution set of the minimization problem $\min\{\xi(\mathbf{x}) + \varphi(\mathbf{x})|\mathbf{x} \in \chi\}$ is nonempty. Then,

$$\mathbf{x}^* \in \arg\min\{\xi(\mathbf{x}) + \varphi(\mathbf{x})|\mathbf{x} \in \chi\}, \tag{24}$$

if and only if

$$\begin{aligned}\mathbf{x}^* \in \chi, \xi(\mathbf{x}) - \xi(\mathbf{x}^*) + \\ (\mathbf{x} - \mathbf{x}^*)^T \bigtriangledown \varphi(\mathbf{x}^*) \geq 0, \ \forall \mathbf{x} \in \chi.\end{aligned} \tag{25}$$

Then, we delve into the convergence analysis of Eq. (15). Assuming that the pair $(\mathbf{x}^{(k)}, \mathbf{y}^{(k)})$ represents the optimal solution to Eq. (8) and Eq. (10), we can obtain the following equations in the form of mixed monotone variational inequalities [14] according to **Lemma 1**:

$$\begin{cases} f(\mathbf{x}) - f(\mathbf{x}^{(k)}) + (\mathbf{x} - \mathbf{x}^{(k)})^\top(\\ \qquad \mathbf{\Phi}^\top\mathbf{y}^{(k-1)} + \alpha^{(k)}(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)})) \geq 0, \\ h^*(\mathbf{y}) - h^*(\mathbf{y}^{(k)}) + (\mathbf{y} - \mathbf{y}^{(k)})^\top(\\ \qquad -\mathbf{\Phi}\mathbf{v}^{(k)} + \beta^{(k)}(\mathbf{y}^{(k)} - \mathbf{y}^{(k-1)})) \geq 0. \end{cases} \tag{26}$$

By defining $\mathbf{u} = (\mathbf{x}, \mathbf{y})^\top$ and $\zeta(\mathbf{u}) = f(\mathbf{x}) + h^*(\mathbf{y})$, we rewrite Eq. (26) as follows:

$$\begin{aligned}&\zeta(\mathbf{u}) - \zeta(\mathbf{u}^{(k)}) + (\mathbf{u} - \mathbf{u}^{(k)})^\top \\ &\{\begin{pmatrix}\mathbf{\Phi}^\top\mathbf{y}^{(k-1)} \\ -\mathbf{\Phi}\mathbf{v}^{(k)}\end{pmatrix} + \begin{pmatrix}\alpha^{(k)}(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}) \\ \beta^{(k)}(\mathbf{y}^k - \mathbf{y}^{(k-1)})\end{pmatrix}\} \geq 0.\end{aligned} \tag{27}$$

Substituting Eq. (9) into Eq. (27), we obtain:

$$\begin{aligned}&\zeta(\mathbf{u}) - \zeta(\mathbf{u}^{(k)}) + (\mathbf{u} - \mathbf{u}^{(k)})^\top\{\begin{pmatrix}\mathbf{\Phi}^\top\mathbf{y}^{(k)} \\ -\mathbf{\Phi}\mathbf{x}^{(k)}\end{pmatrix} + \\ &\begin{pmatrix}\alpha^{(k)}(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}) - \mathbf{\Phi}^\top(\mathbf{y}^{(k)} - \mathbf{y}^{(k-1)}) \\ -\mathbf{\Phi}(\mathbf{x}^{(k)} - \mathbf{x}^{(k-1)}) + \beta^{(k)}(\mathbf{y}^k - \mathbf{y}^{(k-1)})\end{pmatrix}\} \geq 0.\end{aligned} \tag{28}$$

Hence, we can reformulate Eq. (28) as follows:

$$\zeta(\mathbf{u}) - \zeta(\mathbf{u}^{(k)}) + (\mathbf{u} - \mathbf{u}^{(k)})^\top \{\mathbf{G}(\mathbf{u}^{(k)}) + \mathbf{Q}(\mathbf{u}^{(k)} - \mathbf{u}^{(k-1)})\} \geq 0, \tag{29}$$

where

$$\mathbf{G} = \begin{pmatrix} 0 & \mathbf{\Phi}^\top \\ -\mathbf{\Phi} & 0 \end{pmatrix}, \tag{30}$$

and

$$\mathbf{Q} = \begin{pmatrix} \alpha^{(k)}\mathbf{I} & -\mathbf{\Phi}^\top \\ -\mathbf{\Phi} & \beta^{(k)}\mathbf{I} \end{pmatrix}. \tag{31}$$

Note that Eq. (29) is the generalized form of PPA [8, 14], and in the case of $\alpha^{(k)}\beta^{(k)} > \|\mathbf{\Phi}^\top\mathbf{\Phi}\|_{\ell_2}$, $\mathbf{Q}$ is symmetric and positive definite, thus the sequence $\{x^{(k)}\}$ and $\{y^{(k)}\}$ converge to the solution point of the Eq. (6).

3. More Experiments

3.1. Experimental Settings

We employed a dataset of 41,000 images, with 40,000 for training and 1,000 for validation, randomly selected from the COCO2017 unlabeled images dataset [12]. During training, we applied random cropping to resize images to 96×96 and employed data augmentation techniques, such as scaling and rotation. We first pretrained the CPP-Net at a CS ratio of 0.25. Subsequently, we fine-tuned the CPP-Net at various CS ratios based on the pretrained CPP-Net at 0.25 CS ratio. LPIPS was obtained using VGG as the base network. See Tab. 1 for detailed configurations.

Table 1. The configurations of pretraining and fine-tuning.

Configurations	Pretraining	Fine-tuning
base learning rate	8e-5	4e-5
min learning rate	1e-6	1e-6
optimizer	AdamW	AdamW
weight decay	0.05	0.05
optimizer momentum	0.9, 0.999	0.9, 0.999
batch size	8	8
training epochs	400	400
warmup epochs	10	10
warmup schedule	linear	linear
learning rate schedule	cosine annealing	cosine annealing
block size	32	32
stages conut	8	8
channel count	32	32
weight init	trunc.normal (0.2)	-
$\eta^{(k)}$	1e-3	-
$\delta^{(k)}$	1e-3	-
$\lambda^{(k)}$	1.0	-
sampling matrix init	Guassian random matrix	-
implementation	Pytorch 1.12.0	
CPU	Intel(R) Xeon(R) Gold 6226 CPU	
GPU	Tesla V100 PCIe 32GB	

3.2. More Comparisons Results

In this subsection, we evaluated the performance of our CPP-Net with more competing methods and more CS ratios to demonstrate the effectiveness of our proposed CPP-Net across various CS ratios. The best and second-best results in the tables are highlighted in red and blue colors, respectively. First, we compared the performance of our CPP-Net with 21 competing methods, including pure DL-based methods (ReconNet [11], SCSNet [17], CSNet$^+$ [18], DPA-Net [23], MAC-Net [2], BCS-Net [33], NL-CSNet [4], AutoBCS [6], CSformer [25] and TCS-Net [7]), and deep unfolding networks (DUNs) (ISTA-Net$^+$ [27], OPINE-Net$^+$ [28], AMP-Net [31], COAST [26], CASNet [1], DGUNet$^+$ [13], FSOINet [3], TransCS [16], CT-Net+[30], DPC-DUN [20], MAPUN [19] and OCTUF [21]) on the Set11 [11] at seven CS ratios. The results are summarized in Tab. 2. It can be seen that our CPP-Net outperforms other methods with higher PSNR and SSIM on Set11 across all tested CS ratios. We then evaluated the performance of our CPP-Net with several recent methods on General100 [5] and Urban100 [9] at three high CS ratios, as shown in Tab. 3. Moreover, the LPIPS performance of recent methods at different CS ratios is presented in Tab. 4. Our CPP-Net still outperforms recent methods with higher image quality and human perception quality across all tested CS ratios on the General100 and Urban100. In summary, our CPP-Net can achieve superior image quality and improved human perception quality at both low and high CS ratios.

3.3. Performance under Salt-and-Pepper Noise

We introduced varying proportions (denoted as ς) of salt-and-pepper noise to the images in Set11 and assessed the performance of our CPP-Net in handling noisy images. The variable ς denotes the ratio of pixels in the images affected by salt-and-pepper noise. As shown in Tab. 5, ς of salt-and-pepper noise is negatively correlated with the performance of each method, but our CPP-Net still outperforms other methods across all tested proportions of salt-and-pepper noise.

3.4. More Comparisons Results of CS-MRI

In this subsection, we evaluated the performance of our CPP-Net on the single-coil knee dataset available within the fastMRI dataset [10] using the 1D Cartesian masks. The competing methods include DC-CNN [15], RDN [22], ISTA-Net$^+$ [27], CDDN [32], ADMM-CSNet [24], and HiTDUN [29]. Tab. 6 summarizes the results and Fig. 1 illustrates the visual comparisons of different CS-MRI methods. Compared with other methods, our CPP-Net performs better and reconstructs MRI images with more accurate details and less error.

Table 2. Comparisons of average PSNR (dB)/SSIM performance of various CS methods on Set11 at various CS ratios.

Dataset	Methods	0.01	0.04	0.10	0.25	0.30	0.40	0.50	Avg.
Set11	ReconNet (CVPR 2016)	17.43/0.4017	20.93/0.5897	24.38/0.7301	28.44/0.8531	29.09/0.8693	30.60/0.9020	32.25/0.9177	26.16/0.7519
	SCSNet (CVPR 2019)	21.04/0.5562	24.29/0.7589	28.52/0.8616	33.43/0.9373	34.64/0.9511	36.92/0.9666	39.01/0.9769	31.12/0.8584
	CSNet$^+$ (TIP 2020)	20.69/0.5238	24.54/0.7445	28.12/0.8664	32.20/0.9337	33.70/0.9495	36.41/0.9677	38.28/0.9771	30.56/0.8518
	DPA-Net (TIP 2020)	18.05/0.5011	23.50/0.7205	26.99/0.8354	31.74/0.9238	33.35/0.9425	35.04/0.9565	36.73/0.9670	29.34/0.8353
	MAC-Net (ECCV 2020)	18.26/0.4003	24.22/0.6982	27.68/0.8182	32.91/0.9244	33.96/0.9372	35.94/0.9560	37.67/0.9668	30.09/0.8144
	BCS-Net (TMM 2021)	20.81/0.5427	24.90/0.7531	29.36/0.8650	34.20/0.9408	35.40/0.9527	36.52/0.9640	39.58/0.9734	31.54/0.8560
	NL-CSNet (TMM 2021)	21.96/0.6005	26.26/0.8108	30.05/0.8995	- / -	35.68/0.9606	- / -	- / -	- / -
	AutoBCS (TCYB 2023)	19.63/0.5605	24.73/0.7871	28.44/0.8827	33.56/0.9481	34.48/0.9549	36.51/0.9680	36.88/0.9732	30.60/0.8678
	CSformer (TIP 2023)	21.86/0.6071	26.41/0.8058	30.09/0.8925	34.99/0.9534	- / -	- / -	40.23/0.9802	- / -
	TCS-Net (TCI 2023)	21.09/0.5505	25.46/0.7863	29.04/0.8834	33.94/0.9508	34.34/0.9541	35.27/0.9601	37.36/0.9723	30.93/0.8654
	ISTA-Net$^+$ (CVPR 2018)	17.45/0.4131	21.56/0.6240	26.49/0.8036	32.44/0.9237	33.70/0.9382	36.02/0.9579	38.07/0.9706	29.39/0.8044
	OPINE-Net$^+$ (J-STSP 2020)	20.02/0.5362	25.52/0.7879	29.81/0.8904	34.81/0.9514	36.00/0.9600	38.31/0.9724	40.18/0.9800	32.09/0.8683
	AMP-Net (TIP 2021)	20.20/0.5581	25.26/0.7722	29.40/0.8779	34.63/0.9481	36.03/0.9586	38.28/0.9715	40.34/0.9807	32.02/0.8667
	COAST (TIP 2021)	12.40/0.2637	23.55/0.7158	28.70/0.8609	33.96/0.9405	35.09/0.9504	37.09/0.9645	38.92/0.9743	29.96/0.8100
	CASNet (TIP 2022)	21.76/0.6019	26.25/0.8118	30.29/0.9005	35.65/0.9592	36.90/0.9663	39.03/0.9760	40.93/0.9827	32.97/0.8855
	DGUNet$^+$ (CVPR 2022)	22.15/0.6114	26.83/0.8230	30.93/0.9088	36.18/0.9616	36.72/0.9661	38.99/0.9766	41.24/0.9837	33.29/0.8902
	FSOINet (ICASSP 2022)	21.73/0.5937	26.37/0.8119	30.44/0.9018	35.80/0.9595	37.00/0.9665	39.14/0.9764	41.08/0.9832	33.08/0.8847
	TransCS (TIP 2022)	20.15/0.5066	25.41/0.7883	29.54/0.8877	35.06/0.9548	35.62/0.9588	38.46/0.9737	40.49/0.9815	32.10/0.8645
	CT-Net+ (KBS 2023)	- / -	- / -	30.16/0.8966	35.58/0.9570	36.73/0.9642	38.49/0.9752	40.90/0.9822	- / -
	DPC-DUN (TIP 2023)	18.03/0.4601	24.38/0.7498	29.42/0.8801	34.75/0.9483	35.88/0.9570	37.98/0.9694	39.84/0.9778	31.47/0.8489
	MAPUN (IJCV 2023)	- / -	- / -	30.19/0.9014	35.73/0.9602	37.08/0.9676	39.22/0.9775	40.98/0.9834	- / -
	OCTUF (CVPR 2023)	21.75/0.5934	26.45/0.8126	30.70/0.9030	36.10/0.9604	37.21/0.9673	39.41/0.9773	41.34/0.9838	33.28/0.8854
	CPP-Net (Our Method)	**22.19/0.6135**	**27.23/0.8337**	**31.27/0.9135**	**36.35/0.9631**	**37.55/0.9696**	**39.53/0.9781**	**41.39/0.9842**	**33.64/0.8937**

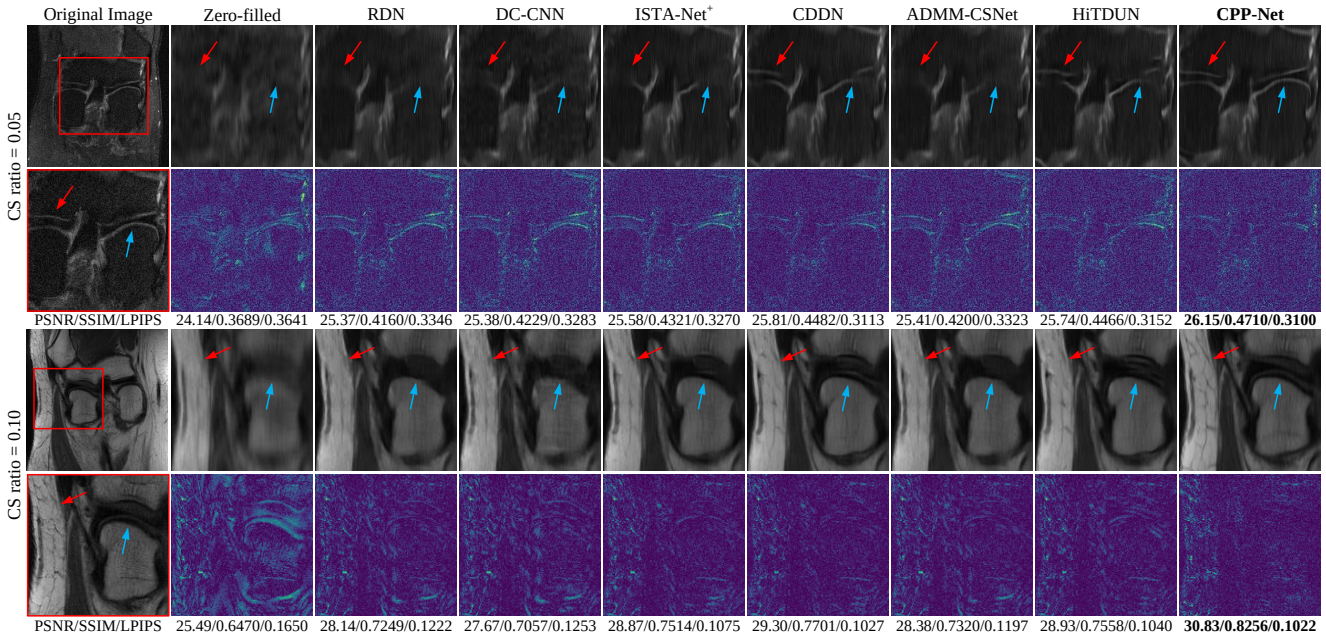

Figure 1. Comparisons of visualization, error maps and corresponding PSNR (dB)/SSIM/LPIPS performance of various CS-MRI methods on the fastMRI dataset using 1D Cartesian masks at CS ratios of 0.05 and 0.10. The arrows point to details in the reconstructed image for better comparison.

3.5. Additional Visualization Results

Fig. 2 shows the visualization of feature maps within the third stage of the different cases in the ablation study, which demonstrates that the integration of DPFB and IFS improves the ability of the CPP-Net to capture and extract fine texture details from the images.

References

[1] Bin Chen and Jian Zhang. Content-aware scalable deep compressed sensing. *IEEE Trans. Image Process.*, 31:5412–5426, 2022. 2

[2] Jiwei Chen, Yubao Sun, Qingshan Liu, and Rui Huang. Learning memory augmented cascading network for compressed sensing of images. In *Eur. Conf. Comput. Vis.*, pages

Table 3. Comparisons of average PSNR (dB)/SSIM performance of recent CS methods on General100 and Set14 at high CS ratios.

Datasets	Methods	0.30	0.40	0.50	Avg.
General100	CASNet	39.32/0.9730	41.56/0.9827	43.74/0.9887	41.54/0.9815
	DGUNet$^+$	38.87/0.9724	41.07/0.9821	43.26/0.9884	41.07/0.9810
	FSOINet	39.40/0.9735	41.62/0.9831	43.69/0.9887	41.57/0.9818
	OCTUF	39.54/0.9740	41.76/0.9833	43.95/0.9892	41.75/0.9822
	CPP-Net	**39.65/0.9746**	**41.92/0.9838**	**44.07/0.9894**	**41.88/0.9826**
Urban100	CASNet	33.35/0.9509	35.46/0.9668	37.46/0.9773	35.42/0.9650
	DGUNet$^+$	33.16/0.9510	35.24/0.9666	37.65/0.9785	35.35/0.9654
	FSOINet	33.84/0.9540	35.93/0.9688	37.80/0.9777	35.86/0.9668
	OCTUF	34.21/0.9555	36.25/0.9669	38.29/0.9797	36.25/0.9674
	CPP-Net	**34.58/0.9586**	**36.50/0.9714**	**38.33/0.9801**	**36.47/0.9700**

Table 4. Comparisons of average LPIPS performance of recent CS methods on General100 and Set14 at various CS ratios.

Datasets	CS Ratios	CASNet	DGUNet$^+$	FSOINet	OCTUF	**CPP-Net**
General100	0.01	0.2198	0.2265	0.2184	0.2252	**0.2130**
	0.04	0.1207	0.1208	0.1178	0.1194	**0.1108**
	0.10	0.0690	0.0653	0.0655	0.0659	**0.0629**
	0.25	0.0311	0.0282	0.0289	0.0287	**0.0273**
	0.30	0.0247	0.0231	0.0224	0.0233	**0.0214**
	0.40	0.0153	0.0142	0.0135	0.0137	**0.0130**
	0.50	0.0091	0.0089	0.0079	0.0077	**0.0075**
	Avg.	0.0700	0.0696	0.0678	0.0691	**0.0651**
Urban100	0.01	0.2866	0.2846	0.2914	0.2969	**0.2707**
	0.04	0.1731	0.1660	0.1729	0.1748	**0.1533**
	0.10	0.1070	0.0962	0.1015	0.1027	**0.0909**
	0.25	0.0520	0.0425	0.0474	0.0449	**0.0408**
	0.30	0.0424	0.0412	0.0382	0.0358	**0.0328**
	0.40	0.0280	0.0278	0.0249	0.0258	**0.0224**
	0.50	0.0185	0.0191	0.0183	0.0170	**0.0150**
	Avg.	0.1011	0.0968	0.0992	0.0997	**0.0894**

Table 5. Average PSNR (dB)/SSIM comparisons of various methods under different proportions ς of salt-and-pepper noise on Set11 at CS ratios of 0.04 and 0.10.

CS Ratios	ς	CASNet	DGUNet$^+$	FSOINet	OCTUF	**CPP-Net**
0.04	0.01	25.85/0.7855	26.16/0.7920	25.82/0.7800	25.85/0.7814	**26.45/0.7969**
	0.02	25.45/0.7615	25.65/0.7619	25.34/0.7542	25.40/0.7558	**25.98/0.7702**
	0.04	24.71/0.7172	24.91/0.7214	24.61/0.7102	24.71/0.7149	**25.14/0.7248**
	0.08	23.69/0.6538	23.74/0.6545	23.43/0.6425	23.56/0.6447	**23.91/0.6575**
0.10	0.01	27.93/0.8203	28.21/0.8252	27.98/0.8194	28.15/0.8207	**28.52/0.8286**
	0.02	26.81/0.7687	27.07/0.7699	26.86/0.7635	27.01/0.7644	**27.29/0.7720**
	0.04	25.31/0.6855	25.49/0.6857	25.26/0.6797	25.42/0.6833	**25.64/0.6863**
	0.08	23.39/0.5777	23.43/0.5796	23.26/0.5734	23.33/0.5725	**23.54/0.5798**

Table 6. Average PSNR (dB)/SSIM performance comparisons of recent CS-MRI methods on fastMRI dataset using Cartesian masks at different CS ratios.

Methods	0.05	0.10	0.15	Avg.
Zero-filled	23.23/0.4564	25.38/0.5314	26.31/0.5763	24.97/0.5214
DC-CNN	26.81/0.5432	27.07/0.5899	27.92/0.6624	27.27/0.5985
RDN	27.00/0.5630	27.32/0.5967	28.89/0.6911	27.74/0.6169
ISTA-Net$^+$	27.31/0.5594	27.82/0.6128	29.27/0.7027	28.13/0.6250
CDDN	27.45/0.5649	28.04/0.6208	29.46/0.7094	28.32/0.6317
ADMM-CSNet	27.06/0.5512	27.43/0.5999	28.79/0.6898	27.76/0.6136
HiTDUN	27.40/0.5630	28.02/0.6208	29.32/0.7053	28.25/0.6297
CPP-Net	**27.74/0.5764**	**28.90/0.6498**	**29.67/0.7161**	**28.77/0.6474**

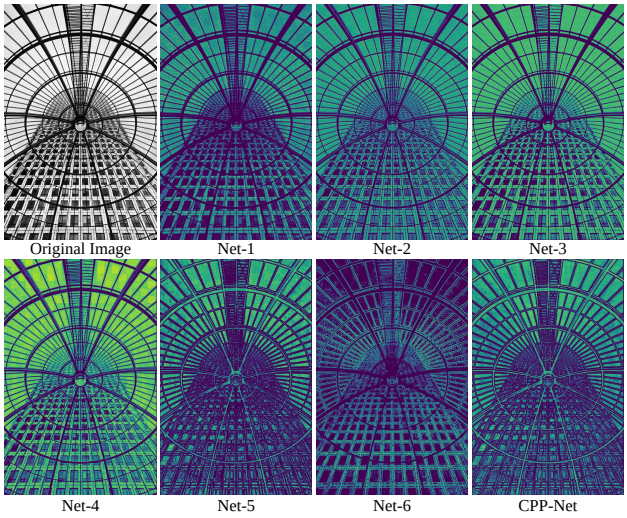

Figure 2. Visualization of feature maps for third stage of different ablation study cases at a CS ratio of 0.10. The integration of DPFB and IFS enhances CPP-Net's attention to detail.

513–529, 2020. 2

[3] Wenjun Chen, Chunling Yang, and Xin Yang. FSOINET: Feature-space optimization-inspired network for image compressive sensing. In *IEEE Int. Conf. Acoust. Speech Signal Process.*, pages 2460–2464, 2022. 2

[4] Wenxue Cui, Shaohui Liu, Feng Jiang, and Debin Zhao. Image compressed sensing using non-local neural network. *IEEE Trans. Multimedia*, 25:816–830, 2021. 2

[5] Chao Dong, Chen Change Loy, and Xiaoou Tang. Accelerating the super-resolution convolutional neural network. In *Eur. Conf. Comput. Vis.*, pages 391–407. 2016. 2

[6] 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 Trans. Cybern.*, 53(4):2558–2571, 2023. 2

[7] Hongping Gan, Minghe Shen, Yi Hua, Chunyan Ma, and Tao Zhang. From patch to pixel: A transformer-based hierarchical framework for compressive image sensing. *IEEE Trans. Comput. Imaging*, 9:133–146, 2023. 2

[8] Bingsheng He and Xiaoming Yuan. Convergence analysis of primal-dual algorithms for a saddle-point problem: From contraction perspective. *SIAM J. Imaging Sci.*, 5(1):119–149, 2012. 2

[9] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *IEEE Conf. Comput. Vis. Pattern Recognit.*, pages 5197–5206, 2015. 2

[10] Florian Knoll, Jure Zbontar, Anuroop Sriram, Matthew J. Muckley, Mary Bruno, Aaron Defazio, Marc Parente, Krzysztof J. Geras, Joe Katsnelson, Hersh Chandarana, Zizhao Zhang, Michal Drozdzalv, Adriana Romero, Michael Rabbat, Pascal Vincent, James Pinkerton, Duo Wang, Nafissa Yakubova, Erich Owens, C. Lawrence Zitnick, Michael P. Recht, Daniel K. Sodickson, and Yvonne W. Lui. fastMRI: A publicly available raw k-space and dicom dataset of knee images for accelerated MR image reconstruc-

tion using machine learning. *Radiology: Artif. Intell.*, 2(1): e190007, 2020. 2

[11] Kuldeep Kulkarni, Suhas Lohit, Pavan Turaga, Ronan Kerviche, and Amit Ashok. ReconNet: Non-iterative reconstruction of images from compressively sensed measurements. In *IEEE Conf. Comput. Vis. Pattern Recognit.*, pages 449–458, 2016. 2

[12] 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 *Eur. Conf. Comput. Vis.*, pages 740–755, 2014. 2

[13] Chong Mou, Qian Wang, and Jian Zhang. Deep generalized unfolding networks for image restoration. In *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pages 17378–17389, 2022. 2

[14] R. Tyrrell Rockafellar. Monotone operators and the proximal point algorithm. *SIAM J. Control Optim.*, 14(5):877–898, 1976. 1, 2

[15] Jo Schlemper, Jose Caballero, Joseph V. Hajnal, Anthony N. Price, and Daniel Rueckert. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Trans. Med. Imag.*, 37(2):491–503, 2018. 2

[16] Minghe Shen, Hongping Gan, Chao Ning, Yi Hua, and Tao Zhang. TransCS: A transformer-based hybrid architecture for image compressed sensing. *IEEE Trans. Image Process.*, 31:6991–7005, 2022. 2

[17] Wuzhen Shi, Feng Jiang, Shaohui Liu, and Debin Zhao. Scalable convolutional neural network for image compressed sensing. In *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, pages 12290–12299, 2019. 2

[18] Wuzhen Shi, Feng Jiang, Shaohui Liu, and Debin Zhao. Image compressed sensing using convolutional neural network. *IEEE Trans. Image Process.*, 29:375–388, 2020. 2

[19] Jiechong Song, Bin Chen, and Jian Zhang. Deep memory-augmented proximal unrolling network for compressive sensing. *Int. J. Comput. Vis.*, 131(6):1477–1496, 2023. 2

[20] Jiechong Song, Bin Chen, and Jian Zhang. Dynamic path-controllable deep unfolding network for compressive sensing. *IEEE Trans. Image Process.*, 32:2202–2214, 2023. 2

[21] Jiechong Song, Chong Mou, Shiqi Wang, Siwei Ma, and Jian Zhang. Optimization-inspired cross-attention transformer for compressive sensing. In *IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2023. 2

[22] Liyan Sun, Zhiwen Fan, Yue Huang, Xinghao Ding, and John Paisley. Compressed sensing MRI using a recursive dilated network. In *AAAI Conf. Artif. Intell.*, 2018. 2

[23] Yubao Sun, Jiwei Chen, Qingshan Liu, Bo Liu, and Guodong Guo. Dual-path attention network for compressed sensing image reconstruction. *IEEE Trans. Image Process.*, 29: 9482–9495, 2020. 2

[24] Yan Yang, Jian Sun, Huibin Li, and Zongben Xu. ADMM-CSNet: A deep learning approach for image compressive sensing. *IEEE Trans. Pattern Anal. Mach. Intell.*, 42(3):521–538, 2020. 2

[25] Dongjie Ye, Zhangkai Ni, Hanli Wang, Jian Zhang, Shiqi Wang, and Sam Kwong. CSformer: Bridging convolution and transformer for compressive sensing. *IEEE Trans. Image Process.*, 32:2827–2842, 2023. 2

[26] Di You, Jian Zhang, Jingfen Xie, Bin Chen, and Siwei Ma. COAST: Controllable arbitrary-sampling network for compressive sensing. *IEEE Trans. Image Process.*, 30:6066–6080, 2021. 2

[27] Jian Zhang and Bernard Ghanem. ISTA-Net: Interpretable optimization-inspired deep network for image compressive sensing. In *IEEE Conf. Comput. Vis. Pattern Recognit.*, pages 1828–1837, 2018. 1, 2

[28] Jian Zhang, Chen Zhao, and Wen Gao. Optimization-inspired compact deep compressive sensing. *IEEE J. Sel. Top. Signal Process.*, 14(4):765–774, 2020. 2

[29] Jian Zhang, Zhenyu Zhang, Jingfen Xie, and Yongbing Zhang. High-throughput deep unfolding network for compressive sensing MRI. *IEEE J. Sel. Top. Signal Process.*, 16 (4):750–761, 2022. 2

[30] Tianfang Zhang, Lei Li, and Zhenming Peng. Optimization-inspired cumulative transmission network for image compressive sensing. *Knowl.-Based Syst.*, 279:110963, 2023. 2

[31] Zhonghao Zhang, Yipeng Liu, Jiani Liu, Fei Wen, and Ce Zhu. AMP-Net: Denoising-based deep unfolding for compressive image sensing. *IEEE Trans. Image Process.*, 30: 1487–1500, 2021. 2

[32] Hao Zheng, Faming Fang, and Guixu Zhang. Cascaded dilated dense network with two-step data consistency for MRI reconstruction. In *Adv. Neural Inform. Process. Syst.*, 2019. 2

[33] Siwang Zhou, Yan He, Yonghe Liu, Chengqing Li, and Jianming Zhang. Multi-channel deep networks for block-based image compressive sensing. *IEEE Trans. Multimedia*, 23: 2627–2640, 2021. 2