Improving the Learning Capability of Small-size Image Restoration Network by Deep Fourier Shifting

Man Zhou

Aerospace Information Research Institute, Chinese Academy of Sciences

University of Science and Technology of China

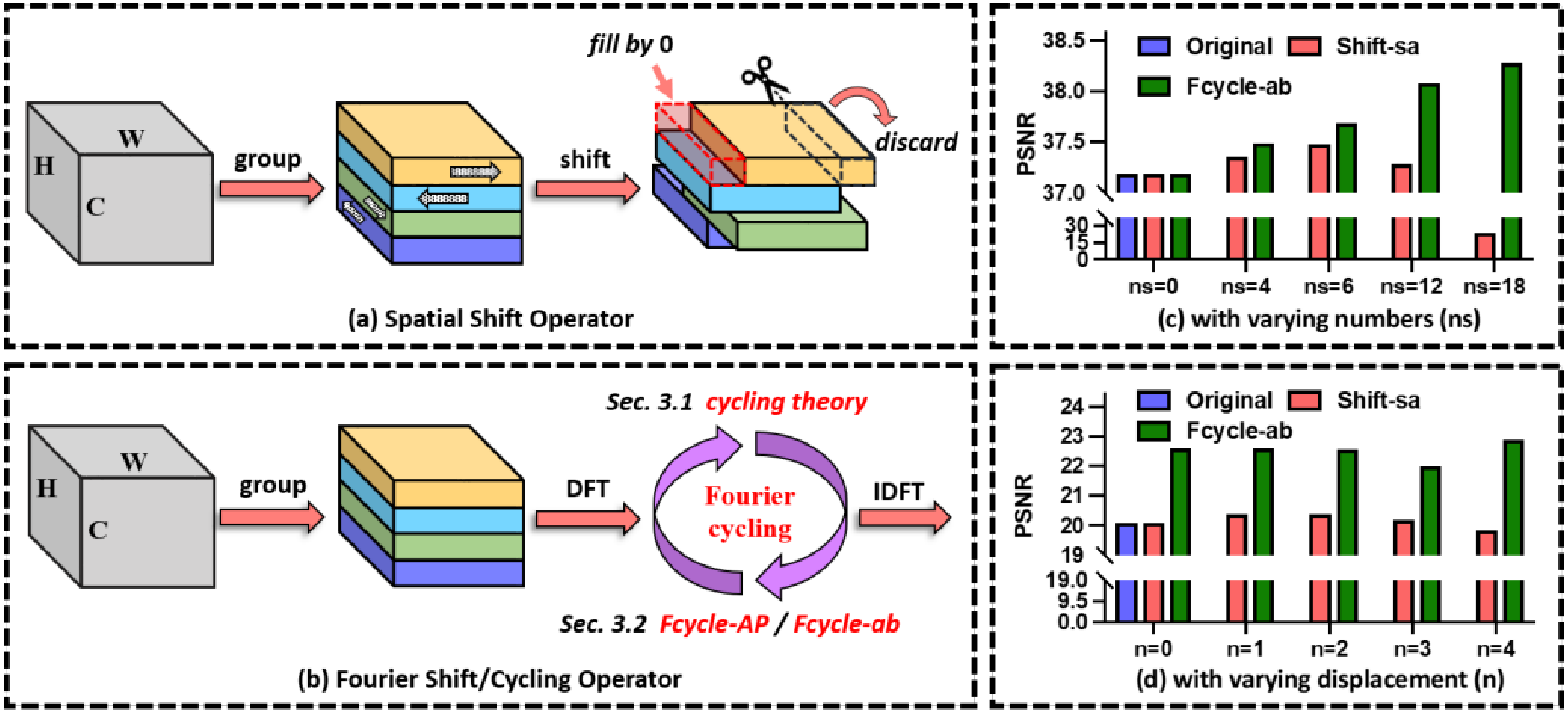


Figure 1: **Comparison between the spatial shift operator and the proposed deep Fourier shift operator.** (a) Traditional spatial shift operator involves a spatial shift mechanism that moves each channel of the input tensor in a distinct spatial direction, thus suffering from severe region-aware information loss and conflicting with the requirements of image restoration tasks. (b) Deep Fourier Shifting/Cycling operator is a more ingenious information-lossless operator, which is tailored for image restoration tasks. (c), (d) Deep Fourier shifting achieves a more stable performance gain than the spatial shifting mechanism with varying “ns” shift displacements and “n” basic units over image de-noising task where the cut-off is for compressing the vertical axis scale to better illustrate the contrast effect clearly.

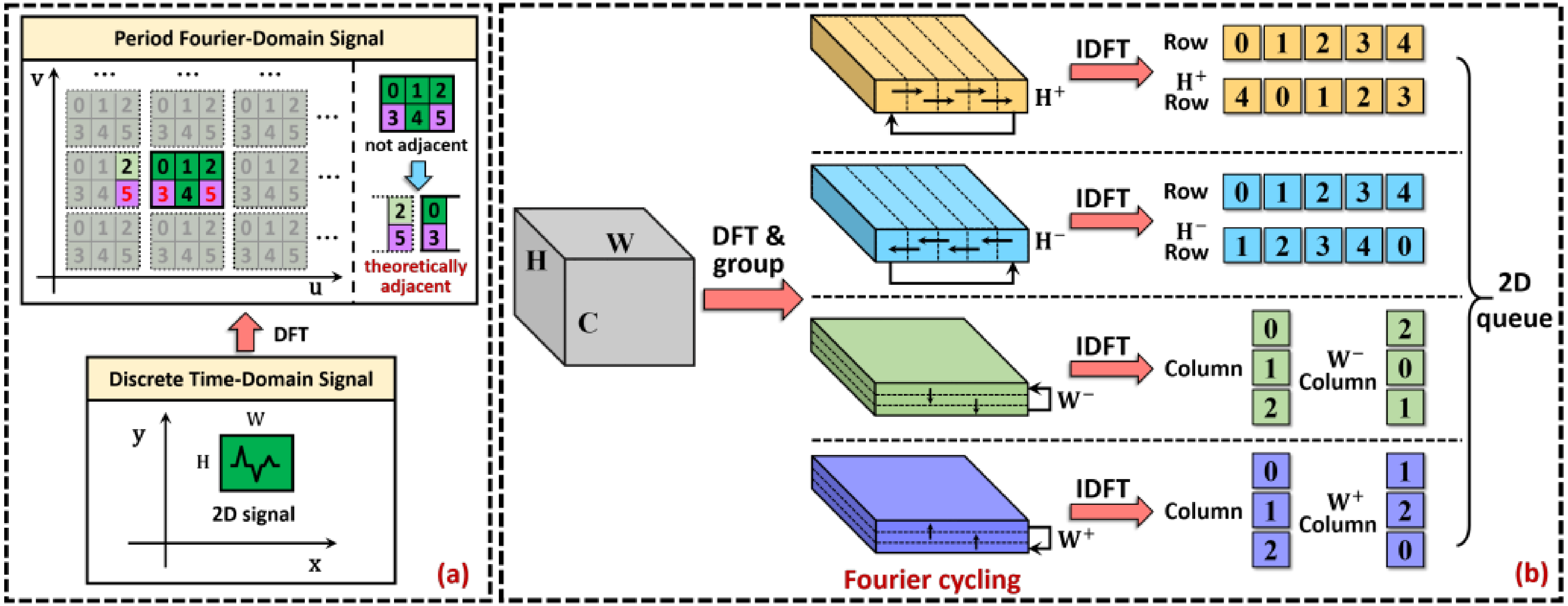


Figure 2: **(a) The information-lossless cycling mechanism.** The discrete Fourier transform of a signal exhibits period-extended and cycling properties. Specifically, in the Fourier domain, the two pixels in sequence beginning and end may not appear adjacent, but due to the period property, they are actually considered adjacent, as indicated by the upper right corner. This inherent period-extended and cycling behavior of the Fourier transform enables us to model the shifting mechanism in a manner that is information-lossless, making it well-suited for image restoration tasks. Consider the Fourier transform of a discrete time-domain signal, represented as $\left(\begin{smallmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \end{smallmatrix}\right)$. It may appear that the values 3 and 5 are not adjacent within the main period. However, owing to the property of period extension, the 3 from the previous period and the 5 from the current period are theoretically considered adjacent. It is reasonable to move the removed area from the end to the front, meeting the cycling mechanism. **(b) Our deep Fourier shifting operator.** Our operator borrows the principle of the spatial shifting mechanism and models the shifting mechanism in information-lossless Fourier cycling rules. The cycling is coded as 2D queue rolling.

Definitions. $f(x, y) \in \mathbb{R}^{\mathrm{H}\times\mathrm{W}\times\mathtt{C}}$ is the spatial signal and $F(u, v) \in \mathbb{R}^{\mathrm{H}\times\mathrm{W}\times\mathtt{C}}$ denotes its Fourier transform where (x, y) and (u, v) represent the space coordinates and Fourier spectrum, respectively.

Theorem. *The Fourier transform of a discrete signal is a period-extended and cycling:* $F(u, v) = F(u + n\mathrm{H}, v) = F(u, v + m\mathrm{W}) = F(u + n\mathrm{H}, v + m\mathrm{W})$ *where* $u = 0, 1, 2, \ldots, \mathrm{H} - 1,\ v = 0, 1, 2, \ldots, \mathrm{W} - 1$ *and* $n, m \in \mathbb{N}$. $\mathbb{N}$ *is the set of positive integers starting from zero.*

2.1 Proof: The Fourier Transform of a Discrete Signal is Period-Extended and Cycling

We show the periodicity and cycling properties of the Fourier transform of a discrete signal, as illustrated in Figure 2(a). Note that the Fourier transform $F(u, v)$ of $f(x, y)$ is expressed as

$$F(u, v) = \frac{1}{\mathrm{HW}} \sum_{x=0}^{\mathrm{H}-1} \sum_{y=0}^{\mathrm{W}-1} f(x, y) e^{-j2\pi(\frac{ux}{\mathrm{H}} + \frac{vy}{\mathrm{W}})}. \tag{1}$$

Then, we show the periodicity of $F(u, v) \in \mathbb{R}^{\mathrm{H}\times\mathrm{W}}$ with H and W. It means $F(u, v) = F(u + n\mathrm{H}, v) = F(u, v + m\mathrm{W}) = F(u + n\mathrm{H}, v + m\mathrm{W})$ where $u = 0, 1, 2, \ldots, \mathrm{H} - 1,$

$v = 0, 1, 2, \ldots, \mathrm{W}-1$ and $n, m \in \mathbb{N}$ that records the set of non-negative integers. We take the $F(u,v) = F(u+n\mathrm{H}, v+m\mathrm{W})$ for example and recall Eq. (1) as

$$\begin{aligned}
&F(u+n\mathrm{H}, v+m\mathrm{W}) \\
&= \frac{1}{\mathrm{HW}} \sum_{x=0}^{\mathrm{H}-1} \sum_{y=0}^{\mathrm{W}-1} f(x,y).e^{-j2\pi(\frac{(u+n\mathrm{H})x}{\mathrm{H}} + \frac{(v+m\mathrm{W})y}{\mathrm{W}})} \\
&= \frac{1}{\mathrm{HW}} \sum_{x=0}^{\mathrm{H}-1} \sum_{y=0}^{\mathrm{W}-1} f(x,y) e^{-j2\pi(\frac{ux}{\mathrm{H}} + \frac{vy}{\mathrm{W}})} e^{-2j\pi m} e^{-2j\pi n} \\
&= F(u,v) e^{-2j\pi m} e^{-2j\pi n},
\end{aligned} \tag{2}$$

where for any integer z, it has $e^{-2j\pi z} = 1$.

Further, $e^{-2j\pi n} = 1$ and $e^{-2j\pi m} = 1$ for $n, m \in \mathbb{N}$. Therefore,

$$F(u+n\mathrm{H}, v+m\mathrm{W}) = F(u,v) e^{-2j\pi m} e^{-2j\pi n} = F(u,v). \tag{3}$$

Similarly, we can prove the periodicity of $F(u,v)$ as well.

$$\begin{aligned}
F(u,v) &= F(u+n\mathrm{H}, v+m\mathrm{W}) \\
&= F(u+n\mathrm{H}, v) = F(u, v+m\mathrm{W}).
\end{aligned} \tag{4}$$

Furthermore, deep Fourier transform can be expressed in Cartesian and polar coordinates by an equivalent form as

$$F(u,v) = \mathrm{A}e^{j\mathrm{P}} = a + bj. \tag{5}$$

The period-extended and cycling property holds over the amplitude-phase and real-imaginary format.

```
def DFS_AP(X):
 # X: input with shape [N, C, H, W]
 # A and P are the amplitude and phase
   A.e{jP} = FFT(X)

   # Fourier shifting transform rules
   A_g = torch.spilt(A, 4, dim=1)
   P_g = torch.spilt(P, 4, dim=1)
   A_fc = Fourier-cycling(A_g)
   P_fc = Fourier-cycling(P_g)
   A_fc = Convs_1x1(A_fc)
   P_fc = Convs_1x1(P_fc)

   # Inverse Fourier transform
   Y = iFFT(A_fc, P_fc)

   Return Y #[N, C, H, W]
```

```
def DFS_ab(X):
 # X: input with shape [N, C, H, W]
 # a and b are the real and imaginary part
   a+bj = FFT(X)

   # Fourier shifting transform rules
   a_g = torch.spilt(a, 4, dim=1)
   b_g = torch.spilt(b, 4, dim=1)
   a_fc = Fourier-cycling(a_g)
   b_fc = Fourier-cycling(b_g)
   A_fc = Convs_1x1(a_fc)
   P_fc = Convs_1x1(b_fc)

   # Inverse Fourier transform
   Y = iFFT(a_fc, b_fc)

   Return Y #[N, C, H, W]
```

Figure 3: **Pseudo-code of the two variants of the proposed deep Fourier shifting.** The left is the *amplitude-phase variant* while the right is the *real-imaginary variant*.

Table 1: Quantitative comparisons on low-light image enhancement. The arrow $\rightarrow$ denotes the generalization setting by training on the data before the arrow and testing directly on the data after the arrow.

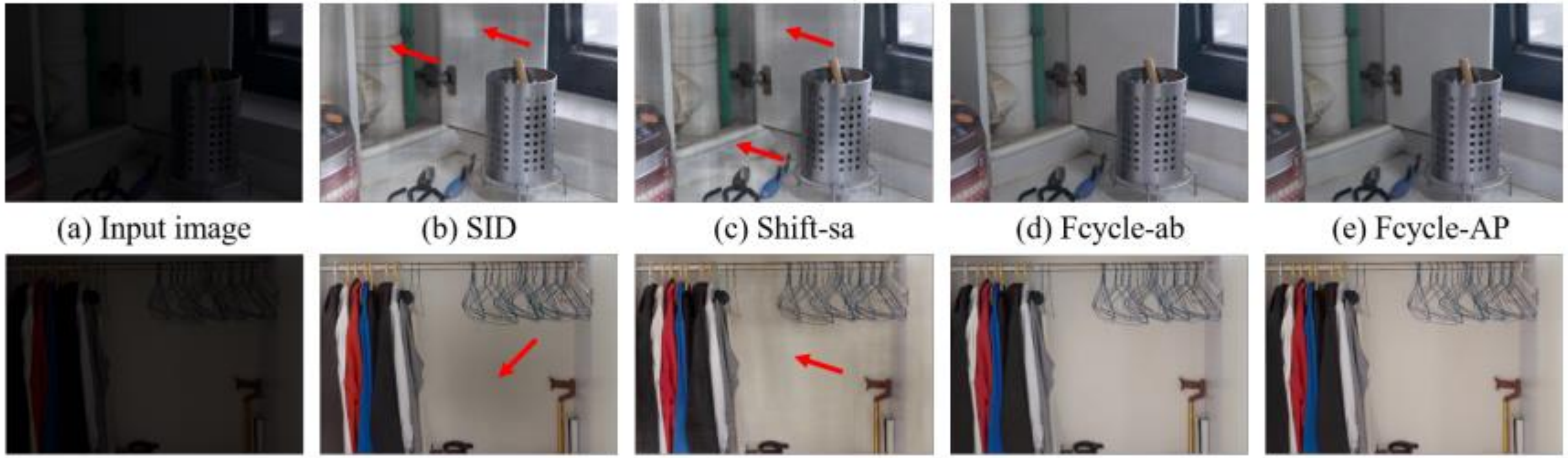
Model	Config	LOL $\rightarrow$		$\rightarrow$ Huawei		Huawei $\rightarrow$		$\rightarrow$ LOL		#Paras
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
DRBN	Original	19.7931	0.8361	17.7929	0.6247	20.1549	0.6851	18.0856	0.7543	0.55M
	Shift-sa	19.7072	0.8343	17.6221	0.6071	20.2165	0.6873	17.9112	0.7532	0.41M
	Fcycle-AP	**22.4274**	**0.8448**	**19.3252**	**0.6472**	20.5855	0.6872	18.8666	0.7587	0.41M
	Fcycle-ab	22.2054	0.8429	19.3125	0.6431	**20.6651**	**0.6876**	**19.1535**	**0.7681**	0.41M
SID	Original	20.1062	0.7895	16.5874	0.5925	20.1742	0.6659	18.5468	0.7441	7.76M
	Shift-sa	20.0148	0.7911	16.8214	0.5911	20.1517	0.6651	18.4998	0.7434	7.53M
	Fcycle-AP	**22.8565**	**0.8019**	19.1707	0.6238	20.9068	**0.6708**	**18.8161**	**0.7494**	7.53M
	Fcycle-ab	22.6313	0.7995	**19.2471**	**0.6242**	**20.9271**	0.6691	18.5741	0.7443	7.53M

Table 2: Comparisons on image denoising.

Dataset	Metric	DNCNN			
		Original	Shift-sa	Fcycle-AP	Fcycle-ab
SIDD	PSNR	37.1992	37.2247	37.6891	**38.1837**
	SSIM	0.8954	0.8980	**0.9013**	0.9066
DND	PSNR	38.33	38.42	**38.69**	38.65
	SSIM	0.8974	0.8985	0.942	**0.949**
	#Paras	1.51M	1.43M	1.43M	1.43M

Table 3: Comparisons on image deblurring.

Dataset	Metric	DeepDeblur			
		Original	Shift-sa	Fcycle-AP	Fcycle-ab
GoPro	PSNR	28.9423	28.9037	**29.2123**	29.1939
	SSIM	0.8716	0.8712	**0.8777**	0.8765
HIDE	PSNR	26.9770	26.9991	**27.2860**	27.2547
	SSIM	0.8468	0.8476	**0.8541**	0.8536
	#Paras	11.72M	10.61M	10.61M	10.61M



(a) Input image (b) SID (c) Shift-sa (d) Fcycle-ab (e) Fcycle-AP

(a) Input image (b) DRBN (c) Shift-sa (d) Fcycle-ab (e) Fcycle-AP

Figure 5: Visual comparison over image enhancement.

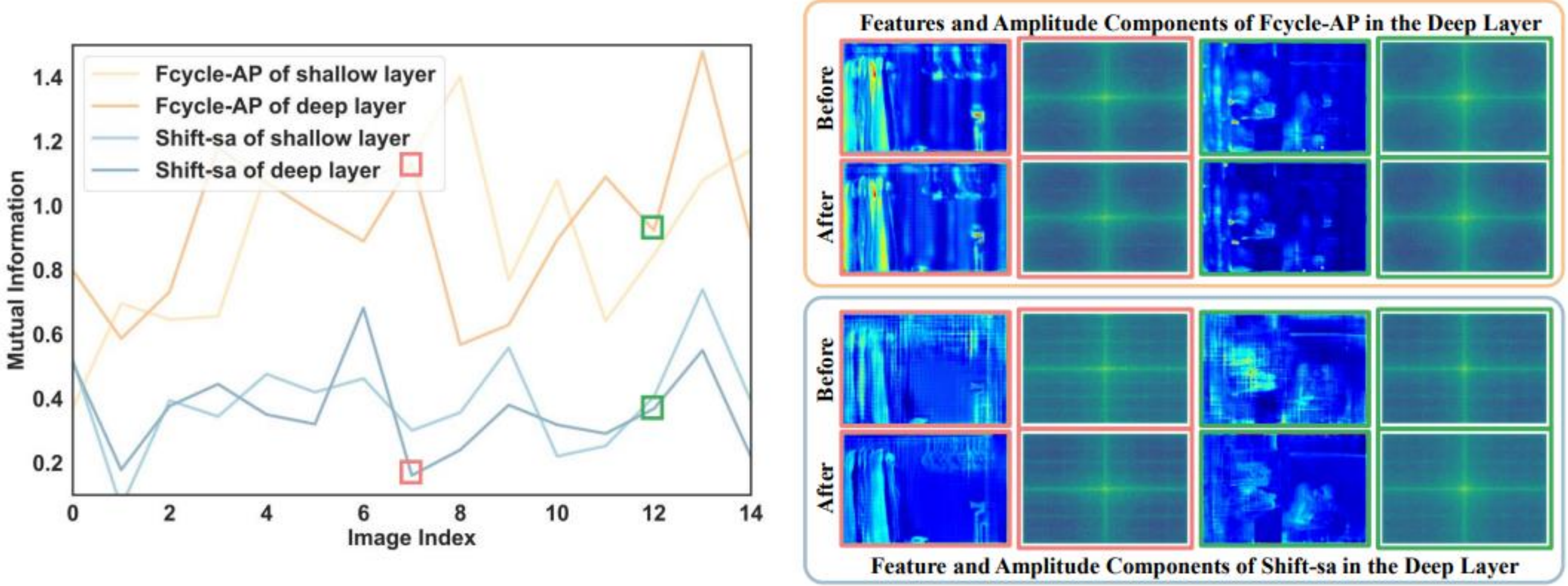


Figure 6: The effectiveness of information preservation. **Left:** we compare mutual information levels before and after employing Fcycle-ab and Shift-sa operators on the LOL test set, respectively. Our operator exhibits significantly higher mutual information than Shift-sa, showcasing its efficacy in information preservation. **Right:** we visualize feature maps and their amplitude components before and after operations. This demonstrates that our Fcycle-AP promotes frequency information

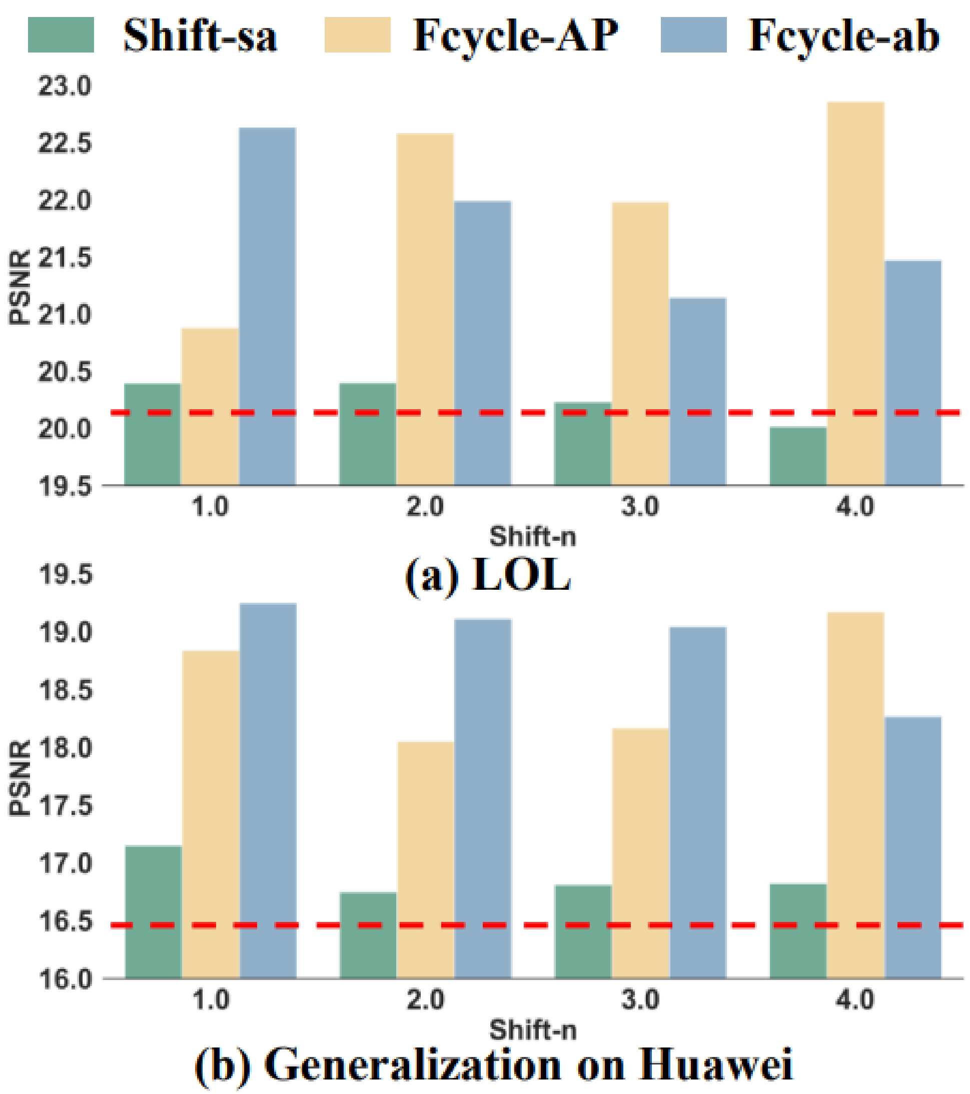


Figure 7: The effect of shifting displacement shift-n on SID.

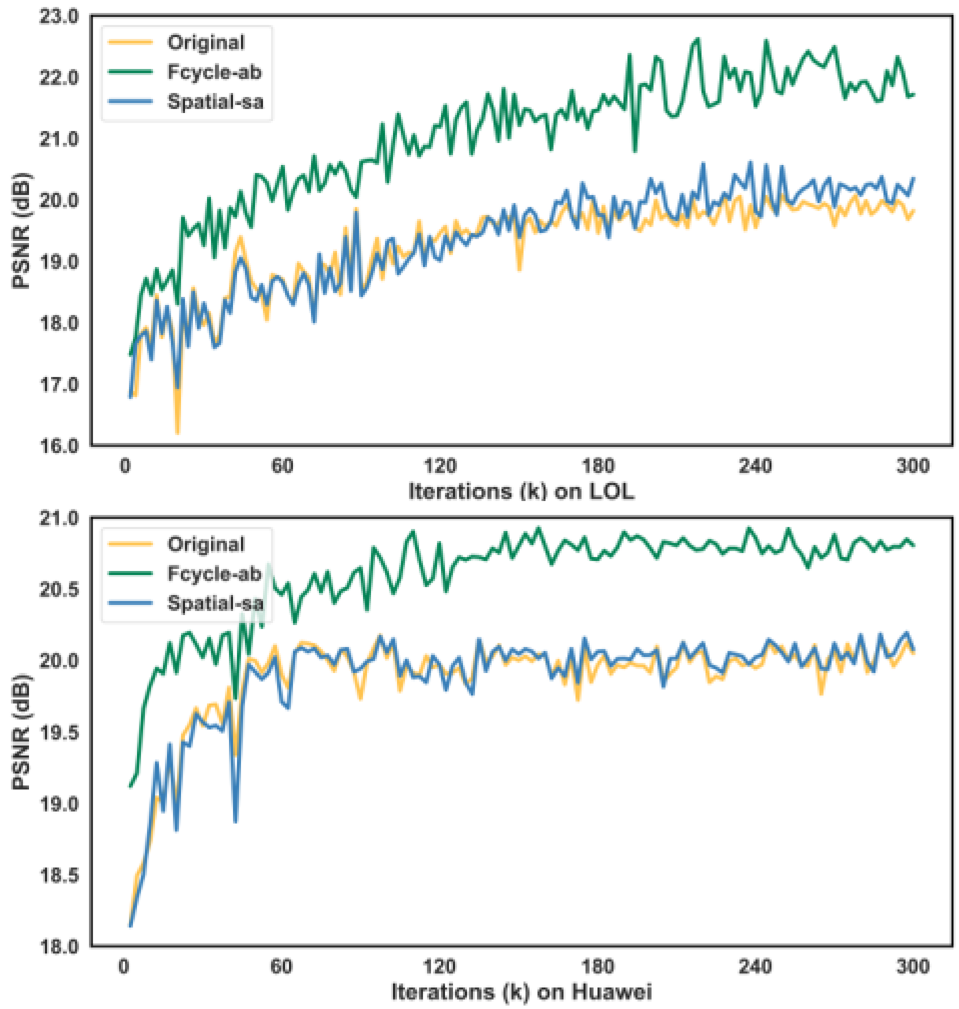


Figure 4: The proposed operators improve the training performance. It shows the training PSNR on the image enhancement task on the LOL and Huawei datasets in the top and bottom.

Thanks for your attention!