

Supplementary Materials: Learning Enriched Features via Selective State Spaces Model for Efficient Image Deblurring

Anonymous Authors

1 DATASETS AND EXPERIMENTAL DETAILS

In this section, we elaborate on the datasets and training settings utilized for five image restoration tasks. A summary of the datasets is presented in Table 1.

1.1 Image Motion Deblurring

In line with established methodologies [4, 12], we utilize the GoPro [8] dataset to train our ALGNet. To assess the generalization capability of ALGNet, we evaluate our GoPro-trained model directly on the test sets of the HIDE [10] and RealBlur [9] datasets. While both the GoPro and HIDE datasets are synthetically generated, the RealBlur dataset consists of image pairs captured under real-world conditions. This dataset comprises two subsets: (1) RealBlur-J, which obtained as camera JPEG outputs, and (2) RealBlur-R, generated offline by applying white balance, demosaicking, and denoising operations to RAW images. Each subset comprises 980 images. We adopt the training strategy outlined in MPRNet [12].

1.2 Single-Image Defocus Deblurring

Following the methodology of previous approaches [2, 11], we employ the recently introduced dataset DPDD [1] for evaluation purposes. The training strategy follows that of IRNext [3].

2 MORE ABLATION STUDIES

We provide more ablation studies on the GoPro dataset [8].

2.1 Loss function

To investigate the influence of the loss function, we conduct experiments by varying the combination of loss functions. The results are summarized in Table 2. In our baseline, we use the L1 loss. However, when employing the proposed loss function (L), we achieve optimal results, with a 0.17 dB improvement over using only the L1 loss.

2.2 Multi-input and Multi-output

To assess the impact of multi-scale input and output on the experimental results, we compared them with single-input and single-output configurations, as shown in Table 3. The use of multi-input and multi-output resulted in a 0.12 dB improvement. Additionally, we observed that the training process was more stable when employing the multi-input and multi-output mode.

3 VISUALIZATION

To demonstrate the effectiveness of our FA, we visualize intermediate feature maps by plotting the resulting feature maps of ALGBlock for models with and without FA in Figure 1. It is evident that when equipped with FA, the model recovers more image details compared to the model without FA.

Table 1: Dataset description for various image deblurring.

Tasks	Datasets	Train Samples	Test Samples
Motion Deblurring	GoPro [8]	2130	1111
	HIDE [10]	0	2025
	RealBlur [9]	0	1960
Defocus Deblurring	DDPD [1]	350	76

Table 2: The impact of loss function.

L_{char}	L_{freq}	L_{edge}	PSNR
✗	✗	✗	33.32
✓	✗	✗	33.38
✓	✓	✗	33.45
✓	✓	✓	33.49

Table 3: The impact of multi-input and multi-output.

Multi-input and Multi-output	PSNR
✗	33.37
✓	33.49

4 ADDITIONAL VISUAL RESULTS

In this part, we provide additional visual results with state-of-the-art methods to demonstrate the effectiveness of the proposed method, organized as follows:

Image motion deblurring: Figure 2 3 4.

Single-image defocus deblurring: Figure 5.

REFERENCES

- [1] Abdullah Abuolaim and Michael S Brown. 2020. Defocus deblurring using dual-pixel data. In *European Conference on Computer Vision*. Springer, 111–126.
- [2] Yuning Cui, Wenqi Ren, Xiaochun Cao, and Alois Knoll. 2024. Image Restoration via Frequency Selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 46, 2 (2024), 1093–1108. <https://doi.org/10.1109/TPAMI.2023.3330416>
- [3] Yuning Cui, Wenqi Ren, Sining Yang, Xiaochun Cao, and Alois Knoll. 2023. IRNeXt: Rethinking Convolutional Network Design for Image Restoration. In *Proceedings of the 40th International Conference on Machine Learning*.
- [4] Yuning Cui, Yi Tao, Zhenshan Bing, Wenqi Ren, Xinwei Gao, Xiaochun Cao, Kai Huang, and Alois Knoll. 2023. Selective Frequency Network for Image Restoration. In *The Eleventh International Conference on Learning Representations*.
- [5] J. Dong, J. Pan, Z. Yang, and J. Tang. 2023. Multi-scale Residual Low-Pass Filter Network for Image Deblurring. In *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. 12311–12320.
- [6] Hang Guo, Jinmin Li, Tao Dai, Zhihao Ouyang, Xudong Ren, and Shu-Tao Xia. 2024. MambaR: A Simple Baseline for Image Restoration with State-Space Model. *arXiv preprint arXiv:2402.15648* (2024).
- [7] Lingshun Kong, Jiangxin Dong, Jianjun Ge, Mingqiang Li, and Jinshan Pan. 2023. Efficient Frequency Domain-based Transformers for High-Quality Image Deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 5886–5895.
- [8] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. 2016. Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), 257–265.

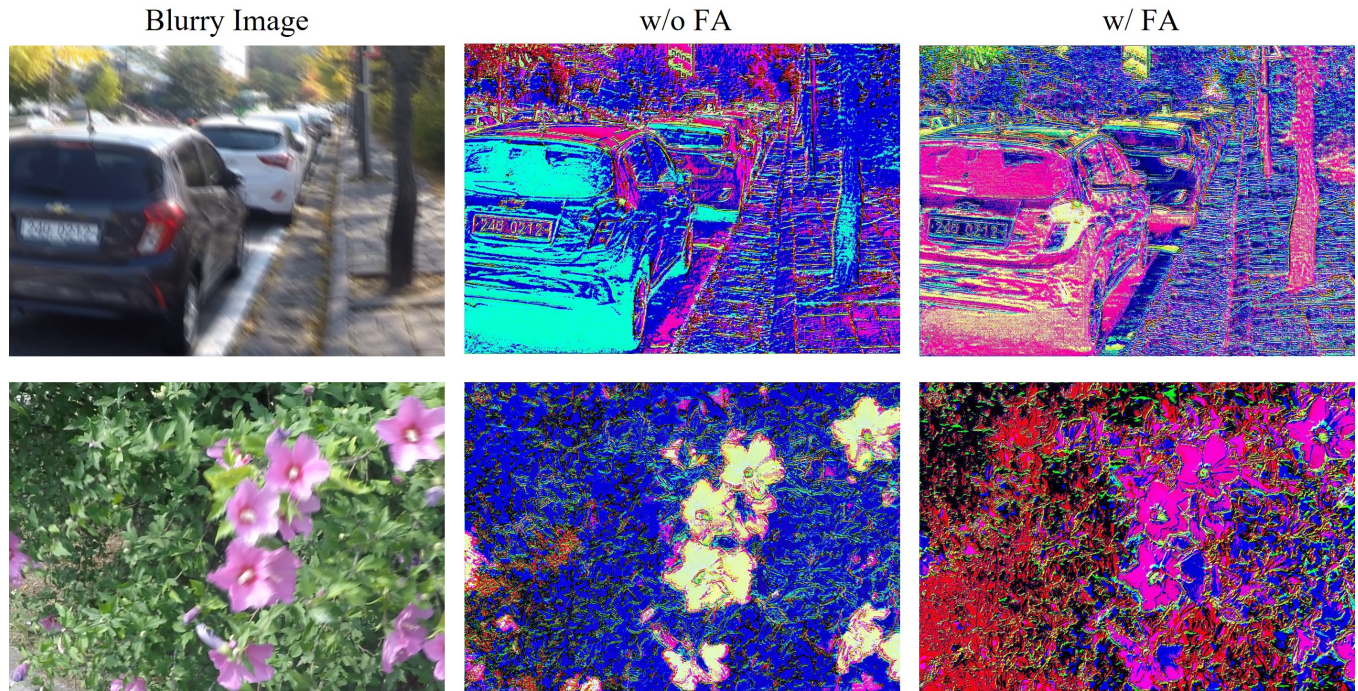


Figure 1: Visualization of intermediate feature maps of models with and without FA.

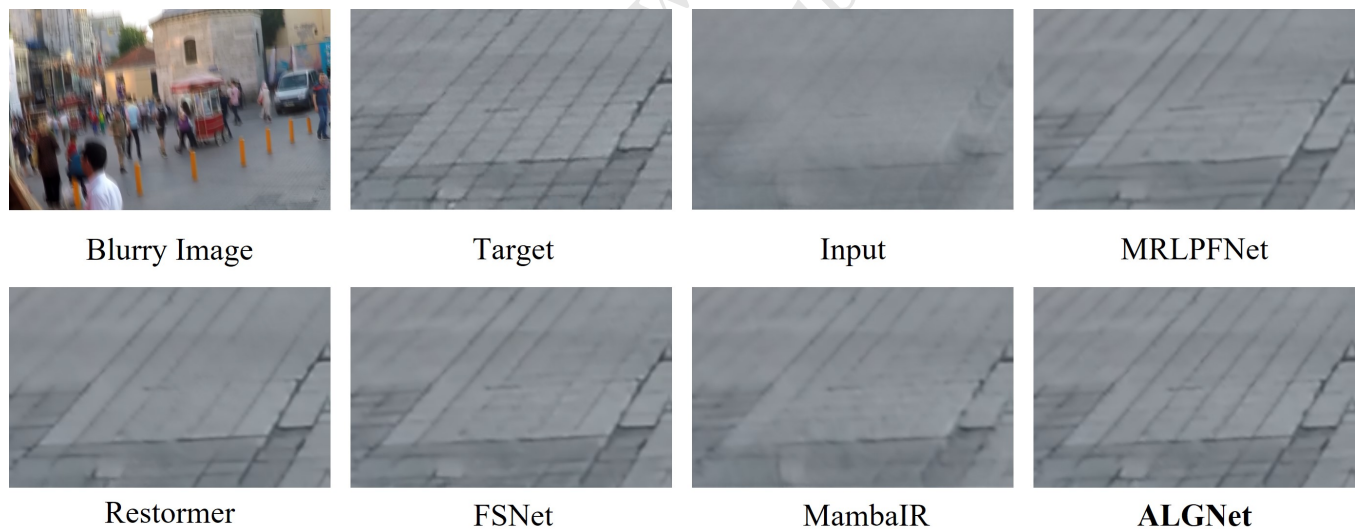


Figure 2: Comparison of image motion deblurring on the GoPro dataset [8] among MRLPFNet [5], Restormer [11], FFTformer [7], MambaIR [6], and our ALGNet.

- [9] Jaesung Rim, Haeyun Lee, Jucheol Won, and Sunghyun Cho. 2020. Real-World Blur Dataset for Learning and Benchmarking Deblurring Algorithms. In *Proceedings of the European Conference on Computer Vision (ECCV)*.
- [10] Ziyi Shen, Wenguan Wang, Xiankai Lu, Jianbing Shen, Haibin Ling, Tingfa Xu, and Ling Shao. 2019. Human-Aware Motion Deblurring. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)* (2019), 5571–5580.

- [11] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. 2022. Restormer: Efficient Transformer for High-Resolution Image Restoration. In *CVPR*.
- [12] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. 2021. Multi-Stage Progressive Image Restoration. In *CVPR*.



Figure 3: Comparison of image motion deblurring on the HIDE dataset [10] among MRLPFNet [5], Restormer [11], FFTformer [7], MambaIR [6], and our ALGNet.



Figure 4: Comparison of image motion deblurring on the RealBlur dataset [9] among MRLPFNet [5], Restormer [11], FFTformer [7], MambaIR [6], and our ALGNet.



Figure 5: Comparison of single-image defocus deblurring on the DDPD dataset [1] among SFNet [4], Restormer [11], MambaIR [6], and our ALGNet.