Distilling Semantic Priors from SAM to Efficient Image Restoration Models
Supplementary Material

Quan Zhang[1,3,*] Xiaoyu Liu[2,3,*] Wei Li[3] Hanting Chen[3] Junchao Liu[3] Jie Hu[3]
Zhiwei Xiong[2] Chun Yuan[1,†] Yunhe Wang[3,†]

[1]Tsinghua Shenzhen International Graduate School
[2]University of Science and Technology of China [3]Huawei Noah's Ark Lab

1. Comparison of existing SAM priors introduction methods.

To thoroughly validate the superiority of our proposed framework over existing methods that incorporate SAM's priors to enhance the image deblurring task, we conducted further testing for the advanced image restoration (IR) model, NAFNet [1], on two widely used datasets, including the GoPro [5] dataset and the ReLoBlur [3] dataset.

In our experiments, we employ the NAFNet architecture with 32 channels as the image deblurring model. We compare three different training methods: one utilizing the concatenation method (CAT) for incorporating SAM priors as proposed in [2], the second method exploiting only the MAP Unit of SAM-Deblur [4], and finally, the SAM-Deblur framework, which combines both the MAP Unit and mask dropout. SAM-Deblur currently represents the state-of-the-art method that leverages SAM's priors to enhance image restoration tasks.

As illustrated in Table 1, the performance of the aforementioned methods on the datasets mentioned clearly showcases the effectiveness and superiority of our proposed framework compared to existing approaches that rely on SAM's priors. It is worth noting that our framework not only achieves better performance but also preserves the inference efficiency of the original IR models without the need for SAM in the inference stage.

Method	GoPro		ReLoBLur	
	PSNR ↑	SSIM ↑	PSNR ↑	SSIM ↑
NAFNet [1]	32.85	0.960	25.26	0.687
NAFNet + CAT [2]	32.88	0.961	29.77	0.882
NAFNet + MAP [4]	32.82	0.960	30.86	0.897
NAFNet + SAM-Deblur [4]	32.83	0.868	32.29	0.903
NAFNet + Ours	**32.90**	**0.961**	**32.35**	**0.904**

Table 1. Quantitative comparison of different methods that introduce SAM priors to boost the deblurring task.

*Both authors contributed equally to this research, which was done during Quan Zhang and Xiaoyu Liu's internship at Huawei Noah's Ark Lab.
†Corresponding author

2. More visualization.

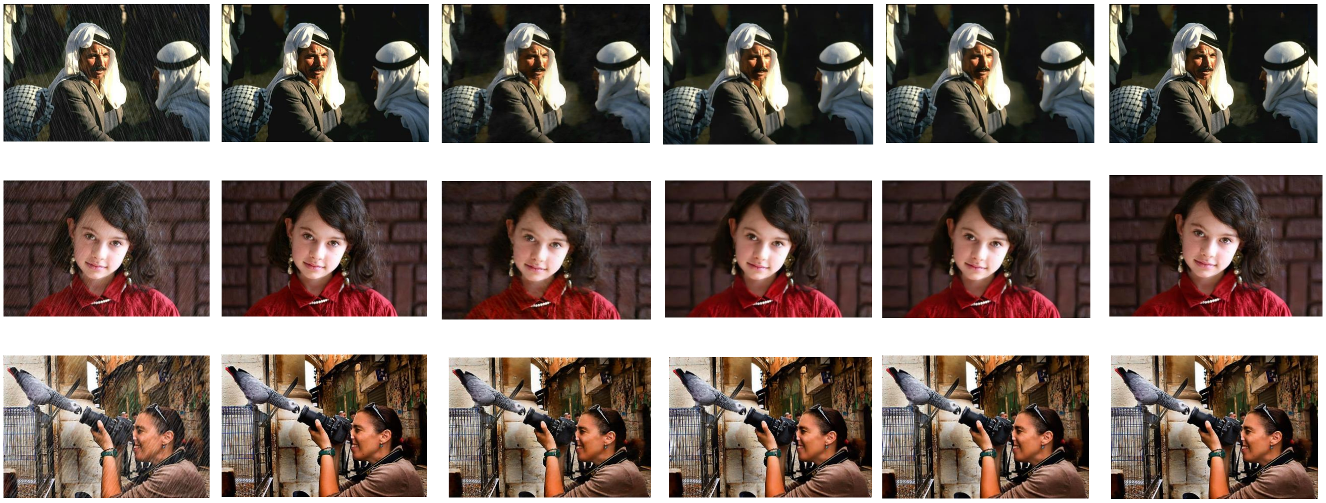

Figure 1. The qualitative comparison of IR models with and without our framework on the DDN dataset for the deraining task.

Figure 2. The qualitative comparison of IR models with and without our framework on the cityscape-syn 100 dataset for the deraining task.

Figure 3. The qualitative comparison of IR models with and without our framework on the cityscape-syn 200 dataset for the deraining task.

Figure 4. The qualitative comparison of IR models with and without our framework on the GoPro dataset for the deblurring task.

References

[1] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *European Conference on Computer Vision*, pages 17–33. Springer, 2022. 1

[2] Zheyan Jin, Shiqi Chen, Yueting Chen, Zhihai Xu, and Huajun Feng. Let segment anything help image dehaze. *arXiv preprint arXiv:2306.15870*, 2023. 1

[3] Haoying Li, Ziran Zhang, Tingting Jiang, Peng Luo, Huajun Feng, and Zhihai Xu. Real-world deep local motion deblurring. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 1314–1322, 2023. 1

[4] Siwei Li, Mingxuan Liu, Yating Zhang, Shu Chen, Haoxiang Li, Hong Chen, and Zifei Dou. Sam-deblur: Let segment anything boost image deblurring. *arXiv preprint arXiv:2309.02270*, 2023. 1

[5] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3883–3891, 2017. 1