Unsupervised Blind Image Deblurring Based on Self-Enhancement

Lufei Chen Xiangpeng Tian Shuhua Xiong Yinjie Lei Chao Ren*
College of Electronics and Information Engineering, Sichuan University, China
{chenlufei, tianxp}@stu.scu.edu.cn, {xiongsh, yinjie, chaoren}@scu.edu.cn

A. Appendix Section

The supplementary material mainly includes the following contents:

- The specific structure of certain used networks.
- The complete structure of the proposed unsupervised framework.
- More detailed explanations of the experiments mentioned in the main document.
- Additional visual comparisons with advanced unsupervised methods.

B. Architecture of the Generator

The structure of the generator used in our network is shown in Fig. 1. Specifically, within the generator, we first extract sharp image features through sharp feature extraction, which are then combined with blurry features obtained from the DGIG module and passed through six residual blocks. Each residual block consists of two 3×3 convolution layers with ReLU activation function. Finally, these features are processed through four convolution layers to synthesize the final blurry image.

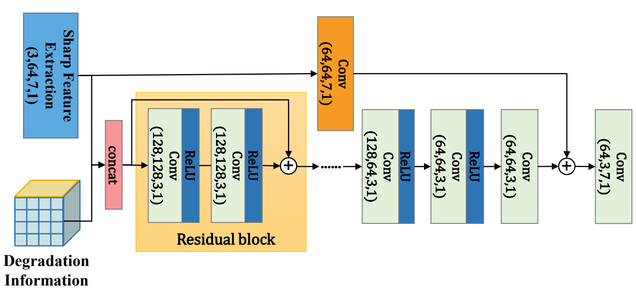

Figure 1. The architecture of the generator.

C. Architecture of the Discriminator

In our network, we use Patch-GAN [11] as the discriminator, as shown in Fig. 2. The input to the discriminator is an image of size 128×128, starting with a 4×4 convolution layer

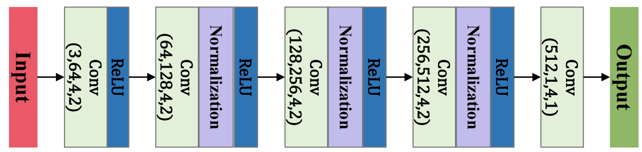

Figure 2. The architecture of the discriminator.

with ReLU activation function. This is followed by three intermediate layers, each of which adds instance normalization between the convolution layer and the ReLU activation function, and ending with a 4×4 convolution layer with a stride of 1. Despite its simplicity, this discriminator structure is capable of focusing on each pixel in the image, leading to more effective training. The adversarial training between the discriminator and the generator ensures that the synthesized blurry images are close to real-world situations.

D. Complete Structure of SEMGUD

In the main document, we describe our multi-generator unsupervised deblurring (MGUD) framework for real world images in detail. In the supplementary material, we elaborate on the implementation of the self-enhancement strategy based on the original MGUD, as shown in Fig. 3. By fixing the reconstructor obtained from the previous training iteration at the input end of the DGIG module, we acquire better degradation guidance information and generate better pseudo-paired data, which further train and refine the reconstructor. Note that, for training stability and fast convergence, we train initial reconstructor weights using the synthetic dataset generated by the kernel estimation network exploited in the main document. After that, the initial reconstructor is used in the real image deblurring framework without the need of real-paired data. As described in the main document, our self-enhancement strategy significantly enhances the reconstructor's performance without the need to alter the network's architecture or increase inference computational complexity.

*Corresponding author

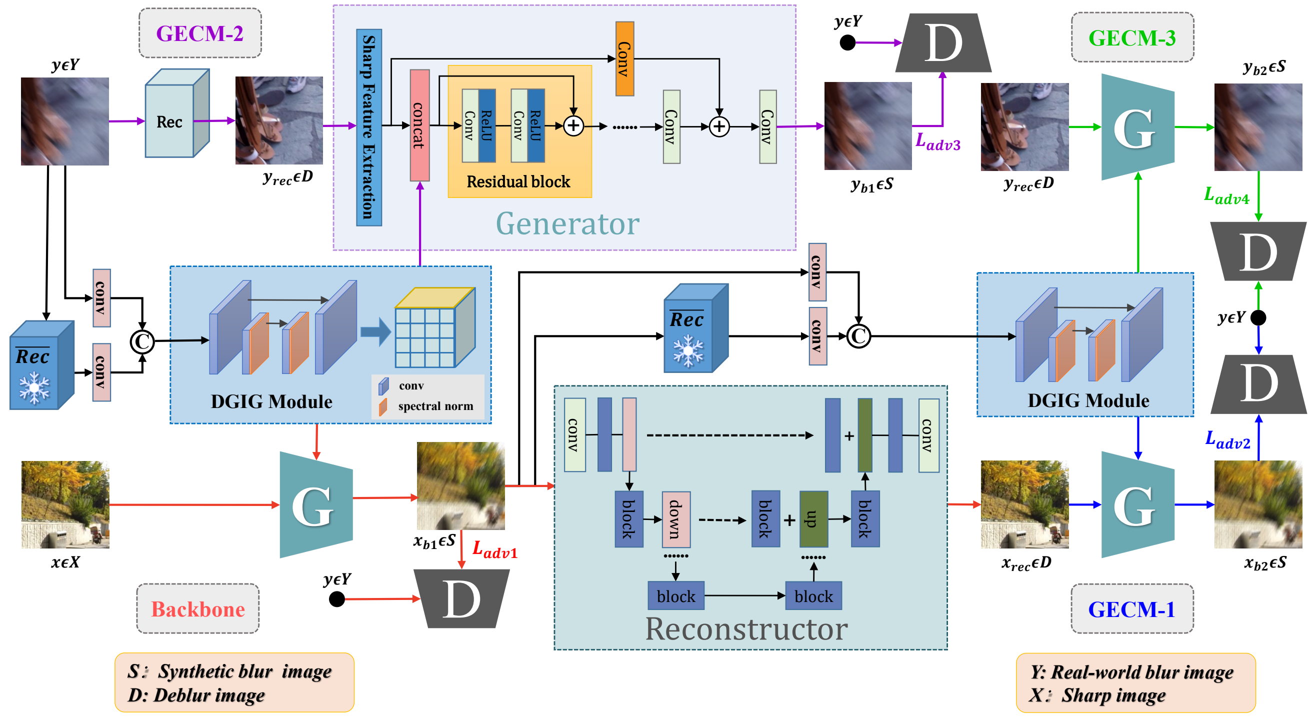

Figure 3. Self-Enhancement Multi-Generator Unsupervised Deblurring (SEMGUD) framework. The whole framework employs four generators and discriminators and uses NAFNet [1] as a reconstructor. The **red arrows** represent the **backbone** of SEMGUD, and the **blue arrows**, **purple arrows**, and **green arrows** respectively represent the different generator complementary constraint modules **GECM-1**, **GECM-2**, and **GECM-3**. DGIG Module denotes the degradation guidance information generation module. $\bar{Rec}$ denotes using the fixed reconstructor trained in the last round at the input of DGIG to synthesize better pseudo-paired data.

E. More Explanation of Ablation Study

Effectiveness of Re-Degradation Principal Component Consistency Loss. In the main document, we conduct a detailed ablation study of the re-degradation principal component consistency RPC^2 loss, and the results are shown in Table 1. The KE denotes the blur kernel estimation network, ϕ denotes the kernel size of the Gaussian filter operator, and ω_ϕ denotes the weights for different values of Gaussian kernels within each set of ϕ. To identify the optimal combination, we experiment with three sets of values for ϕ and two sets for ω_ϕ. We ultimately find that the reconstructor achieves the best performance when ϕ is set to 3, 5, and 7, and ω_ϕ to 1, 0.1, and 0.01. Additionally, the table reveals that using the blur kernel estimation network effectively enhances the reconstructor's performance.

KE	ϕ			ω_ϕ		GoPro		HIDE	
	3,5, and 7	3,7, and 9	3,9, and 15	1,0.2, and 0.04	1,0.1, and 0.01	PSNR↑	SSIM↑	PSNR↑	SSIM↑
✗	✔	✗	✗	✗	✔	28.54	0.919	27.25	0.881
✔	✗	✗	✔	✔	✗	28.72	0.921	27.39	0.885
✔	✗	✗	✔	✗	✔	28.83	0.923	27.48	0.888
✔	✗	✔	✗	✔	✗	28.89	0.922	27.54	0.890
✔	✗	✔	✗	✗	✔	28.97	0.924	27.63	0.892
✔	✔	✗	✗	✔	✗	29.00	0.925	27.64	0.892
✔	✔	✗	✗	✗	✔	**29.06**	**0.927**	**27.64**	**0.892**

Table 1. Ablation study on the hyperparameters ϕ and ω_ϕ of the RPC^2 loss. The first column notes whether the kernel estimation network is used.

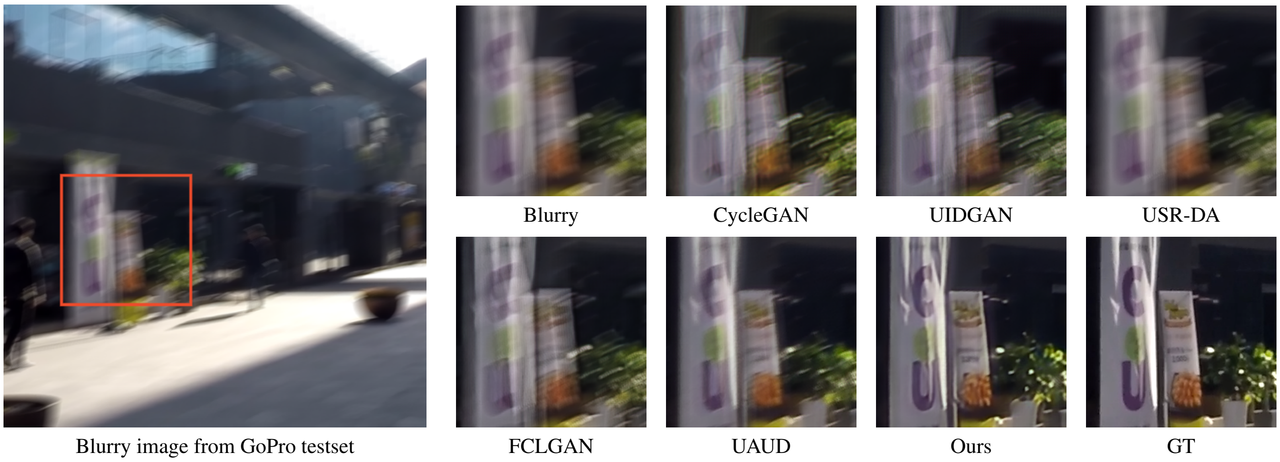

Figure 4. Visual comparisons on the GoPro dataset. From left to right: blurry image, results from CycleGAN [11], UIDGAN [3], USR-DA [8], FCLGAN [10], UAUD [7], SEMGUD (ours), and ground-truth.

F. More Visual Comparison Results

We provide additional visual comparisons on benchmark datasets in Figs. 4 - 12. We compare our SEMUGUD method with several recent state-of-the-art unsupervised image deblurring methods, including CycleGAN [11], UIDGAN [3], USR-DA [8], FCLGAN [10], and UAUD [7]. Since the USDF [2] code is not publicly available at this time, its visual results are not included.

From Figs. 4, 5, and 6, we can observe that the proposed method reconstructs more high-frequency textures on GoPro [4] dataset. In Figs. 7 and 8, it is evident that our method can restore more natural physical characteristics on HIDE [6] dataset. As shown in Figs. 9 and 10, our method achieves good results in removing motion blur on RealBlur [5] dataset. In addition, we also compare the visual results on the RWBI [9] dataset, which contains only real-world blurry images without ground truth. From Figs. 11 and 12, we can see that our method is effective in eliminating real-world blur. We also observe that existing unsupervised methods tend to be ineffective in deblurring more severe motion blur.

References

[1] Liangyu Chen, Xiaojie Chu, Xiangyu Zhang, and Jian Sun. Simple baselines for image restoration. In *European Conference on Computer Vision (ECCV)*, pages 17–33, 2022. 2

[2] Runhua Jiang and Yahong Han. Uncertainty-aware variate decomposition for self-supervised blind image deblurring. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, pages 252–260, 2023. 3

[3] Boyu Lu, Jun-Cheng Chen, and Rama Chellappa. Unsupervised domain-specific deblurring via disentangled representations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 10225–10234, 2019. 3, 4, 5, 6

[4] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3883–3891, 2017. 3

[5] Jaesung Rim, Haeyun Lee, Jucheol Won, and Sunghyun Cho. Real-world blur dataset for learning and benchmarking deblurring algorithms. In *European Conference on Computer Vision (ECCV)*, pages 184–201, 2020. 3

[6] Ziyi Shen, Wenguan Wang, Xiankai Lu, Jianbing Shen, Haibin Ling, Tingfa Xu, and Ling Shao. Human-aware motion deblurring. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 5572–5581, 2019. 3

[7] Xiaole Tang, Xile Zhao, Jun Liu, Jianli Wang, Yuchun Miao, and Tieyong Zeng. Uncertainty-aware unsupervised image deblurring with deep residual prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 9883–9892, 2023. 3, 4, 5, 6

[8] Wei Wang, Haochen Zhang, Zehuan Yuan, and Changhu Wang. Unsupervised real-world super-resolution: A domain adaptation perspective. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4318–4327, 2021. 3, 4, 5, 6

[9] Kaihao Zhang, Wenhan Luo, Yiran Zhong, Lin Ma, Bjorn Stenger, Wei Liu, and Hongdong Li. Deblurring by realistic blurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2737–2746, 2020. 3

[10] Suiyi Zhao, Zhao Zhang, Richang Hong, Mingliang Xu, Yi Yang, and Meng Wang. Fcl-gan: A lightweight and real-time baseline for unsupervised blind image deblurring. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, pages 6220–6229, 2022. 3, 4, 5, 6

[11] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE/CVF In-*

ternational Conference on Computer Vision (ICCV), pages 2223–2232, 2017. 1, 3, 4, 5, 6

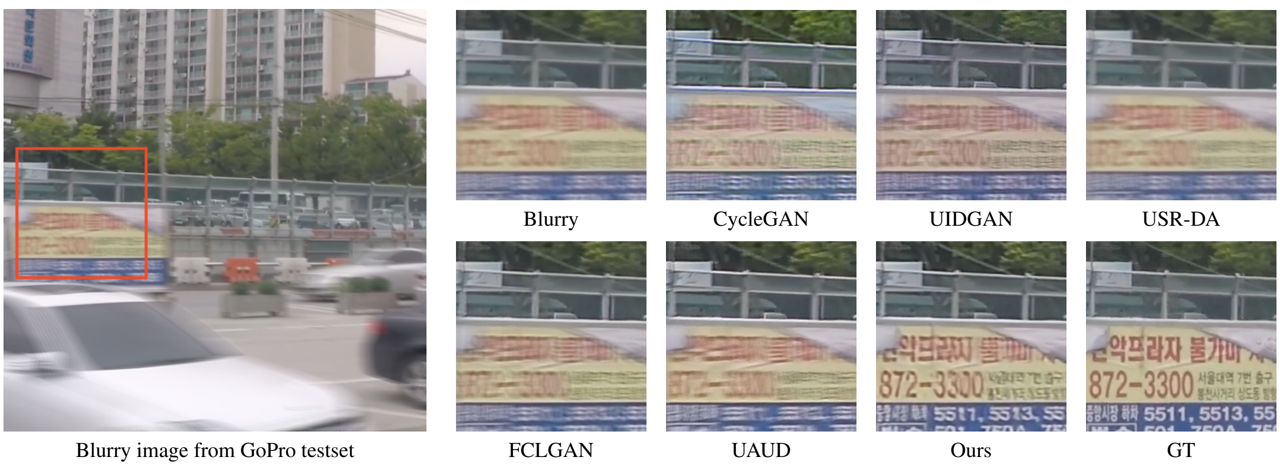

Figure 5. Visual comparisons on the GoPro dataset. From left to right: blurry image, results from CycleGAN [11], UIDGAN [3], USR-DA [8], FCLGAN [10], UAUD [7], SEMGUD (ours), and ground-truth.

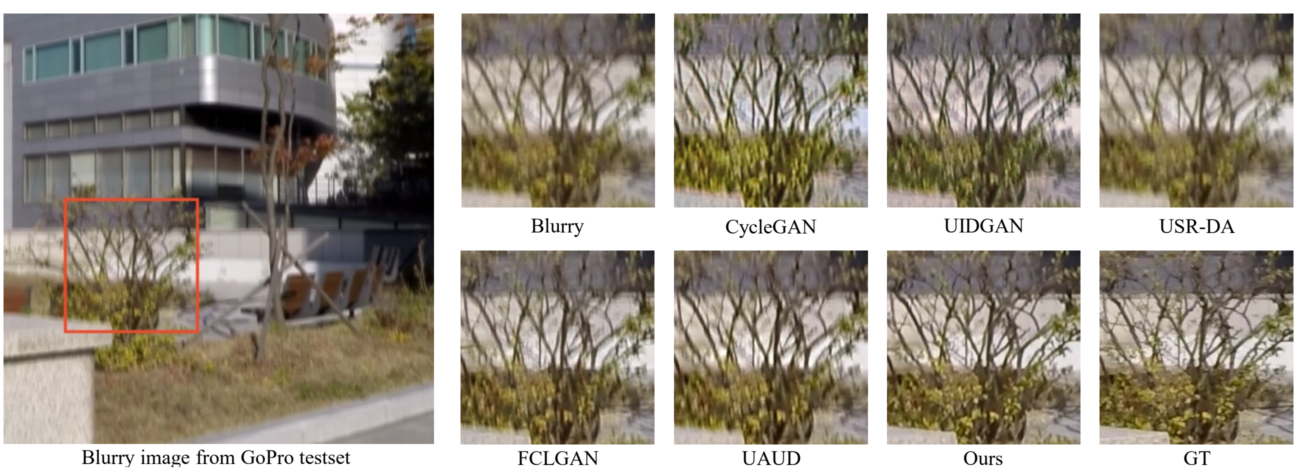

Figure 6. Visual comparisons on the GoPro dataset. From left to right: blurry image, results from CycleGAN [11], UIDGAN [3], USR-DA [8], FCLGAN [10], UAUD [7], SEMGUD (ours), and ground-truth.

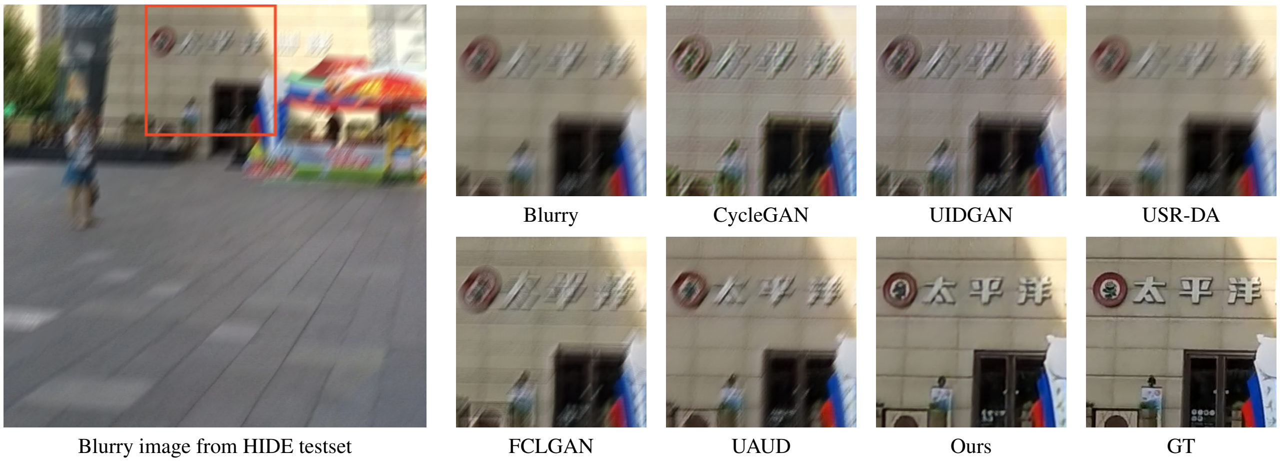

Figure 7. Visual comparisons on the HIDE dataset. From left to right: blurry image, results from CycleGAN [11], UIDGAN [3], USR-DA [8], FCLGAN [10], UAUD [7], SEMGUD (ours), and ground-truth.

Figure 8. Visual comparisons on the HIDE dataset. From left to right: blurry image, results from CycleGAN [11], UIDGAN [3], USR-DA [8], FCLGAN [10], UAUD [7], SEMGUD (ours), and ground-truth.

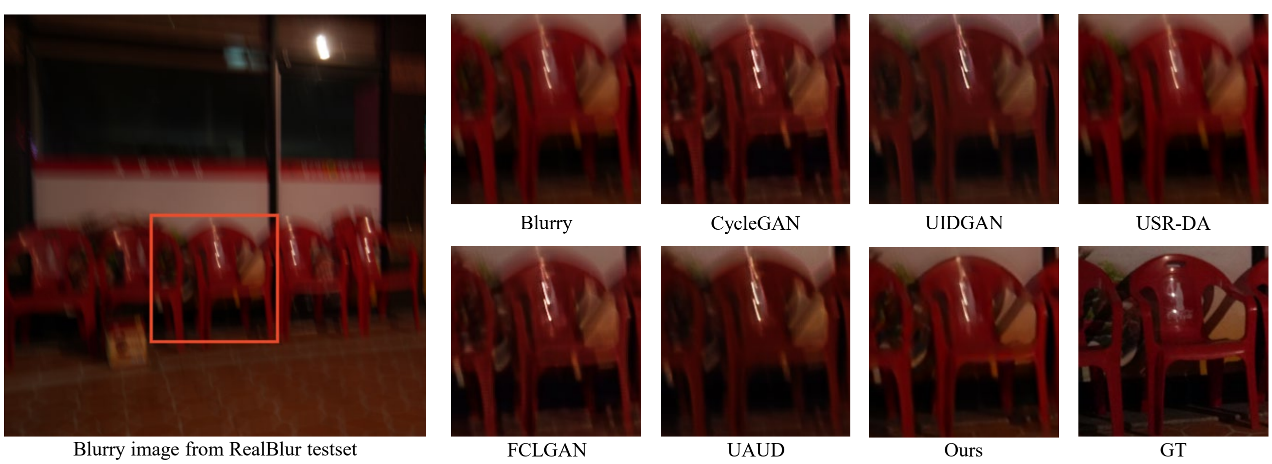

Figure 9. Visual comparisons on the RealBlur dataset. From left to right: blurry image, results from CycleGAN [11], UIDGAN [3], USR-DA [8], FCLGAN [10], UAUD [7], SEMGUD (ours), and ground-truth.

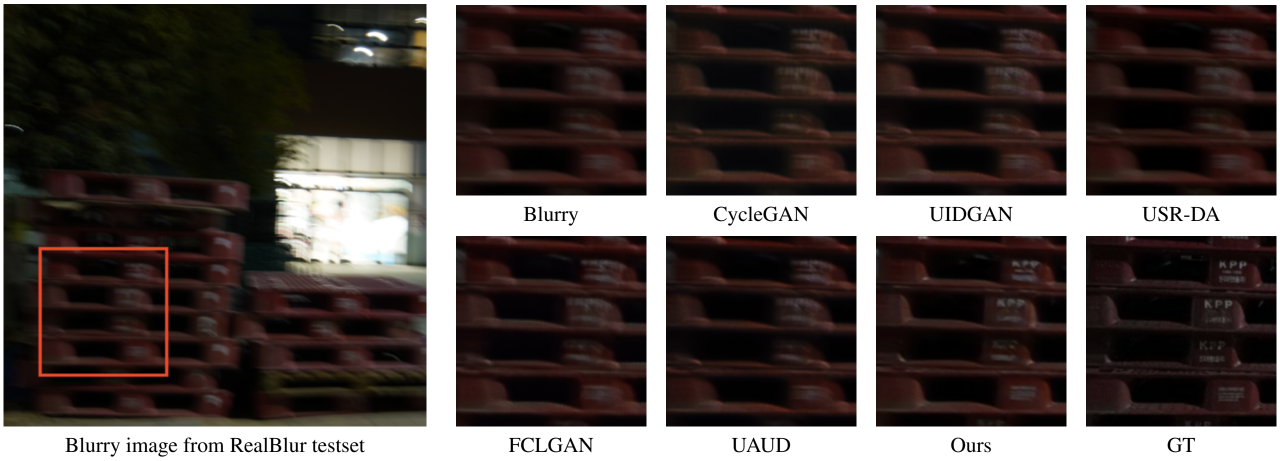

Figure 10. Visual comparisons on the RealBlur dataset. From left to right: blurry image, results from CycleGAN [11], UIDGAN [3], USR-DA [8], FCLGAN [10], UAUD [7], SEMGUD (ours), and ground-truth.

Figure 11. Visual comparisons on the RWBI dataset. From left to right: blurry image, results from CycleGAN [11], UIDGAN [3], USR-DA [8], FCLGAN [10], UAUD [7], and SEMGUD (ours).

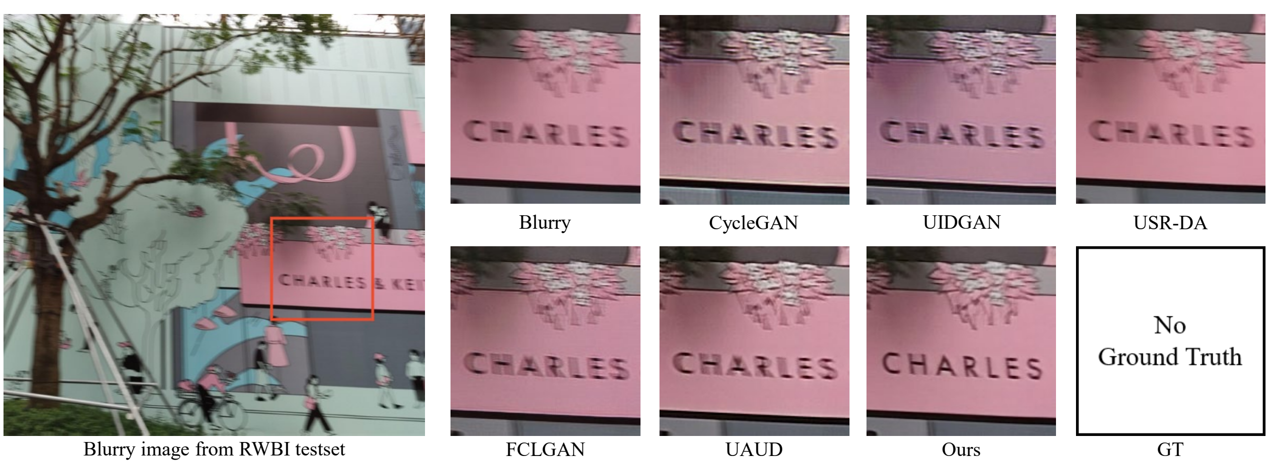

Figure 12. Visual comparisons on the RWBI dataset. From left to right: blurry image, results from CycleGAN [11], UIDGAN [3], USR-DA [8], FCLGAN [10], UAUD [7], and SEMGUD (ours).