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

Abstract

Significant progress in image deblurring has been achieved by deep learning methods, especially the remarkable performance of supervised models on paired synthetic data. However, real-world quality degradation is more complex than synthetic datasets, and acquiring paired data in real-world scenarios poses significant challenges. To address these challenges, we propose a novel unsupervised image deblurring framework based on self-enhancement. The framework progressively generates improved pseudo-sharp and blurry image pairs without the need for real paired datasets, and the generated image pairs with higher qualities can be used to enhance the performance of the reconstructor. To ensure the generated blurry images are closer to the real blurry images, we propose a novel re-degradation principal component consistency loss, which enforces the principal components of the generated low-quality images to be similar to those of re-degraded images from the original sharp ones. Furthermore, we introduce the self-enhancement strategy that significantly improves deblurring performance without increasing the computational complexity of network during inference. Through extensive experiments on multiple real-world blurry datasets, we demonstrate the superiority of our approach over other state-of-the-art unsupervised methods.

1. Introduction

Image deblurring is a classical problem in the field of computer vision, which aims to recover a clear image from its blurred version. The image deblurring tasks can be divided into blind and non-blind deblurring, where the blind image deblurring with unknown degradation is more challenging in general. Conventional model-based methods for blind image deblurring typically involve two main steps: estimating the blur kernel, and then reconstructing the sharp image from the blurred input [10, 26, 40, 42]. These methods' performance is largely constrained by the accuracy of blur kernal
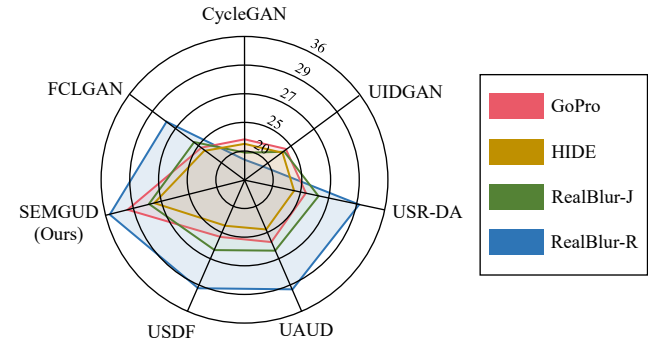
*Corresponding author

Figure 1. Performance comparison of our proposed SEMGUD with other unsupervised methods [12, 23, 36, 39, 48, 50] on different datasets.

estimation. For instance, in [27], the dark channel prior is used to estimate the blur kernel and reconstruct the sharp image. However, the blurry characteristics in real-world scenarios are quite complex, making it challenging to accurately estimate the optimal blur kernel. In addition, these methods often require complex iterative optimisation processes, which may lead to long inference time.

In recent years, with the rapid development of deep learning technology, convolutional neural networks (CNNs) have been widely used in deblurring tasks, achieving significant success. The supervised methods [4–6, 15, 16, 19, 29, 34, 44–46] focus on training deep neural network models using a large number of paired sharp and blurry images. This enables the network to learn the mapping from blurry images to sharp images without the need for blur kernel estimation, achieving end-to-end reconstruction of the blurry and sharp images. For example, DeepDeblur [24] proposes a multi-scale CNN to implement a coarse-to-fine processing pipeline and directly restores sharp images. However, in the real world cases, for the supervised learning methods, collecting paired datasets from the real world is challenging, and manually synthesized datasets are difficult to simulate the complex real image degradation processes.

Compared to supervised deep learning methods, unsupervised deep learning methods [3, 25, 49, 50] for real world

images typically achieve end-to-end image reconstruction without requiring real world paired sharp and blurry images during training. This allows unsupervised methods to more effectively handle complex real-world scenarios, especially when data collection cost is high, or the blur degradation is complex and challenging to model. Unfortunately, due to the high difficulty and challenges of unsupervised deblurring methods in training, there is still very little relevant research [12, 23, 36, 48] in this field. Specifically, because unsupervised methods lack the strong constraints provided by paired datasets, researchers tend to design more extensive deep models and complex inference processes to establish connections between inputs and outputs. This often leads to longer model inference times. Additionally, deploying large models in practical applications presents significant challenges. Therefore, how to synthesize high-quality pseudo-paired datasets and how to improve the performance of unsupervised methods without increasing the complexity of the model have become key research issues.

In this paper, we employ Generative Adversarial Network (GAN) [7] to learn the real-world image degradation distribution, addressing the issue of the lack of real-world paired data. We also introduce a novel re-degradation principal component consistency loss to more accurately synthesize blurred images. Considering that progressively update pseudo paired data can lead to higher performance [21], we propose a self-enhancement deblurring strategy within our unsupervised framework to further enhance the network's performance. This strategy significantly improves network deblurring performance without altering the existing network structure and without increasing inference computational complexity. The main contributions of this paper are as follows:

- We propose a novel self-enhancement based unsupervised deblurring framework. This framework can progressively improve the generated pseudo-paired data and reconstructor without the need for real paired datasets, addressing the issues of insufficient paired data in real deblurring tasks, as well as the substantial complexity increase in the conventional geometric augmentation inference.

- We design a novel loss function called re-degradation principal component consistency (RPC^2) loss. By introducing blur kernel prediction in principal component constraint, the RPC^2 loss enforces the principal components of the synthesized low-quality images to be similar to those of re-degraded images from the original sharp ones. It can decrease the effect of noise interfere in blurring images, and also make the synthetic image has more similar blurring degradation to the real data.

- Extensive experiments are conducted on typical real blurring datasets and the results verify the superior performance and generalization of our method over other existing unsupervised deblurring methods.

2. Related Work

2.1. Deep Supervised Image Deblurring

Recently, deep learning methods have achieved significant success in the field of image restoration [20–22, 24, 25, 28, 30, 34, 41]. For image deblurring, DL-MRF [35] proposes a CNN-based model to estimate the blur kernel and eliminate motion blur. In [2], a CNN is employed to compute an estimate of a clear image blurred by an unknown motion kernel. Thanks to the availability of some paired datasets, learning-based methods have gradually shifted their focus towards learning the mapping from given blurry images to the original sharp images, without explicitly estimating the blur kernel [11, 15, 16, 19, 29, 31, 44, 45]. MPRNet [45] introduces a novel multi-level progressive architecture to generate context-rich and spatially accurate output. DBGAN [47] designs an effective GAN-based model for simulating the synthesis of real-world blurry images. Stripformer [38] develops a transformer-based architecture by constructing horizontal and vertical labels to reweight image features. NAFNet [5] introduces a simple baseline network for image deblurring and denoising tasks. MRLPFNet [8] proposes a simple and effective multi-scale residual low-pass filter network that can better model both low and high frequency information. UFPNet [9] represents the motion blur kernel space in the latent space using normalized flows and designs a CNN to predict latent codes instead of motion blur kernels. Overall, supervised real image deblurring methods are dependent on real paired datasets, yet the scarcity of such data in real-world scenarios frequently limits their applicability.

2.2. Deep Unsupervised Image Deblurring

Compared to supervised methods, unsupervised real image deblurring methods [25, 39, 49, 50] do not involve real paired data during the training process. Consequently, unsupervised methods have weaker constraints between input and output, making it challenging to accurately learn the blur-to-clear mapping. Building on the basic GAN, UID-GAN [23] entangles the content and blur features of blurry images on a domain-specific dataset and then removes the blur from the blurry images. FCLGAN [48] proposes a lightweight and real-time unsupervised single-image blind deblurring baseline. USDF [12] introduces a multi-step self-supervised deblurring framework that iteratively decomposes and reassembles input images, exploiting the uncertainty of blur artifacts to generate a variety of pseudo-blurred and sharp image pairs. UADU [36] proposes an unsupervised semi-blind deblurring model that can effectively recover uniformly blurry image. However, unsupervised real deblurring methods have only received limited research attention in recent years. Existing methods overlook the potential in performance enhancement, and traditional model augmentation tends to increase inference computational complexity. Ad-
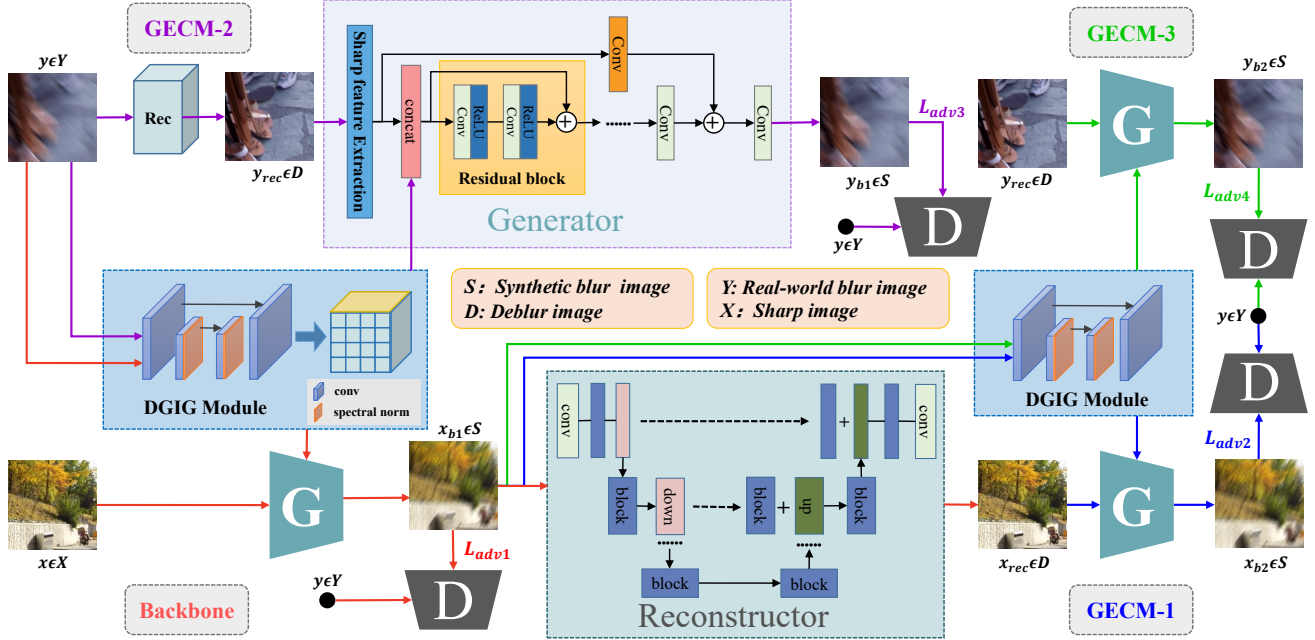
Figure 2. Multi-Generator Unsupervised Deblurring (MGUD) framework. The whole framework employs four generators and discriminators and uses NAFNet [5] as reconstructor. The **red arrows** represent the **backbone** of MGUD, and the **blue arrows**, **purple arrows**, and **green arrows** respectively represent the different generator complementary constraint modules **GECM-1**, **GECM-2**, and **GECM-3**. DGIG Module denotes the degradation guidance information generation module.

ditionally, methods based on synthetic pseudo-real paired images don't address how to continuously improve the quality of synthetic images.

3. Methodology

3.1. Unsupervised Deblurring Framework

We provide a detailed explanation of our proposed unsupervised framework in this subsection. Our approach aims to address the issue of insufficient paired data in real-world deblurring applications, and significantly improve network performance without altering the existing network structure and without increasing inference computational complexity.

3.1.1 Overall Framework of Proposed Method

In response to the lack of paired datasets in the real world, we propose the Multi-Generator Unsupervised Deblurring (MGUD) framework, as illustrated in Fig. 2. The framework employs ResNet-based generator with 6 residual blocks, PatchGAN [50] discriminator, and NAFNet [5] reconstructor. The detailed structures are presented in the supplementary materials.

In generator, we effectively utilize unpaired sharp and blurry images to synthesize pseudo-paired datasets. Inspired by CycleGAN [50], we design a GAN structure that transitions from "original sharp → synthetic blur → reconstructed

sharp → secondary synthetic blur", as shown by the red and blue arrows in Fig. 2. The process $(x, \mathcal{F}(y)) \rightarrow x_{b1}$ transforms original sharp image x to synthetic blurry image x_{b1} according to degradation of blurry image y, where "$\mathcal{F}(\cdot)$" is the degradation guide information generation (DGIG) module, providing ample pseudo-paired data for deblurring. The process $x_{b1} \rightarrow x_{rec}$ achieves deblurring by transforming x_{b1} to reconstructed sharp image x_{rec}. We incorporate a generator complementary constraint module (GECM) following the CycleGAN concept, $(x_{rec}, \mathcal{F}(x_{b1})) \rightarrow x_{b2}$, facilitating the transition from reconstructed sharp image x_{rec} to secondary synthetic blur image x_{b2}, adding additional constraints and enhancing training stability.

Considering that the unsupervised deblurring framework's performance mainly relies on the blur image generator's capability, and also considering the challenges in GAN training, we further introduce two GECMs: $(y_{rec}, \mathcal{F}(y)) \rightarrow y_{b1}$ and $(y_{rec}, \mathcal{F}(x_{b1})) \rightarrow y_{b2}$, which substantially enhance the qualities of the synthetic images.

3.1.2 Generation of Pseudo Paired Datasets and Reconstructor for Deblurring

Compared to supervised methods, the training process of unsupervised methods struggle to accurately learn the mapping from blurry to sharp due to the lack of strong constraints from paired data. GAN is often preferred for unsupervised
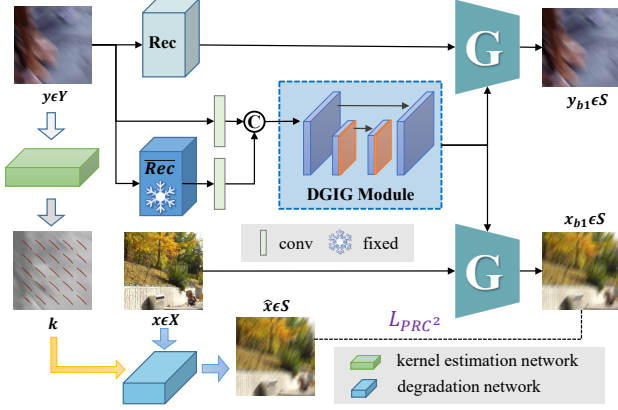
Figure 3. The implementation principles of the proposed self-enhancement strategy and re-degradation principal component consistency loss function. k denotes the blur kernel estimated from the real blurry images. $\bar{Rec}$ denotes using the fixed reconstructor trained in the last round at the input of DGIG to synthesize better pseudo-paired data.

methods due to its superior data synthesis capability. Current mainstream unsupervised methods focus on developing more capable generators to simulate the image blurring process and generate synthetic images that closely resemble real blur. Theoretically, the closer the synthetic blur images are to actual blur, the better the reconstructor is expected to perform. The structure of our generator is depicted in Fig. 2, incorporating a sharp feature extraction module and 6 residual blocks.

The synthesis of high-quality realistic blurred images is challenging due to varying image content influences. Considering many potential factors that could lead to undesired blurring artifacts, we initially process the blurred images y using a DGIG module to obtain blurred features. The DGIG module employs a U-Net architecture, comprising a downsampling layer followed by an upsampling layer. The G learns the blurring characteristics of blurry images to guide the synthesis of sharp images towards realistic blurring. To ensure the generation of more realistic blurred images, we also train a discriminator D to distinguish between synthesized and real-world blurred images, where the generator and discriminator learn collaboratively in an adversarial manner. In the MGUD framework, the adversarial loss $\mathcal{L}_{adv1}$ of the backbone is constrained between the $y\epsilon Y$ and $x_{b1}\epsilon S$, where Y and S denote the real blurred image and synthetic blurred image respectively. Considering that the pseudo-paired images generated by the generator are crucial to the reconstructor's performance in this unsupervised framework, and accurate degradation is difficult to obtain, additional three adversarial losses are further introduced to enhance constraints as shown in Fig. 2:

$$\mathcal{L}_{GAN} = \mathcal{L}_{adv1} + \mathcal{L}_{adv2} + \mathcal{L}_{adv3} + \mathcal{L}_{adv4} \quad (1)$$

For the synthesized blurry image x_{b1}, the deblurred image x_{rec} is obtained through the reconstructor Rec. In our entire framework, we optimize the following loss function to train the reconstructor:

$$\mathcal{L}_{Rec} = \mathcal{L}_{PSNR}(x, x_{rec}) + \mathcal{L}_{SSIM}(x, x_{rec}) \quad (2)$$

where $\mathcal{L}_{PSNR}(\cdot)$ represents the PSNR loss used to constrain the peak signal-to-noise ratio of the image, and $\mathcal{L}_{SSIM}(\cdot)$ represents the SSIM loss used to constrain the structural information of the image.

3.1.3 Re-Degradation Principal Component Consistency Loss

Although adversarial losses can effectively improve the performance of the generator and discriminator, leading to more accurately synthesized realistic blurred images, the data synthesis process is highly susceptible to interference from the inherent information in blurred images, such as content and color. To mitigate this, we draw inspiration from [13] and design a novel loss function called re-degradation principal component consistency (RPC^2) loss. By introducing blur kernel prediction in principal component constraint, this RPC^2 loss ensures the principal components of the synthesized low-quality images align with those of re-degraded images from the original sharp ones, which can decrease the impact of noise interfere in the blurring process. As shown in Fig. 3, we use the kernel estimation network to estimate the blur kernel of the blurry image y, and then utilize this kernel to further re-blur the sharp image x. Subsequently, we maintain the principal component consistency between x and the synthesized blurred image x_{b1} by introducing the L1 norm loss. The process is defined as follows:

$$\mathcal{L}_{RPC^2} = \sum_{\phi=3,5,7} \omega_\phi ||\mathcal{G}_\phi(x \otimes K(y)) - \mathcal{G}_\phi(x_{b1})||_1 \quad (3)$$

where $K(\cdot)$ denotes the blur kernel estimation process, $\otimes$ denotes the blur operation, $\mathcal{G}_\phi(\cdot)$ denotes a Gaussian filter operator with a kernel size of ϕ, and ω_ϕ represents the weight for level ϕ.

Specifically, the RPC^2 loss's role is to guide G in isolating the main degradation component (e.g., motion blur) from y, derived from K, with Gaussian filter acting to prevent G from incorporating residual texture interference (like content or color from y) onto x. Our test results in Section 4.3 show that without using the kernel estimation, the performance will significant decrease, verifying the necessity of the kernel estimation. In addition, we test three well-performed kernel estimation methods [1, 9, 20] and find that they achieve similar final performance. This is due to the introduction of multi-scale Gaussian low-pass filtering, the impact of noise and kernel estimation biases on the image is mitigated to
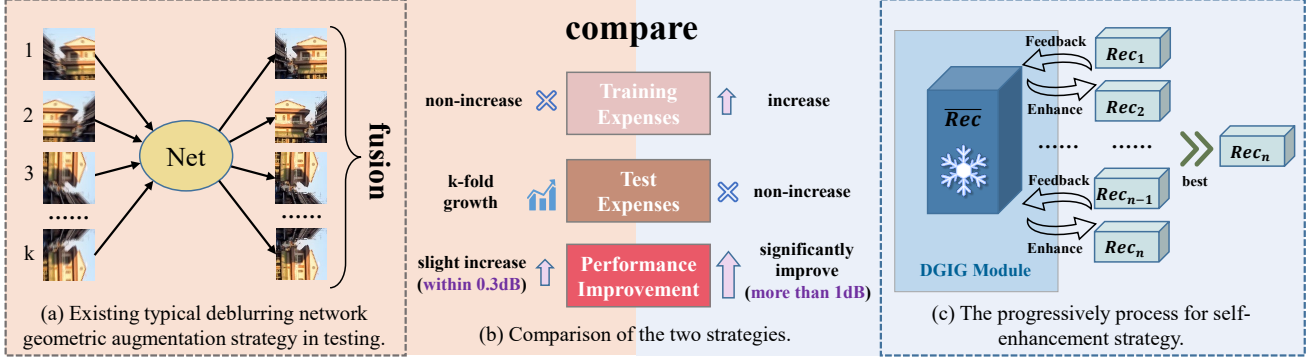
compare

	non-increase ✗	Training Expenses ⇑ increase
	k-fold growth 📈	Test Expenses ✗ non-increase
	slight increase (within 0.3dB) ⇑	Performance Improvement ⇑ significantly improve (more than 1dB)

(a) Existing typical deblurring network geometric augmentation strategy in testing.

(b) Comparison of the two strategies.

(c) The progressively process for self-enhancement strategy.

Figure 4. Comparison of our proposed self-enhancement strategy with the conventional geometric augmentation strategy in terms of training expenses, testing expenses, and performance improvement respectively.

some extent, ensuring good consistency of the main content in the image.

3.2. Inference Complexity Invariant Self-Enhancement Strategy

In order to enable the reconstructor to self-correct and enhance its performance without altering its structure or increasing the complexity of the original network , we propose the Self-Enhancement (SE) strategy. Fig. 3 illustrates the implementation scheme of one of the DGIGs integrating the SE strategy, and the other DGIG with the SE strategy is similar. By employing the SE strategy, a better reconstructor (Rec) can be obtained, thereby learning more accurate degradation. This results in the generation of more realistic synthetic blurred-sharp image pairs and progressively improves the performance of the updated Rec with higher quality synthetic samples. It is exciting to note that the performance of the Rec using the SE strategy shows significant improvement over the one without the SE strategy. The implementation of the SE strategy involves several steps. First, we need to train an initial reconstructor Rec_1. Then, as shown in Fig. 3, keep the Rec_1 parameters fixed to guide the generator to learn degradation information more accurately, generating more realistic pseudo-blurry images. Subsequently, retrain G, D, and Rec until Rec reaches convergence. Finally, repeat the above process, progressively replacing and enhancing the reconstructor until obtain the best-performing Rec_n. Therefore, the complete Rec's loss function is as follows:

$$
\begin{aligned}
\mathcal{L}_{SE-Rec} = &\mathcal{L}_{PSNR}(\bar{Rec}(x_{b1}), x_{rec}) \\
&+ \mathcal{L}_{SSIM}(\bar{Rec}(x_{b1}), x_{rec}) \\
&+ \mathcal{L}_{PSNR}(\bar{Rec}(y), y_{rec}) \\
&+ \mathcal{L}_{SSIM}(\bar{Rec}(y), y_{rec}) + \mathcal{L}_{Rec}
\end{aligned}
\tag{4}
$$

Finally, we obtain the total loss function:

$$
\mathcal{L} = \underset{G}{min}\,\underset{D}{max}\,\mathcal{L}_{GAN} + \Phi_{RPC^2}\mathcal{L}_{RPC^2} + \mathcal{L}_{SE-Rec} \tag{5}
$$

Strategies	Original	Geometric Augmentation	Self-Enhancement
PSNR	27.68	27.79	29.06
PSNR Gains	—	0.11	1.38

Table 1. Comparing the performance enhancement of the self-enhancement strategy and the geometric augmentation strategy on the GoPro dataset.

where Φ_{RPC^2} represents the hyperparameter for the re-degradation principal component consistency loss.

The basic idea of the SE strategy is to use the results of the previous phase as feedback information to guide and improve the training of subsequent stages. A comparative analysis with existing typical geometric augmentation methods, as presented in Fig. 4, reveals that the SE strategy markedly boosts performance with slightly additional training cost, without imposing extra complexity or testing overhead. As Table 1 indicates, our method improves the PSNR by more than 1dB after incorporating the SE strategy, which demonstrates the effectiveness of SE strategy in image deblurring.

4. Experiments

4.1. Datasets and Implementation Details

Datasets. Following the state-of-the-art image deblurring methods [5, 9], we first divide the GoPro dataset [24] (consists of 2103 pairs of blurred and sharp images) into separate sharp and blurred image parts and further constitute unpaired blurred datasets to train the algorithm proposed in this paper. We evaluate our method on the GoPro dataset [24], RealBlur dataset [32], and HIDE dataset [33]. The GoPro dataset consists of 1111 test images, the RealBlur dataset includes 980 images for testing, and the HIDE dataset provides 2025 test images.

Implementation Details. We follow the experimental settings described in [5]. We adopt the Adam optimizer [14] ($\beta_1 = 0.9$, $\beta_2 = 0.999$), the initial learning rate is set to 10^{-4},

	Conference/Journal	Methods	GoPro		HIDE		RealBlur-R		RealBlur-J	
			PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑
Deep supervised	CVPR 2018	DeepDeblur [24]	29.23	0.916	25.73	0.874	32.51	0.841	27.87	0.827
	CVPR 2018	SRN [37]	30.26	0.934	28.36	0.915	35.66	0.947	28.56	0.867
	CVPR 2021	HINet [4]	32.71	0.959	30.32	0.932	35.75	0.949	28.17	0.849
	ECCV 2022	Stripfromer [38]	33.08	0.962	31.03	0.940	36.07	0.952	28.82	0.876
	ECCV 2022	MSDI-Net [17]	33.28	0.964	31.02	0.940	35.88	0.952	28.59	0.869
	ECCV 2022	NAFNet [5]	33.69	0.967	31.32	0.943	35.50	0.953	28.32	0.857
	ICCV 2023	icDPM [31]	33.20	0.963	30.96	0.938	N/A	N/A	28.81	0.872
	CVPR 2023	UFPNet [9]	34.06	0.968	31.74	0.947	36.25	0.953	29.87	0.884
Deep unsupervised	ICCV 2017	CycleGAN [50]	22.54	0.720	21.81	0.690	12.38	0.242	19.79	0.633
	ICCV 2017	DualGAN [43]	22.86	0.722	N/A	N/A	N/A	N/A	N/A	N/A
	CVPR 2019	UIDGAN [23]	23.56	0.738	22.70	0.715	16.64	0.323	22.87	0.671
	ACM MM 2022	FCLGAN [48]	24.84	0.771	23.43	0.732	28.37	0.663	25.35	0.736
	ICCV 2021	USR-DA [39]	25.49	0.787	23.91	0.756	32.32	0.821	26.39	0.784
	ACM MM 2023	USDF [12]	25.58	0.857	23.93	0.829	32.57	0.923	26.59	0.881
	CVPR 2023	UAUD [36]	26.12	0.869	24.37	0.837	32.91	0.885	26.84	0.792
	—	**SEMGUD (Ours)**	**29.06**	**0.927**	**27.64**	**0.892**	**35.51**	**0.946**	**28.01**	**0.844**

Table 2. The comparison results on the benchmark datasets. All the models are trained on the GoPro dataset.
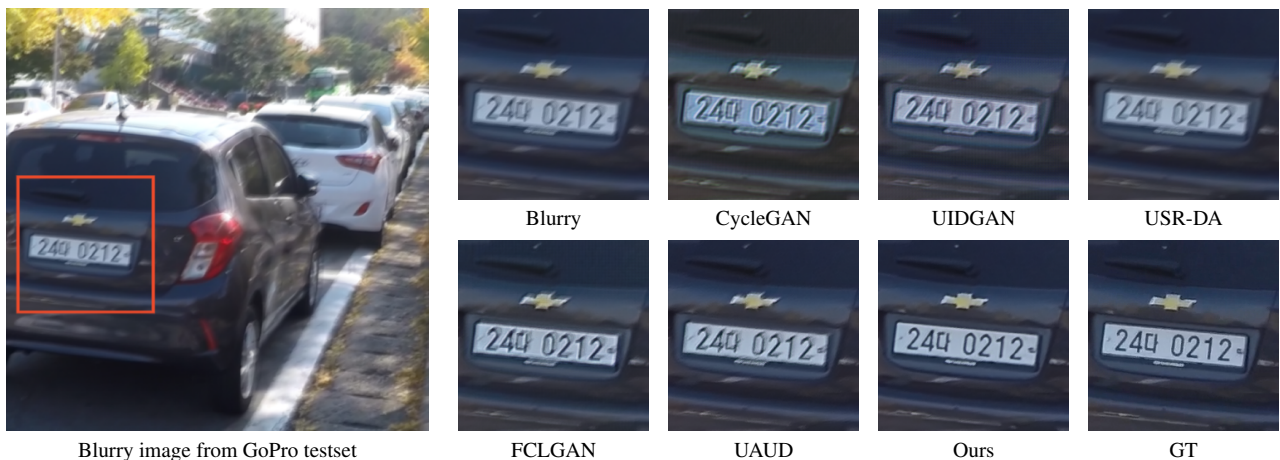

Figure 5. Visual comparisons on the GoPro dataset. From left to right: blurry image, results from CycleGAN [50], UIDGAN [23], USR-DA [39], FCLGAN [48], UAUD [36], SEMGUD (ours), and ground-truth.

and the training patch size is 128×128. For the RPC^2 loss hyperparameter (Φ_{RPC^2}) in the loss function of Eq. 5 is set to 3. More details are presented in the supplementary materials.

4.2. Experimental Results

We compare our method with the most recent unsupervised methods from recent years, including [12, 23, 36, 39, 43, 48, 50]. Additionally, we also list the state-of-the-art supervised methods based on paired images, including [4, 5, 9, 17, 24, 31, 37, 38]. Note that the related models for comparison are limited, since only a few unsupervised deblurring models have been proposed in the field. We evaluate the effectiveness of each method using performance metrics (PSNR and SSIM). The results are directly cited from the original papers or generated using the official models provided by the authors.

Quantitative comparison. Table 2 presents the PSNR and SSIM results for various single-image deblurring test methods on the GoPro, HIDE, and RealBlur test datasets. It is evident that our proposed method outperforms other unsupervised methods. On the GoPro dataset, our method achieves 2.94dB improvement in PSNR over the state-of-the-art unsupervised methods. To validate the effectiveness and generalization of our method, we compare the results on various real-world blurry datasets. Fig. 1 intuitively demonstrates the performance improvement of our method compared to other unsupervised methods. It is well known that supervised methods usually outperform unsupervised methods. But on the RealBlur-R dataset, the performance of
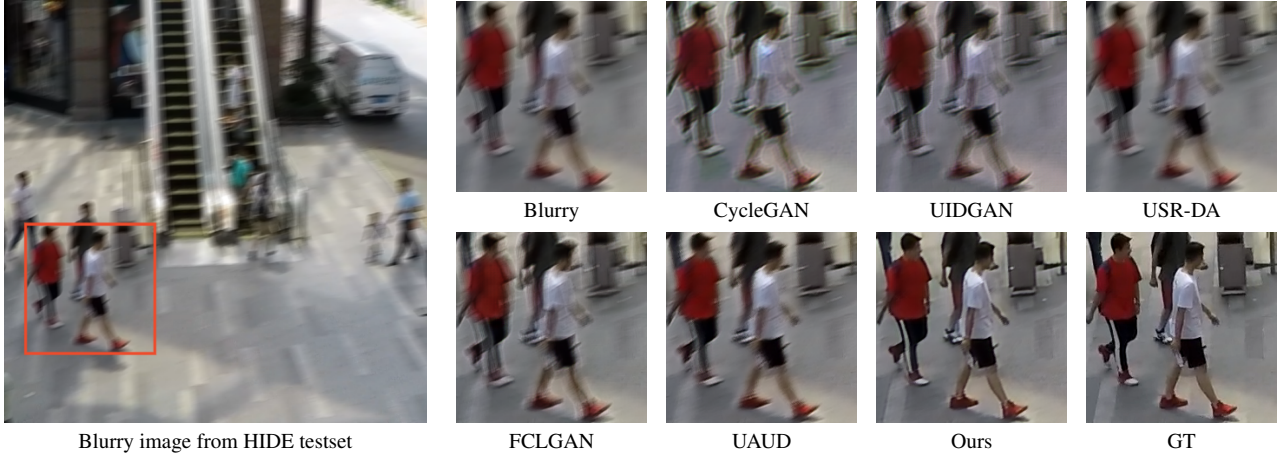
Figure 6. Visual comparisons on the HIDE dataset. From left to right: blurry image, results from CycleGAN [50], UIDGAN [23], USR-DA [39], FCLGAN [48], UAUD [36], SEMGUD (ours), and ground-truth.
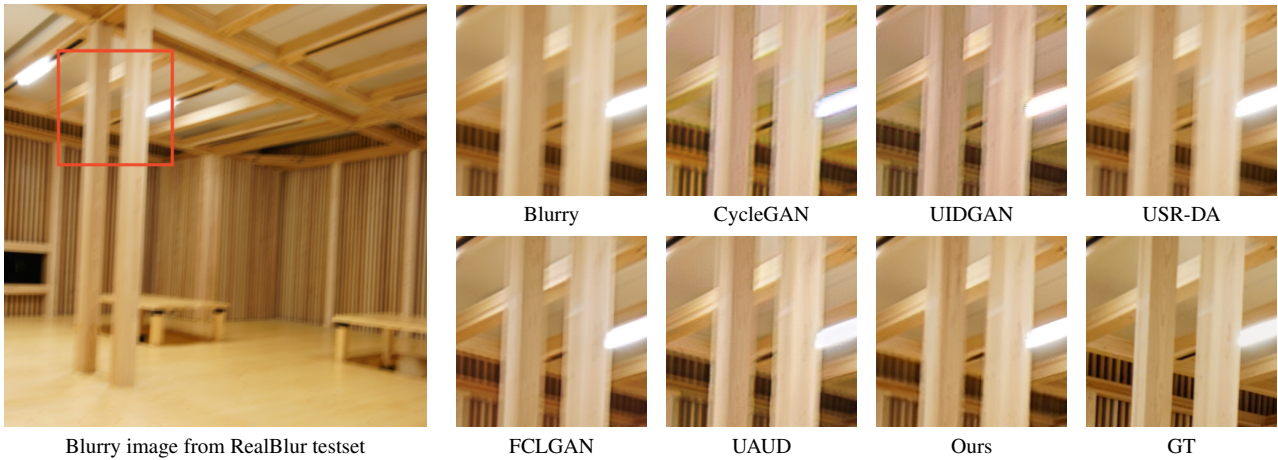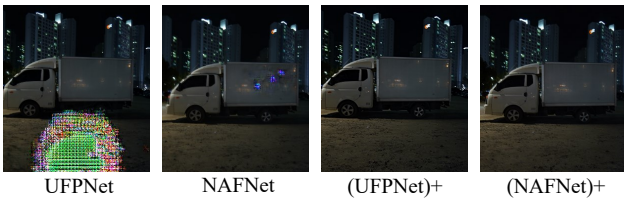

Figure 7. Visual comparisons on the RealBlur dataset. From left to right: blurry image, results from CycleGAN [50], UIDGAN [23], USR-DA [39], FCLGAN [48], UAUD [36], SEMGUD (ours), and ground-truth.

Figure 8. The "mode collapse" in NAFNet [5] and UFPNet [9]: trained on the GoPro dataset may output anomalous pixel regions during testing on the RealBlur-J dataset. "(UFPNet)+" and "(NAFNet)+" denote the results obtained through training with our SEMGUD framework.

KE	ϕ			ω_ϕ		GoPro	
	3,5,7	3,7,9	3,9,15	1,0.2,0.04	1,0.1,0.01	PSNR↑	SSIM↑
✗	✔	✗	✗	✗	✔	28.54	0.919
✔	✗	✗	✔	✔	✗	28.72	0.921
✔	✗	✗	✔	✗	✔	28.83	0.923
✔	✗	✔	✗	✔	✗	28.89	0.922
✔	✗	✔	✗	✗	✔	28.97	0.924
✔	✔	✗	✗	✔	✗	29.00	0.925
✔	✔	✗	✗	✗	✔	**29.06**	**0.927**

Table 3. Ablation study on the hyperparameters ϕ and ω_ϕ of the RPC^2 loss. The first column indicates whether the kernel estimation network is used.

our method is even higher than NAFNet [5]. This suggests that our method's effectiveness is improving and it's even becoming competitive with certain supervised methods. It's worth noting that all the mentioned models are trained on the GoPro dataset, which further demonstrates the excellent generalization performance of our method from GoPro to other real-world datasets.
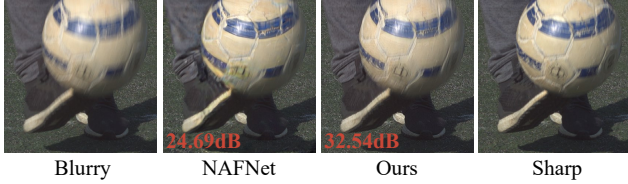
Figure 9. Crop the blurred region from the ReloBlur dataset for visual comparison.

Methods	SE	GoPro		HIDE	
		PSNR↑	SSIM↑	PSNR↑	SSIM↑
FCLGAN [48]	✗	24.70	0.768	23.49	0.733
	✔	25.11	0.772	23.79	0.741
UFPNet [9]	✗	27.37	0.903	25.91	0.862
	✔	28.41	0.916	26.67	0.875
NAFNet [5] (Ours)	✗	27.68	0.911	26.24	0.859
	✔	29.06	0.927	27.64	0.892

Table 4. Ablation study on the superiority and generalization of the self-enhancement strategy for different reconstructors.

Visual comparison. As shown in Figs. 5, 6, and 7, we compare the visual deblurring results of different unsupervised methods on the GoPro, HIDE, and RealBlur datasets, respectively. Note that although UAUD [36] effectively restores images with uniform blur, such performance may not extend to images with non-uniform blur. It can be seen that our proposed method achieves good results in removing motion blur from real blurry images. Additionally, we note that existing unsupervised methods often perform poorly in dealing with more severe blurring as shown in Fig. 6. Furthermore, we observe that while some existing advanced supervised methods like NAFNet and UFPNet achieve the best performance when trained on the GoPro paired datasets, they exhibit "mode collapse" when directly applied to other datasets, as illustrated in Fig. 8. In contrast, unsupervised models trained with our framework do not exhibit this phenomenon. Additionally, we directly use our GoPro-trained model to test the ReloBlur [18] dataset, which contains Real World Partly-blurred images, and find it to be effective for some of the images, as shown in Fig. 9.

kernel estimation networks, e.g., [1] can be used for testing. Therefore, our ablation study on the RPC^2 loss primarily focuses on the blur kernel estimation network and the settings of hyperparameters ϕ and ω in Eq. 3. As shown in Table 3, we observe that the blur kernel estimation network significantly improves the performance of our method.

The Self-Enhancement Strategy. As shown in Table 4, the effect is improved by 1.38dB on the GoPro dataset with the SE strategy. Furthermore, to demonstrate the superiority and versatility of our proposed self-enhancement strategy, we also conduct experiments incorporating SE strategy on several deblurring methods, including NAFNet [5], FCL-GAN [48], UFPNet [9]. Fig. 10 shows in detail the different network performance improvements before and after using the self-enhancement strategy. Note that due to FCLGAN do not provide the complete network weights, the experimental results in the table are obtained from our retraining. This strongly validates the superiority and significant potential of the self-enhancement strategy.

5. Conclusion

In this paper, we introduce a novel unsupervised framework for image deblurring. Our approach incorporates a re-degradation principal component consistency loss, ensuring that the principal components of the synthetically blurred images align closely with those from the re-degraded versions of the original sharp images. We also put forward a self-enhancement strategy that substantially improves performance without modifying the model's architecture or adding to the computational cost in inference. Comprehensive tests on standard datasets reveal that our method surpasses the existing state-of-the-art unsupervised methods with strong generalization capabilities across various real-world blurry image datasets.
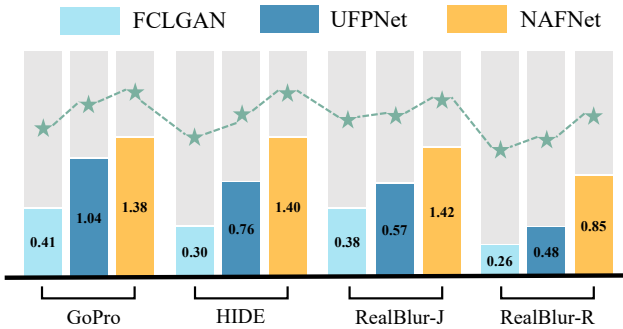
Figure 10. Performance improvement (PSNR gain) of different methods on various datasets driven by the self-enhancement strategy.

4.3. Ablation Study

The Re-Degradation Principal Component Consistency Loss. In fact, we find that if we remove the RPC^2 loss, the generator always struggles to generate satisfactory blurry images. Due to the minimal impact of different kernel estimation networks on the final deblurring performance as illustrated in Section 3.1.3, we can use any well-performed

References

[1] Guillermo Carbajal, Patricia Vitoria, Mauricio Delbracio, Pablo Musé, and José Lezama. Non-uniform motion blur kernel estimation via adaptive decomposition. *arXiv e-prints*, pages arXiv–2102, 2021. 4, 8

[2] Ayan Chakrabarti. A neural approach to blind motion deblurring. In *European Conference on Computer Vision (ECCV)*, pages 221–235, 2016. 2

[3] Huaijin Chen, Jinwei Gu, Orazio Gallo, Ming-Yu Liu, Ashok Veeraraghavan, and Jan Kautz. Reblur2deblur: Deblurring videos via self-supervised learning. In *IEEE International Conference on Computational Photography (ICCP)*, pages 1–9, 2018. 1

[4] Liangyu Chen, Xin Lu, Jie Zhang, Xiaojie Chu, and Chengpeng Chen. Hinet: Half instance normalization network for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pages 182–192, 2021. 1, 6

[5] 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, 3, 5, 6, 7, 8

[6] Sung-Jin Cho, Seo-Won Ji, Jun-Pyo Hong, Seung-Won Jung, and Sung-Jea Ko. Rethinking coarse-to-fine approach in single image deblurring. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4641–4650, 2021. 1

[7] Antonia Creswell, Tom White, Vincent Dumoulin, Kai Arulkumaran, Biswa Sengupta, and Anil A Bharath. Generative adversarial networks: An overview. *IEEE signal processing magazine*, pages 53–65, 2018. 2

[8] Jiangxin Dong, Jinshan Pan, Zhongbao Yang, and Jinhui Tang. Multi-scale residual low-pass filter network for image deblurring. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 12345–12354, 2023. 2

[9] Zhenxuan Fang, Fangfang Wu, Weisheng Dong, Xin Li, Jinjian Wu, and Guangming Shi. Self-supervised non-uniform kernel estimation with flow-based motion prior for blind image deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 18105–18114, 2023. 2, 4, 5, 6, 7, 8

[10] Amit Goldstein and Raanan Fattal. Blur-kernel estimation from spectral irregularities. In *European Conference on Computer Vision (ECCV)*, pages 622–635, 2012. 1

[11] Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, and Aaron C Courville. Improved training of wasserstein gans. *Advances in neural information processing systems*, 30, 2017. 2

[12] 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. 1, 2, 6

[13] Xin Jin, Zhibo Chen, Jianxin Lin, Zhikai Chen, and Wei Zhou. Unsupervised single image deraining with self-supervised constraints. In *IEEE International Conference on Image Processing (ICIP)*, pages 2761–2765, 2019. 4

[14] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 5

[15] Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, and Jiří Matas. Deblurgan: Blind motion deblurring using conditional adversarial networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8183–8192, 2018. 1, 2

[16] Orest Kupyn, Tetiana Martyniuk, Junru Wu, and Zhangyang Wang. Deblurgan-v2: Deblurring (orders-of-magnitude) faster and better. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 8878–8887, 2019. 1, 2

[17] Dasong Li, Yi Zhang, Ka Chun Cheung, Xiaogang Wang, Hongwei Qin, and Hongsheng Li. Learning degradation representations for image deblurring. In *European Conference on Computer Vision (ECCV)*, pages 736–753, 2022. 6

[18] 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*, pages 1314–1322, 2023. 8

[19] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, pages 1833–1844, 2021. 1, 2

[20] Jingyun Liang, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Mutual affine network for spatially variant kernel estimation in blind image super-resolution. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 4096–4105, 2021. 2, 4

[21] Xin Lin, Chao Ren, Xiao Liu, Jie Huang, and Yinjie Lei. Unsupervised image denoising in real-world scenarios via self-collaboration parallel generative adversarial branches. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 12642–12652, 2023. 2

[22] Xiao Liu, Xiangyu Liao, Xiuya Shi, Linbo Qing, and Chao Ren. *Efficient Information Modulation Network for Image Super-Resolution*. 2023. 2

[23] 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. 1, 2, 6, 7

[24] 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. 1, 2, 5, 6

[25] Thekke Madam Nimisha, Kumar Sunil, and AN Rajagopalan. Unsupervised class-specific deblurring. In *European Conference on Computer Vision (ECCV)*, pages 353–369, 2018. 1, 2

[26] Jinshan Pan, Zhe Hu, Zhixun Su, and Ming-Hsuan Yang. Deblurring text images via l0-regularized intensity and gradient prior. In *Proceedings of the IEEE/CVF Conference*

on Computer Vision and Pattern Recognition (CVPR), pages 2901–2908, 2014. 1

[27] Jinshan Pan, Deqing Sun, Hanspeter Pfister, and Ming-Hsuan Yang. Blind image deblurring using dark channel prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1628–1636, 2016. 1

[28] Yizhong Pan, Xiao Liu, Xiangyu Liao, Yuanzhouhan Cao, and Chao Ren. Random sub-samples generation for self-supervised real image denoising. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 12150–12159, 2023. 2

[29] Dongwon Park, Dong Un Kang, Jisoo Kim, and Se Young Chun. Multi-temporal recurrent neural networks for progressive non-uniform single image deblurring with incremental temporal training. In *European Conference on Computer Vision (ECCV)*, pages 327–343, 2020. 1, 2

[30] Chao Ren, Xiaohai He, and Truong Q Nguyen. Adjusted non-local regression and directional smoothness for image restoration. *IEEE Transactions on Multimedia*, pages 731–745, 2018. 2

[31] Mengwei Ren, Mauricio Delbracio, Hossein Talebi, Guido Gerig, and Peyman Milanfar. Multiscale structure guided diffusion for image deblurring. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 10721–10733, 2023. 2, 6

[32] 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. 5

[33] 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. 5

[34] Maitreya Suin, Kuldeep Purohit, and AN Rajagopalan. Spatially-attentive patch-hierarchical network for adaptive motion deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3606–3615, 2020. 1, 2

[35] Jian Sun, Wenfei Cao, Zongben Xu, and Jean Ponce. Learning a convolutional neural network for non-uniform motion blur removal. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 769–777, 2015. 2

[36] 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. 1, 2, 6, 7, 8

[37] Xin Tao, Hongyun Gao, Xiaoyong Shen, Jue Wang, and Jiaya Jia. Scale-recurrent network for deep image deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 8174–8182, 2018. 6

[38] Fu-Jen Tsai, Yan-Tsung Peng, Yen-Yu Lin, Chung-Chi Tsai, and Chia-Wen Lin. Stripformer: Strip transformer for fast image deblurring. In *European Conference on Computer Vision (ECCV)*, pages 146–162, 2022. 2, 6

[39] 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. 1, 2, 6, 7

[40] Oliver Whyte, Josef Sivic, Andrew Zisserman, and Jean Ponce. Non-uniform deblurring for shaken images. *International journal of computer vision*, pages 168–186, 2012. 1

[41] Lin X, Yue J, and Ren C. Unlocking low-light-rainy image restoration by pairwise degradation feature vector guidance. In *arXiv*, 2023. 2

[42] Li Xu, Shicheng Zheng, and Jiaya Jia. Unnatural l0 sparse representation for natural image deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1107–1114, 2013. 1

[43] Zili Yi, Hao Zhang, Ping Tan, and Minglun Gong. Dualgan: Unsupervised dual learning for image-to-image translation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 2849–2857, 2017. 6

[44] Yuan Yuan, Wei Su, and Dandan Ma. Efficient dynamic scene deblurring using spatially variant deconvolution network with optical flow guided training. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3555–3564, 2020. 1, 2

[45] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 14821–14831, 2021. 2

[46] Hongguang Zhang, Yuchao Dai, Hongdong Li, and Piotr Koniusz. Deep stacked hierarchical multi-patch network for image deblurring. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5978–5986, 2019. 1

[47] 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. 2

[48] 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. 1, 2, 6, 7, 8

[49] Suiyi Zhao, Zhao Zhang, Richang Hong, Mingliang Xu, Haijun Zhang, Meng Wang, and Shuicheng Yan. Crnet: Unsupervised color retention network for blind motion deblurring. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, pages 6193–6201, 2022. 1, 2

[50] 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 International Conference on Computer Vision (ICCV)*, pages 2223–2232, 2017. 1, 2, 3, 6, 7