

# Supplementary: Blind image deblurring with noise-robust kernel estimation

Chanseok Lee<sup>1</sup>, Jeongsol Kim<sup>1</sup>, Seungmin Lee<sup>1</sup>, Jaehwang Jung<sup>2</sup>, Yunje Cho<sup>2</sup>, Taejoong Kim<sup>2</sup>, Taeyong Jo<sup>2</sup>, Myungjun Lee<sup>2</sup>, and Mooseok Jang<sup>1,3</sup>

<sup>1</sup> Department of Bio and Brain Engineering, KAIST, Daejeon, South Korea

<sup>2</sup> Advanced Process Development Team 4, Semiconductor R&D Center, Samsung Electronics, South Korea

<sup>3</sup> KAIST Institute for Health Science and Technology, KAIST, Daejeon, South Korea  
 {cslee, mooseok}@kaist.ac.kr

## S1. Analysis of noise-robust kernel estimation function

As we have discussed in the main text, the proposed noise-robust kernel estimation function is described as follows:

$$\hat{h} = C \left( \mathcal{F}^{-1} \left\{ \frac{\overline{\mathcal{F}\{x\}}}{|\mathcal{F}\{x\}|^2 + k} \mathcal{F}\{y_m\} \right\} \right) = C(\mathcal{W}(y_m, x, k)) = F(y_m, x, k) \quad (1)$$

Here, we further analyze the working behavior of the kernel estimation function  $F$  by investigating the property of the inverse filter  $\hat{w}$ .

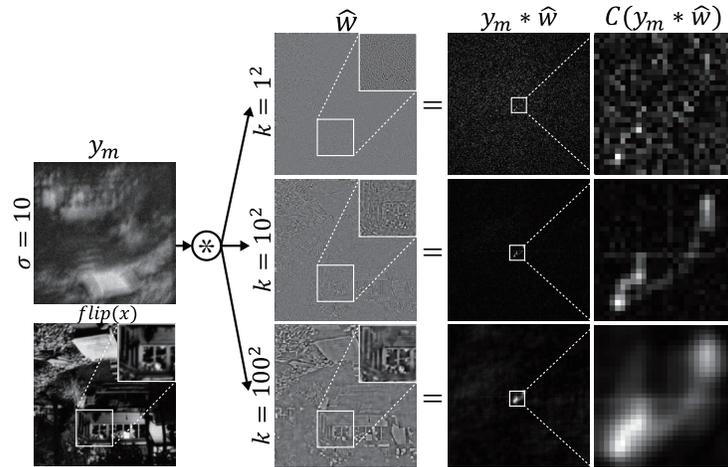
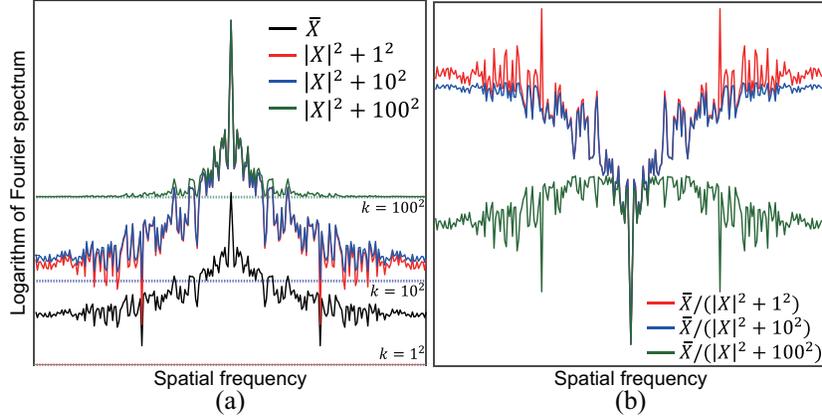


Fig. S1: Visualization of the working behavior of the kernel estimation function.

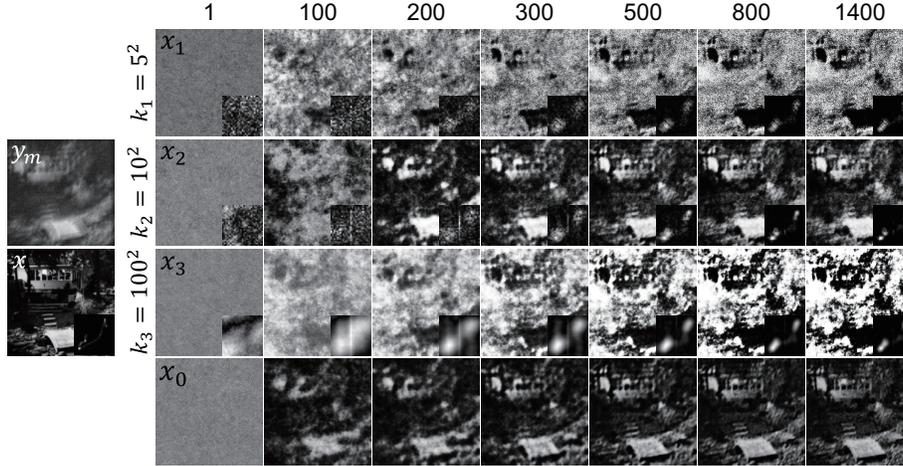


**Fig. S2:** Visualization of the logarithm of Fourier spectrum of .

Fig. S1 shows an  $\hat{w}$  with different  $k$ , inverse filtered results  $y_m * \hat{w}$ , and the final estimated blur kernel  $\hat{h}$ . As it can be inferred from Eq. (1),  $\hat{w}$  is a flipped  $x$ , which is normalized by the sum of the power spectrum of the  $x$  and  $k$ . As can be seen from Fig. S1,  $\hat{w}$  appears to be a high-pass filtered image of  $flip(x)$  due to the effect of the normalization term, *i.e.*  $|\mathcal{F}\{x\}|^2 + k$ . Also, it can be observed that as  $k$  increases from  $1^2$  to  $100^2$ , the low-frequency components of  $\hat{h}$  are gradually recovered. This occurs because  $k$  serves as both an offset and a thresholding operation to  $|X|^2$  (Fig. S2 (a)), and this can be expressed as follows:

$$L = \begin{cases} k, & \text{if } k \gg |X|^2. \\ |X|^2 + k, & \text{otherwise.} \end{cases} \quad (2)$$

, where  $L$  represents the normalization term of  $\hat{W}$ . Fig. S2 (a) illustrates the logarithm of Fourier spectrum of  $L$  and  $\bar{X}$ , and (b) shows the effect of  $L$  to  $\bar{X}$ . Here, we only present the line passing through the center of the Fourier spectrum for effective visualization. As shown in Fig. S2 (b), when  $k$  is small, the high-frequency spectrum is significantly stronger, whereas as  $k$  increases, the low-frequency component gradually becomes stronger than the high-frequency, and this is the reason for the appearance of  $\hat{w}$ . During the inverse filtering process, as the value of  $k$  increases, emphasizing low-frequency components of  $\hat{w}$ ,  $\hat{h}$  gradually becomes more blurred. Simultaneously, the low-frequency components of  $\hat{w}$  contribute to effectively reducing noise by averaging out the zero-mean Gaussian noise. Conversely, when  $k$  is small, high frequencies dominate in  $\hat{w}$ , allowing for the restoration of sharp kernels during the inverse filtering process, although the noise removal effect is significantly reduced.



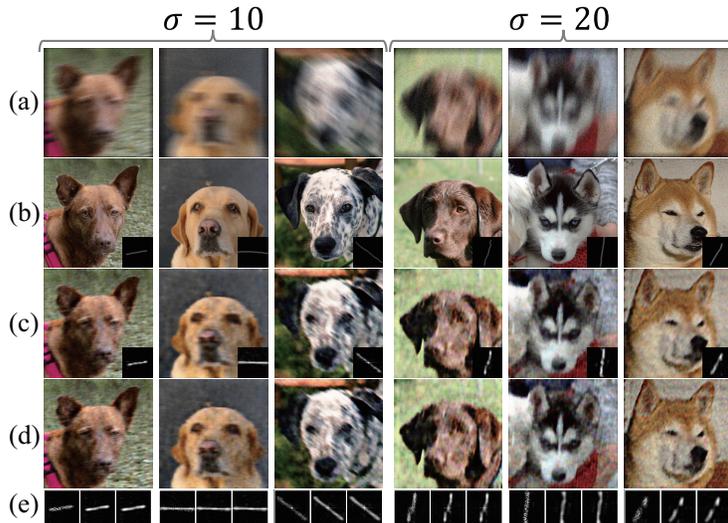
**Fig. S3:** Visualization of the intermediate results of  $x_i$ ,  $h_i$ , and  $x_0$  (iteration=1, 100, 200, 300, 500, 800, 1400).  $y_m$  is degraded by the noise level of  $\sigma = 10$ .

## S2. Analysis of multiple kernel estimation scheme

In this section, we will conduct a comprehensive analysis of the multiple kernel estimation scheme. Firstly, as described in the main text, we introduced an intermediate image to mitigate errors arising from the suboptimal setting of  $k$ . To further investigate this aspect, in Sec. S2.1, we will explore the estimation process of the intermediate image and blur kernel during optimization. Additionally, in Sec. S2.2, we will assess the performance of the multiple kernel estimation scheme and compare it with single kernel estimation, which is performed under the assumption of precisely known noise level.

### S2.1 Analysis of intermediate images

We introduced an intermediate image  $x_i$  to estimate the blur kernel  $h_i$ , aiming to alleviate errors stemming from suboptimal settings of the noise suppression factor  $k_i$ . By doing so, the errors arising from estimating the  $h_i$  are incorporated into the intermediate image  $x_i$ , while the  $h_i$  and  $x_0$  are estimated as desired. To verify our statement, we visualize the intermediate results of  $x_0$ ,  $h_i$ , and  $x_i$ . As illustrated in Fig. S3 where  $k_1 = 5^2$ , initially, due to the smaller  $k$  setting,  $h_1$  appears highly noisy. Nevertheless, as the iterations progress, it is observed that while  $x_1$  is estimated to be highly noisy,  $h_1$  closely approximates the ground truth. On the other hand, when  $k$  is set to a value larger than the optimal value, as illustrated in Fig. S3 where  $k_3 = 100^2$ , initially,  $h_3$  appears notably blurred due to the larger  $k$  setting. However, as the iterations proceed, it is observed that  $x_3$  is intensely saturated to compensate for the blurriness, ultimately resulting in the successful estimation of  $h_3$ .



**Fig. S4:** Qualitative comparison of deblurring performance on motion blurred AFHQ-dog datasets. (a) Blur, (b) Ground truth, (c) Single kernel estimation (SKE) with  $k = 10^2$  for  $\sigma = 10$  and  $k = 20^2$  for  $\sigma = 20$ , (d) Multiple kernel estimation (MKE), and (e) estimated blur kernel with  $k_1 = 5^2$ ,  $k_2 = 15^2$ , and  $k_3 = 25^2$ .

Also, inversion in the brightness of the estimated intermediate images ( $x_1$  and  $x_3$  in Fig. S3) may occur, leading to a  $\pi$ -phase shift in its Fourier component. This results in negative values in the estimated kernel values. We observed that approximately half of the estimated kernel values might initially be negative, and attempting to suppress them leads to optimization failure. This issue is effectively managed by applying a non-negativity constraint (Eq. (1))

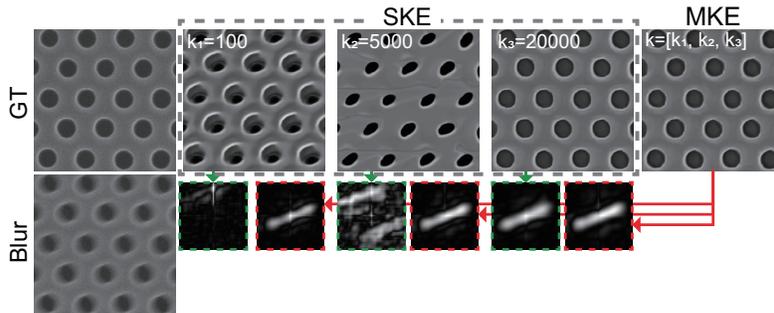
It has been demonstrated that introducing the intermediate images allows for handling the error in blur kernel estimation caused by the suboptimal setting of  $k$ . Another important aspect of the multiple kernel estimation scheme is that estimating both a noisy kernel, which contains sharp edges, and a blurry kernel, which encompasses the overall shape, simultaneously enables the averaging out of kernel estimation errors, facilitating the discovery of the optimal kernel. This ultimately enables the stable estimation of the deblurred image.

## S2.2 Single kernel estimation v.s. Multiple kernel estimation

If we assume that the noise level is accurately known, we can set the noise suppression factor optimally, enabling single kernel estimation. Here, we conduct a qualitative and quantitative comparison of the blind deblurring performance between multiple kernel estimation (MKE) and single kernel estimation (SKE). When performing SKE, we set  $k$  as  $\sigma^2$ . Also, we utilize the AFHQ datasets explained in the main text. As shown in Fig. S4, the MKE demonstrates comparable deblurring performance compared to the SKE. Considering that MKE

**Table 1:** Quantitative comparison of deblurring results from MKE and SKE on AFHQ datasets [7]. **Bold:** best.

Method	AFHQ-dog				AFHQ-cat			
	$\sigma = 10$		$\sigma = 20$		$\sigma = 10$		$\sigma = 20$	
	PSNR $\uparrow$	SSIM $\uparrow$						
MKE	<b>24.23</b>	<b>0.632</b>	<b>22.90</b>	<b>0.539</b>	<b>23.33</b>	<b>0.590</b>	<b>22.11</b>	<b>0.497</b>
SKE	23.94	0.627	22.85	0.527	23.09	0.586	21.98	0.481

**Fig. S5:** Comparison of deblurring results from SKE and MKE on a SEM image.

performs deblurring under the assumption of an unknown noise level, it indicates the remarkable effectiveness of the proposed method. This aspect is further demonstrated in Fig. S5, where SKE notably fails to estimate the clear image and blur kernel when  $k$  deviates from optimal conditions, whereas MKE successfully reconstructs both the clear image and the blur kernel. Furthermore, as shown in Tab. 1, it can be observed that MKE achieves slightly higher or nearly similar PSNR and SSIM scores compared to SKE. Through this analysis, we demonstrate the suitability of the proposed method for blind deblurring in scenarios where the noise level is unknown.

### S3. Implementation details of the proposed method

The proposed method is implemented using the Pytorch library and all the experiments described here are conducted on a NVIDIA Geforce RTX 3080. The fixed tensor  $z$  is sampled from uniform distribution and perturbed at each iteration of the optimization process. The architecture of the generative network  $G_\theta$  is adopted from [21], with modifications made only to the output layer (Sec. S6.). To optimize the network, we use Adam optimizer [11] with an initial learning rate of  $10^{-4}$  which remains fixed for 600 iterations before being halved every 200 iterations. The network parameters are optimized for  $N_{iter} = 1400$  iterations.

## S4. Implementation details of comparison method

**Sanghvi et al. [22].** The method introduces kernel estimation framework by leveraging low-dimensional and differentiable representation of blur kernel, and integrate it into an iterative deblurring framework which employs non-blind solver to address photon-limited blind deblurring problem. We used official implementation<sup>4</sup> and pretrained network provided by the authors.

**Xu et al. [27].** The method presents a novel approach to motion deblurring by introducing a two-phase kernel estimation method. The first phase involves initial kernel estimation using edge information to capture high-frequency details, and the second phase refines this kernel through iterative support detection method. We used the official implementation<sup>5</sup>.

**Pure-let [15].** The method introduces linear expansion of threshold (LET) incorporated into the Poisson unbiased risk estimate (Pure) to optimize Wiener filtering processes. We used official implementation<sup>6</sup>.

**SelfDeblur [21].** SelfDeblur is a study that utilizes the Double-DIP [9] method to perform blind deconvolution. We conducted experiments using almost all of the settings provided in the official implementation<sup>7</sup>. However, as we focused on blind deblurring under noisy conditions the number of iterations is reduced from 5000 to 2000 to prevent overfitting to noise. Optimization involves training with MSE loss for 1000 iterations and minimizing SSIM loss for the remaining iterations.

**WDIP [2].** WDIP enhances SelfDeblur by integrating the conventional Wiener deconvolution method, which is utilized to validate both the deblurred clear image and the estimated blur kernel. We used the official implementation<sup>8</sup>, while reducing the total number of iterations to 2000 to mitigate the overfitting to noise. All of the other hyperparameters are followed by the default settings.

**$l_1/l_2$  [12].** The method utilizes scale-invariant regularization function  $l_1/l_2$  to prevent the diminishment of high-frequency components and to stabilize the optimization process. We used official implementation<sup>9</sup>, and hyperparameters are set as follows:  $\text{min\_lambda} = 250$ ,  $\text{k\_reg\_wt} = 1$ ,  $\text{delta} = 1 \times 10^{-3}$ . All of the other hyperparameters are followed by the default settings.

**Pan-DCP [19].** The method leverages the sparsity of the dark channel in the clear image and incorporates it as a regularization function. We used official implementation<sup>10</sup>, and hyperparameters are set as follows:  $\lambda_{\text{dark}} = 4 \times 10^{-3}$ ,  $\lambda_{\text{grad}} = 4 \times 10^{-3}$ ,  $\lambda_{\text{tv}} = 3 \times 10^{-3}$ ,  $\lambda_{l_0} = 5 \times 10^{-4}$ .

**Pan- $l_0$  [17].** The method utilizes  $l_0$  regularization function to enforce sparsity of

<sup>4</sup> Sanghvi et al.: <https://github.com/sanghviyashiitb/structured-kernel-cvpr23>

<sup>5</sup> Xu et al.: [https://www.cse.cuhk.edu.hk/~leojia/projects/robust\\_deblur/index.html](https://www.cse.cuhk.edu.hk/~leojia/projects/robust_deblur/index.html)

<sup>6</sup> Pure-let: <https://github.com/hijizhou/PureLetDeconv>

<sup>7</sup> SelfDeblur: <https://github.com/csdwren/SelfDeblur>

<sup>8</sup> WDIP: [https://github.com/gbredell/W\\_DIP/tree/main](https://github.com/gbredell/W_DIP/tree/main)

<sup>9</sup>  $l_1/l_2$ : <https://dilipkay.wordpress.com/blind-deconvolution/>

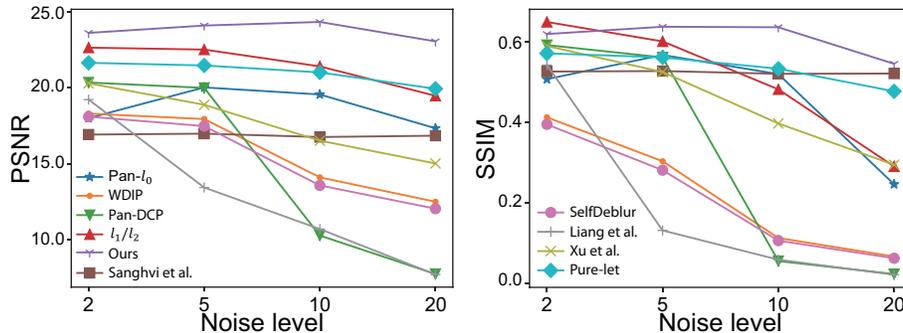
<sup>10</sup> Pan-DCP: <https://jspan.github.io/projects/dark-channel-deblur/index.html>

both clear image and blur kernel. We used official implementation<sup>11</sup>, and hyperparameters are set as follows:  $\lambda_{pixel} = 4 \times 10^{-3}$ ,  $\lambda_{grad} = 4 \times 10^{-3}$ ,  $\lambda_{tv} = 2 \times 10^{-3}$ ,  $\lambda_{l_0} = 2 \times 10^{-4}$

**Liang *et al.* [5].** The method introduces the enhanced sparsity regularization function which is developed from  $l_0$  and  $l_1$  for the gradient of a clear image. We used official implementation<sup>12</sup>, and hyperparameters are set as follows:  $\lambda_{pixel} = 4 \times 10^{-3}$ ,  $\lambda_{grad} = 4 \times 10^{-3}$ ,  $\lambda_{tv} = 2 \times 10^{-3}$ ,  $\lambda_{l_0} = 2 \times 10^{-4}$ .

## S5. Further experimental result

Here, further experimental results are presented. Fig. S6 quantitatively compares deblurring performance for various noise levels. To showcase the effectiveness even with a significantly large blur kernel size, deblurring results on the Lai *et al.* datasets [13] are presented in Figure S7. Further Deblurring results on AFHQ datasets where the noise level of  $\sigma = 10$  is presented in Fig. S8 and Fig. S9, and the noise level of  $\sigma = 20$  is presented in Fig. S10 and Fig. S11. Also, further deblurring results on Sun *et al.* datasets are presented in Fig. S12 and Fig. S13. Lastly, Fig. S14 shows deblurring results on SEM images of square-shaped test targets.



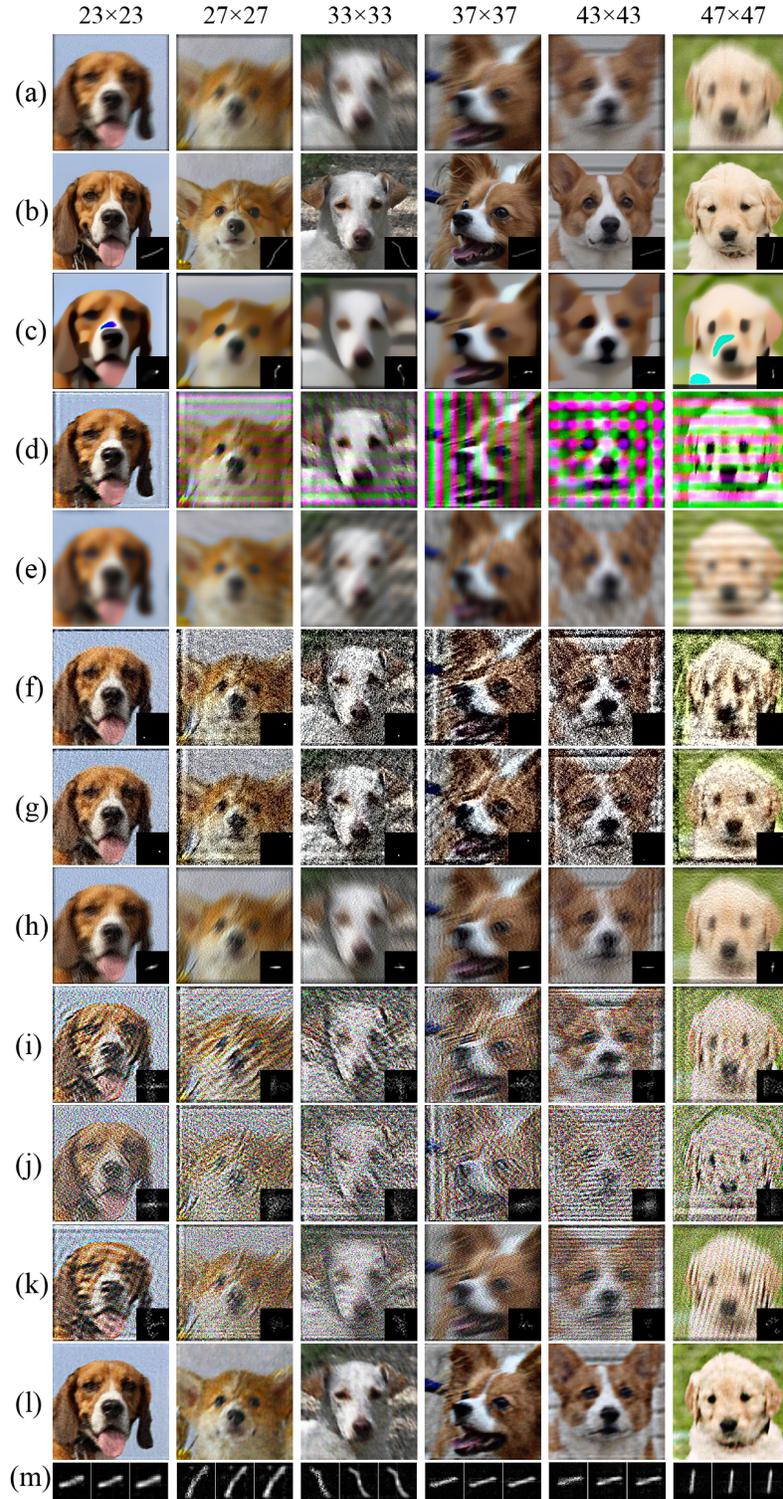
**Fig. S6:** Quantitative comparison of blind deblurring results on AFHQ-dog datasets ( $256 \times 256$ ) with noise level of  $\sigma = 2$ ,  $\sigma = 5$ ,  $\sigma = 10$ , and  $\sigma = 20$ .

<sup>11</sup> Pan- $l_0$ : <https://jspan.github.io/projects/text-deblurring/index.html>

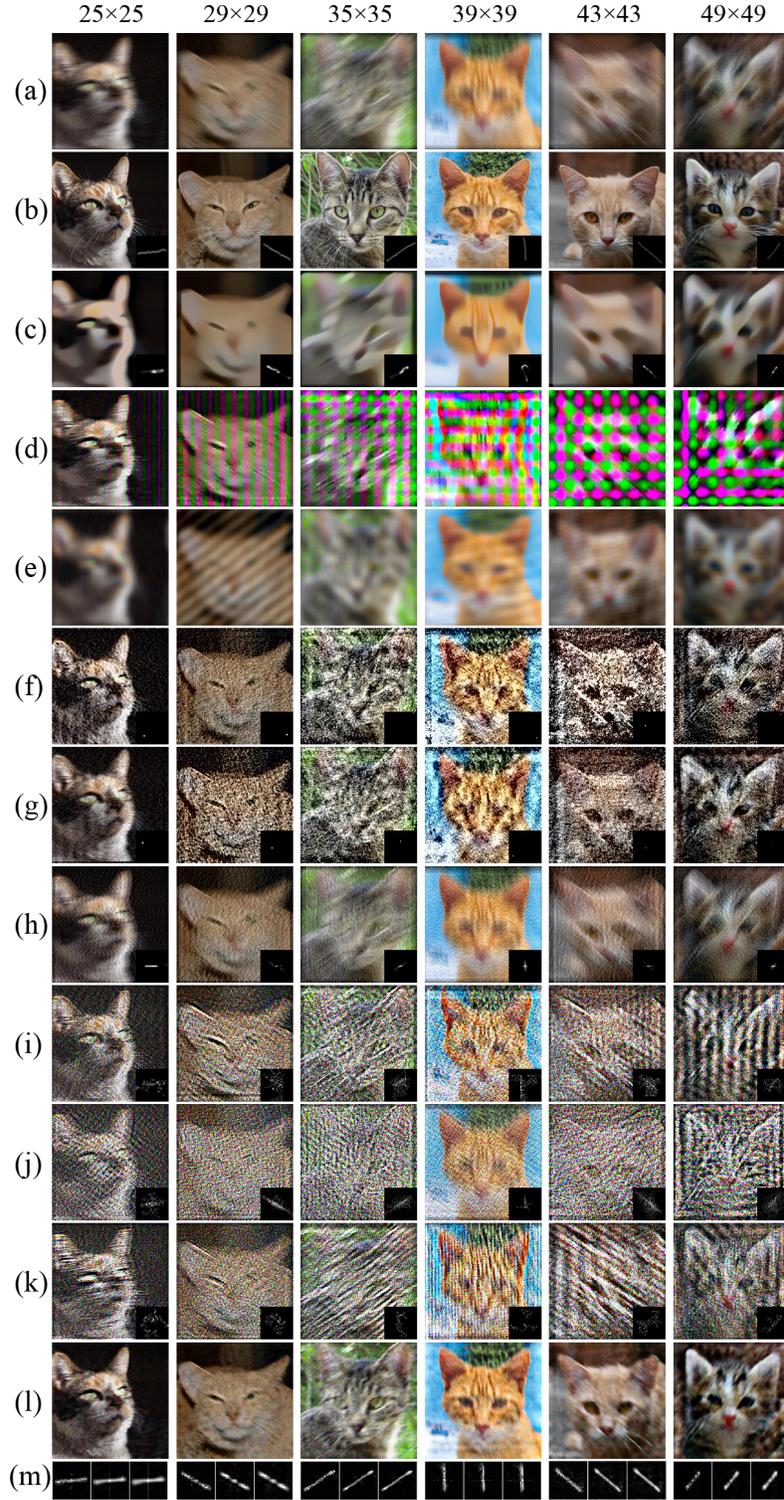
<sup>12</sup> Liang *et al.*: <https://liangchen527.github.io/>



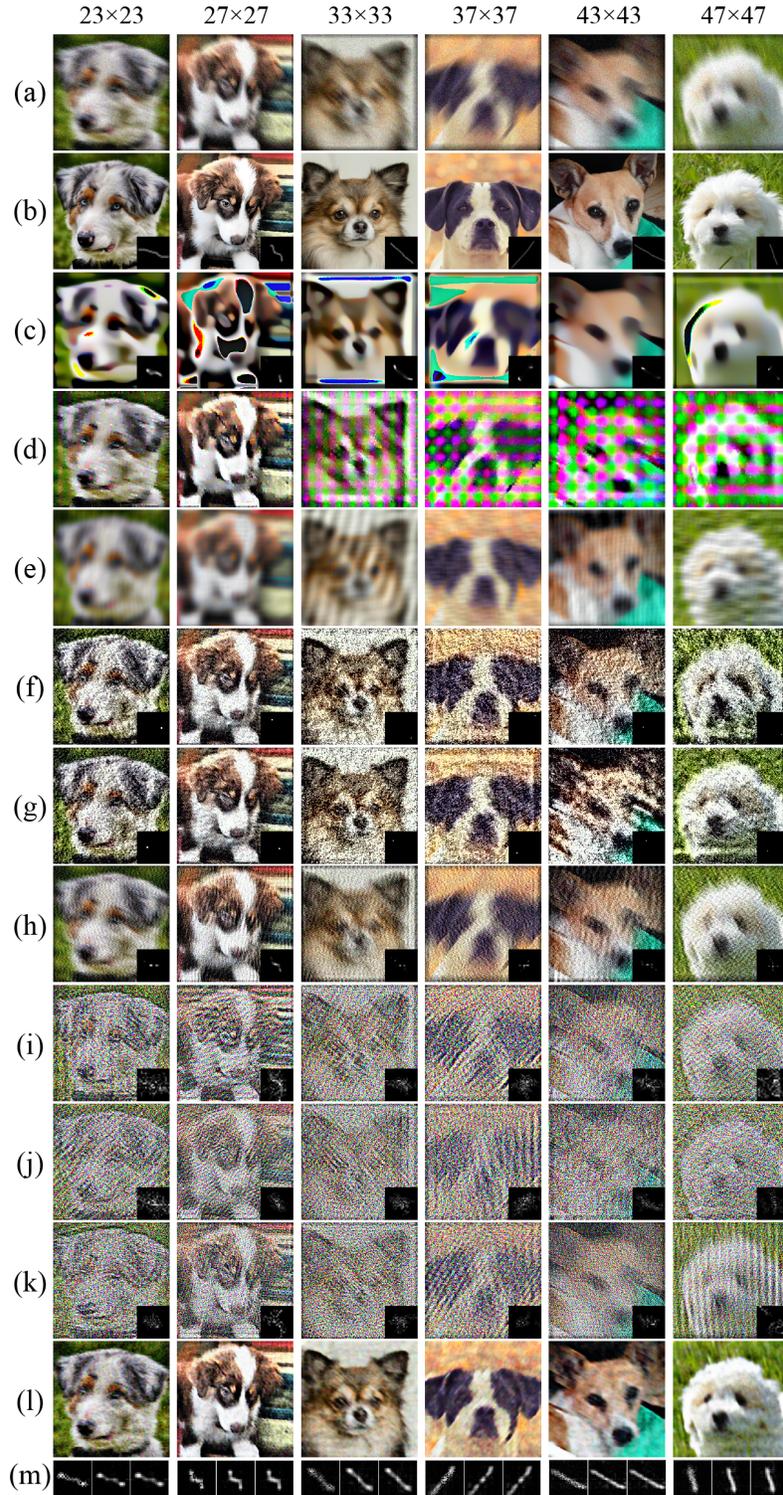
**Fig. S7:** Visual comparison of blind deblurring results on Lai *et al.* datasets with the noise level of  $\sigma = 10$  and  $\sigma = 20$ . (a) Ground truth, (b) Blur ( $\sigma = 10$ ), (d) Blur ( $\sigma = 20$ ), and (c, e) Ours with  $k_1 = 5^2$ ,  $k_2 = 15^2$ , and  $k_3 = 25^2$ .



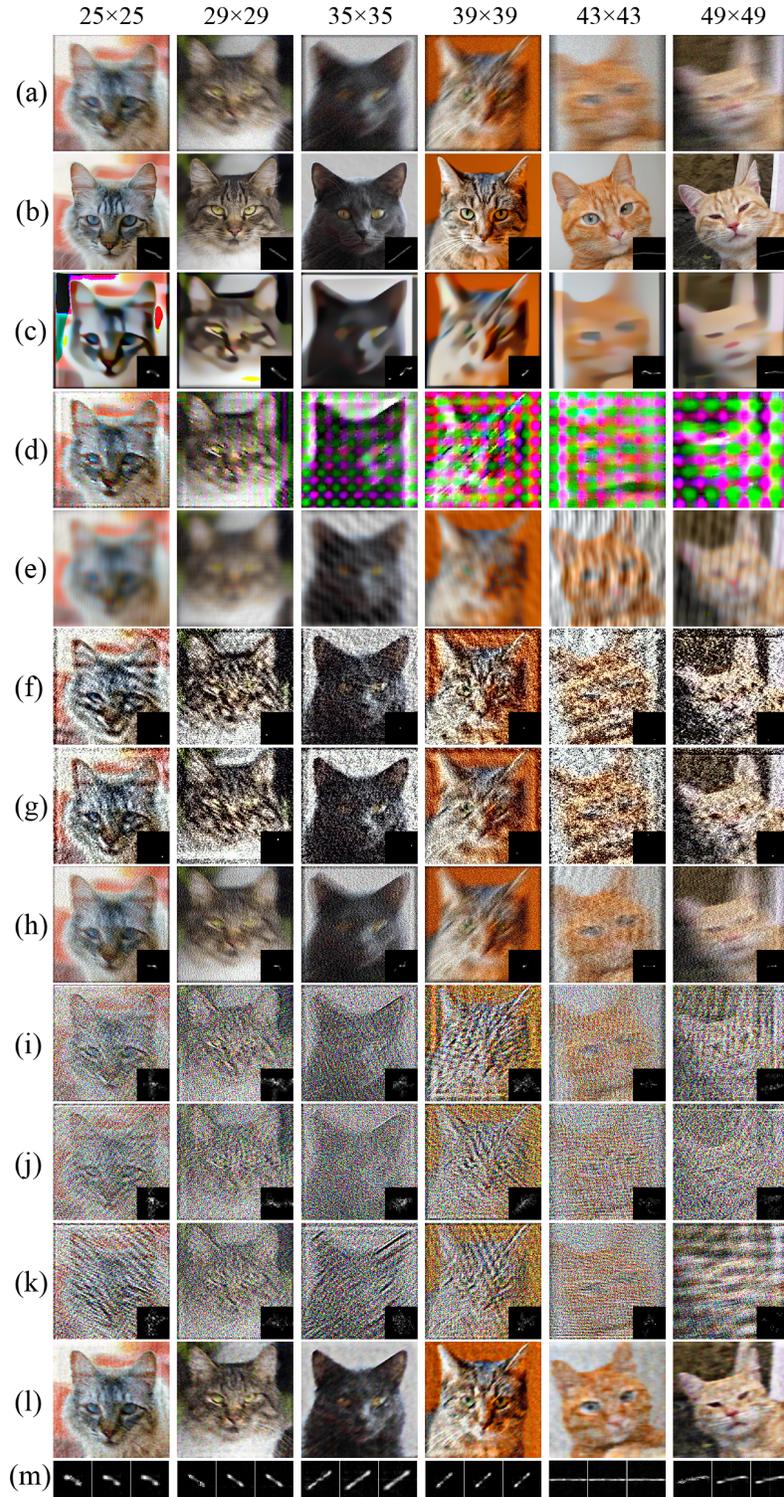
**Fig. S8:** Visual comparison of blind deblurring results on AFHQ-dog datasets ( $256 \times 256$ ) with different kernel sizes and noise level of  $\sigma = 10$ . (a) Blur, (b) Ground truth, (c) Sanghvi *et al.* (d) Xu *et al.* (e) Pure-let (f) SelfDeblur, (g) WDIP, (h)  $l_1/l_2$ , (i) Pan-DCP, (j) Pan- $l_0$  (k) Liang *et al.* (l) Ours, and (m) estimated kernel with  $k_1 = 5^2$ ,  $k_2 = 15^2$ , and  $k_3 = 25^2$ .



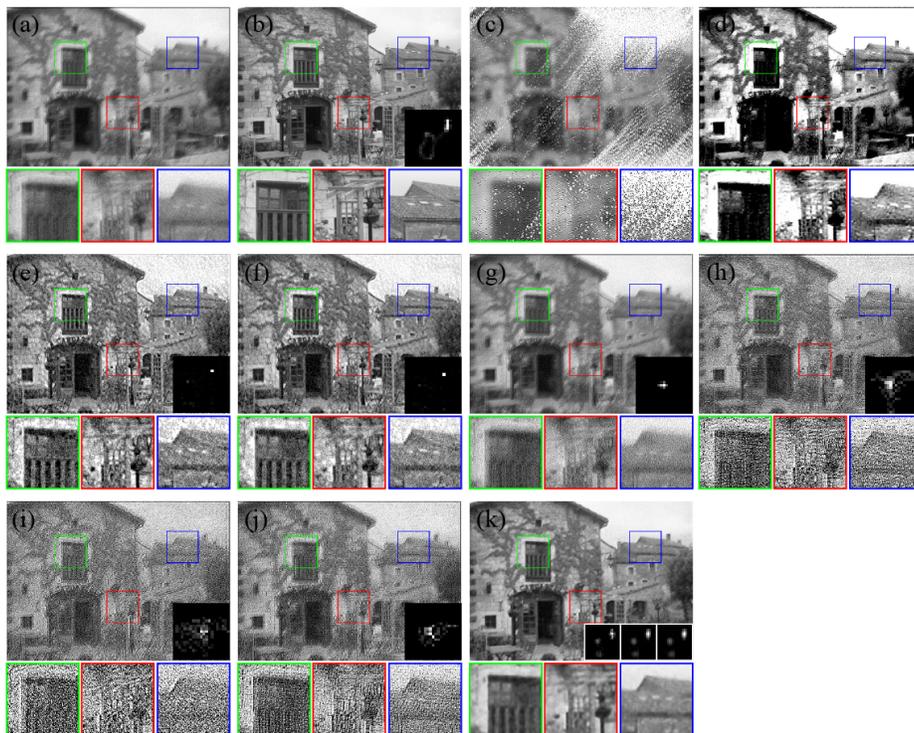
**Fig. S9:** Visual comparison of blind deblurring results on AFHQ-cat datasets ( $256 \times 256$ ) with different kernel sizes and noise level of  $\sigma = 10$ . (a) Blur, (b) Ground truth, (c) Sanghi *et al.* (d) Xu *et al.* (e) Pure-let (f) SelfDeblur, (g) WDIP, (h)  $l_1/l_2$ , (i) Pan-DCP, (j) Pan- $l_0$  (k) Liang *et al.* (l) Ours, and (m) estimated kernel with  $k_1 = 5^2$ ,  $k_2 = 15^2$ , and  $k_3 = 25^2$ .



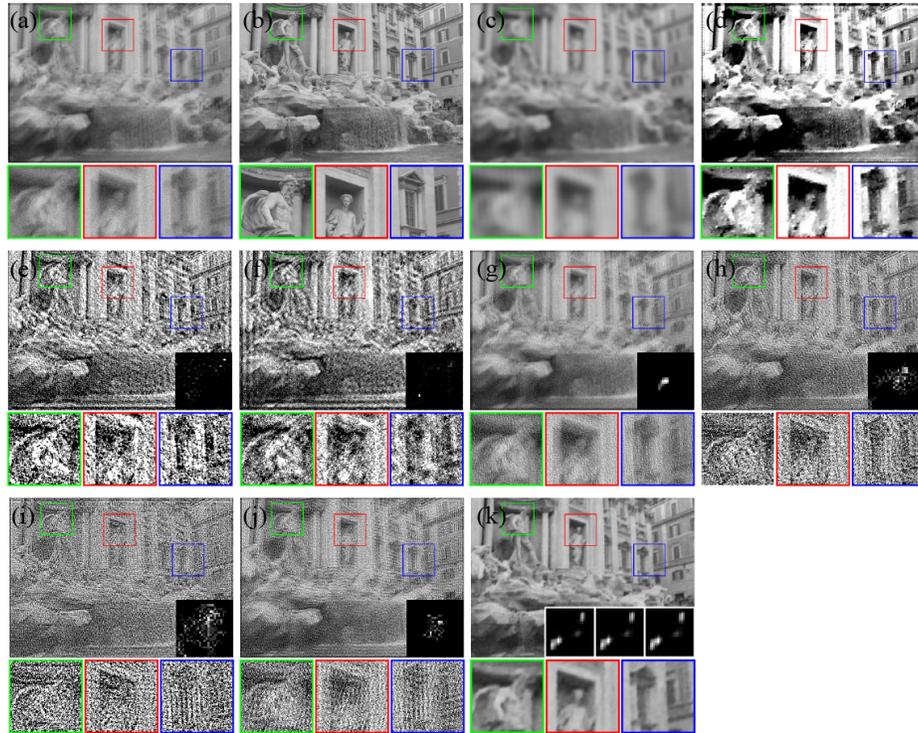
**Fig. S10:** Visual comparison of blind deblurring results on AFHQ-dog datasets ( $256 \times 256$ ) with different kernel sizes and noise level of  $\sigma = 20$ . (a) Blur, (b) Ground truth, (c) Sanghvi *et al.* (d) Xu *et al.* (e) Pure-let (f) SelfDeblur, (g) WDIP, (h)  $l_1/l_2$ , (i) Pan-DCP, (j) Pan- $l_0$  (k) Liang *et al.* (l) Ours, and (m) estimated kernel with  $k_1 = 5^2$ ,  $k_2 = 15^2$ , and  $k_3 = 25^2$ .



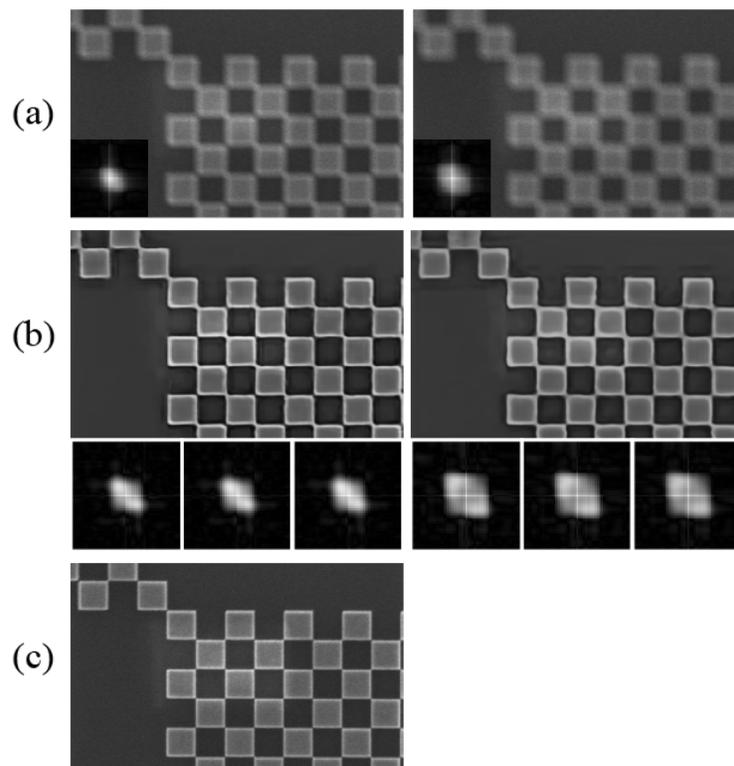
**Fig. S11:** Visual comparison of blind deblurring results on AFHQ-cat datasets ( $256 \times 256$ ) with different kernel sizes and noise level of  $\sigma = 20$ . (a) Blur, (b) Ground truth, (c) Sanghvi *et al.* (d) Xu *et al.* (e) Pure-let (f) SelfDeblur, (g) WDIP, (h)  $l_1/l_2$ , (i) Pan-DCP, (j) Pan- $l_0$  (k) Liang *et al.* (l) Ours, and (m) estimated kernel with  $k_1 = 5^2$ ,  $k_2 = 15^2$ , and  $k_3 = 25^2$ .



**Fig. S12:** Visual comparison of blind deblurring results on Sun *et al.* datasets with the noise level of  $\sigma = 10$ . (a) Blur, (b) Ground truth, (c) Pure-let (d) Xu *et al.* (e) SelfDeblur, (f) WDIP, (g)  $l_1/l_2$ , (h) Pan-DCP, (i) Pan- $l_0$  (j) Liang *et al.* (k) Ours with  $k_1 = 5^2$ ,  $k_2 = 15^2$ , and  $k_3 = 25^2$ .



**Fig. S13:** Visual comparison of blind deblurring results on Sun *et al.* datasets with the noise level of  $\sigma = 20$ . (a) Blur, (b) Ground truth, (c) Pure-let (d) Xu *et al.* (e) SelfDeblur, (f) WDIP, (g)  $l_1/l_2$ , (h) Pan-DCP, (i) Pan- $l_0$  (j) Liang *et al.* (k) Ours with  $k_1 = 5^2$ ,  $k_2 = 15^2$ , and  $k_3 = 25^2$ .



**Fig. S14:** Blind deblurring results on SEM image ( $320 \times 512$ ) of square-shaped test targets. The ground truth PSFs ( $49 \times 49$ ) are obtained by the proposed kernel estimation function with  $k_1 = 10000$ ,  $k_2 = 20000$ , and  $k_3 = 30000$ . (a) Blur, (b) Ours, and (c) Ground truth.

## S6. Network architecture

The architecture of the generative network  $G_\theta$  is adopted from [21], with modifications made to the output layer. The output layer of  $G_\theta$  consists of  $f_0 + N_k \times f_i$  channels. These channels are split into  $1 + N_k$  groups and then convolved with  $g_{\phi_0}$  and  $g_{\phi_i}$  convolution filters, *i.e.* group convolution, to generate clear image and intermediate images, respectively. During experiments conducted on SEM and 2PEF datasets, skip connections are removed, and the number of convolution filters is modified, gradually increasing for the encoder and decreasing for the decoder.

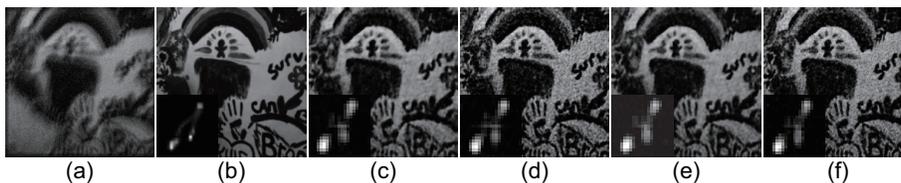
## S7. Ablation study and limitation

In this section, we will present the results of the ablation study and discuss the limitations of the proposed method. Firstly, we conducted a study on how the introduction of the intermediate image contributes to optimizing the proposed method reliably. Secondly, the impact of regularization terms on clear image and blur kernel is investigated, demonstrating that the key enabler of our work is multiple kernel estimation scheme. Next, we investigated how the performance of the proposed method changes depending on the setting of the range of the noise suppression factors. Lastly, we experimented with how the performance of the proposed method varies depending on the number of noise suppression factors. The ablation study was performed using Levin *et al.* datasets [14], consisting of 4 clear images and 8 blur kernels, resulting in the generation of 32 blurry images.

### S7.1 Ablation study

**The effect of regularizations for clear image and blur kernel.** Here, we confirmed that the key enabler for robust deblurring at an unknown and high level of noise is our proposed scheme of Wiener kernel estimation at multiple levels of noise suppression factors, as presented in Fig. S15 and Tab. 2. This feature makes the key distinction with the previous MAP approaches based on regularizations at both clear image and kernel domains. We introduced the L2 norm on the estimated blur kernel and the TV norm on the clear image, respectively, for suppressing the values in the peripheral area of the kernel and for preventing overfitting of the estimated image to noise whose effect is confirmed to be marginal compared to the proposed Wiener kernel estimation scheme. Our claim is further supported by Fig. S5, where SKE fails in optimization despite using both regularization terms.

**The effect of intermediate images.** As we have discussed in Sec. S2., the introduction of intermediate images results in the stable optimization of the proposed method. To verify this, we compared the proposed method with an approach that estimates the blur kernel from a single clear image without utilizing intermediate images. To implement this, the blur kernel was estimated using  $h_i = F(y_m, x_0, k_i)$ . In this experiment, the noise suppression factors were set to



**Fig. S15:** Visualization of ablation study on the impact of regularizations. (a) Blur, (b) ground truth, (c) TV + L2, (d) L2, (e) TV, (f) None.

**Table 2:** Quantitative evaluation of blind deblurring results on Levin *et al.* datasets to assess the impact of intermediate images and regularization terms. **Bold:** best.

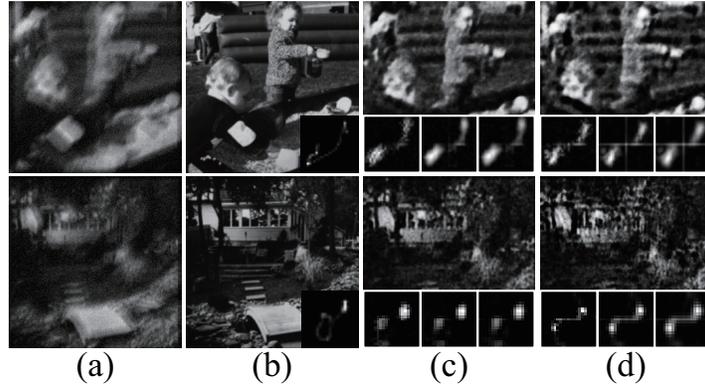
Intermediate image	$\sigma = 10$				$\sigma = 20$	
	TV	L2	PSNR $\uparrow$	SSIM $\uparrow$	PSNR $\uparrow$	SSIM $\uparrow$
○	○	○	23.56	0.635	<b>22.36</b>	<b>0.553</b>
○		○	21.51	0.501	18.91	0.347
○	○		<b>23.71</b>	<b>0.642</b>	21.65	0.479
○			21.91	0.512	19.03	0.354
	○	○	21.87	0.564	21.22	0.495

$k_1 = 5^2$ ,  $k_2 = 15^2$ , and  $k_3 = 25^2$  for both cases. As shown in the Tab. 2, the proposed method outperforms significantly in terms of performance compared to the method that does not utilize intermediate images. Additionally, Fig. S16 illustrates the stability of the proposed method in accurately estimating both the blur kernel and the clear image (Fig. S16 (c)), whereas the compared method exhibits some ringing artifacts in the estimated clear image or incorrect kernel estimation (Fig. S16 (d)), attributed to the suboptimal setting of  $k$ .

**Setting of the noise suppression factors.** As discussed in the main text, due to the wide range of suboptimal noise suppression factors that can reduce the data-fidelity loss, it is advisable to choose them appropriately considering the noise levels present in nature. To delve deeper into this, we conducted experiments to examine how the setting of noise suppression factors affects the deblurring results. As illustrated in Tab. 3, when there is a significantly small

**Table 3:** Quantitative comparison of the proposed method for different values of  $k$  when the noise level is  $\sigma = 10$ . **Bold:** best.

Metric	$k_1, k_2, k_3$				
	$1^2, 2^2, 5^2$	$1^2, 10^2, 20^2$	$5^2, 10^2, 20^2$	$5^2, 20^2, 50^2$	$50^2, 75^2, 100^2$
PSNR $\uparrow$	19.65	20.46	23.33	<b>23.75</b>	23.51
SSIM $\uparrow$	0.423	0.466	0.627	<b>0.641</b>	0.632



**Fig. S16:** Visualization of ablation study on the impact of intermediate images. (a) Blur, (b) Ground truth, (c) w/ intermediate image, and (d) w/o intermediate image.

value of  $k$  relative to the noise level, such as  $k_1 = 1^2$ , the optimization process encounters difficulties due to the strong influence of noise. However, in the remaining cases, it is observed that even if the optimal blur kernel is not found, successful estimation of clear images is achieved. While it may seem beneficial to set the noise suppression factor generally larger, it can pose challenges in estimating high-frequency components of the blur kernel and make it difficult to find sharp and complex kernels, such as those caused by handshake. Therefore, it is best to set the noise suppression factor to match an approximately estimated noise level.

**The number of the estimated blur kernels.** Finally, we conducted experiments to investigate how the deblurring performance changes with the number of estimated kernels. We gradually increased the number of estimated kernels  $N_k$  from 2 to 5, setting  $k_1 = 5^2$  and  $k_{N_k} = 50^2$ . Also, other noise suppression factors are set as follows:  $k_2 = 20^2$  for  $N_k = 3$ ,  $k_2 = 20^2$  and  $k_3 = 35^2$  for  $N_k = 4$ , and  $k_2 = 15^2$ ,  $k_3 = 25^2$ , and  $k_4 = 35^2$  for  $N_k = 5$ . As can be seen from Tab. 4, the best score is achieved when  $N_k = 3$ . The reason why the deblurring performance decreases as  $N_k$  increases is that the proposed method estimates both intermediate images and clear images from a single neural network. As  $N_k$  increases, the network loses balance between estimating the clear image and the intermediate images. On the other hand, in this experiment, the wide range of  $k$  settings compared to the optimal condition leads to a slightly lower deblurring performance when  $N_k = 2$ . Although we utilized  $N_k = 3$  in all experiments, if one can know that the noise level lies within a certain limited range, considering the minimum and maximum levels of noise to set the noise suppression factors, *i.e.*  $N_k = 2$ , could also be an alternative.

**Table 4:** Quantitative evaluation of blind deblurring results by varying the number of estimated kernels when the noise level is  $\sigma = 10$ . **Bold:** best.

Metric	$N_k = 2$	$N_k = 3$	$N_k = 4$	$N_k = 5$
PSNR $\uparrow$	23.65	<b>23.75</b>	23.21	23.17
SSIM $\uparrow$	0.639	<b>0.641</b>	0.621	0.617

## S7.2 Limitation

Firstly, the proposed method is not suitable for addressing the non-uniform deblurring problem [1,3,8,10,20,24,26,28]. This problem arises when different parts of an image experience varying degrees of blurring, as opposed to the entire image undergoing the same blurring operation. The proposed method necessitates Fourier operations across the entire image to derive the blur kernel, complicating the process of operating on individual image regions. While one potential solution might entail dividing the image into sections and applying the proposed method to each section before merging them, this approach presents numerous challenges, including determining the optimal segmentation of the image into regions.

Furthermore, the proposed method encounters difficulties when handling images containing numerous outliers, such as saturation or hot pixels, or when dealing with images captured in low-light conditions where Poisson noise is prevalent. This challenge stems from the assumption of zero-mean Gaussian noise in the proposed kernel estimation function. Nevertheless, the main goal of the proposed method is to introduce a technique for kernel estimation from noisy blurred images, rendering it highly amenable to integration with other approaches. For example, the integration of additional regularization functions or optimization techniques designed to address outliers [6,18,25] or Poisson noise [4,16,23] could enhance its effectiveness in mitigating such noise or outliers.

Finally, as the proposed method employs the deep image prior (DIP) under the assumption of scenarios where acquiring training data, such as from scanning electron microscopy (SEM) or two-photon excited fluorescence (2PEF), is difficult, there could be a loss of information regarding fine details in the estimated clear image. However, it is worth considering the adoption of pre-trained generative models, such as generative adversarial networks (GANs), conditional GANs, flow-based models, or diffusion models, which have shown outstanding external generalization and performance in recent advancements, as alternatives to the DIP for enhancing image reconstruction capabilities.

## References

1. Bahat, Y., Efrat, N., Irani, M.: Non-uniform blind deblurring by reblurring. In: Proceedings of the IEEE international conference on computer vision. pp. 3286–3294 (2017)
2. Bredell, G., Erdil, E., Weber, B., Konukoglu, E.: Wiener guided dip for unsupervised blind image deconvolution. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. pp. 3047–3056 (2023)
3. Carbajal, G., Vitoria, P., Delbracio, M., Musé, P., Lezama, J.: Non-uniform blur kernel estimation via adaptive basis decomposition. arXiv preprint arXiv:2102.01026 (2021)
4. Carlván, M., Blanc-Féraud, L.: Sparse poisson noisy image deblurring. IEEE Transactions on Image Processing **21**(4), 1834–1846 (2011)
5. Chen, L., Fang, F., Lei, S., Li, F., Zhang, G.: Enhanced sparse model for blind deblurring. In: European Conference on Computer Vision. pp. 631–646. Springer (2020)
6. Cho, S., Wang, J., Lee, S.: Handling outliers in non-blind image deconvolution. In: 2011 international conference on computer vision. pp. 495–502. IEEE (2011)
7. Choi, Y., Uh, Y., Yoo, J., Ha, J.W.: Stargan v2: Diverse image synthesis for multiple domains. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (2020)
8. Fang, Z., Wu, F., Dong, W., Li, X., Wu, J., Shi, G.: 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. pp. 18105–18114 (2023)
9. Gandelsman, Y., Shocher, A., Irani, M.: " double-dip": unsupervised image decomposition via coupled deep-image-priors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 11026–11035 (2019)
10. Hirsch, M., Schuler, C.J., Harmeling, S., Schölkopf, B.: Fast removal of non-uniform camera shake. In: 2011 International Conference on Computer Vision. pp. 463–470. IEEE (2011)
11. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
12. Krishnan, D., Tay, T., Fergus, R.: Blind deconvolution using a normalized sparsity measure. In: CVPR 2011. pp. 233–240. IEEE (2011)
13. Lai, W.S., Huang, J.B., Hu, Z., Ahuja, N., Yang, M.H.: A comparative study for single image blind deblurring. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1701–1709 (2016)
14. Levin, A., Weiss, Y., Durand, F., Freeman, W.T.: Understanding and evaluating blind deconvolution algorithms. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 1964–1971. IEEE (2009)
15. Li, J., Luisier, F., Blu, T.: Pure-let image deconvolution. IEEE Transactions on Image Processing **27**(1), 92–105 (2017)
16. Ma, L., Moisan, L., Yu, J., Zeng, T.: A dictionary learning approach for poisson image deblurring. IEEE Transactions on medical imaging **32**(7), 1277–1289 (2013)
17. Pan, J., Hu, Z., Su, Z., Yang, M.H.: Deblurring text images via l0-regularized intensity and gradient prior. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2901–2908 (2014)
18. Pan, J., Lin, Z., Su, Z., Yang, M.H.: Robust kernel estimation with outliers handling for image deblurring. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 2800–2808 (2016)

19. Pan, J., Sun, D., Pfister, H., Yang, M.H.: Blind image deblurring using dark channel prior. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1628–1636 (2016)
20. Park, D., Kang, D.U., Kim, J., Chun, S.Y.: Multi-temporal recurrent neural networks for progressive non-uniform single image deblurring with incremental temporal training. In: European Conference on Computer Vision. pp. 327–343. Springer (2020)
21. Ren, D., Zhang, K., Wang, Q., Hu, Q., Zuo, W.: Neural blind deconvolution using deep priors. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3341–3350 (2020)
22. Sanghvi, Y., Mao, Z., Chan, S.H.: Structured kernel estimation for photon-limited deconvolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 9863–9872 (2023)
23. Setzer, S., Steidl, G., Teuber, T.: Deblurring poissonian images by split bregman techniques. *Journal of Visual Communication and Image Representation* **21**(3), 193–199 (2010)
24. Sun, J., Cao, W., Xu, Z., Ponce, J.: Learning a convolutional neural network for non-uniform motion blur removal. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 769–777 (2015)
25. Whyte, O., Sivic, J., Zisserman, A.: Deblurring shaken and partially saturated images. *International journal of computer vision* **110**, 185–201 (2014)
26. Whyte, O., Sivic, J., Zisserman, A., Ponce, J.: Non-uniform deblurring for shaken images. *International journal of computer vision* **98**, 168–186 (2012)
27. Xu, L., Jia, J.: Two-phase kernel estimation for robust motion deblurring. In: Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5–11, 2010, Proceedings, Part I 11. pp. 157–170. Springer (2010)
28. Zhang, H., Wipf, D.: Non-uniform camera shake removal using a spatially-adaptive sparse penalty. *Advances in Neural Information Processing Systems* **26** (2013)