Self-supervised defocus blur kernel estimation with symmetry and smoothness constraints

Yunfei Tian, Yunping Zhang, Yuxing Li, Shuo Zhu, and Edmund Y. Lam*

Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam, Hong Kong SAR, China

ABSTRACT

In microscopic imaging systems, the extremely shallow depth of field often causes defocus-induced blurring as a result of slight vertical movements of biological samples. Accurate blur kernel estimation enables clearer image recovery through deconvolution. Traditional methods for blur kernel estimation and image deblurring rely on idealized assumptions and lack robustness to noise. Meanwhile, deep learning approaches require large datasets, labeled training data, and significant computational resources, limiting their applicability. To address these challenges, we propose a self-supervised neural network that incorporates symmetry and smoothness constraints as loss functions for blur kernel estimation and defocus image restoration. This method leverages feature extraction capabilities of neural networks while using handcrafted priors to guide optimization. It requires no data collection, labeling, or pre-training, making it efficient and adaptable while enabling the handling of complex blur kernels. Experimental results show improved kernel estimation accuracy and enhanced image clarity, with potential applications in medical imaging and biological diagnostics.

Keywords: self-supervised neural network, defocus, blur kernel estimation, symmetry and smoothness constraints

1. INTRODUCTION

Microscopes are widely used in biomedical, materials science, and other fields for high-resolution observation of microstructures.[1,2] However, due to the extremely shallow depth of field in optical systems and slight axial movements of samples, defocus blur often occurs during microscope imaging, severely degrading image quality and the accuracy of subsequent analysis.[3,4] Therefore, how to recover clear images from defocus-blurred images has long been a critical research problem in the field of microscopic imaging.[5]

To address these limitations, computational imaging has emerged as a powerful paradigm that integrates optical hardware with algorithmic post-processing to enhance image quality beyond the physical limits of traditional systems.[6–8] By modeling and compensating for optical degradations such as defocus blur, computational imaging enables more flexible and robust image reconstruction.[9,10] This is particularly promising in microscopy, where computational imaging can overcome the constraints of limited depth of field and enable accurate visualization of complex biological or material structures without demanding significant hardware modifications.[10]

Existing image deblurring methods within the computational imaging framework can be broadly classified into two categories: traditional blind deconvolution methods and deep learning-based approaches. Traditional methods such as maximum a posteriori (MAP) estimation typically rely on strong prior assumptions, suffer from limited robustness, and are prone to local optima in complex blurring scenarios.[11,12] In recent years, deep learning methods have achieved remarkable progress in image restoration but generally depend on large quantities of paired training data, which incurs high acquisition costs and limits their generalization ability—particularly in high-precision, low-sample scenarios like microscopic imaging.[13–15]

To address these challenges, we propose a self-supervised microscopic image deblurring method. Without relying on any supervised signals from ground-truth clear images or blur kernels, this method jointly optimizes two generative networks to estimate the latent sharp image and corresponding blur kernel, respectively. We design two regularization loss functions based on the physical characteristics of optical imaging: rotational

Send correspondence to Edmund Y. Lam: elam@eee.hku.hk

Multimodal Sensing and Artificial Intelligence for Sustainable Future, edited by Francesco Soldovieri, Pascal Picart, Vittorio Bianco, Claas Falldorf, Proc. of SPIE Vol. 13570, 135700X
2025 Published by SPIE · 0277-786X · doi: 10.1117/12.3062402

symmetry constraints and spatial smoothness constraints. Embedded as loss terms in the training process, these prior constraints effectively guide the blur kernel estimation toward more natural and structurally reasonable solutions, thereby significantly enhancing the image recovery quality and detail representation capability.

2. METHOD

2.1 Self-supervised Network Design

We employ two generative networks to generate the refocused latent clean image and the blur kernel, respectively.[16] The result of convolving the latent clean image with the blur kernel should closely match the captured blurred image, thereby enabling reverse optimization of the generative networks. The network structure design is illustrated in Fig. 1 (a), where G_I and G_k represent two end-to-end generative networks. I_x denotes a randomly initialized noise image, while I_y represents the latent clean image output after passing through the network G_I. K_x is a randomly initialized noise blur kernel, and K_y is the blur kernel estimated by the network G_k.

Among them, G_I adopts the Deep Image Prior (DIP) approach,[17] which operates as a self-supervised framework without requiring external training data. It can leverage the inherent capabilities of neural networks to capture the statistical characteristics of natural images. It employs an encoder-decoder network structure with skip connections between the encoder and decoder to enhance the network's ability to recover details.[18] Given the small size and relatively simple features of the blur kernel, G_k employs a fully connected network (FCN). Before being input to the network, the two-dimensional randomly initialized noise blur kernel k_x is flattened into a one-dimensional vector. Similarly, after processing by the network G_k, it is reshaped back into a two-dimensional blur kernel.

Given a blurred image I_B collected by a microscope, the reconstruction loss is defined as:

$$\mathcal{L}_{\text{rec}} = \left\| \hat{I}_y \otimes \hat{k}_y - I_B \right\|^2, \tag{1}$$

where $\hat{I}_y = G_I(I_x)$ is the latent clear image predicted by the generator network G_I with parameters θ_I, $\hat{k}_y = G_k(k_x)$ is the blur kernel predicted by the kernel generator G_k with parameters θ_k, $\otimes$ denotes 2D convolution.

This loss encourages the two networks G_I and G_k to produce a clean image and a kernel such that their convolution closely approximates the observed defocus blurred image. During training, the gradients of $\mathcal{L}_{\text{rec}}$ are backpropagated through the convolution operation to update the parameters θ_I and θ_k .[18] In this way, the network learns to generate $\hat{I}_y$ and $\hat{k}_y$ that minimize the discrepancy between the synthesized blur and the collected blurred image I_B.

Besides, to improve the perceptual quality of the predicted latent clean image $\hat{I}_y$ and to suppress artifacts or spurious high-frequency noise, it is common to incorporate a Total Variation (TV) regularization term into the loss function during training.[11] This term acts directly on the output of the image generator G_I, promoting spatial smoothness while preserving essential structural details and maintaining sharp transitions at meaningful edges.

The loss function of TV regularization term is:

$$\mathcal{L}_{\text{TV}}(\hat{I}_y) = \sum_{i=1}^{H-1} \sum_{j=1}^{W-1} \sqrt{(\hat{I}_{y_{i+1,j}} - \hat{I}_{y_{i,j}})^2 + (\hat{I}_{y_{i,j+1}} - \hat{I}_{y_{i,j}})^2}, \tag{2}$$

where $\hat{I}_y \in \mathbb{R}^{H \times W}$ is the predicted image, and (i, j) denotes pixel coordinates.

2.2 Symmetry and Smoothness Constraints

In conventional optical microscopy, defocus blur is generated by the diffraction of light through a circular aperture, resulting in the point spread function (PSF) generally being smooth and centrally symmetric.[19] While larger defocus levels may produce more complex kernel shapes, such as disk-like or ring-shaped patterns, central symmetry is typically retained.[20] Even under moderate lens aberrations or imperfect aperture geometries,

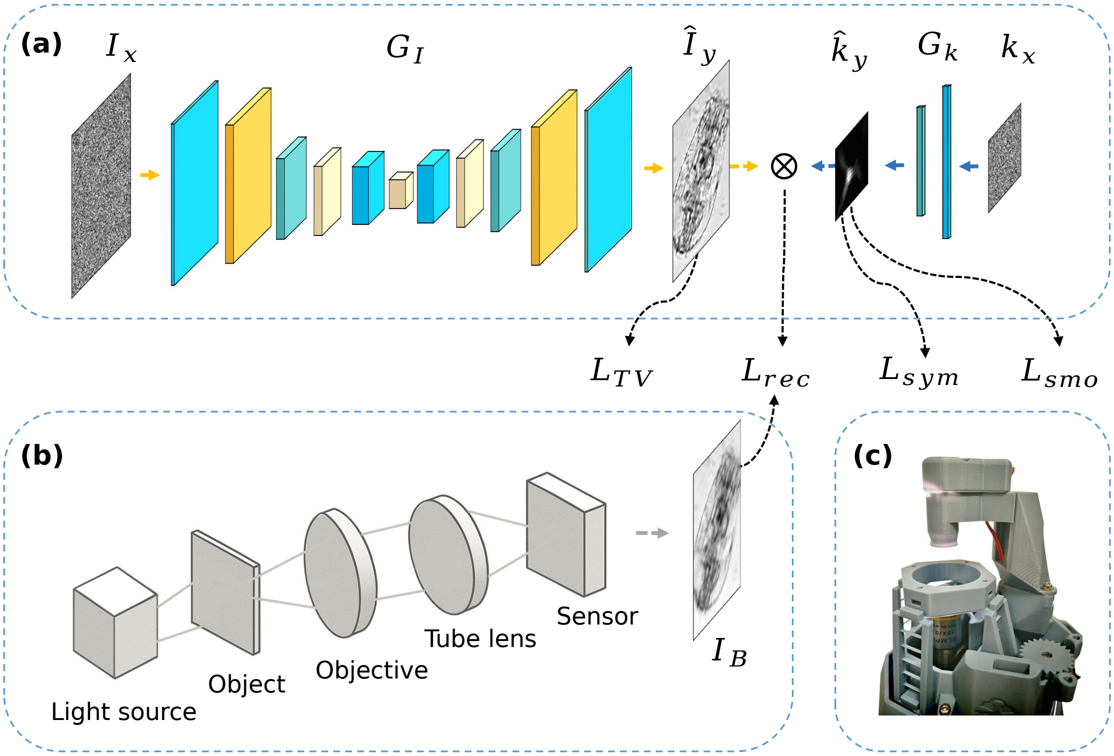

Figure 1. (a) Schematic diagram of the self-supervised network structure and constraint effects, as well as the demonstration of the action positions of different loss functions. (b) Schematic diagram of the optical path of the microscopy system for capturing defocused images. (c)Image of the custom-built microscopy system.

the resulting blur kernel often remains approximately symmetric due to the inherent characteristics of optical defocusing.[21]

To improve the accuracy and physical plausibility of the predicted blur kernels, we embed these optical priors into our kernel estimation process. Specifically, we introduce two regularization losses on the predicted blur kernel $\hat{k}_y = G_k(k_x)$: *rotational symmetry loss* and *spatial smoothness loss*. These loss functions guide the training of the kernel generator G_k by encouraging its output to conform to the expected structural properties of defocus-induced blur.

Rotational symmetry loss. To model the central symmetry property, we penalize differences between the predicted kernel and its rotated versions:

$$\mathcal{L}_{\text{sym}}(\hat{k}_y) = \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} \left\| \hat{k}_y - R_{\alpha_i}(\hat{k}_y) \right\|^2, \tag{3}$$

where $R_{\alpha_i}(\cdot)$ denotes a rotation by angle $\alpha_i \in \{\alpha_1, \alpha_2, \ldots, \alpha_N\}$. Only the valid overlapping regions are used in the loss computation. This constraint helps the generator produce kernels with approximate radial symmetry, as expected in optical defocus.

Smoothness loss. Defocus blur kernels, being physically derived from the PSF of a circular aperture, are also characteristically smooth. To encourage this, we define a gradient-based smoothness loss:

$$\mathcal{L}_{\text{smo}}(\hat{k}_y) = \sum_{i,j} \left[(\hat{k}_{y_{i+1,j}} - \hat{k}_{y_{i,j}})^2 + (\hat{k}_{y_{i,j+1}} - \hat{k}_{y_{i,j}})^2 \right]. \tag{4}$$

This loss penalizes abrupt changes in the spatial domain of the kernel and encourages locally consistent kernel surfaces, which aligns with the continuous energy distribution of defocused light.

Soft constraints. These two regularization terms are grounded in the physics of optical image formation. In cases of ideal or near-ideal circular apertures, the blur kernel is inherently symmetric and smooth. Even when aberrations, optical distortions, or misalignments are present, the deviation from symmetry is often moderate. By applying these constraints as soft priors rather than strict rules, the network retains the flexibility to generate slightly asymmetric kernels when necessary.

We substantiate the claim that the symmetry and smoothness constraints act as soft priors through both network design and empirical evidence. These constraints are incorporated into the loss function as regularization terms with tunable weights, rather than being hard-coded into the model architecture. This allows the network to retain flexibility and generate slightly asymmetric or non-ideal kernels.

This principled design makes our method particularly suitable for self-supervised blind deblurring, where no ground truth kernel is available, and prior knowledge of the blur structure plays a critical role in convergence and accuracy.

Total objective function. The overall loss function used to train our blind deblurring framework combines four components: a reconstruction loss, a TV regularization on the predicted image, and two structural priors on the predicted blur kernel. The full objective is defined as:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{rec}} + \lambda_{\text{TV}} \cdot \mathcal{L}_{\text{TV}} + \lambda_{\text{sym}} \cdot \mathcal{L}_{\text{sym}} + \lambda_{\text{smo}} \cdot \mathcal{L}_{\text{smo}}. \tag{5}$$

The scalar weights $\lambda_{\text{TV}}, \lambda_{\text{sym}}, \lambda_{\text{smo}}$ balance the contribution of each regularization term relative to the main reconstruction objective. $\mathcal{L}_{\text{rec}}$ ensures that the generated sharp image $\hat{I}_y$ and predicted kernel $\hat{k}_y$ produce a convolution result consistent with the observed blurred image I_B. $\mathcal{L}_{\text{TV}}$ is a TV loss applied to $\hat{I}_y$ to suppress noise and encourage spatial smoothness. $\mathcal{L}_{\text{sym}}$ is a rotational symmetry loss that enforces approximate radial symmetry in $\hat{k}_y$. $\mathcal{L}_{\text{smo}}$ penalizes large gradients in $\hat{k}_y$ to encourage smooth kernel structures.

These four loss components are jointly optimized to update both the image generator G_I and the kernel generator G_k. This joint supervision ensures that the predicted clean image and blur kernel are not only consistent with the observed data, but also structurally and physically plausible.

2.3 Experimental Design

Based on the OpenFlexure microscope platform,[22] we constructed a portable microscopy imaging system as shown in Fig. 1 (c), with its optical path diagram presented in Fig. 1 (b). The core optical components include a light source, an objective lens, a tube lens, and a sensor. We conducted experiments on a variety of samples, including resolution test charts, pumpkin stem tissue sections, prepared slides of rotifers, and live rotifers suspended in water. For each sample, slight defocus was manually introduced to produce realistically blurred images that simulate typical defocus conditions in biological microscopy.

Each acquired image is used to train an individual, image specific model. During training, we randomly initialize two independent input tensors of the same spatial resolution with output. These random noise maps, remain fixed during training. The networks G_I and G_k are initialized with random weights and are trained from scratch for each image. They are jointly optimized using the full loss function in Eq. 5. Upon convergence, the generated latent clean image and estimated blur kernel are regarded as the final outputs of our framework. This self-supervised framework enables us to recover a high-quality sharp image and a physically meaningful blur kernel without requiring any external datasets or ground truth references.[23]

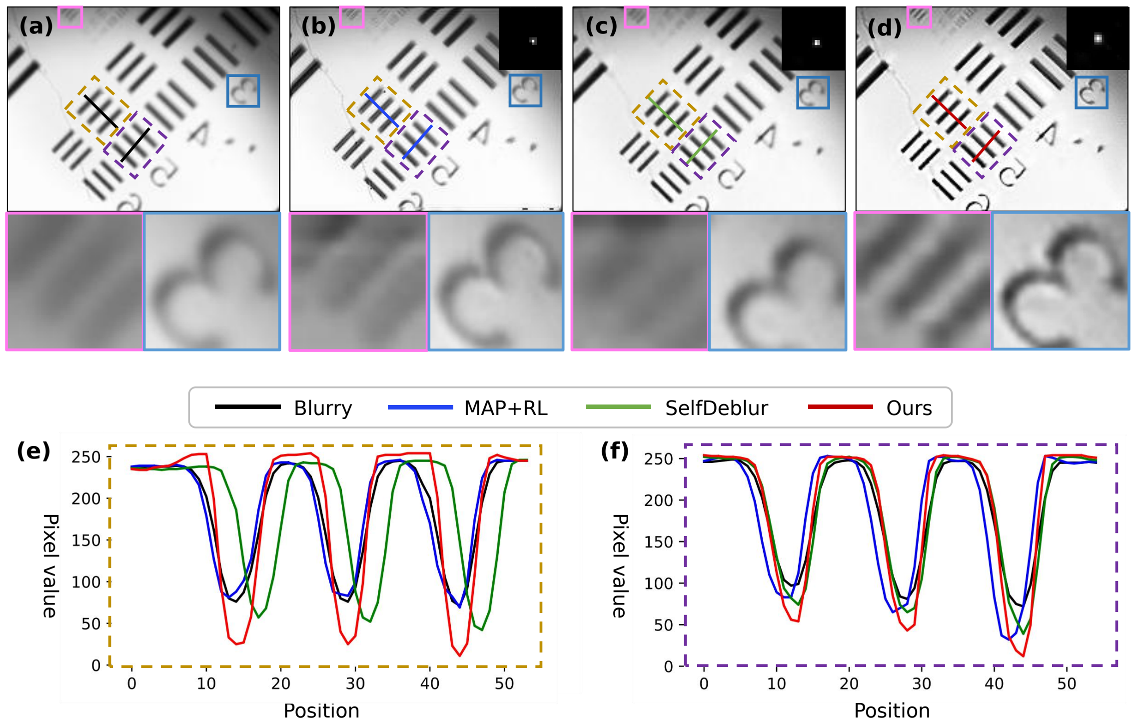

Figure 2. Deblurring results on the resolution chart. (a) Original defocus-blurred image. (b) MAP-based kernel estimation followed by RL deconvolution. (c) Result from our baseline SelfDeblur model without symmetry or smoothness constraints. (d) Result from our full model with both symmetry and smoothness constraints. Restored images and corresponding estimated kernels (top-right) are shown. (e)(f) Pixel intensity profiles along the two marked directions in the main images.

3. RESULTS AND DISCUSSION

Fig. 2 presents the deblurring results on the resolution chart using different methods. Fig. 2(a) shows the original defocus image acquired from the microscope. Fig. 2(b) displays the result obtained using a traditional MAP kernel estimation method, followed by Richardson-Lucy (RL) deconvolution[24] to recover the clean image and the estimated blur kernel (shown in the top-right inset). Fig. 2(c) shows the result of SelfDeblur model,[16] baseline of ours, which does not incorporate the proposed symmetry and smoothness constraints. Fig. 2(d) illustrates the output of our model with symmetry and smoothness regularization applied.

When comparing the estimated blur kernels, our method causes the kernel to be closer to the center, with a more gradual and stable distribution, reducing abrupt changes. The recovered latent clean image exhibits finer structures and significantly improved clarity. For each image, a zoom-in view of a local region is shown below to better visualize the fine structural details. It is observed that our method recovers significantly sharper details and higher local contrast compared to the other approaches.

Fig. 2(e) and Fig. 2(f) further plot the pixel intensity profiles along the two lines marked in the main images, corresponding to different directions. In Fig. 2(e), our method clearly produces steeper gradients and higher contrast compared to the original blurred image. Although the SelfDeblur method also enhances contrast, its intensity profile is noticeably shifted, which is likely due to positional inaccuracies in the estimated blur kernel. In Fig. 2(f), along the orthogonal direction, our method again achieves the most distinct and consistent contrast enhancement.

Since no ground truth images are available, we evaluate the quality of the restored images using several no-reference image sharpness metrics. The *Brenner focus measure* is a simple yet effective indicator of sharpness, calculated using the second-order difference of pixel intensities, typically along the vertical direction. This value

Table 1. Comparison of different deblurring methods using clarity evaluation metrics on the resolution chart.

Metric	Origin	MAP+RL	SelfDeblur	Ours
Brenner focus measure	98.21	163.29	176.12	**404.22**
Tenegrad gradient energy	0.30	0.33	0.35	**0.43**
Laplacian energy	1054.94	2173.32	1815.92	**3381.47**
High-frequency FFT energy	480512.00	1008768.00	847744.00	**1900608.00**

increases with stronger edge contrast and is commonly used to detect defocus blur.[25] The *Tenegrad gradient energy* metric is computed by applying the Sobel operator to extract local gradient magnitudes, and then summing the energy of pixels that exceed a predefined threshold.[26] This approach takes into account both the strength and spatial extent of edge structures, making it well-suited for evaluating image sharpness. The *Laplacian energy* metric is based on the second derivative of image intensity and is sensitive to abrupt changes in pixel values.[27] It provides a measure of fine detail and texture richness by emphasizing high-frequency components. Lastly, the *FFT-based sharpness metric* involves computing the Fourier transform of the image, isolating the high-frequency components corresponding to edges and textures, and summing their total energy.[28] A higher value in this metric reflects the presence of more detailed structures within the image. These metrics allow for quantitative comparison of deblurring performance across different methods, even in the absence of reference images.

Table 1 summarizes the evaluation results for different deblurring methods (corresponding to the outputs shown in Fig. 2). Across all metrics, our method consistently achieves the highest scores, indicating superior restoration quality in terms of detail preservation and sharpness.

These results demonstrate that incorporating symmetry and smoothness constraints during the unsupervised estimation of the blur kernel and latent image helps guide the network to generate more natural and physically plausible kernels, which in turn leads to more accurate and visually sharp restorations.

4. CONCLUSION

In this study, we propose a self-supervised blind deblurring method tailored for defocus degradation in microscopic imaging. The framework employs two lightweight neural networks to jointly estimate a latent sharp image and the corresponding blur kernel, without requiring any ground truth labels or pre-training datasets. To enhance the accuracy of kernel estimation, we incorporate prior knowledge derived from optical imaging principles into the loss function, specifically rotational symmetry and spatial smoothness constraints. These physically motivated priors encourage the generation of more natural and structurally consistent blur kernels, thereby improving the quality of image restoration.

Experiments on various microscopic samples demonstrate that, compared to traditional MAP-based methods and self-supervised approaches without kernel constraints, our method achieves superior performance in both visual quality and no-reference sharpness metrics. The introduction of kernel constraints not only improves the accuracy of blur kernel estimation but also significantly enhances the restoration of fine image details. Therefore, our approach demonstrates promising application prospects in biomedical and other imaging domains.

5. ACKNOWLEDGMENT

The work is supported by the Research Grants Council of Hong Kong (RIF R7003-21).

REFERENCES

[1] Karhana, S., Bhat, M., Ninawe, A., and Dinda, A. K., "Advances in microscopy and their applications in biomedical research," in [*Biomedical Imaging Instrumentation*], 185–212, Elsevier (2022).

[2] Li, Y., Chiu, P. W., Tam, V., Lee, A., and Lam, E. Y., "Dual-mode Imaging System for Early Detection and Monitoring of Ocular Surface Diseases," *IEEE Transactions on Biomedical Circuits and Systems* **18**(4), 783–798 (2024).

[3] Agero, U., Mesquita, L., Neves, B., Gazzinelli, R., and Mesquita, O., "Defocusing microscopy," *Microscopy Research and Technique* **65**(3), 159–165 (2004).

[4] Ge, Z., Wei, H., Xu, F., Gao, Y., Chu, Z., So, H. K.-H., and Lam, E. Y., "Millisecond autofocusing microscopy using neuromorphic event sensing," *Optics and Lasers in Engineering* **160**, 107247 (2023).

[5] Ren, Z., Guan, P., Lam, E. Y., and Zhao, J., "Extended focused imaging in microscopy using structure tensor and guided filtering," *Optics and Lasers in Engineering* **140**, 106549 (2021).

[6] Lam, E. Y., "Artificial intelligence in biophotonics and imaging: Advancing computational reconstruction and inference," in [*Optoelectronics and Communications Conference*], T4A–2, Optica Publishing Group (2021).

[7] Zhu, S., Wang, C., Liu, H., Zhang, P., and Lam, E. Y., "Computational neuromorphic imaging: principles and applications," in [*Computational Optical Imaging and Artificial Intelligence in Biomedical Sciences*], **12857**, 4–10, SPIE (2024).

[8] Zhang, P., Zhu, S., Wang, C., Zhao, Y., and Lam, E. Y., "Neuromorphic imaging with super-resolution," *IEEE Transactions on Circuits and Systems for Video Technology* **35**(2), 1715–1727 (2025).

[9] Wang, C., Zhu, S., Zhang, P., Wang, K., Huang, J., and Lam, E. Y., "Angle-based neuromorphic wave normal sensing," *Laser & Photonics Reviews* **19**(4), 2400647 (2025).

[10] Wang, K., Song, L., Wang, C., Ren, Z., Zhao, G., Dou, J., Di, J., Barbastathis, G., Zhou, R., Zhao, J., et al., "On the use of deep learning for phase recovery," *Light: Science & Applications* **13**(1), 4 (2024).

[11] Chan, T. F. and Wong, C.-K., "Total variation blind deconvolution," *IEEE Transactions on Image Processing* **7**(3), 370–375 (1998).

[12] Xu, Z. and Lam, E. Y., "Maximum a posteriori blind image deconvolution with Huber-Markov random-field regularization," *Optics Letters* **34**, 1453–1455 (May 2009).

[13] Mei, J., Wu, Z., Chen, X., Qiao, Y., Ding, H., and Jiang, X., "Deepdeblur: text image recovery from blur to sharp," *Multimedia Tools and Applications* **78**, 18869–18885 (2019).

[14] Schuler, C. J., Hirsch, M., Harmeling, S., and Schölkopf, B., "Learning to deblur," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **38**(7), 1439–1451 (2015).

[15] Ren, Z., Xu, Z., and Lam, E. Y., "Learning-based nonparametric autofocusing for digital holography," *Optica* **5**(4), 337–344 (2018).

[16] Ren, D., Zhang, K., Wang, Q., Hu, Q., and Zuo, W., "Neural blind deconvolution using deep priors," in [*Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*], 3341–3350 (2020).

[17] Ulyanov, D., Vedaldi, A., and Lempitsky, V., "Deep image prior," in [*Proceedings of the IEEE conference on computer vision and pattern recognition*], 9446–9454 (2018).

[18] Ren, Z., Xu, Z., and Lam, E. Y., "End-to-end deep learning framework for digital holographic reconstruction," *Advanced Photonics* **1**(1), 016004 (2019).

[19] Xu, R., Ou, H., and Lam, E. Y., "Defocus and speckle noise suppression in optical scanning holography," *Optics & Laser Technology* **190**, 113246 (2025).

[20] Li, T., Chen, H., Zhang, M., Liu, S., Xia, S., Cao, X., Young, G. S., and Xu, X., "A new design in iterative image deblurring for improved robustness and performance," *Pattern Recognition* **90**, 134–146 (2019).

[21] Lee, H.-C., "Review of image-blur models in a photographic system using the principles of optics," *Optical Engineering* **29**(5), 405–421 (1990).

[22] Collins, J. T., Knapper, J., Stirling, J., Mduda, J., Mkindi, C., Mayagaya, V., Mwakajinga, G. A., Nyakyi, P. T., Sanga, V. L., Carbery, D., et al., "Robotic microscopy for everyone: the OpenFlexure microscope," *Biomedical Optics Express* **11**(5), 2447–2460 (2020).

[23] Liu, H., Zhang, Y., and Lam, E. Y., "Photon-limited imaging with quanta image sensors via an unsupervised learning framework," in [*2024 IEEE 34th International Workshop on Machine Learning for Signal Processing (MLSP)*], 1–6, IEEE (2024).

[24] Ingaramo, M., York, A. G., Hoogendoorn, E., Postma, M., Shroff, H., and Patterson, G. H., "Richardson–Lucy deconvolution as a general tool for combining images with complementary strengths," *ChemPhysChem* **15**(4), 794–800 (2014).

[25] Brenner, S., "Multigrid methods for the computation of singular solutions and stress intensity factors I: Corner singularities," *Mathematics of Computation* **68**(226), 559–583 (1999).

[26] Hu, J., Wei, Y., Guo, E., and Han, J., "Wavefront shaping for scattering imaging based on no-reference image quality assessment," in [*Nineteenth National Conference on Laser Technology and Optoelectronics*], **13287**, 164–171, SPIE (2024).

[27] Laparra, V., Ballé, J., Berardino, A., and Simoncelli, E. P., "Perceptual image quality assessment using a normalized Laplacian pyramid," *Electronic Imaging* **28**, 1–6 (2016).

[28] Chen, G. Y. and Krzyzak, A., "An FFT-based visual quality metric robust to spatial shift," *International Journal of Wavelets, Multiresolution and Information Processing* **19**(06), 2150024 (2021).