

PHOTONICS Research

Deep learning-based optical aberration estimation enables offline digital adaptive optics and super-resolution imaging

CHANG QIAO,^{1,2,†}  HAoyu CHEN,^{3,4,†} RUN WANG,^{1,†} TAO JIANG,^{3,4} YUWANG WANG,^{5,6} AND DONG LI^{3,4,*}

¹Department of Automation, Tsinghua University, Beijing 100084, China

²Institute for Brain and Cognitive Sciences, Tsinghua University, Beijing 100084, China

³National Laboratory of Biomacromolecules, New Cornerstone Science Laboratory, CAS Center for Excellence in Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, Beijing 100101, China

⁴College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

⁵Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China

⁶e-mail: wang-yuwang@mail.tsinghua.edu.cn

[†]These authors contributed equally to this work.

*Corresponding author: lidong@ibp.ac.cn

Received 27 September 2023; revised 18 December 2023; accepted 20 December 2023; posted 3 January 2024 (Doc. ID 506778); published 1 March 2024

Optical aberrations degrade the performance of fluorescence microscopy. Conventional adaptive optics (AO) leverages specific devices, such as the Shack–Hartmann wavefront sensor and deformable mirror, to measure and correct optical aberrations. However, conventional AO requires either additional hardware or a more complicated imaging procedure, resulting in higher cost or a lower acquisition speed. In this study, we proposed a novel space-frequency encoding network (SFE-Net) that can directly estimate the aberrated point spread functions (PSFs) from biological images, enabling fast optical aberration estimation with high accuracy without engaging extra optics and image acquisition. We showed that with the estimated PSFs, the optical aberration can be computationally removed by the deconvolution algorithm. Furthermore, to fully exploit the benefits of SFE-Net, we incorporated the estimated PSF with neural network architecture design to devise an aberration-aware deep-learning super-resolution model, dubbed SFT-DFCAN. We demonstrated that the combination of SFE-Net and SFT-DFCAN enables instant digital AO and optical aberration-aware super-resolution reconstruction for live-cell imaging. © 2024 Chinese Laser Press

<https://doi.org/10.1364/PRJ.506778>

1. INTRODUCTION

Fluorescence microscopy has been widely used as a powerful tool for visualizing various biological structures and bioprocesses in fixed or live specimens. Ideal optical imaging relies upon the high-quality focusing of excitation light and accurate detection of the emission light from the fluorescent sample. However, both the optics in the microscope and the biological samples being investigated can introduce aberrations, thus causing degradation in resolution, loss of fluorescent photons, and deterioration of signal-to-background ratio, etc. For example, the optics manufacture deficiency or misalignment of optical elements in the imaging system may cause certain aberrations such as spherical and coma aberration, and the refractive index inhomogeneities of biological specimens will bring about more complicated aberrations. Moreover, microscopes with high numerical apertures (NAs), especially the super-resolution microscope, are more sensitive to aberrations, because the

high-NA objectives are more susceptible to high-order aberrations [1]. To detect and correct these optical aberrations, a large number of adaptive optics (AO) technologies have been explored in the last two decades [2].

The implementation of AO generally involves two main components: aberration detection and aberration correction. To measure optics- or sample-induced aberrations, both direct and indirect wavefront sensing methods were developed [1–3]. Direct wavefront sensing methods utilize a dedicated wavefront sensor, mostly the Shack–Hartmann sensor, along with an additional light path for aberration detection. In contrast, the indirect wavefront sensing methods characterize aberrations without specific wavefront sensors but determine them computationally from repetitive acquisitions with either zonal or modal approaches [2]. In recent years, deep neural networks have been applied to directly estimate aberrations from the optical images of point sources [4–6]. However, these methods

are limited to the scenarios where there are guiding stars or single-molecule emitters in the biological sample. Once the aberrations are known, wavefront corrective devices, mostly the spatial light modulators and deformable mirrors, are utilized to compensate for the measured aberrations by reshaping the wavefronts [1–3]. In consequence, conventional AO methods have to rely on additional optical devices or iterative acquisitions to measure and then eliminate the optical aberration, which complicates the optics, imaging procedures, and computation. To overcome these limitations, the development of digital adaptive optics has allowed for the computational detection and correction of optical aberrations for light-field microscopy (LFM) [7,8] in an offline manner, which, however, is only applicable for the certain imaging modality, i.e., LFM.

In optical imaging systems, the image quality as well as the aberrations is typically characterized by their point spread functions (PSFs), which are implicitly encoded in any specimen patch of the microscopic image. Inspired by the understanding, we devised a space-frequency encoding network (SFE-Net), which is trained to directly extract the PSF with aberrations from a single microscopic image. Our results show that the proposed SFE-Net is able to estimate optical aberrations composed of up to 18 Zernike polynomials with high accuracy directly from images of various biological specimens, and the corresponding aberrations can be substantially eliminated via the deconvolution algorithm resorting to the estimated PSF. To further enhance the resolution while removing the optical aberrations for biological images, we integrated the PSF priors into the deep-learning super-resolution (DLSR) neural network architecture design and devised the spatial feature transform-guided (SFT) deep Fourier channel attention network (SFT-DFCAN). We showed that by leveraging PSF information estimated from SFE-Net, the SFT-DFCAN can be trained to digitally eliminate the aberrations and super-resolve the fine structures of specimens directly from the aberrated images, which substantially outperforms its backbone DFCAN architecture [9]. Finally, we demonstrated that the SFE-Net and SFT-DFCAN enable fast, accurate aberration estimation and correction, as well as computational super-resolution, in long-term live-cell imaging experiments.

2. METHODS

A. Training Data Generation

The training data for SFE-Net, SFT-DFCAN, and other deep-learning models compared in this study was generated in a semi-synthetic manner using our previously published dataset BioSR [9]. Specifically, we utilized the ground-truth structured illumination microscopy (GT-SIM) images from BioSR as the biological fluorescence specimens. These images were intentionally degraded according to the optical imaging model, which can be expressed as follows:

$$I = N_{\text{Poisson}}(S * \text{PSF}_{\text{Zernike}}) + G(0, \sigma^2), \quad (1)$$

where I represents the aberrated wide-field (WF) image captured by the optical imaging system; S denotes the biological specimens, i.e., the GT-SIM images; $*$ signifies the convolution operator; $N_{\text{Poisson}}(\cdot)$ represents the Poisson recorrution; $G(0, \sigma^2)$ denotes the Gaussian white noise with a mean of zero

and a variance of σ^2 ; and $\text{PSF}_{\text{Zernike}}$ refers to the aberrated point spread function, whose pupil function is constructed by a weighted summation of Zernike polynomials 4–18 (Wyant ordering). These functions can be mathematically formulated as follows:

$$\text{PSF}_{\text{Zernike}} = \mathcal{F}^{-1} \left\{ A \left(\sum_{n=4}^{18} a_n Z_n \right) \right\}, \quad (2)$$

where Z_n and a_n represent Zernike polynomials and the coefficient of order n , respectively. $A(\cdot)$ denotes a circular apodization function, where the radius is determined by the emission wavelength and detection numerical aperture (NA). $\mathcal{F}^{-1}\{\cdot\}$ denotes the inverse fast Fourier transformation operator.

The coefficient a_n for each order was randomly sampled from a normal distribution with a zero mean and a standard deviation of 0.125. And we set an upper and lower bound of $[-1, 1]$ for all sampled a_n to avoid extremely high or low values that may destabilize the training process. Since a_n obeys a normal distribution $\mathcal{N}(0, 0.125)$ with a loose bound, the root mean square (RMS) of the generated aberration approximately follows a rescaled chi distribution, which can be formulated as

$$\text{RMS} = \sqrt{\sum_{n=4}^{18} a_n^2 / \lambda}, \quad (3)$$

where λ is the emission wavelength. As a result, the total RMS range of the training and testing datasets is $[0, 7.38\lambda]$.

During the training procedure, the aberrated PSF images $\text{PSF}_{\text{Zernike}}$ were used as targets for PSF estimation network models such as SFE-Net, and the GT-SIM images S were used as targets for single image super-resolution (SISR) network models such as SFT-DFCAN.

B. Network Architecture

The architecture of SFE-Net, as shown in Fig. 1(a), consists of a dual-branch encoder and a U-net-based decoder. The encoder network is constituted with two parallel branches: the spatial branch (SB) and frequential branch (FB), which extract deep features in the spatial and frequential domains, respectively. In both branches, a modified residual channel attention network [10] with 4 residual groups \times 4 residual channel attention blocks is employed as their backbone network architecture [Fig. 1(b)]. In contrast with the SB, the FB begins with a fast Fourier transform layer followed by a modulus operator and a logarithm operator in sequence, so as to encode the image feature into the Fourier domain. The output feature maps of the SB and FB are concatenated along the channel and then fed into the U-net-based decoder.

The decoder mainly consists of two parts: a U-net feature extractor and a downscale module. We adopt a relatively deep U-net model [11] that begins with a double convolutional block [Fig. 1(c)] followed by five downscale blocks [Fig. 1(d)] and five upscale blocks [Fig. 1(e)] with five skip connections bridging the features of the same scale. In each downscale block, a max pooling layer and a double convolutional block are employed to downscale and extract features. The upscale blocks use the pixel shuffle layer to upscale the feature channels. The output of the U-net is then passed to the downscale

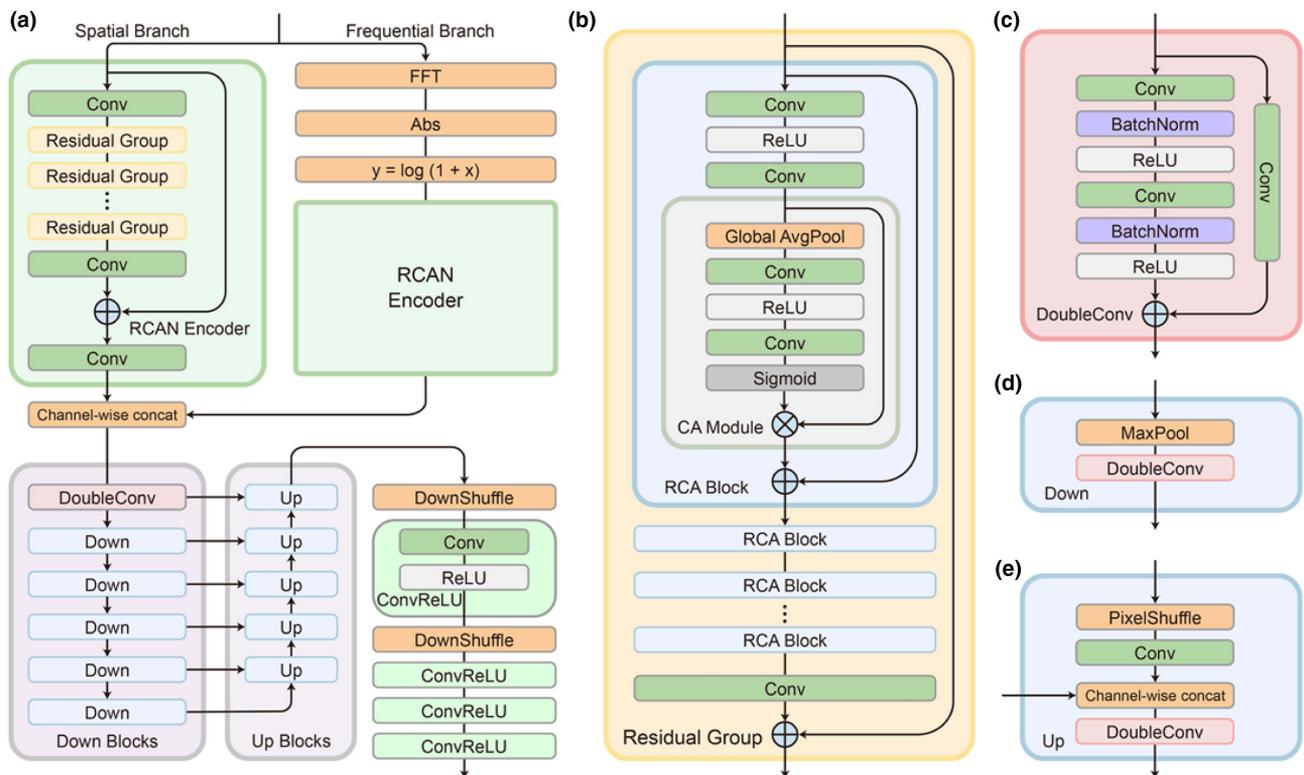


Fig. 1. (a)–(e) Network architecture of space-frequency encoding network. (a) Network architecture of the SFE-Net, (b) residual group, (c) double convolutional block, (d) downscale block, and (e) upscale block.

module, which consists of two $2 \times$ down shuffle layers and four Conv-ReLU blocks. In each Conv-ReLU block, the stride parameter of convolutional layers is set to 2, enabling the downscale module to transform the input feature maps from 132×132 pixels into PSF images of 33×33 pixels.

The overall architecture of SFT-DFCAN is depicted in Fig. 2(a), which is modified from our previously proposed state-of-the-art DLSR model DFCAN [9], which is trained to directly transform a WF image to its SR counterpart. Here, to deliver PSF and aberration information to the image

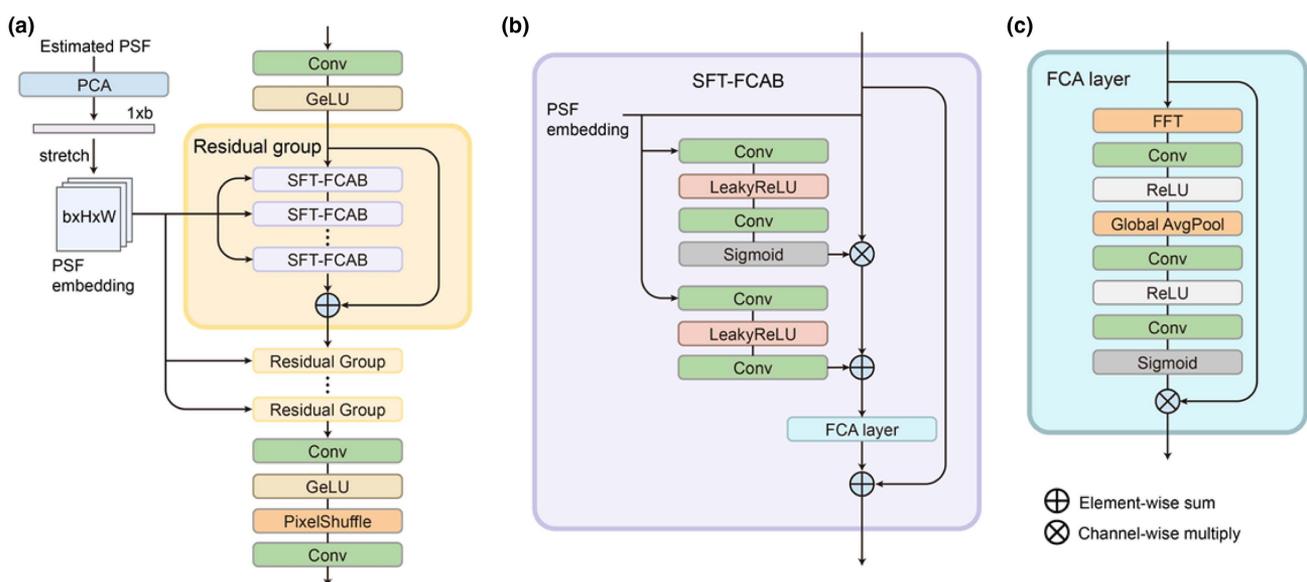


Fig. 2. (a)–(c) Network architecture of spatial feature transform-guided deep Fourier channel attention network (SFT-DFCAN). (a) Network architecture of the SFT-DFCAN, (b) spatial feature transform-guided Fourier channel attention block (FCAB), and (c) Fourier channel attention (FCA) layer.

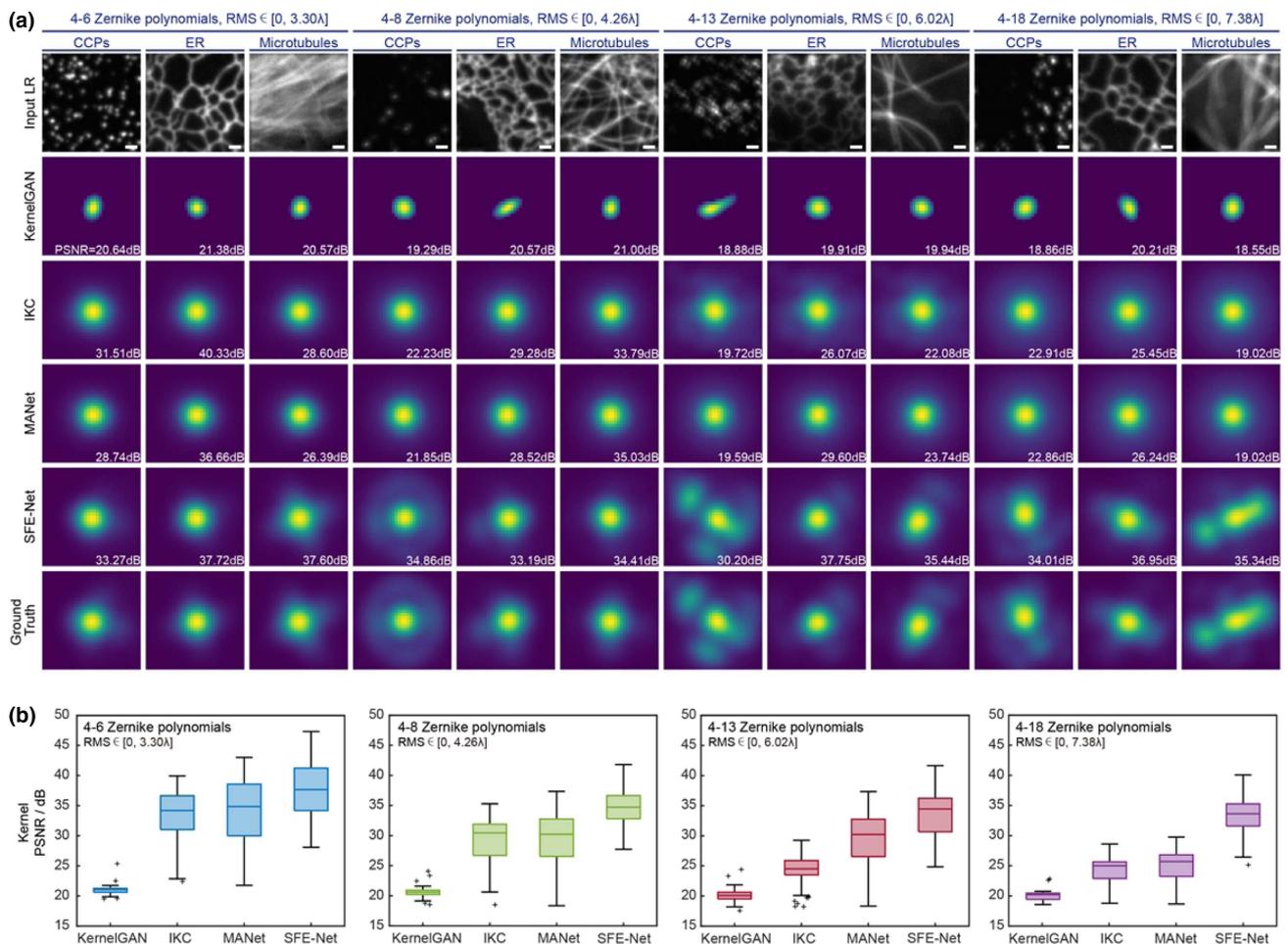


Fig. 4. Optical aberration estimation via SFE-Net. (a) Representative aberrated PSFs estimated by KernelGAN, IKC, MANet, and SFE-Net from WF images of CCPs, ER, and MTs. Four groups of datasets with escalating complexity of aberration were generated, corresponding to Zernike polynomials of orders 4–6, 4–8, 4–13, and 4–18. The top and bottom rows show the input WF images and GT PSF images for reference. Scale bar, 1 μm . (b) Statistical comparisons ($n = 30$) of KernelGAN, IKC, MANet, and SFE-Net in terms of peak signal-to-noise ratio (PSNR) on different training and testing datasets. Center line, medians; limits, 75% and 25%; whiskers, the larger value between the largest data point and the 75th percentiles plus $1.5 \times$ the interquartile range (IQR), and the smaller value between the smallest data point and the 25th percentiles minus $1.5 \times$ the IQR; outliers, data points larger than the upper whisker or smaller than the lower whisker. The same notations for box plots are used in Figs. 6(e) and 7(b).

unsupervised kernel generative adversarial network (KernelGAN) [15,18], iterative kernel correction (IKC) [12], and supervised mutual affine network (MANet) for spatially variant kernel estimation [17]. To evaluate the performance of these methods, we generated four semi-simulated datasets with aberrated PSF following the procedure outlined in Section 2.A, each constituted by different orders of Zernike polynomials (4–6, 4–8, 4–13, and 4–18). The increasing order range reflects the increasing severity of the ill-posedness in the PSF estimation task. The generated datasets were then utilized to evaluate the performance of existing kernel estimation methods. The results, depicted in Fig. 4(a), revealed that the KernelGAN method only generated narrowed anisotropic kernels that significantly deviated from the GT PSF in terms of the shape and size. This discrepancy may arise from the multiple kernel constraints in the algorithm, despite the fact that we have made great efforts to optimize the weighting scalar of each regularization term. In contrast to KernelGAN, both the

IKC and MANet methods consistently produce Gaussian-shaped kernels, regardless of the complexity of the training dataset or biological structures. This indicates that these two methods fail to resolve the optical aberrations from the corresponding WF images. In particular, even when we modified the MANet to focus on estimating a spatially consistent kernel, it still could not generate the correct PSF with notable aberrations, possibly due to its relatively simple network architecture.

To further enhance the feature extraction and representation capability of neural network models in the task of PSF estimation, we devised a novel neural network architecture named SFE-Net. SFE-Net leverages both the spatial features and frequency characteristics of the WF image to estimate the aberrated PSF with high accuracy. As shown in Fig. 4, the SFE-Net is trained in a supervised manner to directly map biological WF images to their corresponding aberrated PSFs. Interestingly, before adopting this supervised training scheme, we have gone

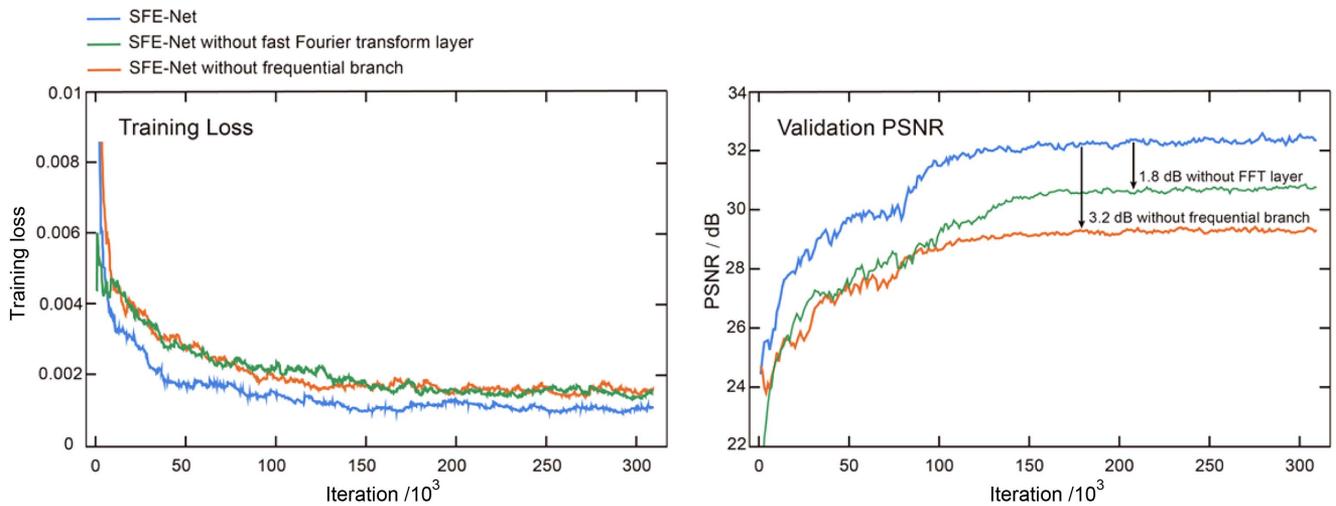


Fig. 5. Progression of training loss and validation PSNR of network model with/without the FFT layer and frequential branch during training process.

through a series of physical model-based PSF estimation approaches, such as using untrained network [19], or modified flow-based kernel priors [20]. However, we found that although straightforward, the data-driven supervised mapping strategy with SFE-Net remarkably outperformed other conceptually more complex ideas.

As is shown in Fig. 4(a), KernelGAN, IKC, and MANet fail to extract the optical aberration from WF images, even when the aberrated PSF is relatively simple, i.e., generated with 4–6 Zernike polynomials. In contrast, our SFE-Net accurately generates complex aberrated PSF constituted of up to 18 orders of Zernike polynomials, with an average peak signal-to-noise ratio (PSNR) higher than 30 dB [Fig. 4(b)]. Furthermore, we performed an ablation study on the frequential branch of SFE-Net to validate the gain of incorporating frequential information in the feature extraction process of the PSF estimation network. Specifically, we trained three versions of the SFE-Net on the same training dataset: a standard SFE-Net, a modified version without the frequential branch, and another modified version without the fast Fourier transform (FFT) layer in the frequential branch. The training loss and validation PSNR curves for these three models shown in Fig. 5 demonstrate that the inclusion of the frequential branch, especially the FFT layer, effectively accelerates network convergence and contributes to an improved PSF estimation performance by 3.2 dB in PSNR.

B. Blind Deconvolution with Accurate PSF Estimation

Integrated with deconvolution algorithms, the accurate PSF estimation through SFE-Net provides a straightforward yet efficient solution for numerically compensating optical aberrations and improving spatial resolution in an unsupervised manner. To systematically evaluate the impact of the PSF on deconvolution algorithms during the processing of images with optical aberrations, we processed WF images of CCPs, ER, and MTs with Richardson–Lucy (RL) deconvolution [21,22] using an ideal Gaussian PSF with theoretically accurate full width at half-maximum (FWHM), as well as PSFs estimated by

KernelGAN, IKC, MANet, and SFE-Net [Figs. 6(a)–6(c)]. Our findings demonstrate that RL deconvolution with accurate PSF, i.e., GT PSFs and PSFs estimated by SFE-Net, substantially removes the artifacts induced by aberrations, such as an anomalous distortion in CCP images and ringing artifacts in ER and MT images. Moreover, it enhances both the resolution and contrast for all biological structures. Conversely, when provided with an incorrect PSF estimated by other methods, the RL deconvolution algorithm fails to remove the aberration-induced artifacts and may even generate anamorphic structures [indicated by red arrows in Figs. 6(a)–6(c)].

We measured the PSNR between deconvolved images obtained using GT PSFs and those obtained using estimated PSFs, with deconvolution iterations ranging from 5 to 100, across various biological structures. Our results demonstrate that the deconvolved images produced using the PSF estimated by SFE-Net consistently exhibit a significantly higher PSNR compared to those generated using PSFs estimated by other existing methods. This holds true regardless of the number of deconvolution iterations [Fig. 6(d)] or the specific biological structure being analyzed [Fig. 6(e)].

C. Aberration-Aware Image Super-Resolution

SISR networks have been developed to instantly enhance the resolution of biological images in an end-to-end manner, irrespective of the image formation model [9]. Recent studies have shown that incorporating physical prior knowledge, such as the PSF, can improve the performance of the super-resolution network. Given the remarkable ability of the proposed SFE-Net to recognize the PSF from low-resolution images, we reasoned that incorporating the prior knowledge of PSF and optical aberrations could benefit the performance of SISR. To validate this hypothesis, we incorporated the SFT layer [12] with our previously proposed DFCA model [9] to devise the SFT-DFCAN, which leverages both the aberrated PSF information and Fourier channel attention mechanism to enhance the performance of image super-resolution. In particular, in SFT-DFCAN, the feature maps are affinely transformed

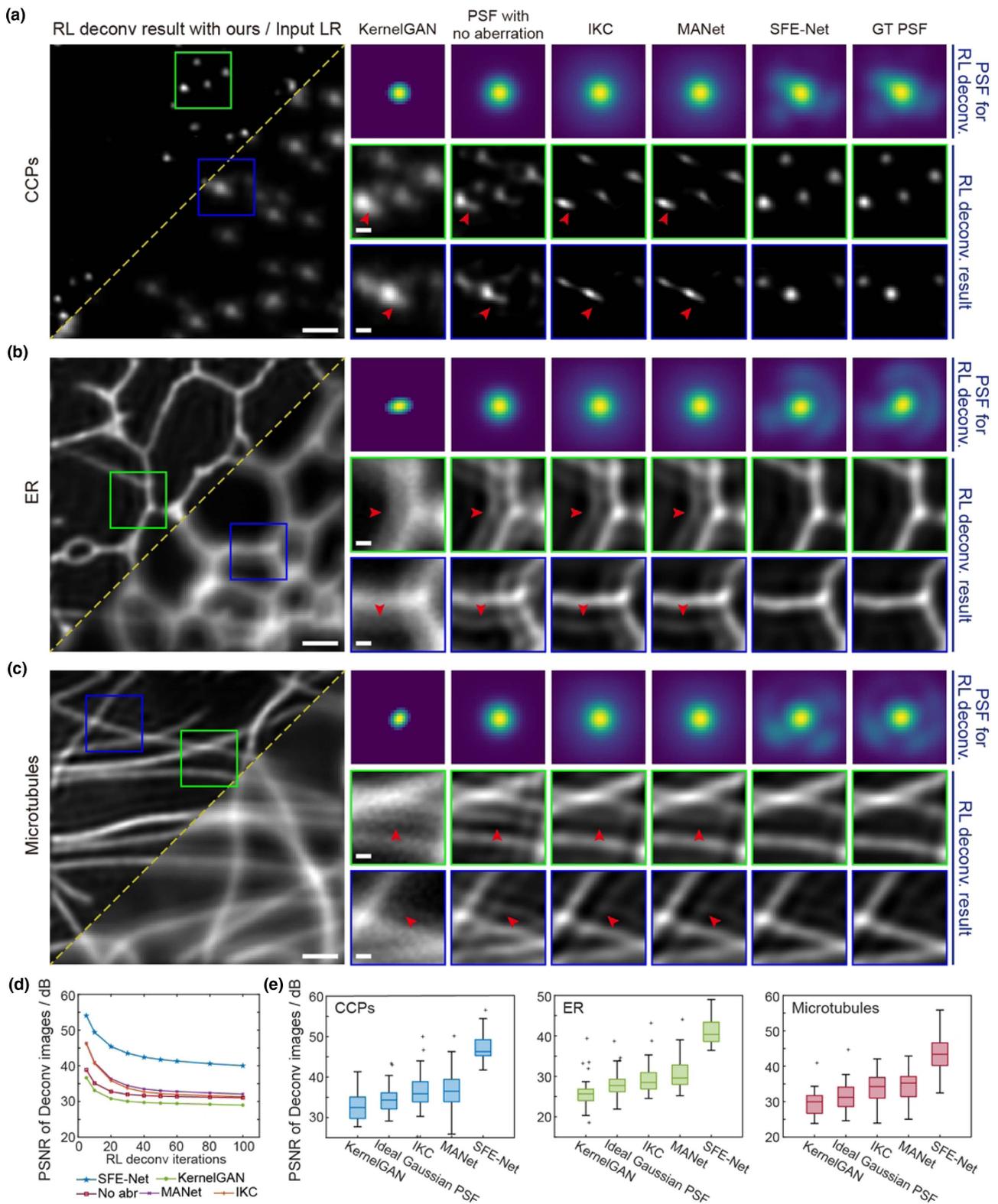


Fig. 6. (a)–(c) Blind deconvolution with the estimated PSF. Representative deconvolved images of (a) CCPs, (b) ER, and (c) MTs processed with the RL deconvolution algorithm using ideal Gaussian PSF and PSF estimated by KernelGAN, IKC, MANet, and SFE-Net. The aberrated WF images (bottom right in the first column), deconvolved images (top left in the first column), and GT PSF images are shown. (d) PSNR curves calculated between RL deconvolved images using GT PSF and estimated PSFs, with the deconvolution iteration ranging from 5 to 100 ($n = 120$). (e) Statistical comparisons of PSNR for testing datasets of CCPs (left), ER (middle), and MTs (right), respectively ($n = 30$). Scale bar, 1 μm [(a)–(c)], 0.25 μm [zoom-in regions of (a)–(c)].

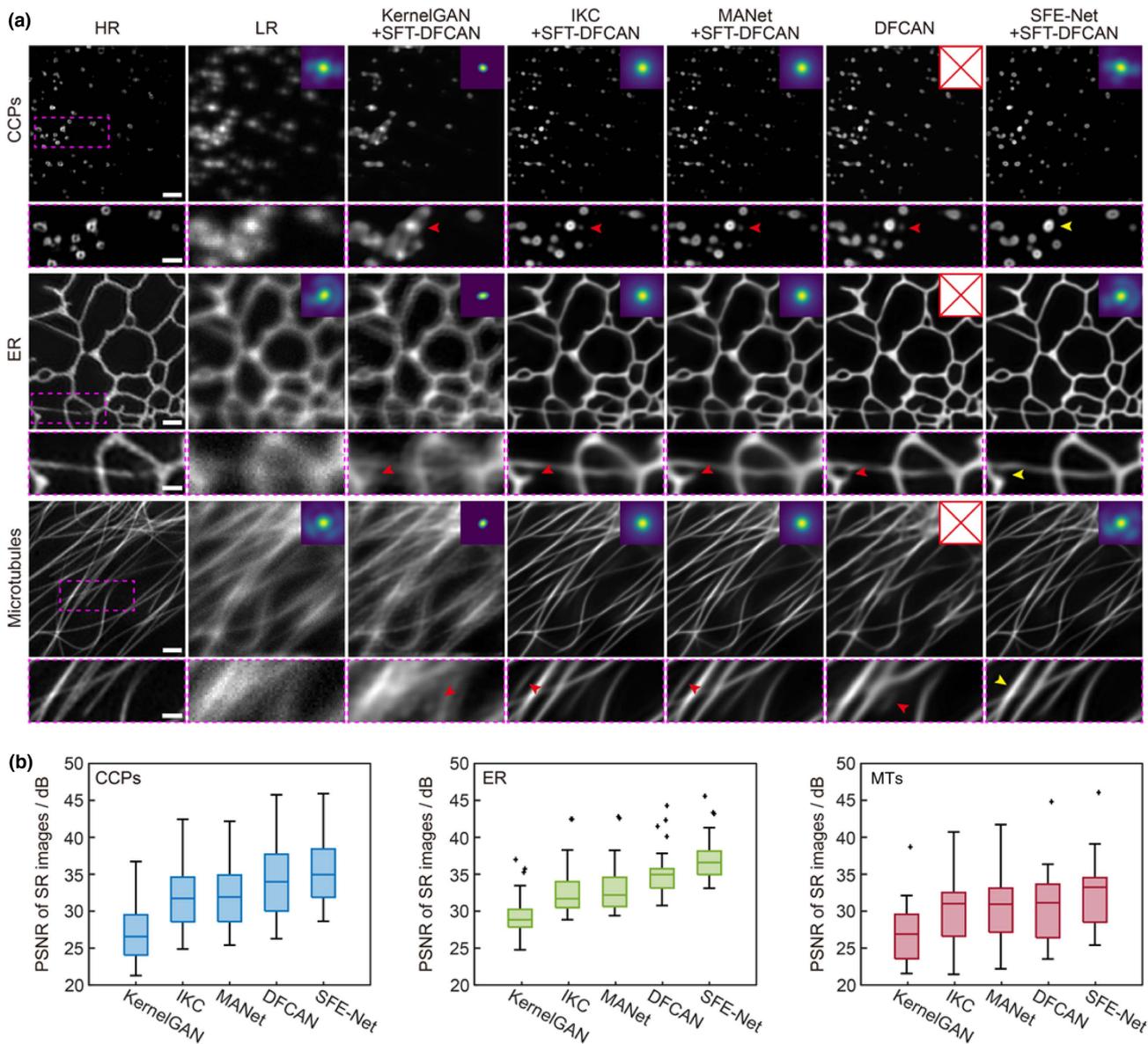


Fig. 7. Aberration-aware image super-resolution reconstruction with the estimated PSF. (a) Representative SR images reconstructed by DFCAN and SFT-DFCAN with PSFs obtained from KernelGAN, IKC, MANet, and SFE-Net. Low-resolution images and high-resolution GT images are provided for reference. The corresponding estimated PSF images are presented in the top right corner of each reconstructed SR image. Scale bar, 1 μm , and 0.5 μm (zoom-in regions). (b) Statistical comparison of PSNR values for the output SR images produced by DFCAN and SFT-DFCAN with PSFs estimated by KernelGAN, IKC, MANet, and SFE-Net ($n = 30$).

through scaling and shifting operations, which are conditioned on the estimated PSF and aberration, thereby enabling adaptive encoding of the PSF information into the neural networks.

Next, we validated the performance of SFT-DFCAN using the dataset generated following the steps outlined in Section 2.A. We trained an SFT-DFCAN model using pairs of low- and high-resolution images along with their corresponding GT PSFs, and a DFCAN model with low- and high-resolution image pairs for comparison. Figure 7(a) displays representative SR images reconstructed using the DFCAN and SFT-DFCAN models with PSFs generated by SFE-Net and other PSF estimation methods. These results show that while a well-trained DFCAN model can partially remove the optical aberration and

reconstruct high-frequency information, it struggles to capture the fine structure of biological specimens, often resulting in the generation of hallucinated structures [indicated by the red arrow in the sixth column of Fig. 7(a)]. In contrast, the SFT-DFCAN model, benefiting from the prior knowledge of PSF and estimated aberration, has the theoretical capability to recover biological structures with higher fidelity. Both of the qualitative and quantitative comparisons [Fig. 7(b)] between DFCAN and SFE-Net-guided SFT-DFCAN indicate that the incorporation of PSF and aberration information rationalizes the training and inference process of DFCAN models and provides substantial improvements in output fidelity and resolution.

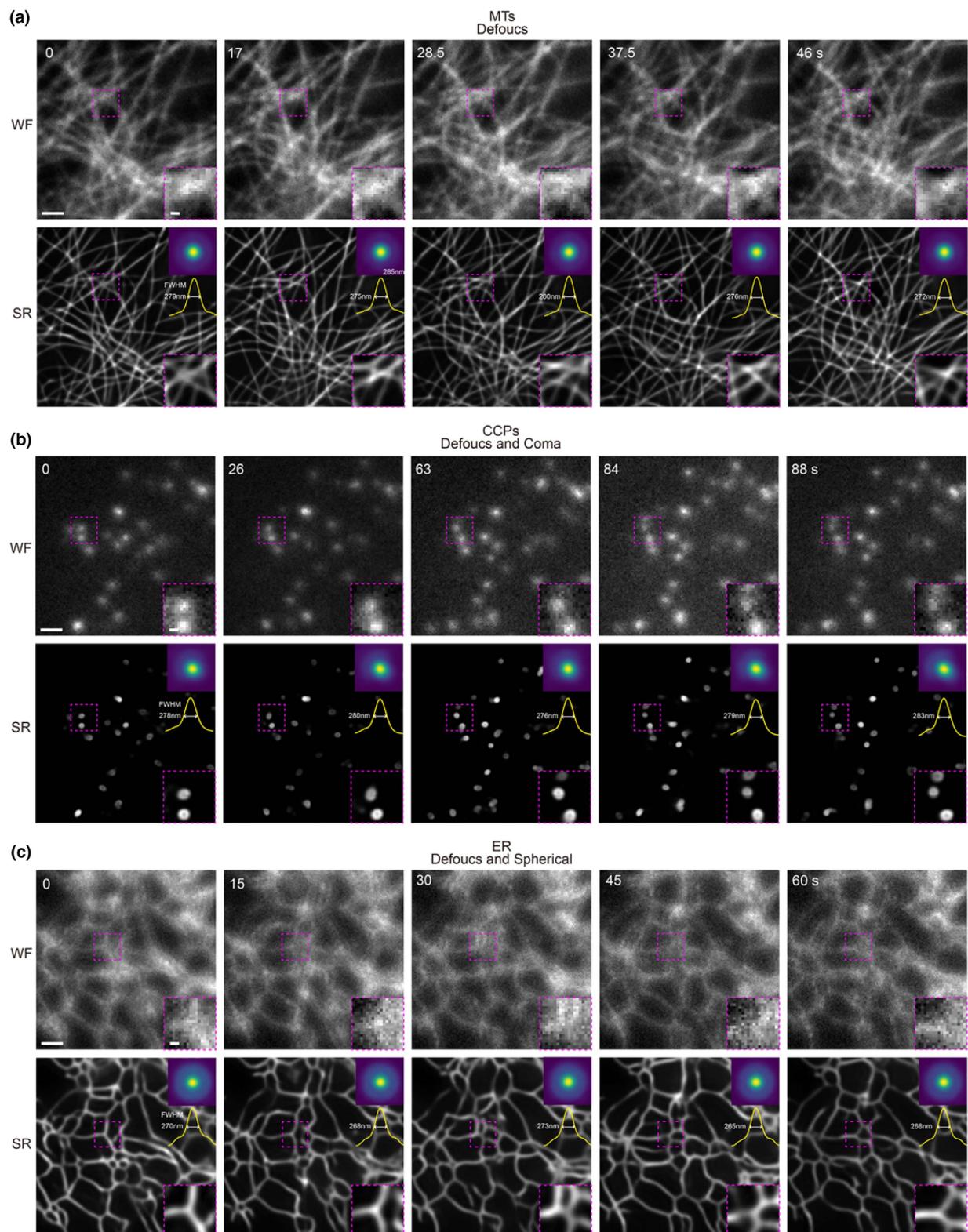


Fig. 8. Digital adaptive optics and super-resolution for live-cell imaging. Time-lapse WF images, estimated PSFs by SFE-Net, and corresponding SR images generated by SFE-Net-facilitated SFT-DFCAN of (a) MTs, (b) CCPs, and (c) ER. During the imaging procedure, the defocus aberration is manually added on MTs (a) data, while a combination of defocus and coma aberrations and a combination of defocus and spherical aberrations are applied on CCPs (b) and ER (c) images, respectively. The PSFs estimated by SFE-Net, along with their corresponding profiles and FWHM values, are displayed in the top right corner of SR images. Scale bar, 1 μm [(a)–(c)], and 0.2 μm [zoom-in regions of (a)–(c)].

On the other hand, existing PSF estimation methods such as KernelGAN, IKC, and MANet tend to produce PSF estimates that deviate significantly from the actual one, thereby misleading the high-frequency feature extraction and reconstruction in SFT-DFCAN models. Nevertheless, with the aberrated PSF estimated by SFE-Net, the SFT-DFCAN successfully recovers the fine structures of CCPs, ER, and MTs, exhibiting high consistency with the high-resolution GT images [Fig. 7(a)]. Additionally, the statistical comparison of the output SR images generated by different methods [Fig. 7(b)] demonstrates that SFE-Net-based SFT-DFCAN outperforms other PSF estimation methods-based SFT-DFCAN by a significant margin, enabling high-quality single-image SR reconstruction from aberrated WF images.

D. Digital Adaptive Optics and Super-Resolution for Live-Cell Imaging

For a well-established optical imaging system, the most common aberrations during live-cell imaging experiments are defocus, coma, and spherical aberrations. These aberrations are typically caused by factors such as the drifting of the focusing plane, axial movement of the samples, tilt of the sample holder, and mismatch of refractive index between the samples and the cover slip. Additionally, there may be subtle changes or misalignment in the imaging system over its service life, which are often unnoticed and cannot always be corrected in time. To address these inherent optical aberrations in an offline manner, we employed the well-trained SFE-Net and SFT-DFCAN models to perform digital adaptive optics and super-resolution reconstruction for time-lapse experimental WF images.

In our experimental setup, we initially captured 100 consecutive frames of a live COS7 cell expressing Ensconsin-mEmerald using our home-build multi-modality SIM system. The time interval between each image was set as 0.5 s. To simulate focus drifting and axial movement, we deliberately introduced a disturbance along the z axis of the motorized sample stage. As depicted in the upper row of Fig. 8(a), the WF images of MTs exhibited slightly varying degrees of defocus aberration at different timepoints. To address this issue, we utilized the proposed SFE-Net to estimate the PSF for each frame. The estimated PSFs were then input into a well-trained SFT-DFCAN model, enabling the reconstruction of SR images based on real-time aberrated PSF information. Consequently, despite severe blurring in the WF images, the SFT-DFCAN equipped with SFE-Net was capable of clearly recovering the densely interlaced MTs [bottom row in Fig. 8(a)].

Subsequently, we proceeded to image a live SUM 159 cell expressing clathrin-EGFP and intentionally adjusted the tilt angle of the sample holder to introduce some coma aberration, which is another common type of aberration in *ex vivo* imaging when the cover slip is not perpendicular to the optical axis. As is shown in Fig. 8(b), the SFE-Net successfully estimated the anisotropic PSFs with coma aberration, facilitated by which the SFT-DFCAN model removed the optical aberrations and clearly resolved the hollow structure of CCPs.

Finally, we imaged another live COS7 cell labelled with calnexin-mEmerald for 100 timepoints, with a time interval of 1 s. Prior to imaging, we deliberately axially offset the sample stage and adjusted the correction collar of the objective to introduce

both defocus and spherical aberrations manually. As anticipated, the SFE-Net reliably estimated the time-varying mixed defocus and spherical aberrations. This estimation facilitated the downstream SFT-DFCAN model in resolving the reticular structure of ER with high resolution and contrast [Fig. 8(c)]. These results illustrate that the proposed SFE-Net has the capability to recognize spatiotemporally varying optical aberrations without any additional hardware, except for the aberrated image itself. This capability enables digital aberration compensation and aberration-aware super-resolution reconstruction for time-lapse live-cell imaging.

4. DISCUSSION

In this paper, we introduced the SFE-Net, a novel method capable of accurately estimating aberrated PSF directly from WF images. One key advantage of SFE-Net over conventional direct wavefront sensing methods is its ability to provide real-time aberration estimation without requiring any additional optical hardware. Additionally, unlike existing indirect wavefront sensing methods that involve time-consuming iterative acquisition and optimization procedures, SFE-Net can estimate aberrations from a single frame at a timescale of ~ 30 ms. This makes it suitable for imaging long-term bioprocesses where optical aberrations vary over time and need to be measured and corrected promptly. By utilizing the PSF generated through SFE-Net, we can effectively address various aberrations and improve spatial resolution in biological images using both the unsupervised RL deconvolution algorithm and the supervised SFT-DFCAN model. Notably, our experiments revealed that by incorporating prior knowledge of aberrated PSF, the SFT-DFCAN model substantially surpassed its backbone model DFCAN where PSF information was not incorporated. Finally, we demonstrated the practical applications of SFE-Net and the facilitated SFT-DFCAN model in digitally correcting optical aberrations and achieving instant image super-resolution in time-lapse live-cell imaging experiments.

More potential applications and extension of SFE-Net are anticipated. First, SFE-Net was trained to estimate aberrated PSF generated with up to 18 orders of Zernike polynomials. However, including higher orders of Zernike polynomials could noticeably degrade the performance of SFE-Net. Upgrading the backbone network architecture of SFE-Net to state-of-the-art models such as Swin-Transformer [23] may expand the application scope. Second, in this paper we primarily conducted principal verification of SFE-Net with a semi-simulated dataset. However, when applying SFE-Net models trained with simulated data to experimental images, there is inevitably a degradation in performance due to the domain shift problem. More ideally, the training dataset of SFE-Net should be acquired via an imaging system with wavefront shaping capability and trained with aberrated images experimentally acquired and corresponding ground-truth aberration applied on the wavefront shaping elements. Third, although we mainly demonstrated the offline digital AO functionality of SFE-Net, it can also be used in various hardware-based AO systems in multiple modalities of microscopes or telescopes by training SFE-Net models based on the physical parameters of the corresponding imaging system. In particular, owing to the temporal sensitivity and image

patch-based estimation scheme of SFE-Net, it can be applied to measure and correct the drastically varying aberrations both spatially and temporally in telescope technologies. We hope that our methods will inspire further developments of next-generation adaptive optics and super-resolution microscopy.

Funding. National Natural Science Foundation of China (31970659, 32125024); National Key Research and Development Program of China (2021YFA1300303); Chinese Academy of Sciences (YSBR-076, ZDBS-LY-SM004); China Postdoctoral Science Foundation (2022M721842, 2023T160365); Tsinghua University (2022SM035); New Cornerstone Science Foundation.

Acknowledgment. The authors thank T. Kirchhausen for the donor plasmids used for genome editing and help in generating the genome-edited cell lines.

Disclosures. D. L., C. Q., and H. C. have a pending patent application on the presented frameworks.

Data Availability. The python codes of SFE-Net and SFT-DFCAN, and the MATLAB codes for generating training and testing datasets with BioSR are publicly accessible on Github [24]. Data underlying the results presented in this paper are available upon reasonable request.

REFERENCES

- N. Ji, "Adaptive optical fluorescence microscopy," *Nat. Methods* **14**, 374–380 (2017).
- Q. Zhang, Q. Hu, C. Berlage, *et al.*, "Adaptive optics for optical microscopy [Invited]," *Biomed. Opt. Express* **14**, 1732–1756 (2023).
- K. M. Hampson, R. Turcotte, D. T. Miller, *et al.*, "Adaptive optics for high-resolution imaging," *Nat. Rev. Methods Primers* **1**, 68 (2021).
- D. Saha, U. Schmidt, Q. Zhang, *et al.*, "Practical sensorless aberration estimation for 3D microscopy with deep learning," *Opt. Express* **28**, 29044–29053 (2020).
- B. P. Cumming and M. Gu, "Direct determination of aberration functions in microscopy by an artificial neural network," *Opt. Express* **28**, 14511–14521 (2020).
- P. Zhang, S. Liu, A. Chaurasia, *et al.*, "Analyzing complex single-molecule emission patterns with deep learning," *Nat. Methods* **15**, 913–916 (2018).
- J. Wu, Y. Guo, C. Deng, *et al.*, "An integrated imaging sensor for aberration-corrected 3D photography," *Nature* **612**, 62–71 (2022).
- J. Wu, Z. Lu, D. Jiang, *et al.*, "Iterative tomography with digital adaptive optics permits hour-long intravital observation of 3D subcellular dynamics at millisecond scale," *Cell* **184**, 3318–3332 (2021).
- C. Qiao, D. Li, Y. Guo, *et al.*, "Evaluation and development of deep neural networks for image super-resolution in optical microscopy," *Nat. Methods* **18**, 194–202 (2021).
- Y. Zhang, K. P. Li, K. Li, *et al.*, "Image super-resolution using very deep residual channel attention networks," in *Proceedings of the European Conference on Computer Vision (ECCV)* (2018), pp. 294–310.
- O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (2015), pp. 1–8.
- J. Gu, H. Lu, W. Zuo, *et al.*, "Blind super-resolution with iterative kernel correction," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2019), pp. 1604–1613.
- D. Hendrycks and K. Gimpel, "Gaussian error linear units (GELUs)," *arXiv*, [arXiv:1606.08415](https://arxiv.org/abs/1606.08415) (2016).
- J. Caballero, C. Ledig, A. Aitken, *et al.*, "Real-time video super-resolution with spatio-temporal networks and motion compensation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2017), pp. 4778–4787.
- S. Bell-Kligler, A. Shocher, and M. Irani, "Blind super-resolution kernel estimation using an internal-GaN," *arXiv*, [arXiv:1909.06581](https://arxiv.org/abs/1909.06581) (2019).
- A. Liu, Y. Liu, J. Gu, *et al.*, "Blind image super-resolution: a survey and beyond," *arXiv*, [arXiv:2107.03055](https://arxiv.org/abs/2107.03055) (2021).
- J. Liang, G. Sun, K. Zhang, *et al.*, "Mutual affine network for spatially variant kernel estimation in blind image super-resolution," in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), pp. 4096–4105.
- A. Shocher, S. Bagon, P. Isola, *et al.*, "InGaN: capturing and remapping the 'DNA' of a natural image," *arXiv*, [arXiv:1812.00231](https://arxiv.org/abs/1812.00231) (2018).
- D. Ren, K. Zhang, Q. Wang, *et al.*, "Neural blind deconvolution using deep priors," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2020), pp. 3341–3350.
- J. Liang, K. Zhang, S. Gu, *et al.*, "Flow-based kernel prior with application to blind super-resolution," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (2021), pp. 10601–10610.
- L. B. Lucy, "An iterative technique for the rectification of observed distributions," *Astron. J.* **79**, 745 (1974).
- W. H. Richardson, "Bayesian-based iterative method of image restoration," *J. Opt. Soc. Am.* **62**, 55–59 (1972).
- Z. Liu, Y. Lin, Y. Cao, *et al.*, "Swin transformer: hierarchical vision transformer using shifted windows," in *Proceedings of the IEEE/CVF International Conference on Computer Vision* (2021), pp. 10012–10022.
- <https://github.com/HypnosRin/SFE-Net>.