Intelligent Computing
A SCIENCE PARTNER JOURNAL

RESEARCH ARTICLE

Universal and High-Fidelity Resolution Extending for Fluorescence Microscopy Using a Single-Training Physics-Informed Sparse Neural Network

Zitong Ye[1], Yuran Huang[1], Jinfeng Zhang[1,2], Yunbo Chen[1,2], Hanchu Ye[1], Cheng Ji[1,2], Luhong Jin[3], Yanhong Gan[1], Yile Sun[1], Wenli Tao[1], Yubing Han[1], Xu Liu[1], Youhua Chen[1,2*], Cuifang Kuang[1,2,4*], and Wenjie Liu[1,4*]

[1]State Key Laboratory of Extreme Photonics and Instrumentation, College of Optical Science and Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China. [2]Ningbo Research Institute, Zhejiang University, Ningbo, Zhejiang 315100, China. [3]Key Laboratory for Biomedical Engineering of Ministry of Education, Department of Biomedical Engineering, Zhejiang University, Hangzhou, Zhejiang 310027, China. [4]Zhejiang Lab, Hangzhou, Zhejiang 311121, China.

*Address correspondence to: wenjieliu@zju.edu.cn (W.L.); cfkuang@zju.edu.cn (C.K.); chenyh21012@zju.edu.cn (Youhua Chen)

Citation: Ye Z, Huang Y, Zhang J, Chen Y, Ye H, Ji C, Jin L, Gan Y, Sun Y, Tao W, et al. Universal and High-Fidelity Resolution Extending for Fluorescence Microscopy Using a Single-Training Physics-Informed Sparse Neural Network. *Intell. Comput.* 2024;3:Article 0082. https://doi.org/10.34133/icomputing.0082

Submitted 12 July 2023
Accepted 9 January 2024
Published 26 February 2024

As a supplement to optical super-resolution microscopy techniques, computational super-resolution methods have demonstrated remarkable results in alleviating the spatiotemporal imaging trade-off. However, they commonly suffer from low structural fidelity and universality. Therefore, we herein propose a deep-physics-informed sparsity framework designed holistically to synergize the strengths of physical imaging models (image blurring processes), prior knowledge (continuity and sparsity constraints), a back-end optimization algorithm (image deblurring), and deep learning (an unsupervised neural network). Owing to the utilization of a multipronged learning strategy, the trained network can be applied to a variety of imaging modalities and samples to enhance the physical resolution by a factor of at least 1.67 without requiring additional training or parameter tuning. Given the advantages of high accessibility and universality, the proposed deep-physics-informed sparsity method will considerably enhance existing optical and computational imaging techniques and have a wide range of applications in biomedical research.

Introduction

The advent and continuous evolution of optical super-resolution (SR) microscopy has considerably expanded the boundaries of human understanding of biological structures, dynamics, and functions at the nanometer level [1–3]. However, it remains difficult for conventional optical SR techniques to acquire the maximum information of samples from multiple dimensions. Consequently, the improvement of spatial resolution typically necessitates the sacrifice of other advantages, such as temporal resolution, imaging depth, field of view, universality, and simplicity, which are equally or even more important in certain applications [4,5].

Computational imaging [6], including conventional computing [7,8] and deep learning inference [9–13], has been considered an effective approach for addressing this issue. For example, an artificial neural network has been introduced into single-molecule localization microscopy to reconstruct an SR image from a remarkably small number of raw frames and/or high-density emitters, thereby reducing acquisition time without damaging spatial resolution [14,15]. The trade-off in point-scanning imaging systems, such as confocal and stimulated emission depletion (STED) microscopy, has been alleviated via the use of deep-learning-based supersampling of sub-Nyquist-sampled data [16,17]. With regard to the spatial resolution of high-speed microscopy, techniques such as total internal reflection fluorescence microscopy, light-sheet microscopy, structured illumination microscopy (SIM), conventional image SR operation [7,8], and supervised and unsupervised deep learning [12,13,16,18–22] have been shown to surpass the physical spatial resolution limitation. Unsupervised deep learning has also been applied to hyperspectral SR tasks to unmix overlapping information at the subpixel level [23–25]. In addition, physics-informed neural networks have demonstrated high efficiency and accuracy in solving partial differential equations with appropriate objective functions and network structures [26–28].

Despite these important advancements, computation-assisted SR imaging methods still suffer from severe challenges in handling

highly ill-posed and multifactor-coupled inverse resolution-improvement problems [9,12,29–32]. Such challenges include lack of accessibility (preparing compatible and high-quality training data, fine-turning parameters, and iterative optimization requires experts and is time-consuming), reproducibility (output result varies with the algorithm parameters, training set, and network architecture), fidelity (artifacts are prone to be introduced because of the pure computational property), and universality (the computation framework is task-specific, data-driven, sample-dependent, and modality-dependent).

To address this persistent issue, we propose a deep-physics-informed sparsity (DPS) neural network wherein the framework is designed holistically to fully synergize the strengths of the forward optics model, prior knowledge, back-end algorithm, and deep learning. In addition to surpassing the resolution limit of existing microscopy techniques (~60 nm in the SIM modality), DPS simultaneously addresses all the other aforementioned issues and effectively benefits from its multipronged learning strategy.

The main contributions of this study are summarized as follows:

1. By embedding fluorescence imaging processes, sample priors, and deep back-projection units into the neural deconvolution strategy, we achieved a ~1.67-fold resolution enhancement without the need for high-quality ground-truth datasets. Compared to state-of-the-art deblurring methods, the pretrained DPS achieved the best performance in terms of structural fidelity.
2. We developed a multistep deep learning inference strategy (Denoise-DPS) that incorporates sequential denoising, conventional reconstruction, and artifact and background removal to holistically improve the performance of the DPS network under high-noise and high-background conditions.
3. We adopted an unsupervised learning strategy to improve the universality of the network. High fidelity and universality were demonstrated by successfully extending the resolution of wide-field, confocal, SIM, and STED images with DPS and Denoise-DPS networks trained using only a single SIM dataset among diverse biological samples including microtubules (MTs), mitochondria, nuclear pore complexes (NPCs), F-actin, and clathrin-coated pits (CCPs), in fixed or live cells.

Methods

Related works

Fluorescence image SR is an ill-posed problem and can be expressed as

$$\arg\min\left\{\left\|y - PSF \otimes x\right\|_2^2\right\}, \quad (1)$$

where *PSF* denotes the point spread function (PSF) of the optical system, y is the raw low-resolution (LR) image, x is the recovered SR result, and $\otimes$ denotes a convolution operation. To recover x, sparse deconvolution utilizes the sparsity and continuity priors provided by specific biological structures, which have been proven to be highly structure-dependent and reliant on parameter tuning [7]. The mean shift SR algorithm (MSSR) compresses the PSF by calculating the intensity gradient and fluorescence density; however, it can lead to negative pixel values and information loss during the iteration of updating consecutive MSSR images, especially for complex structures [8]. With regard to deep-learning-based SR methods, supervised networks with carefully paired training data have been widely adopted to transform diffraction-limited input images into super-resolved images, which are still sample- and modality-dependent, as well as ground-truth-limited [16,18,29]. Recently, unsupervised deep learning has been applied to image deblurring and SR tasks [33,34]. However, these studies focused on real-world applications, such as camera shaking correction, which involves anisotropic blur kernels. Therefore, state-of-the-art methods are unsuitable for deblurring in fluorescence microscopy.

Principles and workflow of DPS framework

In this study, we used an unsupervised neural network to overcome the structural dependency of classic optimization algorithms. Compared with other resolution-enhancing algorithms, an unsupervised neural network can yield more accurate restoration results owing to its ability to build high-level features [35,36]. In addition, continuity and sparse priors constrain the solution of ill-posed problems in the iterative SR microscopy optimization process. Therefore, we introduced unsupervised neural networks and fluorescence microscopy priors into the inverse process to maintain structural integrity.

To ensure resolution enhancement while maintaining structural fidelity, we designed the following physics-informed objective function (Fig. 1A):

$$\arg\min\left\{\mathrm{SSIM}\left(x \otimes srPSF, y\right) + \alpha R_{\mathrm{Hessian}}(x) + \beta\|x\|_1\right\}, \quad (2)$$

where the structure similarity index measure (SSIM) is a perceptually motivated metric that is widely used to reconstruct structural details [37] and the L1 norm ($||\ ||_1$) and Hessian matrix (R_{Hessian}) correspond to the sparsity and continuity constraints, respectively. The 2 constraints were balanced by the weights α and β to ensure optimal outputs. The term *srPSF* in Eq. 2 denotes the equivalent PSF after resolution enhancement. This is in contrast with the commonly used PSF corresponding to the LR scenario in Eq. 1, which improves the resolution by considering the physical imaging model. Hessian-related constraints help avoid possible gridding artifacts and maintain structural fidelity [21,22,38]. In addition, sparsity is an important feature in microscopic imaging [7], and, for a high-resolution result, the smaller the PSF, the higher the sparsity level. Thus, the addition of the sparsity constraint can further extract high-frequency information and improve the resolution by approximately 1.4-fold as compared to reconstruction without the sparsity constraint (Fig. 1B and C). Furthermore, to improve the performance of reserving the high-frequency details of the network, which is necessary for resolution improvement, particularly for high-fidelity reconstruction, we modified the conventional U-NET architecture by introducing an error-feedback unit (see the following 2 sections for network details). A preprocessing step was adopted for effective background removal (Fig. 1A and Figs. S1 and S2). The synergistic effect of these components results in resolution enhancements characterized by high fidelity and broad universality.

Res-U-DBPN architecture

Conventional U-NETs suffer from missing feature map information during the downsampling process, and subsequent

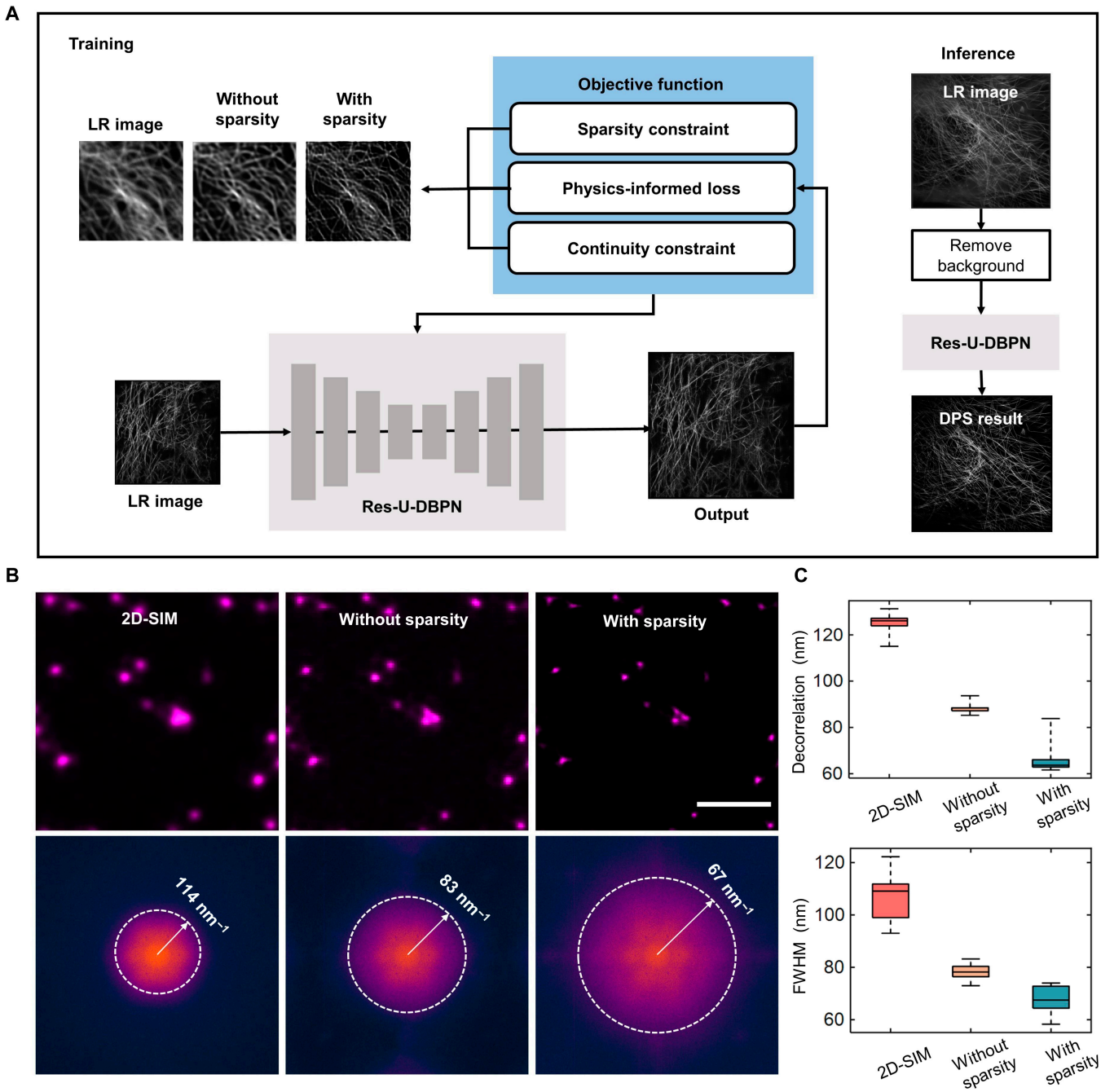

Fig. 1. Principle of DPS framework. (A) Schematic showing the DPS training (left) and inference (right) procedures. (B) Comparison of bead sample reconstruction under different conditions. The spectra of the 2D-SIM image and DPS-SIM reconstruction without and with the sparsity constraint are shown. (C) Statistical resolution comparisons of 2D-SIM and DPS-SIM in the cases shown in (B). Resolutions were measured by decorrelation analysis ($n = 8$, top) and FWHM values ($n = 8$, bottom). Scale bar, 1 μm (B).

upsampling operations cannot retrieve these lost features. Inspired by the deep back-projection network (DBPN) [39], we integrated an error-feedback unit into the U-NET architecture to further improve the performance of the U-shaped network on the SR task.

The proposed network (called Res-U-DBPN) has a structure (Fig. 2A) that is based on Res-UNET [40] and consists of 4 up-projection units and 4 down-projection units (Fig. 2B). In the down-projection units, the original image was first convolved to obtain the original feature map, followed by 2-scale downsampling to acquire the LR feature map. The LR feature map was deconvolved to obtain a high-resolution feature map by upsampling to the original scale. The obtained LR feature map was subtracted from the original LR image, added to the downsampling layer, and then convolved to produce the synthetic LR feature map, which was then added to the previous LR feature map to obtain the final LR feature map.

The process of the up-projection units is analogous to that of the down-projection units. The original feature map was first convolved with a 3×3 kernel and then deconvolved to obtain a high-resolution feature map by upsampling to the 2-scale level. The LR feature map was obtained by 2-scale downsampling from the high-resolution feature map and subtracted from the original feature map. The subtracted feature map was fed into

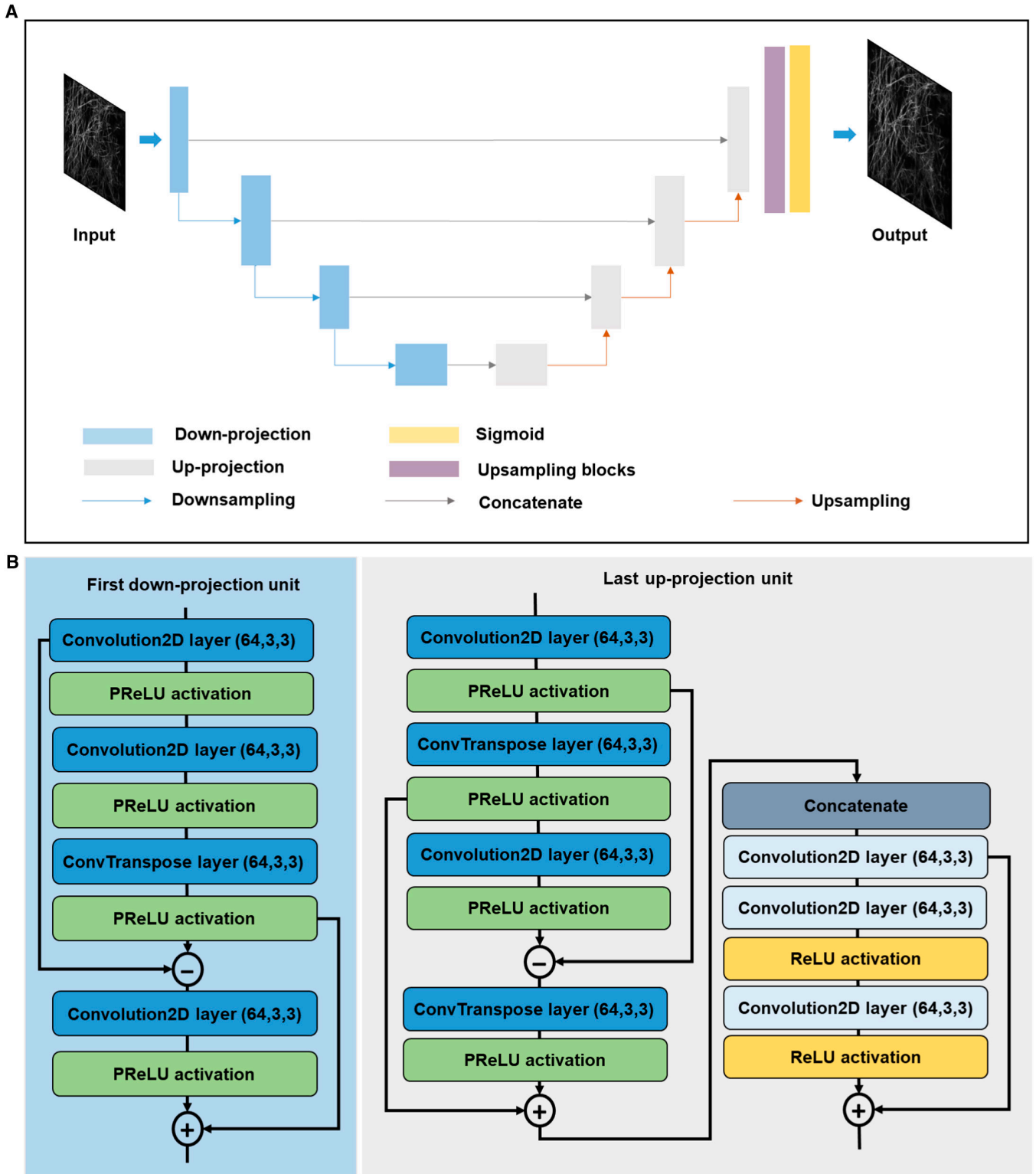

Fig. 2. Res-U-DBPN architecture. (A) Network comprising 4 up-projection units, 4 down-projection units, and 4 scales, with each scale having an identity skip connection between the up-projection units and the down-projection units. The blue and gray blocks represent the down-projection units and up-projection units, respectively. (B) Architectures of the first down-projection unit (left) and the last up-projection unit (right) shown in (A). The architecture of the projection unit mainly consisted of the convolutional layer, parametric rectified linear unit (PReLU) activation function, and transpose convolutional layer, whose stride was set to 2.

the upsampling layer and deconvolved to produce a synthetic high-resolution feature map, which was then added to the high-resolution feature map to obtain the final high-resolution feature map. These projection elements can be regarded as a self-correcting process that feeds projection errors to the sampling layer to improve network performance during iterative optimization (see details in Section S2 and Fig. S3).

As for the complexity of our models, 3 parameters were considered: million floating point operations (MFLOPs), memory, and the number of parameters (NumPara) of the neural network. The results are shown in Table 1.

Network structure and training

The framework of the DPS method is a typical U-NET with skip blocks between the corresponding down-projection and up-projection units, inspired by the DBPN [39]. In the training phase, the 2 scalar weighting factors in Eq. 2 were empirically set to $\alpha = 1 \times 10^{-6}$ and $\beta = 5 \times 10^{-4}$. We first trained the networks on a preliminary training dataset containing only F-actin with a learning rate of 1×10^{-4} for 40,000 minibatch iterations to acquire the initial model and then on the dataset containing both F-actin and CCP for the final model. The 2-step transfer training strategy with specific samples was designed to leverage the structure-preserving and resolution-enhancing capabilities afforded by the F-actin and CCP training datasets, respectively, and to accelerate the training process (Fig. S4). Our 2-step transfer training strategy with specific samples balanced resolution, structure fidelity, and generalization (Fig. S4). Once trained, the Res-U-DBPN network can be used to infer super-resolved details from the LR.

A dual-stage architecture was introduced into the Denoise-DPS framework, which contained 2 major parts: denoising and DPS processes (Figs. S5 and S6). The denoising process includes modules for denoising, conventional reconstruction (if necessary), and artifact removal. The denoising module adopts the Res-UNET [40] structure (Fig. S7) and is trained using raw SIM images acquired under increasing low signal-to-noise ratio (SNR) levels as the input and high-SNR raw SIM images in the same region as the ground truth. The objective function of the denoising network is a combination of the SSIM, L1 norm, and frequency-domain loss.

$$\arg\min\Big\{\big\{\alpha \mathrm{SSIM}(x,y) + \beta \mathrm{L1}(x,y) + \varepsilon \mathrm{L1}\left[\mathrm{FFT}(x) - \mathrm{FFT}(y)\right]\big\}\Big\}. \tag{3}$$

Here, x denotes the output of the network, and y denotes the corresponding ground truth. FFT denotes the fast Fourier transform. In our experiments, the weights α, β, and ε were set to 0.06, 0.6, and 0.4, respectively. The training dataset contained ~7,500 pairs of CCP and F-actin images (128 × 128 pixels) at 7 SNR levels. The training was terminated after 40,000 iterations. Once trained, the network can be applied to other imaging modalities (see Section S1).

As revealed by the experiments, an optional artifact removal module based on Res-UNET may be introduced into the denoising module to further improve the performance. This might aid imaging techniques requiring postreconstruction, such as SIM (Section S1 and Fig. S5). The input was the result processed by the denoising module and conventional reconstruction algorithm in succession, and the ground truth was a high-quality image obtained by the conventional imaging method in the same region. In this module, we used perceptual loss [41] to penalize the differences between the output and target values. Approximately 5,000 pairs of images were generated for the training dataset. After 40,000 iterations, the training was terminated to obtain an available model for various structures.

The network of the conventional denoised SIM reconstruction (termed Res-UNET-SIM in Fig. S6) was based on Res-UNET with an additional upsampling block (Fig. S7). The training input of Res-UNET-SIM comprised 9 raw SIM images at different SNR levels, and the ground truth was the corresponding high-SNR SIM result. The objective function is defined as follows:

$$\arg\min\left\{\alpha SSIM(x,y) + \beta L1(x,y)\right\}, \tag{4}$$

where x denotes the output of the network and y denotes the corresponding ground truth. The function was the combination of SSIM and L1 norm loss. Further, the weights α and β were set to 0.1 and 1, respectively. Approximately ~7,500 pairs of F-actin and CCP images were generated as the training datasets. After 40,000 iterations, the training was terminated to obtain an available model for effect comparison.

The training and inference phases were performed on a computer workstation equipped with a Xeon(R) Gold 6134 CPU (Intel) and an RTX 2080 graphic processing card (NVIDIA) using Python v.3.6, TensorFlow v.2.4.0, and Keras v.2.2.4. All training datasets were generated using an open-source, high-quality SIM dataset [29].

Simulation evaluation of DPS and Denoise-DPS networks

We simulated a set of LR images by convolving the ground-truth targets using the theoretical SR PSF as the input for the resolution-enhancing algorithms. Gaussian noise with different variances may be added to study the influence of noise on reconstruction. To quantitatively assess structural fidelity, we used 2 metrics, namely, the average peak SNR (PSNR) and gradient magnitude similarity deviation (GMSD) [42], calculated on the basis of the ground-truth image and the output of the different resolution-enhancing algorithms. Furthermore, we calculated the error map by analyzing the pixel-wise absolute differences between the synthetic LR images and enhanced results convolved with the theoretical SR PSF.

PSF transformation for cross-modality applications

The PSF of the input images must be equivalent to the PSF used in the loss function (*srPSF* in Eq. 2), which is determined using 3 parameters: the excitation wavelength λ_{srPSF}, pixel size, and numerical aperture (NA). According to Nyquist sampling theory, if the pixel size of the input image is less than $\lambda_{\mathrm{input}}/2\mathrm{NA}_{\mathrm{input}}$, upsampling or downsampling processes could be adopted to transform the PSF of the input image to the request image. The rescaling factor is expressed as follows:

$$\mathrm{factor}_{\mathrm{scale}} = \frac{\mathrm{pixelsize}_{\mathrm{input}}}{\mathrm{pixelsize}_{\mathrm{request}}} = \frac{\lambda_{\mathrm{input}} \cdot \mathrm{NA}_{\mathrm{input}} \cdot \mathrm{pixelsize}_{\mathrm{input}}}{\lambda_{srPSF} \cdot \mathrm{NA}_{srPSF} \cdot \mathrm{pixelsize}_{srPSF}} \left(\mathrm{pixelsize}_{\mathrm{request}} < \frac{\lambda_{\mathrm{input}}}{2\mathrm{NA}_{input}}\right). \tag{5}$$

Table 1. Evaluation of the complexity of the Res-U-DBPN model (input tensor size = 1 × 256 × 256 × 1)

	MFLOPs	NumPara (M)	Memory (MB)
Res-U-DBPN	178.9	89.46	374.4

This theory holds true only when the pixel size of the rescaled image is less than $\lambda_{\text{input}} / 2\text{NA}_{input}$.

Results and Discussion

Extension of spatial resolution of SIM with high fidelity

We first validated the performance of the DPS framework on the SIM modality (DPS-SIM) through simulations and experimental tests considering its high temporal resolution but limited spatial resolution. In the simulations, the LR images were generated by convolving the ground-truth images with the equivalent SIM PSF (Fig. 3 and Fig. S8). The generated LR images were processed using different algorithms for comparison. When reconstructing the standard LR test images (Fig. 3), although all 3 methods—sparse-SIM, MSSR-SIM, and DPS-SIM—showed resolution improvement, DPS-SIM performed substantially better in terms of structural fidelity within the entire field of view. The quantitative assessment results from GMSD [43], PSNR, and the degraded error map further verified and highlighted the effective and reliable resolution improvement achieved by using the proposed method.

The quantitative PSNR comparison results for the standard datasets [44] are presented in Table 2. The Richard–Lucy deconvolution was used as a baseline to evaluate the performance of the algorithm. We compared the proposed DPS with MSSR, sparse deconvolution, W-DIP, and SelfDeblur (non-blind) [7,8,33,34] using the default settings. For comparison, the ground-truth images were convolved with the simulated PSF to obtain blurred images. The results show the strong quantitative advantage of the proposed unsupervised neural network deblurring approach over other algorithms, demonstrating the superiority of the proposed neural network in terms of fidelity. Whereas W-DIP and SelfDeblur had higher PSNR values on the Lena and Peppers datasets, the proposed method achieved a better overall performance.

For the ring-shaped structure simulation, Gaussian noise with different variance levels was added to the LR image before

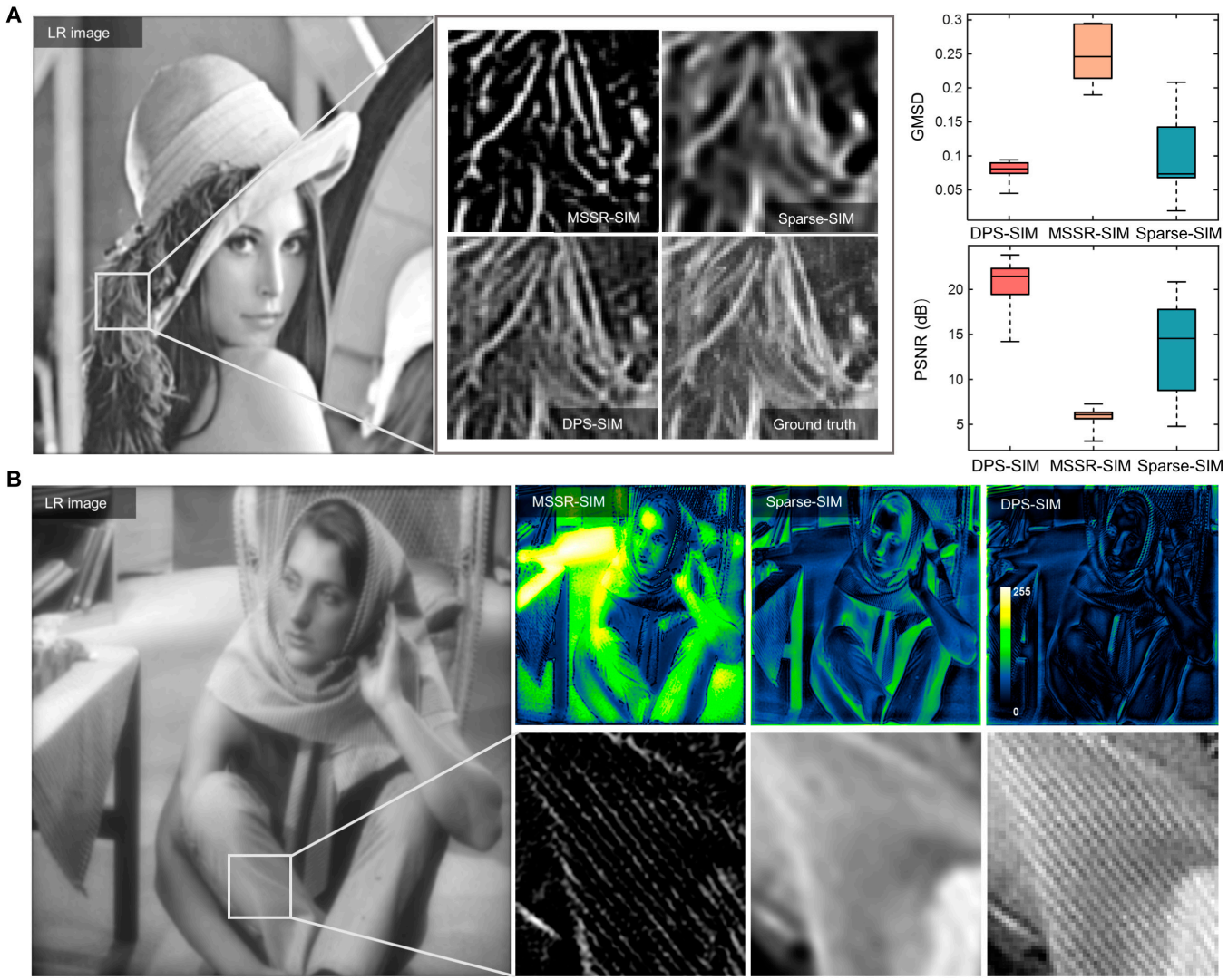

Fig. 3. Reconstruction comparison of different test targets with MSSR, sparse deconvolution, and the DPS framework. (A) Statistical comparisons of MSSR, sparse deconvolution, and DPS in terms of PSNR and GMSD are shown in the right corner ($n = 10$). (B) Error maps calculated by subtracting the ground-truth image from the processed results convolved with SIM PSF are shown in the top right corner with the higher color value denoting the higher error. LR images were generated by convolving the ground-truth images with equivalent SIM PSF.

subsequent reconstruction. Compared with the MSSR-SIM algorithm, DPS-SIM provides high-fidelity reconstructions across the entire SNR and radius ranges (Fig. S8). In addition to providing the highest-quality and most trustworthy SR results, DPS-SIM is approximately 90% faster than the other methods.

To demonstrate the capability of the DPS framework to extend SIM resolution in practical experiments, standard test samples, including Argo-SIM line pairs with a spacing distance of 0 to 390 nm (Fig. 4A) and adjacent 40- and 100-nm fluorescent beads (Fig. 4B), were first imaged using a home-built SIM system [45]. As shown in Fig. 1B, the 60-nm line pair could be resolved as 2 parallel lines after DPS reconstruction, whereas it was obscured in the conventional 2-dimensional (2D) SIM results. The 2 adjacent beads that were indistinguishable using conventional SIM were also clearly separated using DPS-SIM (Fig. 4B).

Next, we validated the superiority of DPS-SIM using more complex and generalized biological samples, including wire-like MTs (Fig. 5A), spherical NPCs (Fig. 5B), reticular F-actin (Fig. 5C), and annular CCPs (Figs. 5D and E and Movie S1). Evidently, DPS-SIM improved the resolution of conventional 2D-SIM by providing more details and smaller structures with high fidelity in all scenarios. For example, 2 actin filaments were successfully resolved after processing the original 2D-SIM image using DPS. The high-fidelity resolution enhancement and live-cell compatibility afforded by DPS-SIM further helped track the more detailed physiological activities of CCPs labeled with clathrin-enhanced green fluorescent protein. The previously invisible inner rings of the CCPs in the 2D-SIM image became visible after DPS-SIM processing, with the smallest distinguishable diameter of ~78 nm (Fig. 5D and E) [46]. The time-lapse DPS-SIM data (Fig. 5E and Movie S1) indicate the event of pore formation and growth, in which the entire CCP, including the inner ring, grew larger over time; however, it was difficult to distinguish the dynamics of the inner ring on the basis of the 2D-SIM results because of insufficient spatial resolution. A similar CCP behavior was also reported for a high-NA SIM [47]. However, because the resolution of the high-NA SIM is still diffraction-limited, only the inner CCP rings at the later growth phase were clearly visible, and the inner dynamics of the early growth phase were missing. Furthermore, the full width at half maximum (FWHM) and decorrelation analyses [48] statistically corroborated that the proposed DPS-SIM can extend the spatial resolution of 2D-SIM to ~60 nm with high fidelity among universal sample structures.

Multistep DPS framework for low-SNR SIM imaging

A well-known limitation of microscopy techniques, including SIM, is that high-SNR raw data are required to achieve high-quality reconstruction. However, noise-induced reconstructed artifacts are inevitable in biological experiments, especially for live-cell imaging, because the light intensity and exposure time must be decreased to reduce photobleaching and phototoxicity. Although current deep learning denoising SIM reconstruction methods can enhance the signal remarkably, they commonly suffer from structural distortion because they directly learn the mapping between low-SNR raw images and high-resolution images [49]. To improve the quality of DPS reconstruction under low-SNR conditions without sacrificing resolution, fidelity, or universality, we further developed a holistic multistep inference strategy (Denoise-DPS-SIM) by dividing the multi-factor-hybrid image degradation problem into 5 steps—a denoising network, conventional reconstruction, an artifact removal network, background removal, and a Res-U-DBPN network—and addressing the major obstacle of each step progressively (Fig. 6A; see also details in Section S1 and Figs. S5 to S7). Compared to conventional Res-UNET-based denoising SIM algorithms that learn the SR results from 9 low-SNR raw images (Fig. S6), the proposed multistep denoising method effectively preserved the high-frequency information (Fig. 6B and C). For example, when processing the 2D-SIM and DPS-SIM MT data (Fig. 6B), details such as the adjacent tubes were successfully preserved by the proposed multistep method; however, artifacts existed in both the Res-UNET-SIM and Res-UNET-DPS-SIM results (Fig. 6C).

Figure 7 shows a representative low-SNR experiment, in which the raw CCP data were reconstructed using both conventional 2D-SIM and DPS-SIM with and without incorporating the abovementioned denoising process. As expected, the 2D-SIM results exhibited obvious background noise and structural distortion [Fig. 7A (left) and B (top)], and Denoise-2D-SIM effectively suppressed noise [Fig. 7A (middle) and B (middle)]. Denoise-DPS-SIM further improved the resolution of the Denoise-2D-SIM image by preserving the hollow structure of the CCP [Fig. 7A (right) and B (bottom)]. Furthermore, we tested the performance of Denoise-DPS-SIM in a low-SNR live-cell experiment. As shown in Fig. 7C and Movie S2, with the assistance of the Denoise-DPS framework, the hollow ring structure was clearly resolved and well maintained over the entire time range, whereas 2D-SIM and Denoise-2D-SIM could not provide sub-100-nm details and DPS-SIM introduced structural deformation owing to the low SNR. By extracting

Table 2. PSNR comparisons between the proposed method, Richard–Lucy (RL) deconvolution, and algorithms in [7,8,33,34]. The values in boldface correspond to the optimal results for each test image.

	Barbara	Room	C.man	Lena	Peppers	Pirate
RL	12.83	14.51	11.17	13.90	12.58	15.11
MSSR	4.91	6.25	5.77	5.91	6.25	6.32
Sparse	14.59	16.91	14.50	13.51	7.91	8.77
W-DIP	18.68	17.15	18.11	19.79	**26.78**	19.02
SelfDeblur	18.76	21.37	19.69	**21.71**	25.85	22.30
Proposed	**19.45**	**21.78**	**22.06**	21.15	22.57	**22.33**

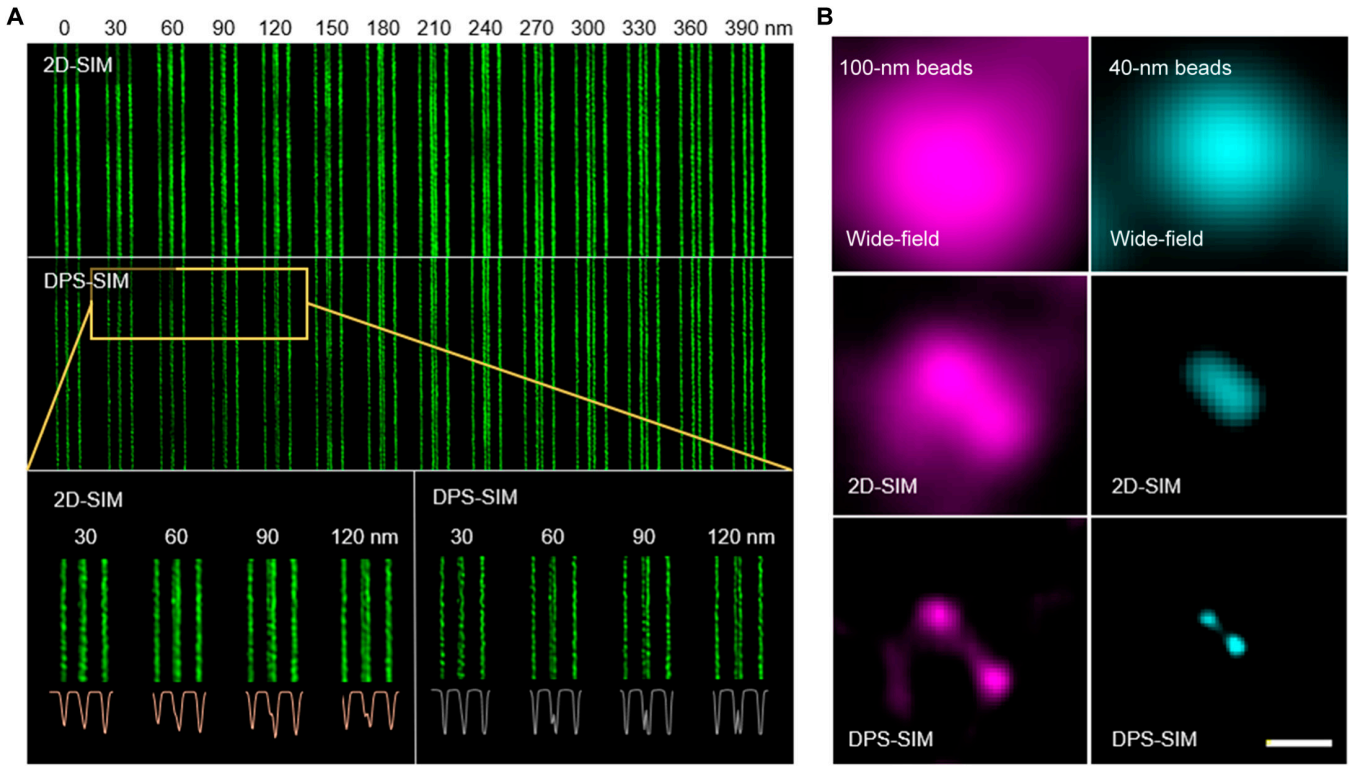

Fig. 4. Performance of DPS framework. (A) Reconstruction comparison of conventional 2D-SIM and DPS-SIM using spaced fluorescent line pattern sample separated by steps of 30 nm (from 0 to 390 nm). The reconstruction results using both methods of the region enclosed by the yellow box are magnified and shown below with corresponding intensity profiles. (B) Fluorescent beads (40 and 100 nm in diameter) recorded by wide-field microscopy, 2D-SIM, and DPS-SIM. Scale bar, 160 nm (B).

high-fidelity details from noisy raw images, the consecutive behaviors of the contact, fusion, and fission of 2 pores were distinctly reconstructed by Denoise-DPS-SIM (Fig. 7C, bottom); these aspects are difficult to capture using existing SR techniques.

Applicability of DPS framework across fluorescence microscopy modalities

Another remarkable characteristic of the proposed DPS network is that, because it was designed and trained in a holistic manner, it was directly applicable to other modalities with high fidelity without extra training or parameter-turning requirements, thus further facilitating the broad application of deep learning algorithms. To achieve PSF matching among the different imaging modalities, we preserved all information in the frequency domain while resizing the images (see Methods for details).

To empirically investigate the universality of the performance of the DPS framework, we used the Argo-SIM test slide on a home-built point-scanning confocal and STED imaging system. The results shown in Fig. 8A demonstrate that confocal microscopy allowed the distinction of 2 parallel rows of fluorophores located at a minimum distance of ~210 nm; this is consistent with the theoretical analysis. With the proposed DPS process, a line pair distance of 150 nm can be distinguished. Therefore, the DPS method provided at least a 1.4-fold resolution enhancement for confocal imaging.

Next, we imaged the MTs (Fig. 8B), mitochondrial outer membrane (Fig. 8C), and NPCs (Fig. 8D) using a confocal system. Among these results, DPS-Confocal not only increased the sharpness and diminished the diameter of the wire-like MT and globular NPC structures but also successfully resolved the hollow structure of the mitochondrial outer membrane with a sharper outline and lower background. This confirmed that the trained DPS network using one SIM dataset could be extended to diverse imaging modalities and fluorescent specimens with high-fidelity resolution enhancement. We also performed long-term live-cell confocal imaging of the MT samples for up to 10 min (see Movie S3). DPS-Confocal imaging not only revealed the SR dynamics of the MT cytoskeleton, such as the contact and detachment of adjacent tubules, but also suppressed the photobleaching caused by the high laser power and dwell time used in common confocal imaging.

Furthermore, we assessed both DPS and Denoise-DPS performance in the STED imaging mode (Fig. 8E). In general, STED has a higher noise value than does confocal microscopy because fewer photons are collected by the detector. For comparison, the corresponding STED, Denoise-STED, DPS-STED, and Denoise-DPS-STED images are shown in Figure 8E. Owing to the noise in the raw STED image, the DPS results presented artifacts, such as visible background noise, tube discontinuity, and distortion. However, these issues were addressed by Denoise-DPS-STED reconstruction, in which 2 adjacent tubes were resolved with high SNR and fidelity (Fig. 8F). The Denoise-DPS network was finally applied to live-cell STED imaging owing to its robust resolution-extending capacity with only a single training (see Movie S4). Overall, the proposed DPS and Denoise-DPS frameworks provide powerful tools to further improve the resolution and image quality of most fluorescent images while ensuring high reliability and broad universality.

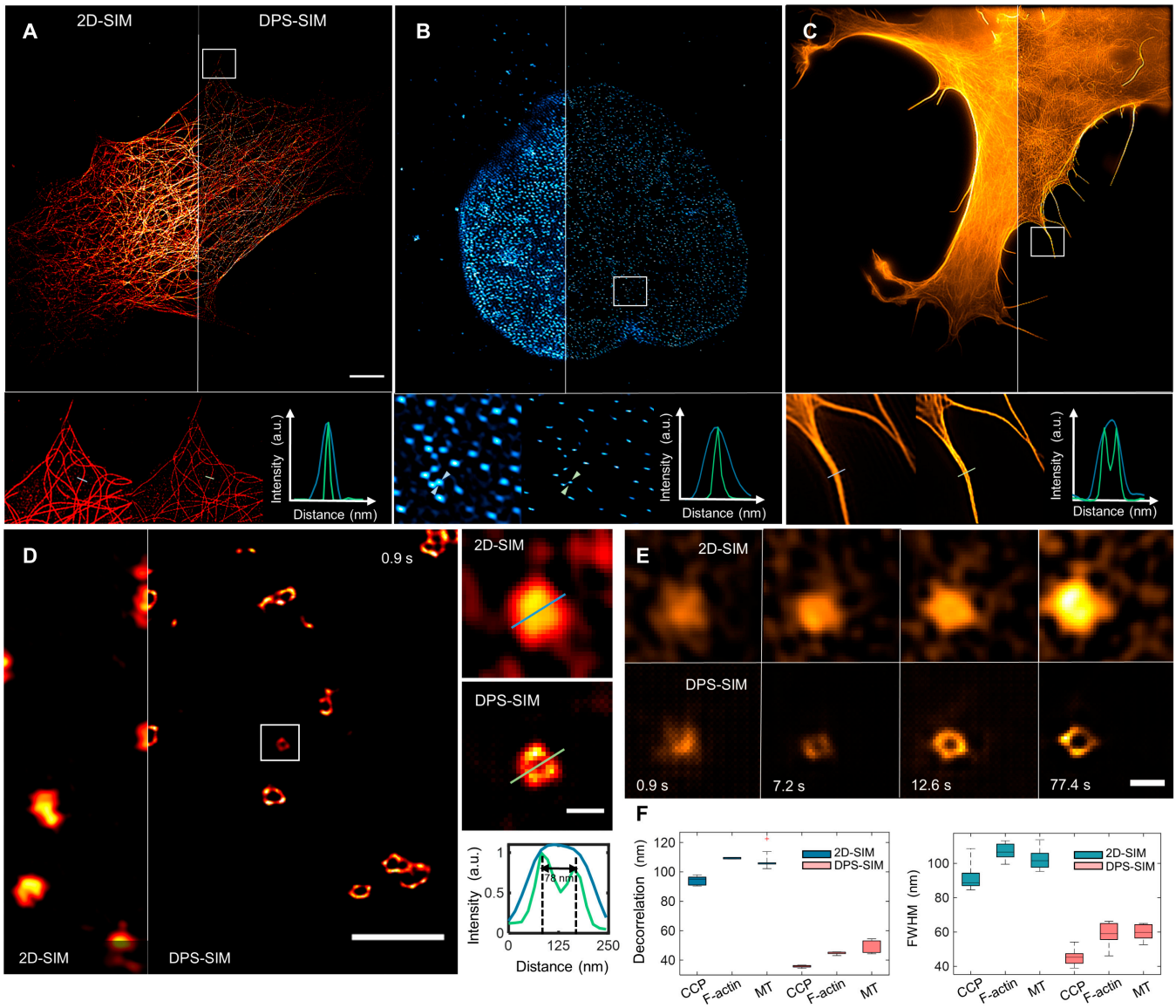

Fig. 5. DPS framework universally extends the spatial resolution of SIM for different biological samples. (A to C) Representative images reconstructed by conventional 2D-SIM and DPS-SIM of MT, NPC, and F-actin samples available in [7]. The regions enclosed by the white boxes are magnified and shown below with intensity profiles along corresponding indicators. (D) A representative CCP image reconstructed by conventional 2D-SIM and DPS-SIM. Magnified views of the region enclosed by the white box are shown on the right with intensity profiles along corresponding lines. The 2-peak distance was calculated using double-peak Gaussian fitting. (E) Dynamic activity of a representative growing CCP observed by 2D-SIM (top) and DPS-SIM (bottom). (F) Statistical resolution comparisons of 2D-SIM and DPS-SIM in the cases of CCP, MT, and F-actin samples. Resolutions were measured using decorrelation analysis ($n = 10$, left) and by calculating the FWHM values ($n = 10$, right). Scale bars, 6 μm (A to C), 1 μm (D, left), 100 nm (D, right), and 200 nm (E). a.u., arbitrary units.

Conclusion

In this study, we devised 2 holistic image postprocessing frameworks, DPS and Denoise-DPS, to substantially boost classical fluorescence microscopy techniques. Compared to conventional computation algorithms, the proposed methods exhibit several remarkable advantages. First, the spatial resolution is no longer limited by the training boundary that exists in state-of-the-art deep learning methods but is surpassed by the DPS network (~60 nm in the DPS-SIM result) as the priors and physics-informed information—the sparsity constraint, forward imaging model, and equivalent SR PSF—are utilized to regulate the unsupervised neural network. Furthermore, the proposed network, referred to as Res-U-DBPN, adopts an error-feedback strategy and performs better in preserving high-frequency information. Second, the aforementioned resolution is improved with considerably high fidelity, even under low-light-dose conditions, as the informative multistep denoising module is integrated into the DPS network. Finally, concurrently considering the front-end optics model and imaging process, physical and mathematical consistency, back-end optimization algorithm, and unsupervised deep learning strategy, the efficacy of the proposed DPS and Denoise-DPS methods, without requiring fine-tuning of the fuzzy parameters, high-quality training pair collection, or hardware modification, was demonstrated. Consequently, the trained network with only a single

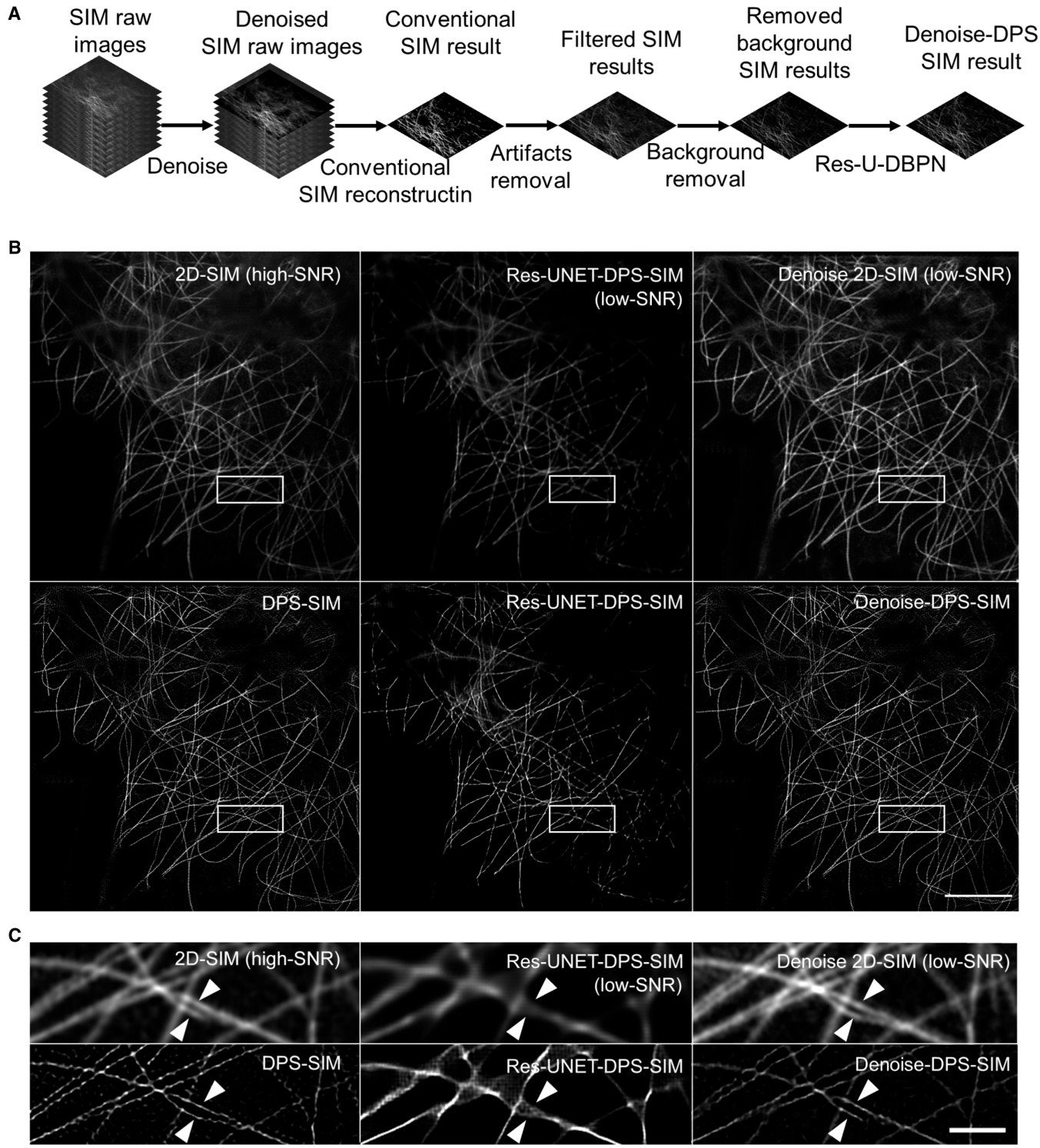

Fig. 6. Comparison of conventional and proposed denoised SIM methods applied to MT data available in [29]. (A) The workflow of the multistep DPS-SIM framework, including denoising, conventional SIM reconstruction, artifact removal, and DPS reconstruction. (B) Top row: High-SNR resolution-limited SIM reconstruction by conventional 2D-SIM (left) is shown for reference. Low-SNR 2D-SIM images were processed using Res-UNET-SIM (middle) and Denoise-2D-SIM (right). Bottom row: Resolution-extended SIM images by DPS-SIM (left) reconstructed with the top left data, Res-UNET-DPS-SIM (middle) reconstructed with the top middle data, and Denoise-DPS-SIM (right) reconstructed with the top right data. (C) Corresponding magnified views of the white-boxed regions in (B) with 2 arrowheads indicating structure preservation achieved using the proposed denoised SIM method. Scale bars, 6 μm (B) and 1 μm (C).

SIM dataset can be directly applied to other imaging modalities and samples to achieve high-fidelity resolution improvement (Fig. S9), thus considerably extending the accessibility, reproducibility, and generalization of the proposed algorithm.

The proposed method exhibited both high structural fidelity and resolution enhancement in fluorescence microscopy image deblurring. However, there are still 2 limitations to the implementation of this method. First, the complexity and high memory

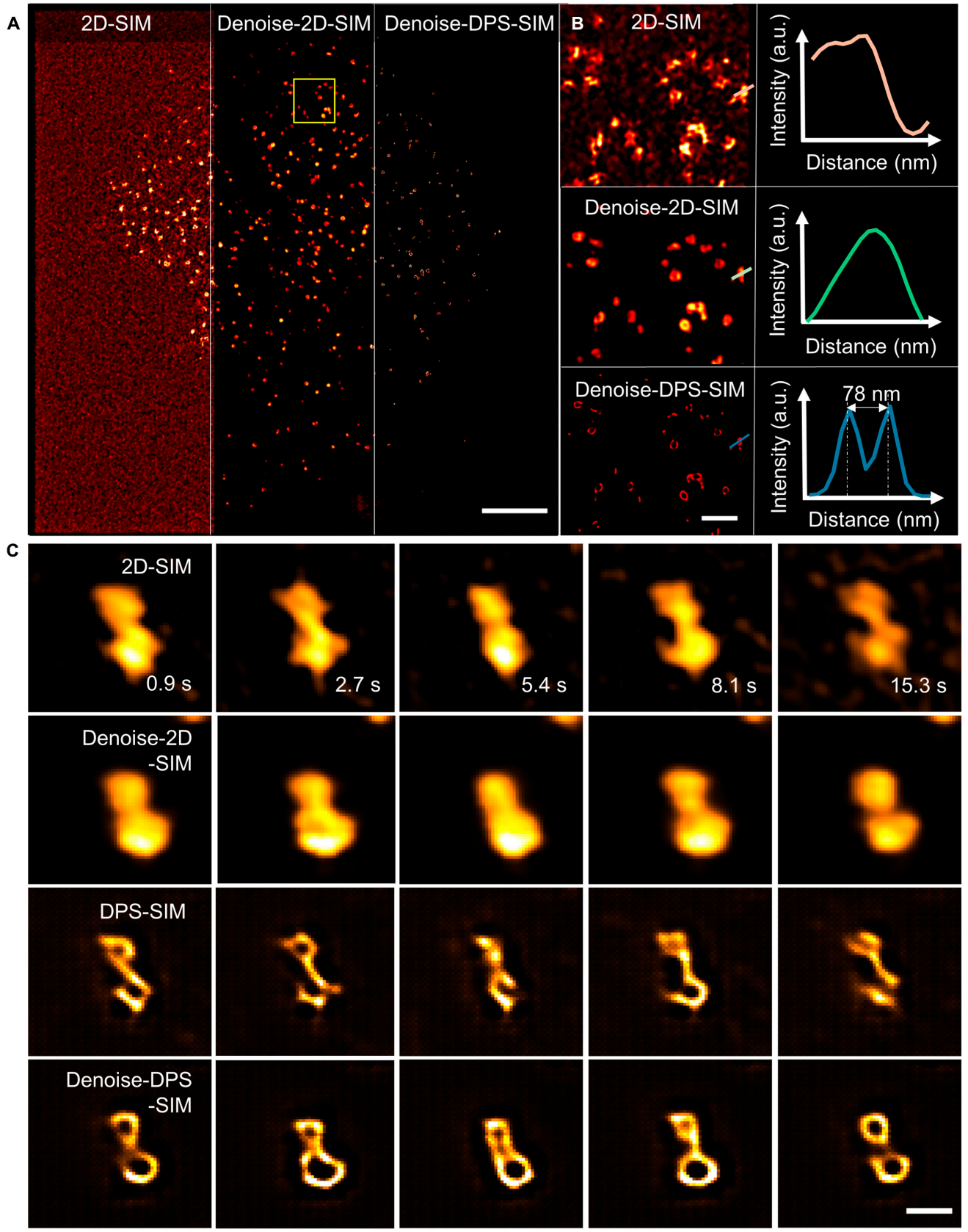

Fig. 7. Denoise-DPS-SIM for higher-fidelity resolution improvement under low-SNR condition. (A) Low-SNR CCP data available in [29] reconstructed with conventional 2D-SIM (left), Denoise-2D-SIM (middle), and Denoise-DPS-SIM (right). Denoise-2D-SIM represents the result processed by the denoise module, conventional 2D-SIM reconstruction, and artifact removal module only. (B) Magnified views of the yellow-boxed region in (A) with the intensity profiles along corresponding lines shown on the right. From top to bottom: 2D-SIM, Denoise-2D-SIM, and Denoise-DPS-SIM. (C) Dynamic activity of a CCP observed by 2D-SIM (top), Denoise-2D-SIM (middle top), DPS-SIM (middle bottom), and Denoise-DPS-SIM (bottom). Scale bars, 4 μm (A), 800 nm (B), and 250 nm (C).

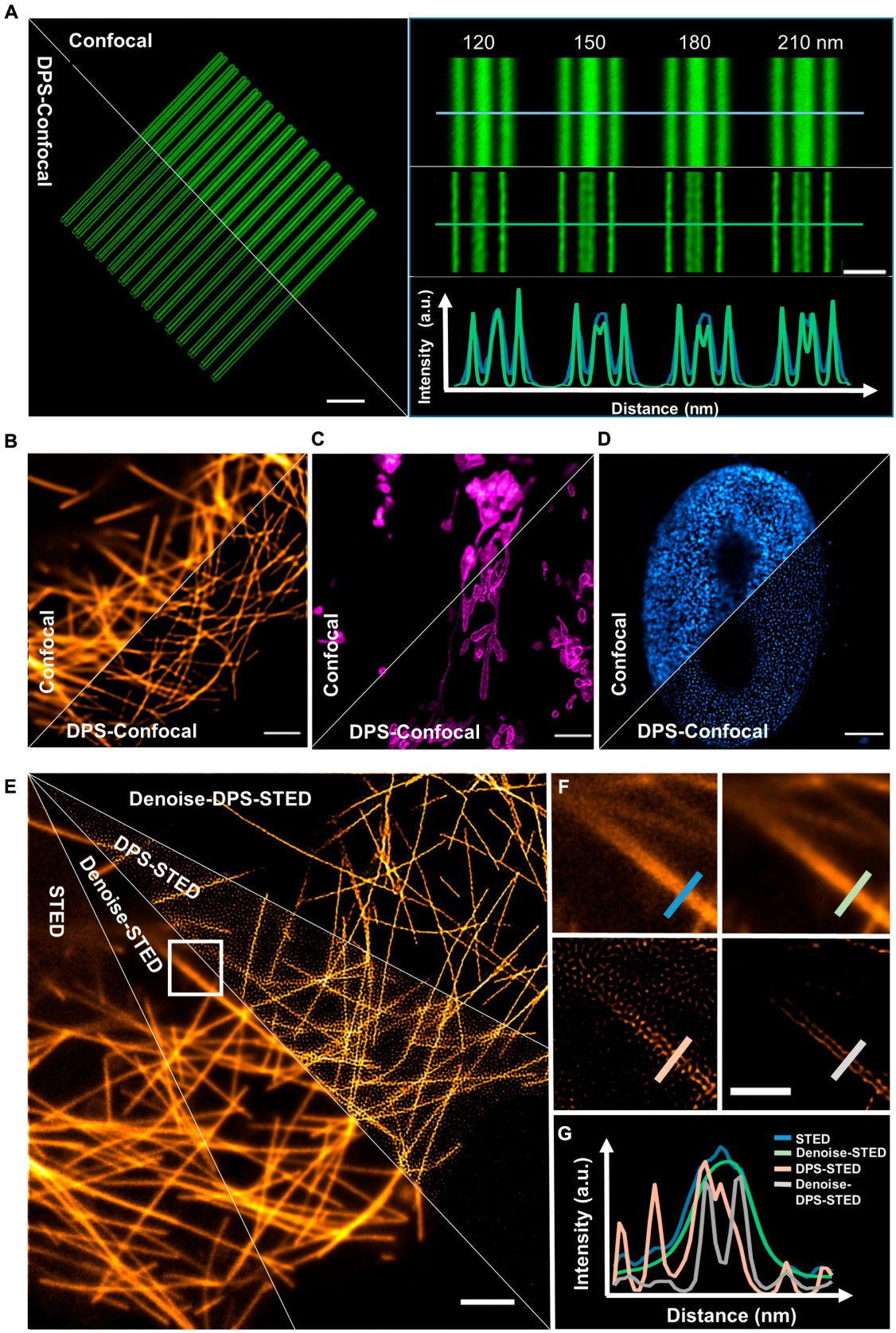

Fig. 8. DPS applications in confocal and STED microscopy modalities. (A) Reconstruction comparison of conventional confocal and DPS-Confocal using spaced line pattern sample separated by steps of 30 nm (from 0 to 390 nm). The region enclosed by the white box is magnified and shown on the right with corresponding intensity profiles. Representative SR images reconstructed by conventional confocal and DPS-Confocal modalities: (B) MTs, (C) mitochondrial outer membrane, and (D) NPCs. (E) A representative MT dataset imaged by conventional STED, Denoise-STED, DPS-STED, and Denoise-DPS-STED. (F) Magnified views of the region enclosed by the white box in (E) are shown on the right. (G) The intensity profiles along corresponding lines in (F). Scale bars, 12 μm (A), 4 μm (B to D), 2 μm (E), and 800 nm (F).

cost of the DPS model limit its high-speed application; for example, a 1,024 × 1,024 field-of-view image requires 6 GB of memory for graphics processing unit. To overcome this restriction, model pruning, model quantization, or knowledge distillation can be introduced to efficiently deploy the DPS model on devices that have limited computational resources. Second, sufficient Nyquist-sampling conditions are required for the cross-modality application of the proposed model because the equivalent PSF processing cannot completely provide the missing frequency information, likely resulting in artifacts in further resolution enhancement operations. A single model solving the SR task of an arbitrary scale factor provides a possible approach to overcoming these constraints [50]. In addition, recent physics-informed studies have emphasized the influence of network complexity and multiobjective loss functions combining data-driven and physics-based losses to enhance accuracy. Training a DPS with an appropriate multiobjective function may further improve its denoising and SR capabilities [26–28].

In conclusion, on the basis of the excellent results obtained using the DPS and Denoise-DPS models, this study highlights the significance of holistic network architecture design considering multitechnique components and imaging processes to handle highly ill-posed tasks (e.g., image degradation), resolving imaging technique trade-off problems (e.g., the trade-offs between spatial resolution, imaging speed, method simplicity, fidelity, and universality) and can thus pave the way to methodological innovations and new biological discoveries.

Acknowledgments

Funding: This work was financially supported by the Natural Science Foundation of Zhejiang Province (LQ23F050010 and LY23F050010); the National Natural Science Foundation of China (62125504, 61827825, and 61975188); the Key Research and Development Program of Zhejiang Province (2020C01116); the Zhejiang Provincial Ten Thousand Plan for Young Top Talents (2020R52001); and the National Key Research and Development Program of China (2021YFF0700302).

Author contributions: W.L. and Z.Y. conceived the study. W.L., Youhua Chen, and C.K. supervised the study. W.L. and Z.Y. designed the DPS and Denoise-DPS algorithms. Yunbo Chen, H.Y., Y.G., and X.L. helped develop the algorithms. W.L., Z.Y., Youhua Chen, J.Z., and C.J. designed the SIM experiments. Z.Y. and Y. Huang designed the confocal and STED experiments. L.J., Y.S., W.T., and Y. Han prepared the samples. Z.Y. composed the figures and videos under the supervision of W.L. W.L. and Z.Y. wrote the manuscript with input from all authors. All the authors discussed the results and commented on the manuscript.

Competing interests: The authors declare that they have no competing interests.

Data Availability

The dataset supporting the conclusions of this study is available at https://github.com/WenjieLab/DPS-framework-for-microscopy. The source code is available at https://github.com/WenjieLab/DPS-framework-for-microscopy.

Supplementary Materials

Figs. S1 to S9
Table S1
Sections S1 to S5
Movies S1 to S4
Supplementary References

References

1. Sahl SJ, Hell SW, Jakobs S. Fluorescence nanoscopy in cell biology. *Nat Rev Mol Cell Biol*. 2017;18(11):685–701.
2. Sigal YM, Zhou R, Zhuang X. Visualizing and discovering cellular structures with super-resolution microscopy. *Science*. 2018;361(6405):880–887.
3. Werner C, Sauer M, Geis C. Super-resolving microscopy in neuroscience. *Chem Rev*. 2021;121(19):11971–12015.
4. Schermelleh L, Ferrand A, Huser T, Eggeling C, Sauer M, Biehlmaier O, Drummen GPC. Super-resolution microscopy demystified. *Nat Cell Biol*. 2019;21(1):72–84.
5. Andrews B, Chang JB, Collinson L, Li D, Lundberg E, Mahamid J, Manley S, Mhlanga M, Nakano A, Schöneberg J, et al. Imaging cell biology. *Nat Cell Biol*. 2022;24(8):1180–1185.
6. Mait JN, Euliss GW, Athale RA. Computational imaging. *Adv Opt Photon*. 2018;10:409–483.
7. Zhao W, Zhao S, Li L, Huang X, Xing S, Zhang Y, Qiu G, Han Z, Shang Y, Sun DE, et al. Sparse deconvolution improves the resolution of live-cell super-resolution fluorescence microscopy. *Nat Biotechnol*. 2022;40(4):606–617.
8. Torres-García E, Pinto-Cámara R, Linares A, Martínez D, Abonza V, Brito-Alarcón E, Calcines-Cruz C, Valdés-Galindo G, Torres D, Jabloñski M, et al. Extending resolution within a single imaging frame. *Nat Commun*. 2022;13(1):7452.
9. Belthangady C, Royer LA. Applications, promises, and pitfalls of deep learning for fluorescence image reconstruction. *Nat Methods*. 2019;16(12):1215–1225.
10. Barbastathis G, Ozcan A, Situ G. On the use of deep learning for computational imaging. *Optica*. 2019;6(8):921–943.
11. Kaderuppan SS, Wong EWL, Sharma A, Woo WL. Smart nanoscopy: A review of computational approaches to achieve super-resolved optical microscopy. *IEEE Access*. 2020;8:214801–214831.
12. von Chamier L, Laine RF, Jukkala J, Spahn C, Krentzel D, Nehme E, Lerche M, Hernández-Pérez S, Mattila PK, Karinou E, et al. Democratising deep learning for microscopy with ZeroCostDL4Mic. *Nat Commun*. 2021;12(1):2276.
13. Huang B, Li J, Yao B, Yang Z, Lam EY, Zhang J, Yan W, Qu J. Enhancing image resolution of confocal fluorescence microscopy with deep learning. *PhotoniX*. 2023;4:2.
14. Ouyang W, Aristov A, Lelek M, Hao X, Zimmer C. Deep learning massively accelerates super-resolution localization microscopy. *Nat Biotechnol*. 2018;36(5):460–468.
15. Speiser A, Müller LR, Hoess P, Matti U, Obara CJ, Legant WR, Kreshuk A, Macke JH, Ries J, Turaga SC. Deep learning enables fast and dense single-molecule localization with high accuracy. *Nat Methods*. 2021;18(9):1082–1090.
16. Wang H, Rivenson Y, Jin Y, Wei Z, Gao R, Günaydın H, Bentolila LA, Kural C, Ozcan A. Deep learning enables cross-modality super-resolution in fluorescence microscopy. *Nat Methods*. 2019;16(1):103–110.
17. Fang L, Monroe F, Novak SW, Kirk L, Schiavon CR, Yu SB, Zhang T, Wu M, Kastner K, Latif AA, et al. Deep learning-based point-scanning super-resolution imaging. *Nat Methods*. 2021;18(4):406–416.

18. Zhao Y, Zhang M, Zhang W, Zhou Y, Chen L, Liu Q, Wang P, Chen R, Duan X, Chen F, et al. Isotropic super-resolution light-sheet microscopy of dynamic intracellular structures at subsecond timescales. *Nat Methods*. 2022;19(3):359–369.
19. Chobola T, Theileis A, Taucher J, Peng T. DELAD: Deep Landweber-guided deconvolution with Hessian and sparse prior. ArXiv. 2022. https://doi.org/10.48550/arXiv.2209.15377
20. Burns Z, Liu Z. Untrained, physics-informed neural networks for structured illumination microscopy. *Opt Express*. 2023;31(5):8714–8724.
21. He Y, Yao Y, He Y, Huang Z, Luo F, Zhang C, Qi D, Jia T, Wang Z, Sun Z, et al. Surpassing the resolution limitation of structured illumination microscopy by an untrained neural network. *Biomed Opt Express*. 2023;14(1):106–117.
22. Qiao C, Zeng Y, Meng Q, Chen X, Chen H, Jiang T, Wei R, Guo J, Fu W, Lu H, et al. Zero-shot learning enables instant denoising and super-resolution in optical fluorescence microscopy. BioRxiv. 2023. https://doi.org/10.1101/2023.02.24.529803
23. Yao J, Hong D, Chanussot J, Meng D, Zhu X, Xu Z. Cross-attention in coupled unmixing nets for unsupervised hyperspectral super-resolution. In: Vedaldi A, Bischof H, Brox T, Frahm JM, editors. *Computer vision – ECCV 2020. ECCV 2020. Lecture notes in computer science*. Cham: Springer; 2020. p. 208–224.
24. Yao J, Zhang B, Li C, Hong D, Chanussot J. Extended vision transformer (ExViT) for land use and land cover classification: A multimodal deep learning framework. *IEEE Trans Geosci Remote Sens*. 2023;61:1–15.
25. Hong D, Yao J, Li C, Meng D, Yokoya N, Chanussot J. Decoupled-and-coupled networks: Self-supervised hyperspectral image super-resolution with subpixel fusion. *IEEE Trans Geosci Remote Sens*. 2023;61:5527812.
26. Roy AM, Guha S. A data-driven physics-constrained deep learning computational framework for solving von Mises plasticity. *Eng Appl Artif Intell*. 2023;122:Article 106049.
27. Zhao X, Gong Z, Zhang Y, Yao W, Chen X. Physics-informed convolutional neural networks for temperature field prediction of heat source layout without labeled data. *Eng Appl Artif Intell*. 2023;117:Article 105516.
28. Roy AM, Bose R, Sundararaghavan V, Arróyave R. Deep learning-accelerated computational framework based on physics informed neural network for the solution of linear elasticity. *Neural Netw*. 2023;162:472–489.
29. Qiao C, Li D, Guo Y, Liu C, Jiang T, Dai Q, Li D. Evaluation and development of deep neural networks for image super-resolution in optical microscopy. *Nat Methods*. 2021;18(2):194–202.
30. Laine RF, Arganda-Carreras I, Henriques R, Jacquemet G. Avoiding a replication crisis in deep-learning-based BioImage analysis. *Nat Methods*. 2021;18(10):1136–1144.
31. Heil BJ, Hoffman MM, Markowetz F, Lee SI, Greene CS, Hicks SC. Reproducibility standards for machine learning in the life sciences. *Nat Methods*. 2021;18(10):1132–1135.
32. Hoffman DP, Slavitt I, Fitzpatrick CA. The promise and peril of deep learning in microscopy. *Nat Methods*. 2021;18(2):131–132.
33. Bredell G, Erdil E, Weber B, Konukoglu E. Wiener guided DIP for unsupervised blind image deconvolution. Paper presented at: IEEE/CVF Winter Conference on Applications of Computer Vision (WACV); 2023 Jan 02–07; Waikoloa, HI, USA.
34. Ren D, Zhang K, Wang Q, Hu Q, Zuo W. Neural blind deconvolution using deep priors. Paper presented at: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2020 Jun 13–19; Seattle, WA, USA.
35. Le QV. Building high-level features using large scale unsupervised learning. Paper presented at: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing; 2013 May 26–31; Vancouver, BC, Canada.
36. He YZ, Yao J, Liu L, Gao Y, Yu J, Ye S, Li H, Zheng W. Self-supervised deep-learning two-photon microscopy. *Photonics Res*. 2023;11(1):1–11.
37. Zhao H, Gallo O, Frosio I, Kautz J. Loss functions for image restoration with neural networks. *IEEE Trans Comput Imaging*. 2017;3(1):47–57.
38. Huang X, Fan J, Li L, Liu H, Wu R, Wu Y, Wei L, Mao H, Lal A, Xi P, et al. Fast, long-term, super-resolution imaging with hessian structured illumination microscopy. *Nat Biotechnol*. 2018;36(5):451–459.
39. Haris M, Shakhnarovich G, Ukita N. Deep back-projection networks for super-resolution. Paper presented at: IEEE/CVF Conference on Computer Vision and Pattern Recognition; 2018 Jun 18–23; Salt Lake City, UT, USA.
40. Xiao X, Lian S, Luo Z, Li S. Weighted res-UNet for high-quality retina vessel segmentation. Paper presented at: 2018 9th international conference on information Technology in Medicine and Education (ITME); 2018 Oct 19–21; Hangzhou, China.
41. Johnson J, Alahi A, Fei-Fei L. Perceptual losses for real-time style transfer and super-resolution. arXiv:1603.08155 (2016).
42. Xue W, Zhang L, Mou X, Bovik AC. Gradient magnitude similarity deviation: A highly efficient perceptual image quality index. *IEEE Trans Image Process*. 2014;23(2):684–695.
43. Descloux A, Grußmayer KS, Radenovic A. Addendum: Parameter-free image resolution estimation based on decorrelation analysis. *Nat Methods*. 2020;17(10):1061–1063.
44. Heide F, Heidrich W, Wetzstein G. Fast and flexible convolutional sparse coding. Paper presented at: IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2015 Jun 07–12; Boston, MA, USA.
45. Chen Y, Liu W, Zhang Z, Zheng C, Huang Y, Cao R, Zhu D, Xu L, Zhang M, Zhang YH, et al. Multi-color live-cell super-resolution volume imaging with multi-angle interference microscopy. *Nat Commun*. 2018;9(1):4818.
46. Fagotto F, Aslemarz A. EpCAM cellular functions in adhesion and migration, and potential impact on invasion: A critical review. *Biochim Biophys Acta Rev Cancer*. 2020;1874(2):Article 188436.
47. Li D, Shao L, Chen BC, Zhang X, Zhang M, Moses B, Milkie DE, Beach JR, Hammer JA III, Pasham M, et al. Extended-resolution structured illumination imaging of endocytic and cytoskeletal dynamics. *Science*. 2015;349(6251):aab3500.
48. Descloux A, Grußmayer KS, Radenovic A. Parameter-free image resolution estimation based on decorrelation analysis. *Nat Methods*. 2019;16(9):918–924.
49. Fang F, Li J, Zeng T. Soft-edge assisted network for single image super-resolution. *IEEE Trans Image Process*. 2020;29:4656–4668.
50. Hu X, Mu H, Zhang X, Wang Z, Tan T, Sun J. Meta-SR: A magnification-arbitrary network for super-resolution. Paper presented at: IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2019 Jun 15–20; Long Beach, CA, USA.