Optics Communications 537 (2023) 129431

Contents lists available at ScienceDirect

Optics Communications

journal homepage: www.elsevier.com/locate/optcom

Reconstruction of structured illumination microscopy with an untrained neural network

Xin Liu [a], Jinze Li [b], Xiang Fang [a], Jiaoyue Li [a], Juanjuan Zheng [a,c], Jianlang Li [a], Nauman Ali [a], Chao Zuo [a,d,*], Peng Gao [a,**], Sha An [a,**]

[a] School of Physics, Xidian University, Xi'an, China
[b] School of Optoelectronic Engineering, Xidian University, Xi'an, 710071, China
[c] State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, 710119, China
[d] Smart Computational Imaging Laboratory (SCILab), School of Electronic and Optical Engineering, Nanjing University of Science and Technology, Nanjing, JS, China

ARTICLE INFO

Keywords:
Structured illumination microscopy
Deep learning
Neural network
Super-resolution
Image reconstruction

ABSTRACT

Structured illumination microscopy (SIM) is one of super-resolution optical microscopic techniques, and it has been widely used in biological research. In this paper, a physics-driven deep image prior framework for super-resolution reconstruction of SIM (entitled DIP-SIM) is proposed. DIP-SIM does not rely on a large number of labeled data, and the output becomes more interpretable due to the intrinsic constraint of a physical model. Both the simulation and experiment verify that DIP-SIM can reconstruct a super-resolution image with a quality comparable to conventional SIM. Of note, it allows for super-resolution reconstruction from three raw images for two-orientation SIM and four raw images for three-orientation SIM, and hence it has a much faster imaging speed and lower photobleaching compared with the traditional SIM. We can envisage that the proposed method can be applied to chemistry and biomedical fields, etc.

1. Introduction

Optical microscopy is an indispensable tool for exploring the unknown microscopic world. Due to the restriction of the diffraction limit, the spatial resolution of conventional optical microscopy is limited to $\lambda/2$, which is generally about 200 nm [1–4]. As one of the key advances in optical microscopy during the past three decades, super-resolution (SR) microscopic approaches have been proposed to enhance the spatial resolution to tens or a few nanometers. Representative SR techniques include single-molecule localization microscopy (PALM/STORM) [5,6], stimulated emission depletion microscopy (STED) [7,8], and structured illumination microscopy (SIM) [9,10]. Among them, SIM has the advantages of fast-speed, non-invasive, low phototoxicity, and photobleaching. Hence SIM has been acting as an essential imaging tool in biomedical research, especially for live cell imaging [11–13].

In SIM, samples are illuminated with periodic sinusoidal patterns of different phase shifts and different orientations, and the intensity images (with moiré pattern) are captured with a 2D sensor. Several algorithms, such as Fair-SIM [14], Open-SIM [15], and CC-SIM [16], have been developed to reconstruct a SR image from the recorded raw images. The majority of the algorithms perform SR reconstruction in the frequency domain. In the implementation, these algorithms decouple the frequency spectra along different diffraction orders of structured illumination, shift them back to their original position, and combine them in the frequency domain. An inverse Fourier transform on the synthesized spectrum yields a SR image. In parallel with the frequency-domain approaches, spatial-domain approaches have been proposed to directly process SIM data in the spatial domain [17,18]. Compared to frequency-domain algorithms, this approach uses the same number of frames to achieve a comparable resolution but with a much faster processing speed. However, for these methods, the incorrect estimation of experimental parameters (optical transfer function, structured illumination direction, phase, etc.) causes the inappropriate superimposition of spectra along different diffraction orders of structured illumination, which degrades the final image quality. Conventional SIM turns nine raw images (three orientations and three phase shifts) into a SR image with isotropic spatial resolution. As a valuable effort to reduce the raw image, a frequency domain SIM-RA based on an ordinary least square technique allows for SR image reconstruction using four raw SIM images [19].

In recent years, deep learning [20,21] has been demonstrated as a powerful tool in solving various inverse problems through training

* Corresponding author at: School of Physics, Xidian University, Xi'an, China.
** Corresponding authors.
E-mail addresses: zuochao@njust.edu.cn (C. Zuo), peng.gao@xidian.edu.cn (P. Gao), ansha@xidian.edu.cn (S. An).

https://doi.org/10.1016/j.optcom.2023.129431
Received 24 November 2022; Received in revised form 7 February 2023; Accepted 12 March 2023
Available online 15 March 2023

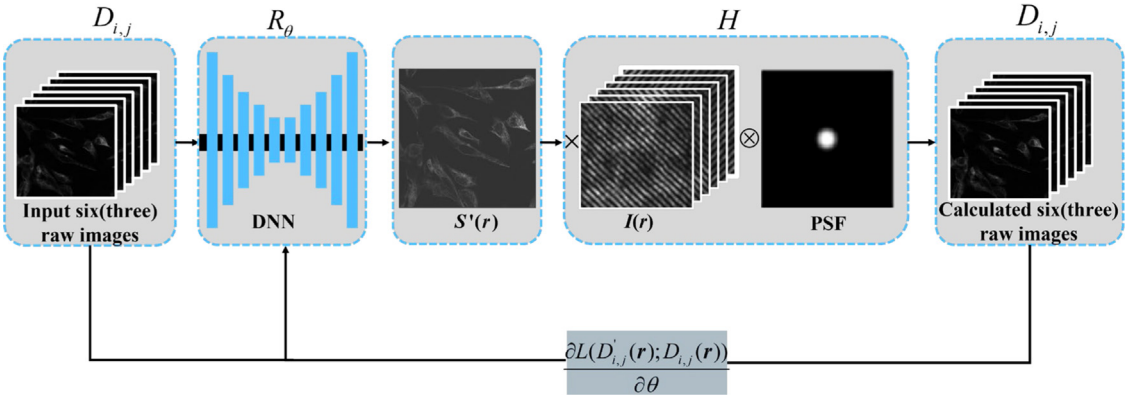

Fig. 1. The schematic diagram of DIP-SIM model. $D_{i,j}(\boldsymbol{r})$ indicates six (three) raw images as input to the network R_θ. R_θ indicates the estimation of a SR image $S'(\boldsymbol{r})$. $D'_{i,j}(\boldsymbol{r})$ are the calculated six (three) raw images. $H\{\cdot\}$ is the operator described with Eq. (3).

a network with a large quantity of paired images. The incorporation of deep learning into SIM has invited many ideas and breakthroughs in SR reconstruction of SIM. To cite a few, Qiao et al. proposed a deep Fourier channel attention network (DFCAN), which exhibited robust image reconstruction under low signal-to-noise ratio (SNR) conditions [22]. Shah et al. proposed and demonstrated super-resolution residual encoder–decoder SIM (SR-REDSIM) and residual encoder–decoder fairSIM (RED-fairSIM) for high-robustness SIM reconstruction under different noise levels. The authors proved that RED-fairSIM has better generalization ability than SR-REDSIM under different SIM imaging conditions [23]. Jin et al. proposed a DL-SIM method, which reduces the number of high-quality raw images required for reconstructing SIM images and ultimately improves the temporal resolution [24]. The majority of deep-learning based SIM reconstruction approaches are based on data-driven neural networks, and therefore, they require a large pair of inputs (raw images) and desired outputs (ground truth) to optimize network parameters. The raw images can be acquired on a real SIM system using various samples and experimental conditions. However, the high-quality outputs corresponding to the raw images are difficult to obtain. Furthermore, the training processes of different network frameworks may take hours or even days. Although simulated data can be used as the training set, it is sometimes difficult to model variants of test samples, and the mapping function is only suitable for objects with the same prior as the simulated training set [25]. Burns et al. proposed a physics-informed neural network (PINN), which used nine sub-frames (3 orientations and 3 phases) for the super-resolution reconstruction of SIM images [26]. Ulyanov et al. designed a Deep Image Prior (DIP) framework that uses an untrained network as a constraint to solve common inverse problems, considering a well-designed network framework has an implicit bias in images [27]. The DIP has a preponderant advantage in that it does not need pre-training with a large amount of labeled data.

Inspired by the idea of DIP, we present a physics-driven deep image prior framework for the super-resolution reconstruction of SIM (entitled DIP-SIM). In DIP-SIM, the neural network parameters are updated by optimizing the loss between the experimental raw images and the one predicted with the imaging model. It allows for SR reconstruction from six or even three frames of raw data for two-orientation SIM, and four raw images for three-orientation SIM. Hence it has a much faster imaging speed and lower photobleaching compared with the traditional SIM.

2. Methods

Structured illumination microscopy utilizes moiré patterns to down-shift high-frequencies of the object wave into the frequency support of the microscope and eventually enhances the resolution of the microscope. In SIM, the intensity distribution of structured illumination can be written as:

$$I_{i,j}(\boldsymbol{r}) = I_0\left[1 + \frac{m}{2}\cos\left(2\pi k_s \cdot \boldsymbol{r} + \varphi_{i,j}\right)\right], \tag{1}$$

where $\boldsymbol{r}$ is the spatial coordinate. $I_0(\boldsymbol{r})$ is the dc term of the periodic illumination. m is the modulation factor. $\varphi_{i,j}$ is the jth phase shifting for the ith fringe orientation. k_s is the spatial frequency. In spatial light modulator (SLM) or digital micromirror device (DMD) based SIM, k_s and $\varphi_{i,j}$ are digitally set by the SLM or DMD, so that they can be measured once for all.

Under the illumination of the fringe pattern $I_{i,j}(\boldsymbol{r})$, a fluorescent labeled sample $S(\boldsymbol{r})$ is imaged by the microscopic system, and the recorded intensity image corresponds to the convolution of $S(\boldsymbol{r}) \cdot I_{i,j}(\boldsymbol{r})$ with the intensity point spread function (PSF) $h_D(\boldsymbol{r})$ of the system:

$$D_{i,j}(\boldsymbol{r}) = \left[S(\boldsymbol{r}) \cdot I_{i,j}(\boldsymbol{r})\right] \otimes h_D(\boldsymbol{r}). \tag{2}$$

Herein, a physics-driven deep image prior framework structured illumination microscopy (entitled DIP-SIM) is proposed to reconstruct a SR image from $D_{i,j}(\boldsymbol{r})$, as shown in Fig. 1.

Firstly, the estimation of a SR image $S'(\boldsymbol{r})$ is performed with a Encoder–Decoder CNN network [27].

$$S'(\boldsymbol{r}) = R_\theta(D_{i,j}(\boldsymbol{r})), \tag{3}$$

where θ indicates the parameter setting of the network R_θ. This network [28] consists of five encoder base units, five decoder base units, and long skip connections implemented via concatenation, as shown in Fig. 2. The encoder base unit mainly contains convolution blocks (3×3 convolution (stride 2) + batch normalization + Leaky ReLU + 3×3 convolution (stride 1) + batch normalization + Leaky ReLU), with which image features are extracted. The down-sampling is performed by setting the sliding stride of the convolutional layer to 2, while up-sampling is accomplished by the Lanczos operator. The decoder basic unit mainly contains convolution blocks (3×3 convolution (stride 1) + batch normalization + Leaky ReLU +1×1 convolution (stride 1) + batch normalization + Leaky ReLU), with which image information is recovered. The long skip connection uses convolution blocks (1×1 convolution (stride 1) + batch normalization + Leaky ReLU). The function of the long skip connection is to fuse shallow information with deep information to provide more semantic information for the decoding process. In the output layer, an estimated high-quality SIM image is reconstructed by using 1×1 convolution and sigmoid.

Then, the output of the neural network, i.e., $S'(\boldsymbol{r})$, is used to calculate the intensity distribution of the raw images $D'_{i,j}(\boldsymbol{r})$ under the structured illumination $I_{i,j}(\boldsymbol{r})$:

$$D'_{i,j}(\boldsymbol{r}) = H(S'(\boldsymbol{r})), \tag{4}$$

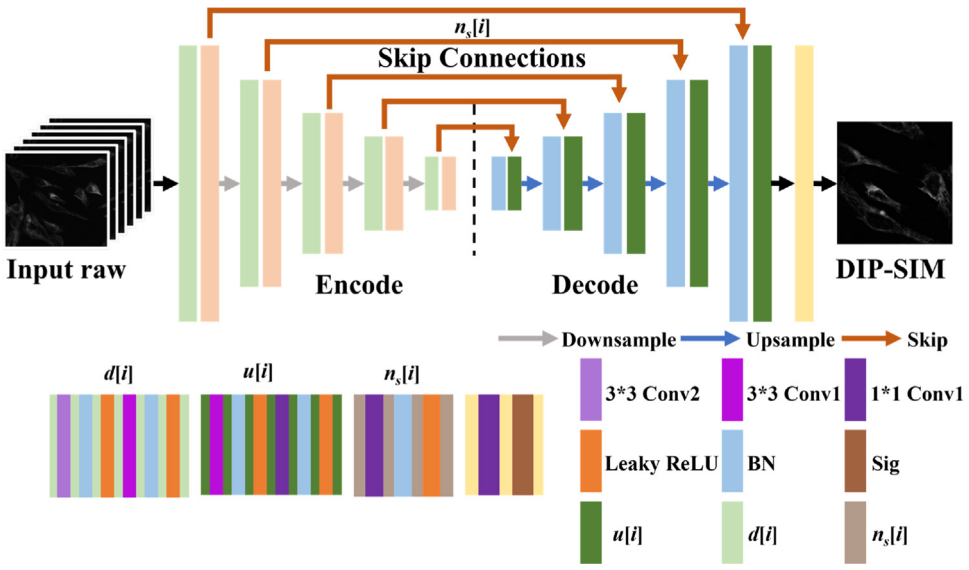

Fig. 2. Encoder–decoder architecture with long skip connections. In the right side of the figure, Conv2 represents the convolutional layer setting the stride to 2, BN represents batch normalization, Sig represents the sigmoid activation function, and Conv1 represents the convolutional layer setting the stride to 1. $d[i]$, $u[i]$, $n_s[i]$ represents at the ith encoder base unit, decoder base unit and skip connections, respectively.

where $H\{.\}$ is the operator to mimic the SIM imaging process governed by the physical model in Eq. (2). Specifically, $S'(\boldsymbol{r})$ is multiplied with the structured illumination $I_{i,j}(\boldsymbol{r})$, and is further convoluted with the PSF $h_D(\boldsymbol{r})$ of the system.

Finally, the neural network parameters are updated by optimizing the loss between the experimental raw images $D_{i,j}(\boldsymbol{r})$ and the calculated raw images $D'_{i,j}(\boldsymbol{r})$. For this purpose, the combination of SSIM and mean square error is used as the loss function of DIP-SIM:

$$R_{\theta^*} = \arg\min \| D'_{i,j}(\boldsymbol{r}) - D_{i,j}(\boldsymbol{r}) \|^2 + \delta \times [1 - SSIM(D'_{i,j}(\boldsymbol{r}), D_{i,j}(\boldsymbol{r}))], \quad (5)$$

where δ is a weighting factor used to balance the relative contribution of the structural similarity index (SSIM) loss function and the mean square error (MSE) loss function, which is set to 0.1 in most cases in this paper. The parameter of the neural network is optimized as the loss gradually decreases. Eventually, the feasible mapping function $R_{\theta*}$ between $S_{SR}(\boldsymbol{r})$ and $D_{i,j}(\boldsymbol{r})$ can be generated:

$$S_{SR}(\boldsymbol{r}) = R_{\theta^*}(D_{i,j}(\boldsymbol{r})). \quad (6)$$

In DIP-SIM, the SR reconstruction of the sample image is obtained by the interplay between $H\{.\}$ and R_θ, which is intrinsically constrained by the priors of $D_{i,j}(\boldsymbol{r})$ and the physical model in the DIP-SIM. It can be seen that Eq. (5) does not contain the ground truth $S(\boldsymbol{r})$ that are often high-resolution images obtained with a high-NA objective. It means that DIP-SIM does not require ground truth $S(\boldsymbol{r})$ during the training process. Therefore, compared to conventional data-driven based end-to-end neural networks [18,22–25,29,30], the proposed DIP-SIM network does not require the data pairs for pre-training since it has implicit biases in images [27]. The network can obtain prior knowledge from several raw images instead of large datasets of images, significantly reducing the experimental data as required by conventional data-driven networks.

Different from the conventional SIM reconstruction, for which nine raw images (with 3 directions × 3 phase shifts) are required to reconstruct a super-resolution SIM image, DIP-SIM can reconstruct a SR image with fewer raw images. In the following, we will demonstrate DIP-SIM reconstruction with six raw images (termed DIM-SIM-6) and three raw images (termed DIP-SIM-3) for two-orientation SIM. Specifically, for DIP-SIM-6, two groups of structured illumination with two orthogonal orientations and three phase shifts of 0, $2\pi/3$, and $4\pi/3$ for each were used. For DIP-SIM-3, three raw images were used. Two raw images were obtained using the structured illumination along the x direction under the phase shift of 0 and $2\pi/3$. The third image is the raw image obtained using the structured illumination along the y direction and the phase shift of 0.

Algorithm 1 SIM super-resolution reconstruction. The paper used default values of N = 1000, $\alpha = 0.01$, $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\delta = 0.1$

Input: The six (three) raw images $D_{i,j}(\boldsymbol{r})$ and H

Output: $S'(\boldsymbol{r}) = R_{\theta^*}(D_{i,j}(\boldsymbol{r}))$

1. initialize: randomly initialize the parameters θ in the neural network R_θ
2. for step = 1, 2,…,N do
3. $D'_{i,j}(\boldsymbol{r}) \leftarrow H(R_\theta(D_{i,j}(\boldsymbol{r})))$
4. $\ell_\theta \leftarrow MSE(D'_{i,j}(\boldsymbol{r}), D_{i,j}(\boldsymbol{r})) + \delta \times [1 - SSIM(D'_{i,j}(\boldsymbol{r}), D_{i,j}(\boldsymbol{r}))]$
5. $\theta \leftarrow Adam(\nabla_\theta \ell_\theta, \alpha, \beta_1, \beta_2)$
6. end for

The network was implemented based on the PyTorch platform (version 1.10.1, coded with Python 3.7.5). We adopted the Adam optimizer with a learning rate of $\alpha = 0.01$, $\beta_1 = 0.9$, and $\beta_2 = 0.999$ to update the parameters in the neural network. The code was deployed on the server with an Intel Core i9-10980XE CPU, 128 GB RAM, and an NVIDIA GeForce RTX 3090 running the CentOS Linux operating system. The main progress is illustrated in Algorithm 1.

3. Results

The performance of DIP-SIM has been demonstrated with both simulation and experiment. The simulation is ideal for quantifying the reconstruction quality of DIP-SIM, while the experiment is suited to verify the practicability of the DIP-SIM.

3.1. Simulation

In the first simulation,we have simulated DIP-SIM imaging for two-orientation SIM, for which the images of various biological samples selected from the BioSR dataset are used as ground truth images [22]. To simulate raw intensity images, the images are multiplied by different structured illumination and convolved with a known PSF, as described in Eq. (2). The parameters of the simulated SIM system are the same

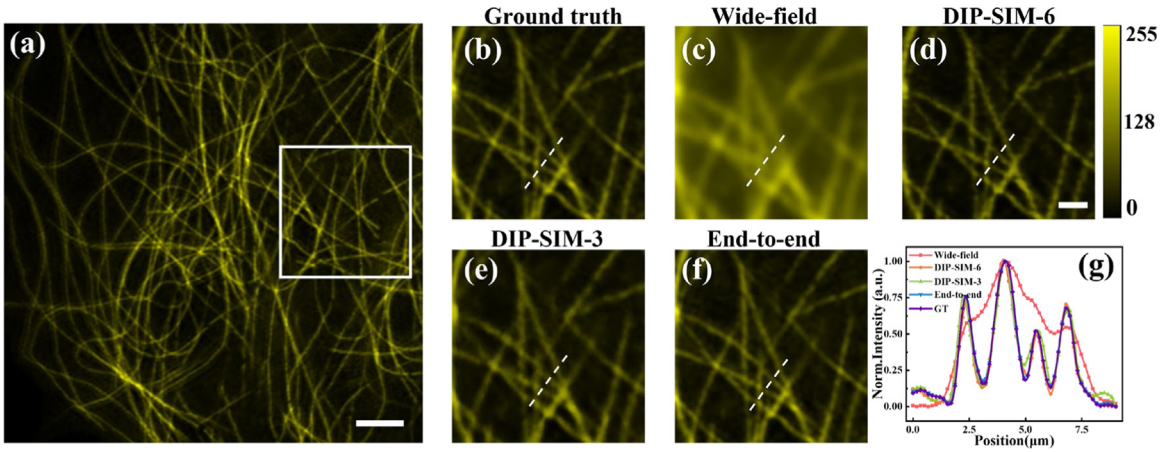

Fig. 3. Wide-field and DIP-SIM-6 of a microtubule image from BioSR dataset. (a) Ground truth (GT). (b) Magnified view of the region in (a) indicated with a white box. (c) Wide-field image, obtained by simply averaging six raw images. (d)~(f) Images reconstructed by DIP-SIM-6 and DIP-SIM-3, and conventional end-to-end network. (g) The intensity profiles along the dash lines in (b–f). Scale bar in (a) is 10 μm. Scale bar in (b–f) is 3 μm.

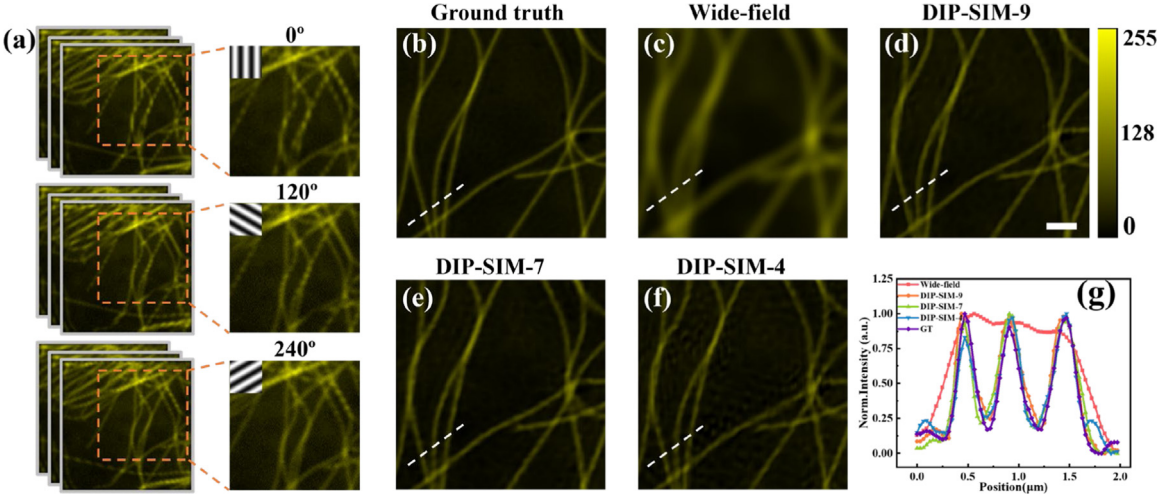

Fig. 4. Three-orientation DIP-SIM imaging of a microtubule image. (a) Raw images obtained with the SI with an azimuth of 0°, 120°, and 240°, respectively. (b–f) Images of ground truth (GT), wide-field, DIP-SIM-9, DIP-SIM-7, DIP-SIM-4. (g) The intensity profiles along the dash lines in (b–f). Scale bar in (d) is 3 μm.

as the experimental setting described in Section 3.2. Specifically, the structured illumination pattern is characterized with $I_0 = 1$, $m = 0.75$, and $k_s = 1.5 \times 10^{-6}$ μm^{-1} [31] obtained by the raw image calculation [15]. Taking the fringe orientations of +45°, −45° and phase shifts of 0, $2\pi/3$, and $4\pi/3$, we can obtain six different structured illumination patterns. The ground-truth images and the simulated raw intensity images are used as training pairs. First of all, a comparison of the proposed DIP-SIM and the conventional data-driven-based end-to-end network approach was carried out. An end-to-end network, which uses the same neural network model but without using a physical model, was used to reconstruct a SR image via a mapping function obtained by fitting the training pairs. The output of the end-to-end network is shown in Fig. 3(f). For performance quantification, Fig. 3(a) shows a high-resolution ground truth image of a microtubule image selected from BioSR dataset. The image of a sub-region indicated with a white box in Fig. 3(a) is magnified and shown in Fig. 3(b). The SR images, reconstructed using six raw images by DIP-SIM-6 and three raw images by DIP-SIM-3, are shown in Fig. 3(d) and (e), respectively. Compared to the wide-field image, both the DIP-SIM-6 and DIP-SIM-3 images have a higher spatial resolution as shown in Fig. 3.

Furthermore, the reconstructed images of DIP-SIM-3 and DIP-SIM-6 were compared with that of the end-to-end network. It turns out that the MSE of the reconstructed images by the end-to-end network, DIP-SIM-3, and DIP-SIM-6 with respect to the ground-truth image is 0.04, 0.06, and 0.04, respectively. Furthermore, a cut-line was extracted from the same position of the images in Fig. 3(b–f), and the intensity distributions along the lines are shown in Fig. 3(g). The comparison confirms that DIP-SIM-6 and DIP-SIM-3 can resolve finer structures of microtubules than the wide-field microscopy. As expected, compared to the end-to-end network approach, both DIP-SIM-6 and DIP-SIM-3 can obtain a high-quality SR-SIM image without a large amount of labeled data.

In the second simulation, we have demonstrated that the DIP-SIM approach can be applied to three-orientation SIM, which has an isotropic resolution enhancement. The microtubule sample was used as the test sample. Three groups of raw intensity images (see Fig. 4(a)) were simulated using the structured illumination (SI) with an azimuth of 0°, 120°, and 240°, and each group contains three phase-shifted raw images with a phase shift increment of $2\pi/3$. The images of the GT, wide-field, DIP-SIM-9, DIP-SIM-7, and DIP-SIM-4 are shown in Fig. 4(b–f), respectively. Compared to the wide-field image obtained by simply averaging nine raw images, the images reconstructed by DIP-SIM-9, DIP-SIM-7, and DIP-SIM-4 have a distinctly-enhanced spatial resolution. The intensity distributions along the line crossing the same position in Fig. 4(b–f) were extracted and compared in Fig. 4(g). It is distinct that the DIP-SIM-9, DIP-SIM-7, and DIP-SIM-4 images show similar structures, among which the two closely-placed tubules with a distance of 0.45 μm from are well separated. While, the quality of the

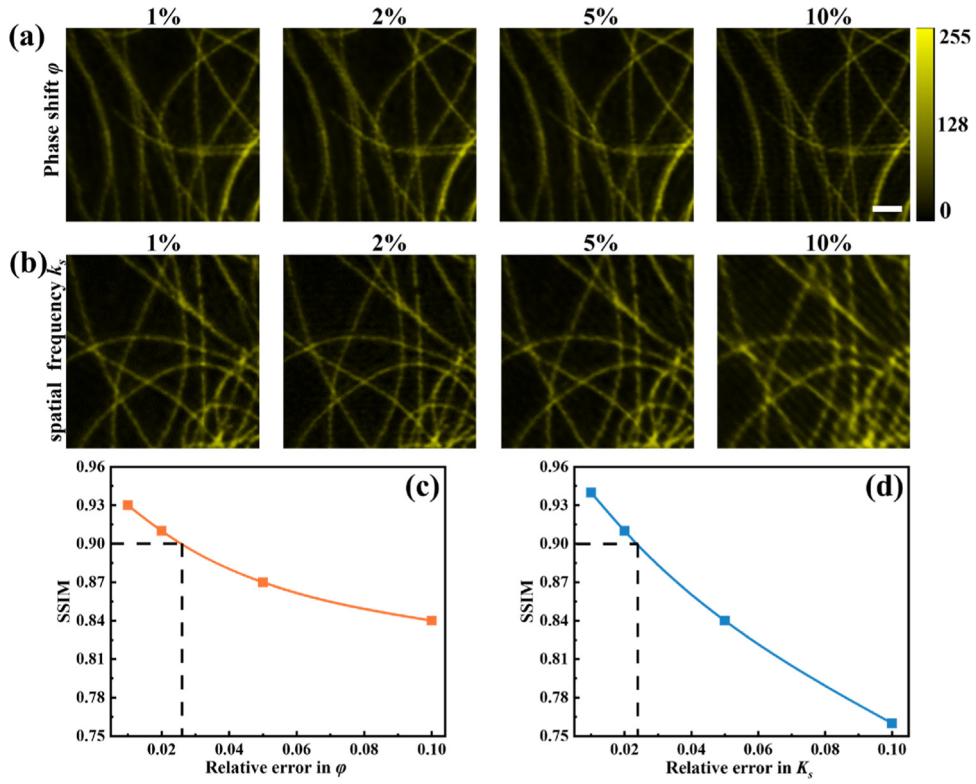

Fig. 5. Effect of the estimation error of SI parameters on DIP-SIM-6 reconstruction. (a)~(b) The DIP-SIM-6 images reconstructed using the phase shifts and the spatial frequency with different error levels. (c)~(d) The SSIM of the DIP-SIM-6 images with respect to the ground truth for different error levels. The true phase shift of 0, $2\pi/3$, and $4\pi/3$ and the spatial frequency of 1.5×10^{-6} m^{-1}. The scale bar in (a) is 3 μm.

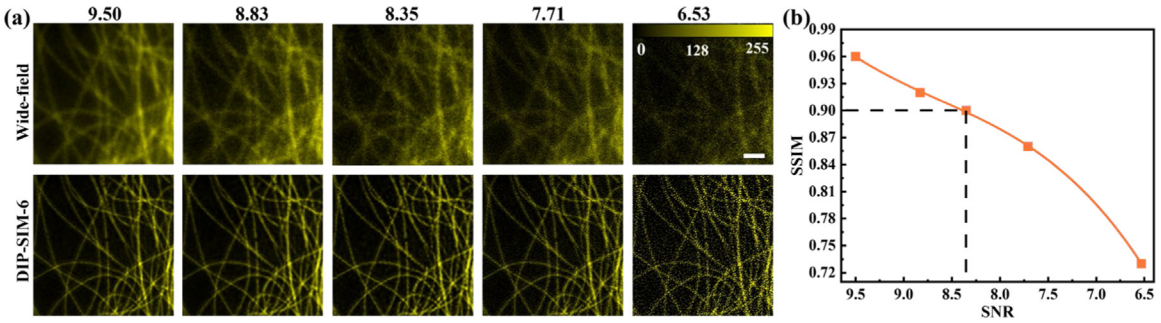

Fig. 6. Effect of Poisson noise on DIP-SIM-6 reconstruction. (a) The wide-field and DIP-SIM-6 images at five different Poisson noise levels. (b) The SSIM versus the SNRs of the raw images. The scale bar in (a) is 3 μm. The curve in (b) is the three-order polynomial fit.

reconstructed images decreases with the raw image number, which is reflected in the fact that DIP-SIM-9, DIP-SIM-7, and DIP-SIM-4 images have their SSIMs of 0.97, 0.95, and 0.93, respectively. If the number of raw images is further reduced, the artifacts of the reconstruction will become overwhelming since the analytic solutions are uncertain (not unique) for the constraint of fewer than four raw images [32].

In the real implementation, the parameters of the structured illumination are calculated from the raw images using the traditional algorithm [15]. Errors may occur when the SNR of raw images is low. We examined the performance of DIP-SIM against the estimation error of the phase shift and spatial frequency. DIP-SIM-6 imaging of the microtubular sample was simulated using Eqs. (1) and (2). Then, DIP-SIM-6 images were reconstructed using the phase shift and the spatial frequency with 1%, 2%, 5%, 10% deviation from their true values, and the results are shown in Fig. 5(a) and (b), respectively. For a better quantitative evaluation, the SSIM of the DIP-SIM-6 images with respect to the ground truth for different error levels of the phase shift and spatial frequency was calculated, as shown in Fig. 5(c–d). It is distinct that the larger the error is, the higher artifacts (residual fringe) appear in DIP-SIM images. DIP-SIM-6 are tolerant to 2.5% and 2% error on the phase shifts and spatial frequency if a SSIM of 0.9 is used as the threshold.

To follow up, the same microtubular sample from BioSR was used to quantify the robustness of the proposed DIP-SIM against background noise. Fig. 6(a) shows the pairs of the wide-field and DIP-SIM-6 images reconstructed from different sets of raw images at five different Poisson noise levels of 9.50, 8.83, 8.35, 7.71, and 6.53, respectively. The noise level was quantified by the signal-to-noise (SNR) in the raw images. The SSIM of the DIP-SIM-6 images at different noise levels with respect to the ground truth was calculated, as shown in Fig. 6(b). The results show that the SSIM decreases with the SNR of the raw images, while the SSIM remains still above 0.9 as long as the signal to noise (SNR) is above 8.4. We find from a further study that DIP-SIM-6 is inferior to the conventional SIM algorithm on suppression of the out-of-focus

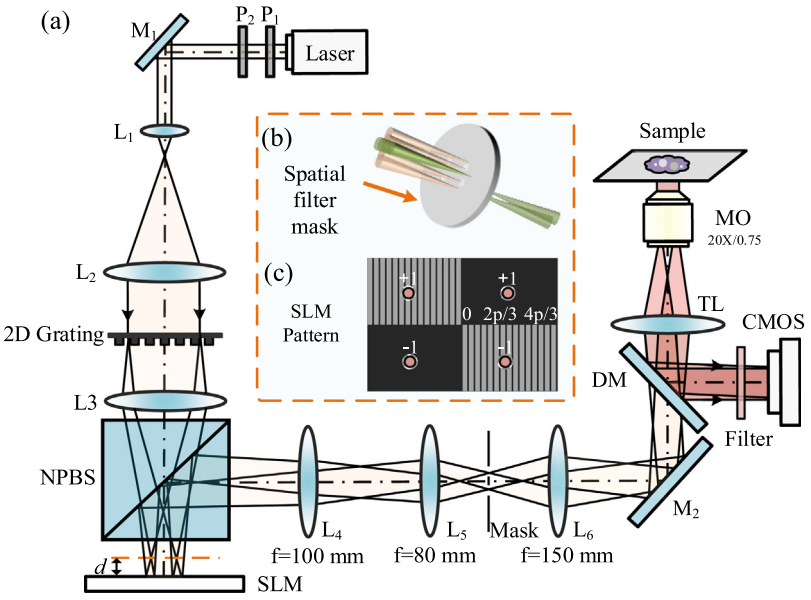

Fig. 7. The schematic diagram of large-field SIM. (a) Experimental setup. (b) Pinhole filtering of the ±1st spectra after loading the binary gratings to SLM. (c) Phases loaded on the SLM to generate fringe patterns. DM, dichroic mirror; L_1–L_6, achromatic lens; M_1–M_2, mirror; MO, micro-objective; NPBS, non-polarizing beam splitter; P_1–P_2, polarizer; SLM, spatial light modulator; TL, tube lens.

artifact, considering the conventional SIM has an intrinsic background removal capability via phase-shifting calculation.

3.2. Experiments

The schematic diagram of two-orientation SIM setup is shown in Fig. 7(a). A 532 nm diode laser (1875-532L, Laserland, Wuhan, China) is used as the light source, and the intensity of the illumination beam is adjusted by rotating the polarizer P_1. A polarizer P_2 located in the beam path converts the illumination into horizontally polarized light, ensuring the maximum modulation efficiency of the spatial light modulator (SLM). The illumination beam is further expanded by a telescope system $L_1 - L_2$ so that a uniform illumination is generated. A 2D grating, of which the ±1st diffraction orders have the highest diffraction efficiency, is placed on the beam path to generate structured illumination. After being Fourier converted by the lens L_3, the spectrum of the illumination appears on the plane with a distance d from a SLM (1920 × 1200 pixels, pixel size 8 μm, HDSLM80R, UPOLabs Co., Ltd., China). The +1st diffraction order of the grating along the x (y) direction is phase shifted by the SLM, while the ±1st diffraction orders along the y (x) are diffracted off by the binary grating loaded on SLM. These spectra are further imaged onto the filter mask by the telescope system $L_4 - L_5$. The mask contains four circular holes (diameter 350 μm, spacing 5.5 mm) that preserve only the spectrum of the ±1st orders along the x and y directions, as shown in Fig. 7(b). The circular holes will block the diffraction orders once being diffracted by the binary grating loaded on SLM. Furthermore, two polarizers with polarization azimuth +45° (−45°) are placed on the spectrum of the ±1st orders in the x and y directions, respectively. Such polarization setting avoids the interference of the beams in the x and y directions. The resulting structured illumination fringes are imaged to the sample plane by two telescope systems ($L_5 - L_6$ and TL-MO). The phase mask in Fig. 7(c) will be used to generate selecting fringe orientation and perform phase-shifting. The grayscale images corresponding to the 0, $2\pi/3$, and $4\pi/3$ phase shifts are sequentially loaded into the SLM to generate structured light illumination with different phase shifts in the x (y) direction. The intensity images of three phase shifts in different directions are recorded by a CMOS (4096 × 3000 pixels, pixel size 3.45 μm, Basler ace acA4112-20um, Basler Vision Technology (Beijing) Co., Ltd., China).

DIP-SIM was verified by two-orientation SIM imaging 240 nm-diameter fluorescent beads (RF240C, emission-peak wavelength 580 nm, Shanghai Huge Biotechnology Co., Ltd, China). The fluorescent beads were imaged with the SIM setup shown in Fig. 8. Two groups of fringe patterns with orthogonal orientation were utilized to illuminate the sample in sequence, and each group of the patterns was shifted three times. The generated six raw intensity images were recorded by the CMOS camera. The numerical aperture (NA = 0.75) of the objective limits the lateral resolution to $\delta = \lambda_{em}/(2\mathrm{NA}) = 0.39$ μm estimated by Abbe criterion, where $\lambda_{em} = 580$ nm is the emission-peak wavelength of the fluorophore used. The period of the structured illumination fringes is $P_{\mathrm{SIM}} = 0.69$ μm, and the theoretical enhancement in the lateral resolution of SIM is $[(2 \times \mathrm{NA}/\lambda_{em}) + 1/P_{\mathrm{SIM}}]/[2 \times \mathrm{NA}/\lambda_{em}] = 1.6$, meaning that the theoretical resolution of large-field SIM of 0.24 μm (estimated by Abbe criterion) [33].

The super-resolution images of the fluorescent beads were reconstructed by using the conventional reconstruction approach [31,34], DIP-SIM-6, and DIP-SIM-3. Fig. 8(a)-left and right show the wide-field image and DIP-SIM-6 image. And, the images of four sub-regions obtained by wide-field mode (obtained by averaging the six raw intensity images), conventional SIM, DIP-SIM-6, DIP-SIM-3 are shown in rows 1–4 of Fig. 8(b), respectively. Compared to the wide-field image, the conventional SIM, DIP-SIM-6, and DIP-SIM-3 images have a much higher spatial resolution. This conclusion is further confirmed by analyzing the intensity distributions along the line crossing the same beads in the images obtained by the four modalities, as shown in Fig. 8(c). Obviously, the two closely-placed beads with a distance of 0.38 μm that are not resolvable in the wide-field image become distinguishable in conventional-SIM, DIP-SIM-6, and DIP-SIM-3 images. To quantitatively assess the lateral resolution of each imaging modality, eighteen fluorescent beads were randomly selected from Fig. 8(a), and the intensity distributions along the line crossing the bead centers were extracted and fitted with Gaussian functions. The full widths at the half maximum (FWHMs) are shown in Fig. 8(d). It is found in Fig. 8(d) that the average FWHM is 0.54 ± 0.03 μm, 0.30 ± 0.03 μm, 0.31 ± 0.03 μm, 0.34 ± 0.03 μm for conventional wide-field microscopy, conventional SIM, DIP-SIM-6, and DIP-SIM-3, respectively. Specifically, a resolution enhancement factor of 1.74 is achieved for DIP-SIM-6 and 1.59 for DIP-SIM-3. The image decorrelation analysis has also been conducted to

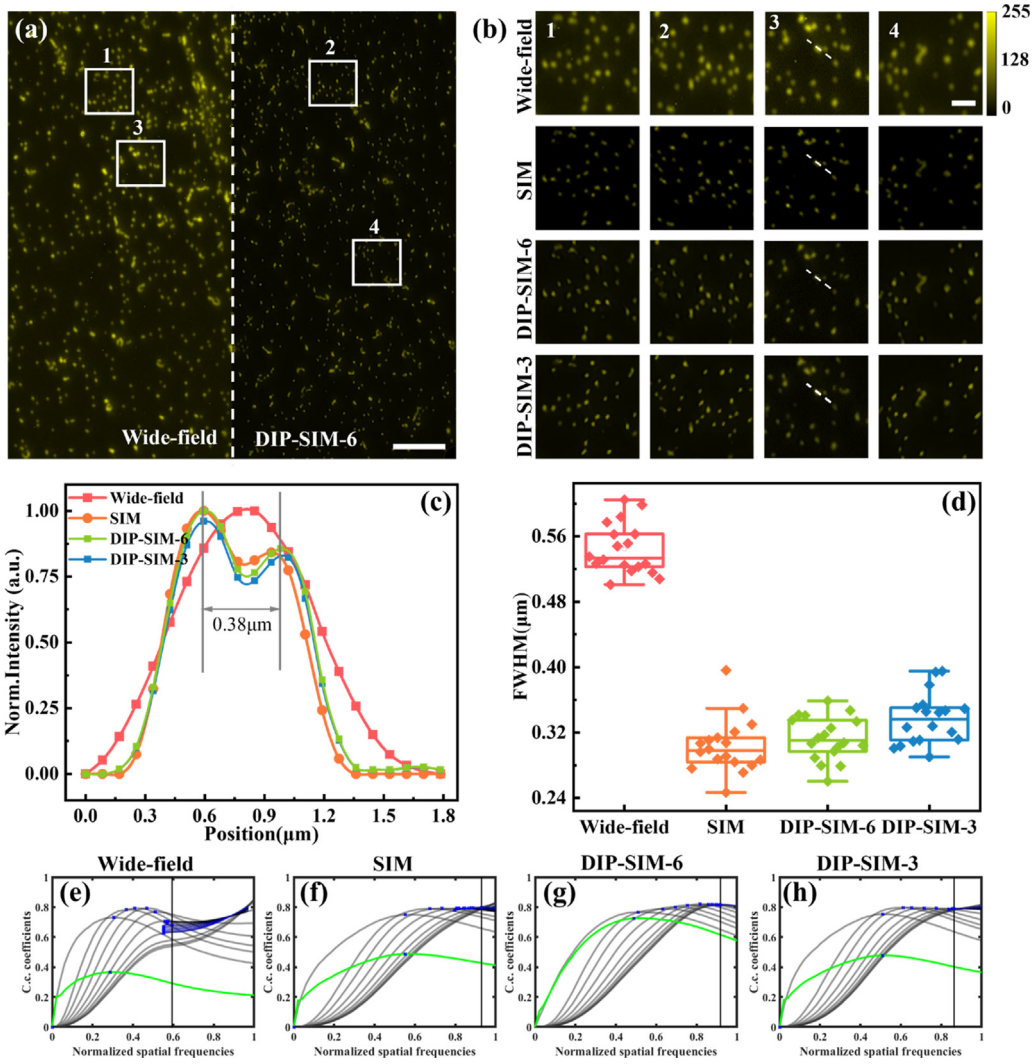

Fig. 8. Wide-field and DIP-SIM-6 imaging on 240 nm-diameter fluorescent beads. (a) Wide-field (left) and DIP-SIM-6 (right). (b) Magnified view of the region in (a) indicated with a white box: row 1: wide-field image, row 2~4: images reconstructed by the conventional SIM, DIP-SIM-6 and DIP-SIM-3. (c) The intensity profiles along the dash lines in (b). (d) Statistic of spatial resolution in term of the FWHM of individual particles. (e)~(h) Decorrelation-based resolution analysis of the wide-field, conventional SIM, DIP-SIM-6, DIP-SIM-3 images. Scale bar in (a) is 10 μm. Scale bar in (b) is 3 μm.

evaluate the spatial resolution of different imaging modalities, as shown in Fig. 8(e–h) [35]. The results reveal that the wide-field, SIM, DIP-SIM-6, and DIP-SIM-3 images have a normalized cut-off frequency of k_c = 0.58 μm^{-1}, 0.93 μm^{-1}, 0.92 μm^{-1}, and 0.86 μm^{-1}, meaning that they have a spatial resolution (calculated by $2\delta/M/k_c = 2 \times 3.45/20/k_c$) of 0.59 μm, 0.37 μm, 0.38 μm, and 0.40 μm, respectively. Here, δ and M are the pixel size of the camera and the magnification of the SIM system.

The comparison implies that DIP-SIM-3 only requires three raw images to achieve a plausibly comparable spatial resolution to DIP-SIM-6, and therefore, it has a much faster imaging speed and a two-fold reduction of photobleaching. However, the DIP-SIM-3 suffers from a burn-in effect (i.e., the illumination patterns do not average out to a uniform illumination), leading to a poorer reconstruction quality compared to the DIP-SIM-6 image. In mathematics, this is due to the fact that the analytical reconstruction becomes uncertain (or not unique) from the equations constrained by three raw images [32].

As is generally known, the excitation light in SIM can cause photobleaching, which will, in turn, prevent SIM from long-term imaging of fluorescent-labeled samples. In principle, the lower the intensity of the excitation light is, the lower the photobleaching can be caused. However, a low excitation light often results in a lower brightness in the recorded image, which challenges the super-resolution reconstruction. Herein, we compared the performance of conventional SIM, DIP-SIM-6, and DIP-SIM-3 when dealing with the data acquired with week-brightness samples. In this experiment, 240 nm-diameter fluorescent beads were used as the sample. To mimic low-brightness raw images obtained with low-intensity excitation light, we recorded raw images with a monotonously reduced exposure time. The brightness of a raw image was characterized by photons per bead (PPB) in the image. Fig. 9 shows the reconstructed wide-field image, conventional SIM, DIP-SIM-6, and DIP-SIM-3 images of the sample at four different brightness with 780, 700, 610, and 280 PPB, respectively. ROIs from the same location in the four-modality images were magnified and shown in the lower-right corner of each image. The intensity distributions along the line crossing two connecting beads in the four-modality images at different brightness conditions were compared in Fig. 9(b–e).

Furthermore, we quantitatively assessed the quality of DIP-SIM-6 and DIP-SIM-3 images, using the conventional SIM image with the highest brightness (i.e., 780 PPB) as the ground truth. Herein, the peak signal-to-noise ratio (PSNR) and SSIM were used as the criteria, and the results are shown in Table 1. It is found that as the brightness of the

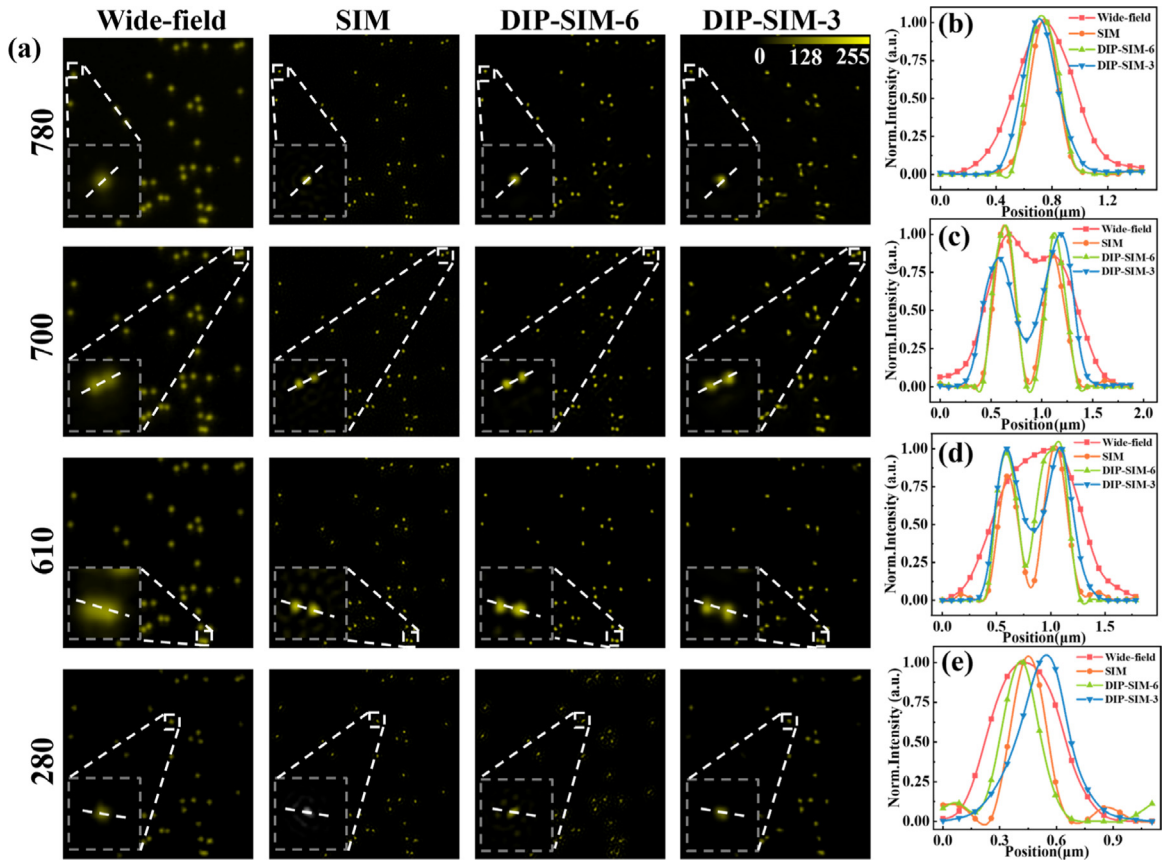

Fig. 9. Wide-field and DIP-SIM images of fluorescent beads at different exposure time. (a) Wide-field, conventional SIM, DIP-SIM-6, and DIP-SIM-3 images. four regions of interest (ROIs) at different locations are magnified for better comparison of wide-field, conventional SIM, DIP-SIM-6, and DIP-SIM-3. (b–e) Intensity distribution for different exposure times. The scale bar in (a) is 10 μm.

raw data decreases, the reconstruction quality of both DIP-SIM-6 and DIP-SIM-3 decrease, and the latter one is slightly worse than the prior one.

In addition, we applied DIP-SIM-6 and DIP-SIM-3 to super-resolution imaging of mouse stem cells labeled with primary antibody (β-Tubulin antibody, #2146, produced in rabbit) and secondary antibody (Goat anti-Rabbit IgG(H + L)-Alexa Fluor 532). Fig. 10(a) shows the wide-field (upper-left) and DIP-SIM-6 (lower-right) images of microtubules in fixed cells. The comparison shows that the latter one has shaper structures with a clearer background. Further, the wide-field, conventional SIM, DIP-SIM-6, and DIP-SIM-3 images of four ROIs, indicated with the four white boxes, are magnified and shown in Fig. 10(b). It is obvious that the DIP-SIM-6, DIP-SIM-3, and conventional SIM images show much finner structures of the microtubules. The intensity distributions (Fig. 10(c)) along the cut lines in Fig. 10(b) exhibit that the two fibers that are 500 nm apart (below the diffraction limit) are not resolvable in the wide-field image. However, they can be well resolved in both the DIP-SIM-6, DIP-SIM-3, and conventional SIM images. In general, the loss decreases gradually with the increase of the iteration number, as shown in Fig. 10(d). While, the neural network of DIP-SIM-6 and DIP-SIM-3 can provide a high-quality SR-SIM image.

Table 1
Comparison of DIP-SIM-6 and DIP-SIM-3 reconstruction from raw data of different brightness.

Photons per particle	DIP-SIM-6	DIP-SIM-3
	PSNR/SSIM	
780	31.52/0.95	29.62/0.93
700	30.82/0.93	29.47/0.89
610	29.11/0.88	27.73/0.86
280	25.94/0.86	25.65/0.83

4. Conclusion and discussion

In conclusion, a physics-driven deep image prior framework for the super-resolution reconstruction of SIM (entitled DIP-SIM) is proposed and demonstrated. The incorporation of a physical model releases the neural network from the dependency on a large number of labeled data. DIP-SIM can reconstruct a super-resolution image from six images for two-orientation and from nine raw images for three-orientation SIM, respectively. The reconstruction has a quality comparable to conventional SIM. Furthermore, the proposed DIP-SIM allows for the reconstruction of a SR image with fewer raw images, such as three raw images for two-orientation SIM and four raw images for three-orientation SIM, despite the analytical reconstruction becomes underdetermined for fewer than four raw images [35]. This is due to the fact that the iteration of DIP-SIM can better extract and utilize high-frequency information encoded in the raw images as long as the parameter is realistic. However, DIP-SIM may provide artificial SR structures (artifacts) where the underdetermined problem occurs due to missing bright-fringe coverage. As another disadvantage, DIP-SIM requests pre-known parameters of structured illumination and PSF, and the relative error on the parameter estimation should be lower than 2%. After all, DIP-SIM has a much faster imaging speed and lower photobleaching compared with the traditional SIM. We can envisage that DIP-SIM will be widely applied in biomedical imaging.

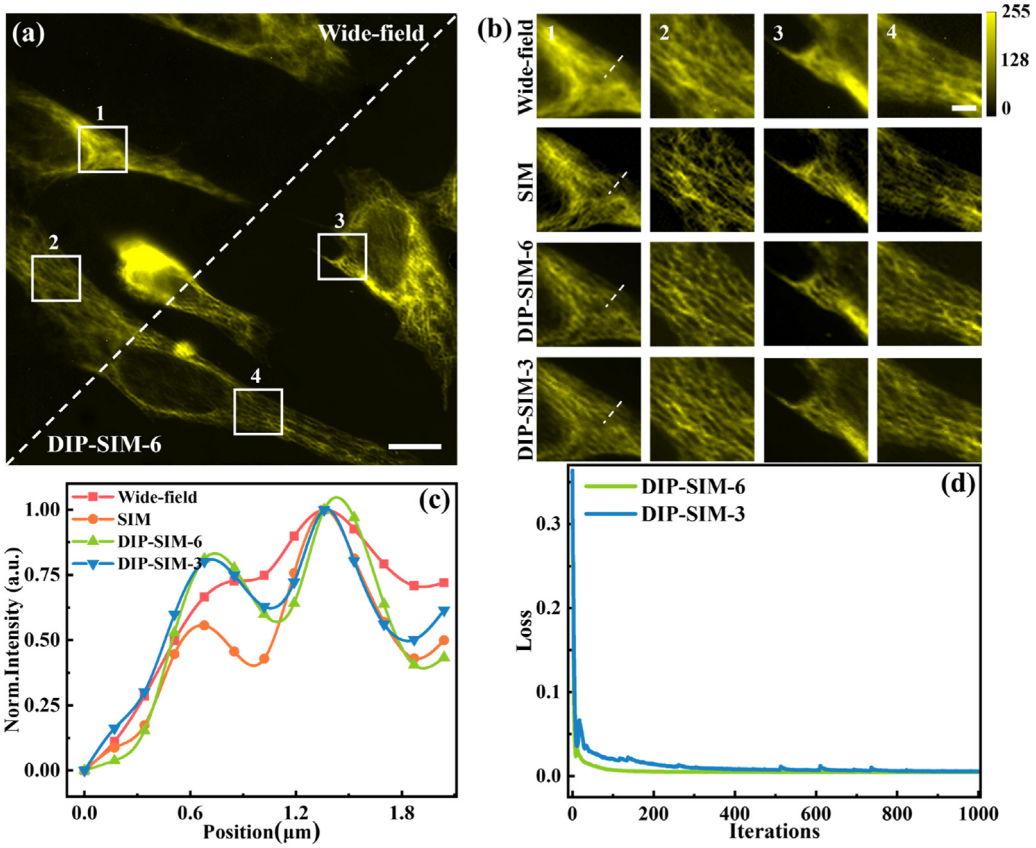

Fig. 10. Wide-field and DIP-SIM-6 images of the microtubules in fixed mouse stem cells labeled with Alexa fluor 532. (a) Wide-field image (upper-left) and DIP-SIM-6 image (lower-right) of the sample. (b) Magnified view of the four ROIs in (a) indicated with a white box: row 1∼4: Wide-field, conventional SIM, DIP-SIM-6, and DIP-SIM-3 images. (c) The intensity profiles along the dash lines in (b). (d) Loss changes versus the number of iterations. The scale bar in (a) is 10 μm, and the scale bar in (b) is 3 μm.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

X. L. and J. L. set up the neuron network and performed data analysis. X. F. performed the experiment. S. A., J. L., J. Z., J. L., N. A., and C. Z. contributed to data analysis. X. L. wrote the manuscript. P. G. supervised the subject and wrote the manuscript with X. L. All the authors edited the manuscript.

Funding

The authors acknowledge the support from the National Key Research and Development Program of China (2022YFE0100700 and 2021YFF0700303); the National Natural Science Foundation of China (NSFC 62075177); the Natural Science Foundation of Shaanxi Province, China (2023JCQN0731, 2023JCYB518); Exchange Program Between China-and Poland (2021–2022); Key Laboratory of Wuliangye-flavor Liquor Solid-state Fermentation, China National Light Industry (2019JJ 012); Natural Science Basic Research Program of Shaanxi, China (No. 2023-JC-QN-0765).

References

[1] M. Bates, B. Huang, G.T. Dempsey, X. Zhuang, Multicolor super-resolution imaging with photo-switchable fluorescent probes, Science 317 (2007) 1749–1753.
[2] M.G. Gustafsson, Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy, J. Microsc. 198 (2000) 82–87.
[3] M.G. Gustafsson, L. Shao, P.M. Carlton, C.R. Wang, I.N. Golubovskaya, W.Z. Cande, D.A. Agard, J.W. Sedat, Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination, Biophys. J. 94 (2008) 4957–4970.
[4] L. Schermelleh, R. Heintzmann, H. Leonhardt, A guide to super-resolution fluorescence microscopy, J. Cell Biol. 190 (2010) 165–175.
[5] M.J. Rust, M. Bates, X. Zhuang, Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM), Nature Methods 3 (2006) 793–796.
[6] Y.M. Sigal, R. Zhou, X. Zhuang, Visualizing and discovering cellular structures with super-resolution microscopy, Science 361 (2018) 880–887.
[7] S.W. Hell, J. Wichmann, Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy, Opt. Lett. 19 (1994) 780–782.
[8] P. Gao, B. Prunsche, L. Zhou, K. Nienhaus, G.U. Nienhaus, Background suppression in fluorescence nanoscopy with stimulated emission double depletion, Nat. Photon. 11 (2017) 163–169.
[9] L. Shao, P. Kner, E.H. Rego, M.G. Gustafsson, Super-resolution 3D microscopy of live whole cells using structured illumination, Nature Methods 8 (2011) 1044–1046.
[10] M. Saxena, G. Eluru, S.S. Gorthi, Structured illumination microscopy, Adv. Opt. Photon. 7 (2015) 241–275.
[11] Y. Wu, H. Shroff, Faster, sharper, and deeper: structured illumination microscopy for biological imaging, Nature Methods 15 (2018) 1011–1019.
[12] P. Kner, B.B. Chhun, E.R. Griffis, L. Winoto, M.G. Gustafsson, Super-resolution video microscopy of live cells by structured illumination, Nature Methods 6 (2009) 339–342.
[13] M.F. Langhorst, J. Schaffer, B. Goetze, Structure brings clarity: structured illumination microscopy in cell biology, Biotechnol. J.: Healthc. Nutr. Technol. 4 (2009) 858–865.

[14] M. Müller, V. Mönkemöller, S. Hennig, W. Hübner, T. Huser, Open-source image reconstruction of super-resolution structured illumination microscopy data in ImageJ, Nature Commun. 7 (2016) 1–6.
[15] A. Lal, C. Shan, P. Xi, Structured illumination microscopy image reconstruction algorithm, IEEE J. Sel. Top. Quantum Electron. 22 (2016) 50–63.
[16] K. Wicker, O. Mandula, G. Best, R. Fiolka, R. Heintzmann, Phase optimisation for structured illumination microscopy, Opt. Express 21 (2013) 2032–2049.
[17] D. Dan, Z. Wang, X. Zhou, M. Lei, T. Zhao, J. Qian, X. Yu, S. Yan, J. Min, P.R. Bianco, Rapid image reconstruction of structured illumination microscopy directly in the spatial domain, IEEE Photonics J. 13 (2021).
[18] C. Ling, C. Zhang, M. Wang, F. Meng, L. Du, X. Yuan, Fast structured illumination microscopy via deep learning, Photon. Res. 8 (2020) 1350–1359.
[19] A. Lal, C. Shan, K. Zhao, W. Liu, X. Huang, W. Zong, L. Chen, P. Xi, A frequency domain SIM reconstruction algorithm using reduced number of images, IEEE Trans. Image Process. 27 (2018) 4555–4570.
[20] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (2015) 436–444.
[21] G. Ongie, A. Jalal, C.A. Metzler, R.G. Baraniuk, A.G. Dimakis, R. Willett, Deep learning techniques for inverse problems in imaging, IEEE J. Sel. Areas Inf. Theory 1 (2020) 39–56.
[22] C. Qiao, D. Li, Y. Guo, C. Liu, T. Jiang, Q. Dai, D. Li, Evaluation and development of deep neural networks for image super-resolution in optical microscopy, Nature Methods 18 (2021) 194–202.
[23] Z.H. Shah, M. Müller, T.-C. Wang, P.M. Scheidig, A. Schneider, M. Schüttpelz, T. Huser, W. Schenck, Deep-learning based denoising and reconstruction of super-resolution structured illumination microscopy images, Photon. Res. 9 (2021) B168–B181.
[24] L. Jin, B. Liu, F. Zhao, S. Hahn, B. Dong, R. Song, T.C. Elston, Y. Xu, K.M. Hahn, Deep learning enables structured illumination microscopy with low light levels and enhanced speed, Nature Commun. 11 (2020) 1–7.
[25] C.N. Christensen, E.N. Ward, M. Lu, P. Lio, C.F. Kaminski, ML-SIM: universal reconstruction of structured illumination microscopy images using transfer learning, Biomed. Opt. Express 12 (2021) 2720–2733.
[26] Z. Burns, Z. Liu, Untrained, physics-informed neural networks for structured illumination microscopy, 2022, arXiv preprint arXiv:2207.07705.
[27] D. Ulyanov, A. Vedaldi, V. Lempitsky, Deep image prior, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 9446–9454.
[28] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.
[29] M.A. Boland, E.A. Cohen, S.R. Flaxman, M.A. Neil, Improving axial resolution in Structured Illumination Microscopy using deep learning, Phil. Trans. R. Soc. A 379 (2021) 20200298.
[30] E. Xypakis, G. Gosti, T. Giordani, R. Santagati, G. Ruocco, M. Leonetti, Deep learning for blind structured illumination microscopy, Sci. Rep. 12 (2022) 1–7.
[31] J. Zheng, X. Fang, K. Wen, J. Li, Y. Ma, M. Liu, S. An, J. Li, Z. Zalevsky, P. Gao, Large-field lattice structured illumination microscopy, Opt. Express 30 (2022) 27951–27966.
[32] F. Ströhl, C.F. Kaminski, Speed limits of structured illumination microscopy, Opt. Lett. 42 (2017) 2511–2514.
[33] P.T. Brown, R. Kruithoff, G.J. Seedorf, D.P. Shepherd, Multicolor structured illumination microscopy and quantitative control of polychromatic light with a digital micromirror device, Biomed. Opt. Express 12 (2021) 3700–3716.
[34] K. Wen, X. Fang, Y. Ma, M. Liu, S. An, J. Zheng, T. Kozacki, P. Gao, Large-field structured illumination microscopy based on 2D grating and a spatial light modulator, Opt. Lett. 47 (2022) 2666–2669.
[35] A. Descloux, K.S. Grußmayer, A. Radenovic, Parameter-free image resolution estimation based on decorrelation analysis, Nature Methods 16 (2019) 918–924.