PHOTONICS Research

Deep-learning based denoising and reconstruction of super-resolution structured illumination microscopy images

Zafran Hussain Shah,[1] Marcel Müller,[2] Tung-Cheng Wang,[2] Philip Maurice Scheidig,[1] Axel Schneider,[1] Mark Schüttpelz,[2] Thomas Huser,[2,3] and Wolfram Schenck[1,4]

[1]Faculty of Engineering and Mathematics, Bielefeld University of Applied Sciences, 33619 Bielefeld, Germany
[2]Faculty of Physics, Bielefeld University, 33615 Bielefeld, Germany
[3]e-mail: thomas.huser@physik.uni-bielefeld.de
[4]e-mail: wolfram.schenck@fh-bielefeld.de

Super-resolution structured illumination microscopy (SR-SIM) provides an up to twofold enhanced spatial resolution of fluorescently labeled samples. The reconstruction of high-quality SR-SIM images critically depends on patterned illumination with high modulation contrast. Noisy raw image data (e.g., as a result of low excitation power or low exposure time), result in reconstruction artifacts. Here, we demonstrate deep-learning based SR-SIM image denoising that results in high-quality reconstructed images. A residual encoding–decoding convolutional neural network (RED-Net) was used to successfully denoise computationally reconstructed noisy SR-SIM images. We also demonstrate the end-to-end deep-learning based denoising and reconstruction of raw SIM images into high-resolution SR-SIM images. Both image reconstruction methods prove to be very robust against image reconstruction artifacts and generalize very well across various noise levels. The combination of computational image reconstruction and subsequent denoising via RED-Net shows very robust performance during inference after training even if the microscope settings change. © 2021 Chinese Laser Press

https://doi.org/10.1364/PRJ.416437

1. INTRODUCTION

Fluorescence microscopy remains one of the most powerful tools for imaging cell biology samples because of its ability to specifically label molecular structures and to visualize them in different fluorescence color channels. It also offers exceptionally high sensitivity and can visualize molecular processes even below the optical diffraction limit. Several methods that enable imaging of fluorescently labeled samples down to the nanoscale have been developed during the last two decades [1]. Super-resolution structured illumination microscopy (SR-SIM) is a particularly compelling method, because it works with the majority of samples and fluorophores commonly used in cell biology without imposing specific requirements on sample preparation [2,3]. Therefore, it can even be applied to living samples [4–12]. SR-SIM in its most common form uses a series of sinusoidal illumination patterns with a pattern periodicity at or near the diffraction limit. This patterned excitation light is phase-shifted laterally and rotated to different discrete angles to acquire a series of raw images, which are then passed on to an image reconstruction algorithm to obtain the final super-resolved SR-SIM image [13]. Several implementations of

reconstruction algorithms that operate in the frequency space were developed and a number of open access tools are now available that aim to enhance the speed and spatial resolution, and minimize reconstruction artifacts in the final reconstructed image (i.e., fairSIM [14], OpenSIM [15], SIMToolbox [16], and CC-SIM [17]). The most common property among all these image reconstruction algorithms is that they require a series of high-quality raw images to be able to reconstruct a high-quality super-resolution image. However, reconstruction in the frequency space fails to reliably reconstruct SR-SIM images if the signal-to-noise ratio (SNR) is too low (e.g., because the laser excitation power level was too low), the sample exposure time chosen was too short, or the sample has already undergone irreversible photobleaching [18]. Recently, several developments that aim to reduce reconstruction artifacts from SR-SIM images have been undertaken. Huang *et al.* used a fully analytical approach to reduce noise and minimize reconstruction artifacts using the Hessian matrix theory [10]. Hoffman and Betzig proposed the reconstruction of SIM images in lower pixel count tiles with subsequent merger to reduce reconstruction artifacts [19]. Jin *et al.*, on the other hand, used deep neural networks to reconstruct the cropped

regions of SR-SIM images [20]. Christensen *et al.* used a deep learning architecture for the reconstruction from synthetic raw SIM images [21] with subsequent testing on real microscope images. Ling *et al.* relied on a special type of convolutional neural network, a CycleGAN, for the same purpose [22]. And Weigert *et al.* used deep learning algorithms to enhance isotropic resolution and the signal-to-noise ratio (SNR) of fluorescence microscopy images in general [23].

Here, we provide a comprehensive study on several deep-learning based approaches to denoise and reconstruct SR-SIM images. First, as starting point, we include an end-to-end deep learning architecture and workflow that is related to the existing literature combining SIM and deep learning [20–22] in that a single network is used that receives raw SIM images as input and produces a super-resolution image as the output. This approach is named super-resolution residual encoder–decoder structured illumination microscopy (SR-REDSIM) in the following discussion. In this case, the entire SIM reconstruction process is performed by the deep convolutional neural network.

Second, for what we believe, to the best of our knowledge, is an entirely novel workflow, we combine classical computational SIM reconstruction with a deep learning network. This workflow is called a residual encoder–decoder fairSIM (RED-fairSIM). RED-fairSIM is a combination of the fairSIM image reconstruction package [14], which performs image reconstruction using commonly used frequency domain algorithms and a deep convolutional neural network for subsequent artifact reduction and denoising. Finally, we also show the results for a workflow where deep learning is first applied to the raw SIM images for denoising with subsequent classical computational SIM reconstruction; thus, exactly the other way around, as in RED-fairSIM. This approach is named preRED-fairSIM.

For the main network architecture for all three approaches we use the residual encoder–decoder network (RED-Net) [24], which is comparatively lightweight. For training, raw image data directly from the microscope is used, which offers a straightforward, practical solution without the need to create synthetic data or carry out substantive preprocessing in contrast to the existing literature. We found that the first two methods are robust in their ability to significantly improve the quality of the reconstructed SR-SIM images. We show, for what we believe is the first time, that the trained networks of these two methods generalize well to real microscopic SIM data with different SNRs. The novel RED-fairSIM workflow shows superior performance in this regard. Furthermore, we also demonstrate that RED-fairSIM generalizes better even if the microscope settings are changed after training.

2. MATERIALS AND METHODS

A. Training Data Generation

Deep learning methods rely on training data, which, in our case, consist of noisy raw SIM images for the input and ideally noise and artifact free, super-resolved SIM reconstructions for the output. Thus, we first need to generate such a data set that is large enough to effectively train the network, but also captures all aspects of the SIM imaging process (sample behavior,

instrument imperfections, and data processing artifacts) well enough.

In principle, the data acquisition process of a SIM microscope can be simulated. In this case, the expected output represents the ground truth data upon which the simulation is based, and which is known without SNR or resolution limits. In addition, the amount of available training data would only be limited by the processing time, as the generation would be fully automated and not rely on access to a microscope system. However, we decided against this pure *in silico* approach. Although the basic effects of structured illumination, Poisson-distributed noise, and even basic optical imperfections are rather easy to simulate, modeling the response of a full structured illumination microscope correctly is very complex. Additionally, such a simulation would likely have to be adjusted to reflect the properties of a specific SR-SIM instrument to capture changes (e.g., when switching to a different manufacturer or even a specific installation of an SR-SIM microscope). The same argument holds true for the fluorescent samples themselves. Although some simulations that provide perfect ground-truth data exist (e.g., for single molecule localization microscopy [25,26]), they, again, do not capture all of the variability found in real-world samples.

The option chosen for data generation for the work presented here is to use real microscope data from standard biological samples. This approach naturally captures all aspects and imperfections of the samples and of the specific instrument in question [23], but also poses constraints. Because data collection requires both instrument time and manual sample preparation and handling, the amount of training data is naturally limited. There also is no perfect ground truth available. To acquire the high-quality reference images presented to the networks as desired output, we adjusted the instrument to provide high-SNR raw frames and processed those with the classical, frequency-domain-based image reconstruction algorithm. While these images are low in noise and reconstruction artifacts, they are never completely devoid of them. For this reason, we refer to them by the term "reference image" instead of "ground truth." To acquire noisy, low-SNR images as input, the samples were then photo-bleached by continued exposure and acquisition of raw SIM images, which naturally reduces the fluorescent light output over time and results in a series of images with steadily decreasing SNR. (See Section 2.C for details.)

B. Sample Preparation and Data Acquisition

U2OS cells were cultured in DMEM supplemented with 10% FBS and grown on round coverslips of 170 ± 5 nm thickness (No. 1.5H). Cells were fixed with 4% PFA for 15 min, followed by PBS washes, and permeabilization with 0.5% Triton-X100 for 3 min. Another two rounds of PBS washes were done prior to blocking with 3% BSA. For the immunolabeling, the microtubuli, cells were stained with anti-tubulin Ab (Invitrogen Cat. No. 322500) 1:400 for 2 h at room temperature, followed by a PBS wash and one additional hour of incubation with Alexa 488-conjugated anti-mouse IgG 1:400. Cells were then briefly washed with PBS before Vectashield was applied to embed the coverslip onto the standard slide glass for imaging.

The DeltaVision|OMX (GE Healthcare, Chicago, IL, USA) was used to acquire 3D-SIM raw images. A total of 101 randomly selected fields of view were acquired by exposing each field-of-view to the full laser power of the 488 nm excitation laser and the exposure time was set at 20 ms for each of the 15 raw image frames. A total of 200 image repetitions were collected at each position without delays. Taking into account the camera readout and pattern switching time, acquiring 15 raw SIM images, making up one timestamp in the raw image stack, takes approximately 375 ms.

C. Data Preprocessing

In our work, we used 101 different cell structures (fields of view). Out of these, 81 were selected for the training data and the remaining 20 for the test data. Each cell structure was captured for 200 repetitions. During each repetition, 15 frames were recorded, iterating the phase and orientation of the sinusoidal SIM illumination pattern, yielding an image stack of size $15 \times 512 \times 512$ (frames, width, height). During this time-lapse acquisition, the samples underwent photobleaching, which reduces the amount of active fluorescent emitters and thus the amount of emitted photons. Therefore, less and less light is captured and the SNR steadily decreases during the acquisition of such a time series.

The cell structures from the timestamps 175 to 200 are therefore considered as noisy training and test input while the samples from timestamp 0 are considered as clean output images. All 15 clean raw SIM images of the 101 cell structures from timestamp 0 are used to reconstruct high-resolution reference SIM images of size 1024×1024 pixels by using fairSIM, which employs a classic frequency-domain-based reconstruction. In this work, the input dimension is $15 \times 512 \times 512$ (frames, width, height) whereas the output dimension is 1024×1024 (width, height) pixels. A total of 2525 samples were further divided into training and test data. The training data contains 2025 images of the first 81 cell structures; the test data is composed of 500 test images that are created from the remaining 20 cell structures.

The only preprocessing step involved in our work is the linear scaling of the training and test data to match the overall brightness between the input and output. In addition, we tested an image augmentation approach to double the amount of training data by rotating each image by an angle of 180°.

The data from each time series over 200 repetitions were subdivided into different noise levels. Noise level 0 stands for the highest SNR in our data at timestamp 0. In our work, this is our reference data. The image data from timestamps 175–200 represent the highest noise level 4, the data from timestamps 125–150 represent noise level 3, the data from timestamps 75–100 are noise level 2, and the data from timestamps 25–50 are noise level 1. In this study, data from noise level 4 are only used in the training process whereas data from noise levels 1, 2, 3, and 4 are used in the test phase.

D. Architecture and Training of SR-REDSIM

In the first deep-learning based SR-SIM image reconstruction method, named SR-REDSIM, the reconstruction and denoising of noisy raw SIM images are both performed by a single deep learning model. This model is a modified version of the RED-Net. RED-Net is an encoding–decoding framework with symmetric convolutional–deconvolutional layers along with skip-layer connections. It was previously used to accomplish different image restoration tasks such as image denoising, image super-resolution, and image inpainting [24]. The original RED-Net architecture is only composed of encoding–decoding blocks with the size of the network input being the same as the size of the network output. Therefore, super-resolution with this architecture has to rely on explicit image pre-upsampling [24]. In contrast, our modified RED-Net architecture contains an additional upsampling block after the encoding–decoding blocks. This upsampling block inside our model has the advantage that the input images are first denoised in their lower-dimensional space, which reduces the training time and effort.

Most of the super-resolution architectures such as the enhanced deep super-resolution network (EDSR) [27] or the residual channel attention network (RCAN) [28] are very deep and require a significant amount of training data. In comparison, our architecture is comparably lightweight.

SR-REDSIM is based on a modified version of RED-Net. The complete pipeline of this approach is shown in Fig. 1(a), whereas the architecture of SR-REDSIM and details about the model parameters are given in Fig. 2(a). The SR-REDSIM architecture consists of three blocks: the encoder, the decoder, and the upsampling block. SR-REDSIM contains a total of 44 convolutional and deconvolutional layers with symmetric skip connections. The encoder block is composed of 21 convolutional layers, whereas the decoder contains 21 deconvolutional layers. The upsampling block consists of two deconvolutional layers that perform the upsampling task by adjusting the size of the stride. The SR-REDSIM model provides the best results after training the model for 100 epochs. The SR-REDSIM model is trained only with high-level noise data from timestamps 175 to 200. During the training process, the ADAM optimizer and the L2 loss function, also known as least squares error, are used, so

$$\text{Loss} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2. \tag{1}$$

In Eq. (1), y_i represents the true pixel intensity, $\hat{y}_i$ represents the predicted pixel intensity, and n is the number of pixels (in the image).

E. Architecture and Training of RED-fairSIM

The RED-fairSIM is combination of fairSIM and RED-Net. The pipeline of RED-fairSIM is shown in Fig. 1(b). In this approach, first fairSIM is used to transform the raw SIM images into a super-resolved output image by employing a classic, frequency-domain SIM reconstruction algorithm. This output image now contains noise, which, due to the frequency-domain algorithm, takes a SIM-specific form, and might show other reconstruction artifacts. It is subsequently processed by RED-Net.

1. fairSIM

The fairSIM reconstruction is performed in three steps: parameter estimation, reconstruction, and filtering. Mathematical and algorithmic details are provided in the original publication
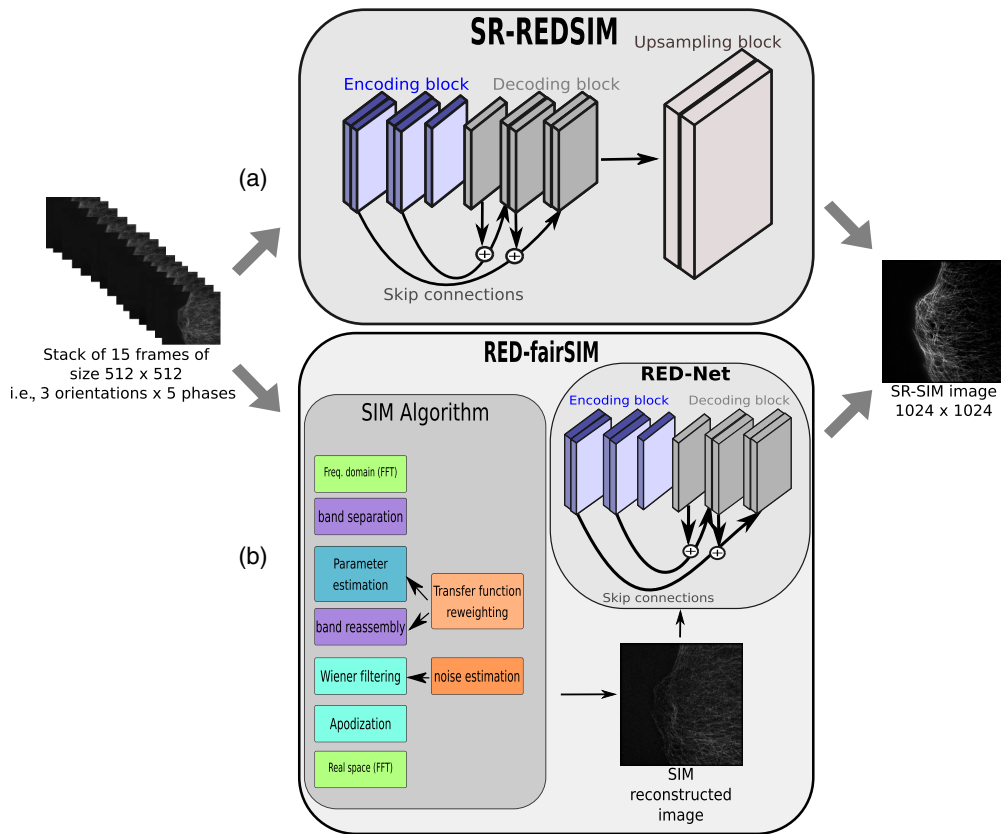
Fig. 1. Schematics of the deep learning CNN architecture of the two different SR-SIM image denoising (RED-fairSIM) and image denoising and reconstruction (SR-REDSIM) methods. In both approaches, a stack of 15 raw (noisy) SIM images (three angles with five phases each) is used as input. The output is the reconstructed SR-SIM image. (a) SR-REDSIM is composed of three main blocks. The encoding block contains mainly the convolutional layers whereas the decoding block consists of deconvolutional layers and the upsampling block of deconvolutional upsampling layers. (b) In the RED-fairSIM method, fairSIM is first used to computationally reconstruct noisy SR-SIM images that are then further propagated into the RED-Net for denoising. The architecture of RED-Net is composed of the encoder and the decoder blocks.

[14]. A synthetic optical transfer function, with NA = 1.4, $\lambda = 525$ nm, and $a = 0.31$ (a is a compensation parameter [14,29]), is used. A total of 500 counts of background are subtracted per pixel. SIM reconstruction parameters (pattern orientation and global phase, for example) are automatically determined by the fairSIM standard in an iterative cross-correlation approach. Filter parameters are set to a generalized Wiener filter strength of $w = 0.05$, apodization is set at 1.9× the resolution limit with a bend of 0.8, and a notch-style filter is implemented as OTF attenuation with a strength of 0.995 and an FWHM of 1.2 μm^{-1} is in use. For the detailed meaning and influence of these parameters, please refer to the fairSIM source code, its accompanying publication [14], and this general guide to SIM reconstruction parameters [29].

2. RED-Net Architecture

The architecture of RED-Net used in RED-fairSIM consists of 15 convolutional and 15 deconvolutional layers along with symmetric skip connections, as shown in Fig. 2(b). The output of fairSIM is propagated into the RED-Net to denoise the reconstructed noisy sample. During the training phase, the noisy SR-SIM images of size 1024 × 1024 pixels along with the reference SR-SIM images of the same size are used as input–output

pairs for the RED-Net. The network is trained for 100 epochs with the ADAM optimizer and L2 loss.

F. Architecture and Training of U-Net-fairSIM

U-Net is also a popular deep learning architecture that is extensively used in the domain of image restoration such as image denoising and super-resolution [23,30,31]. In U-Net-fairSIM, we simply replaced the RED-Net from the RED-fairSIM approach with the U-Net architecture. The U-Net is also trained for 100 epochs with the ADAM optimizer and L2 loss.

3. RESULTS

A. SR-REDSIM: SR-SIM Image Denoising and Reconstruction Using the Super-resolution REDSIM Method

SR-REDSIM is an entirely deep-learning based, end-to-end method. The complete pipeline of SR-REDSIM is shown in Fig. 1(a) and the architecture of SR-REDSIM, as shown in Fig. 2(a), is explained in more detail in Section 2.D. During the training process, we used all 15 raw noisy SIM images (three angles with five phases each) of size 512 × 512 pixels [i.e., stack dimensions were 15 × 512 × 512 (frames, width, height)] as input along with the reconstructed super-resolved SIM image
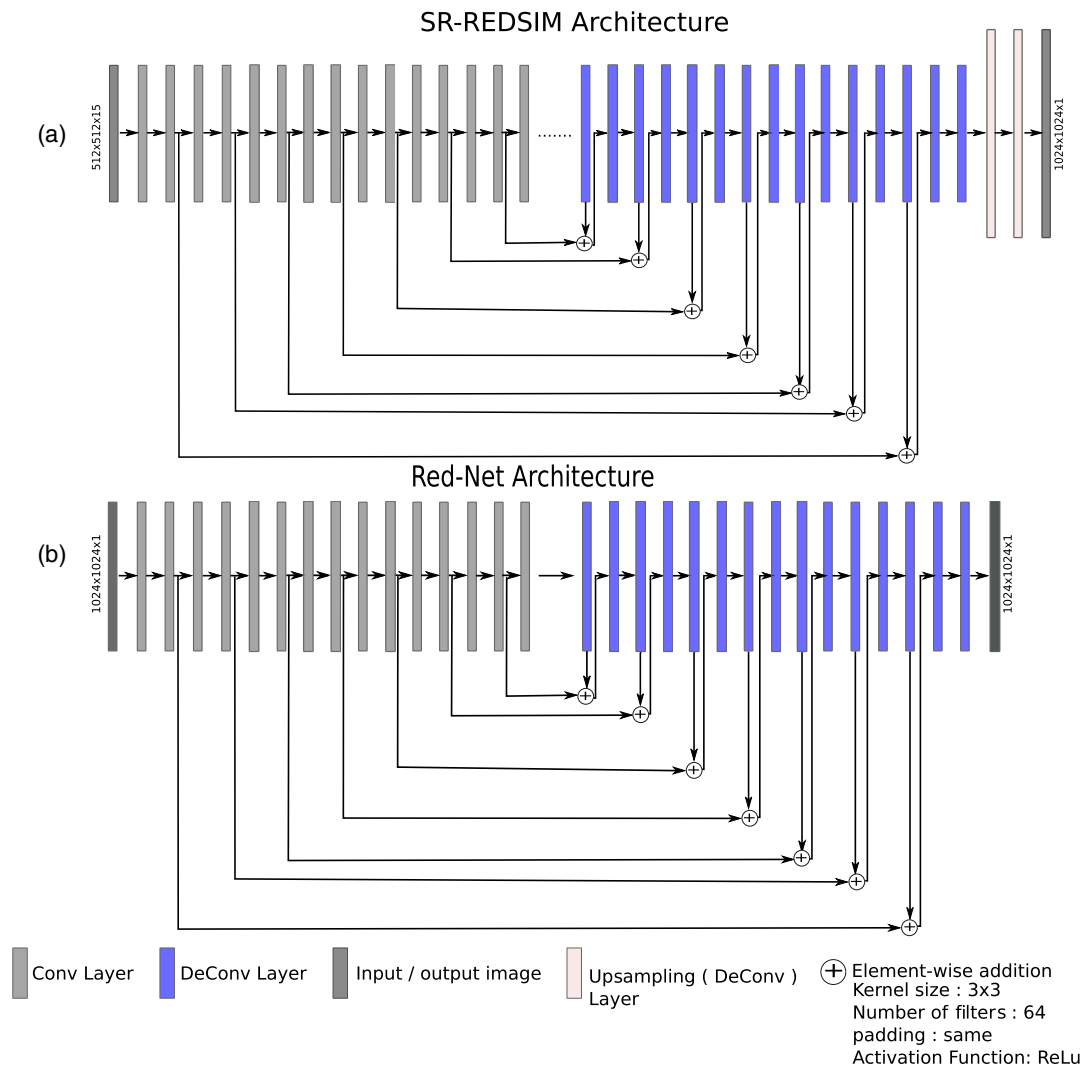
Fig. 2. The architecture of the networks used in this work. (a) SR-REDSIM architecture is composed of three different blocks. The encoding and decoding blocks contain 21 convolutional and deconvolutional layers, respectively, whereas the upsampling blocks consist only of two upsampling layers. This architecture was used in the SR-REDSIM method to denoise and reconstruct the raw SIM images. (b) The complete RED-Net architecture contains 15 convolutional and 15 deconvolutional layers along with the additive symmetric skip connection layers. This architecture was used in the RED-fairSIM and preRED-fairSIM methods for denoising (in preRED-fairSIM, the input and output have a size of only $512 \times 512 \times 1$).

of size 1024×1024 pixels as output. The output was generated by the fairSIM software from raw SIM images recorded with the highest SNRs, while the input images were taken from noise level 4. Note that Section 2.C offers an explanation of the noise levels. The trained network was tested afterward on unseen test data from noise level 4. The super-resolution images obtained during this test are depicted in column 2 of Fig. 3, whereas columns 1 and 5 show the results of noisy fairSIM (reconstructed by fairSIM from noisy raw SIM images; noise level 4) and reference fairSIM (reconstructed by fairSIM from raw SIM images with the highest SNR). The comparison of these images and of specific regions of interest (ROIs) between fairSIM in Fig. 3 (column 1, all rows) and SR-REDSIM (column 2, all rows) clearly shows that the noise is completely removed by SR-REDSIM. However, in the reconstruction by SR-REDSIM, fine cell structures are partly

suppressed compared to the reference output (compare column 2, row 2, ROI 1 in Fig. 3 with column 5, row 2, ROI 1). In rows 3/4 and 5/6 of Fig. 3, the structure of the cell is well denoised and reconstructed by SR-REDSIM. Moreover, the evaluation of the SR-REDSIM method on the basis of peak SNR (PSNR) and structural similarity index measurement (SSIM) [32] values in Table 1 shows a significant improvement compared to fairSIM.

B. RED-fairSIM: SR-SIM Reconstruction of Noisy Input Data by Using a Combination of fairSIM and RED-Net

One of the well-known open-source reconstruction algorithm implementations is fairSIM. It is widely used for super-resolution tasks among the other tools in SIM microscopy. However, it cannot reconstruct a clean, high-quality super-resolution
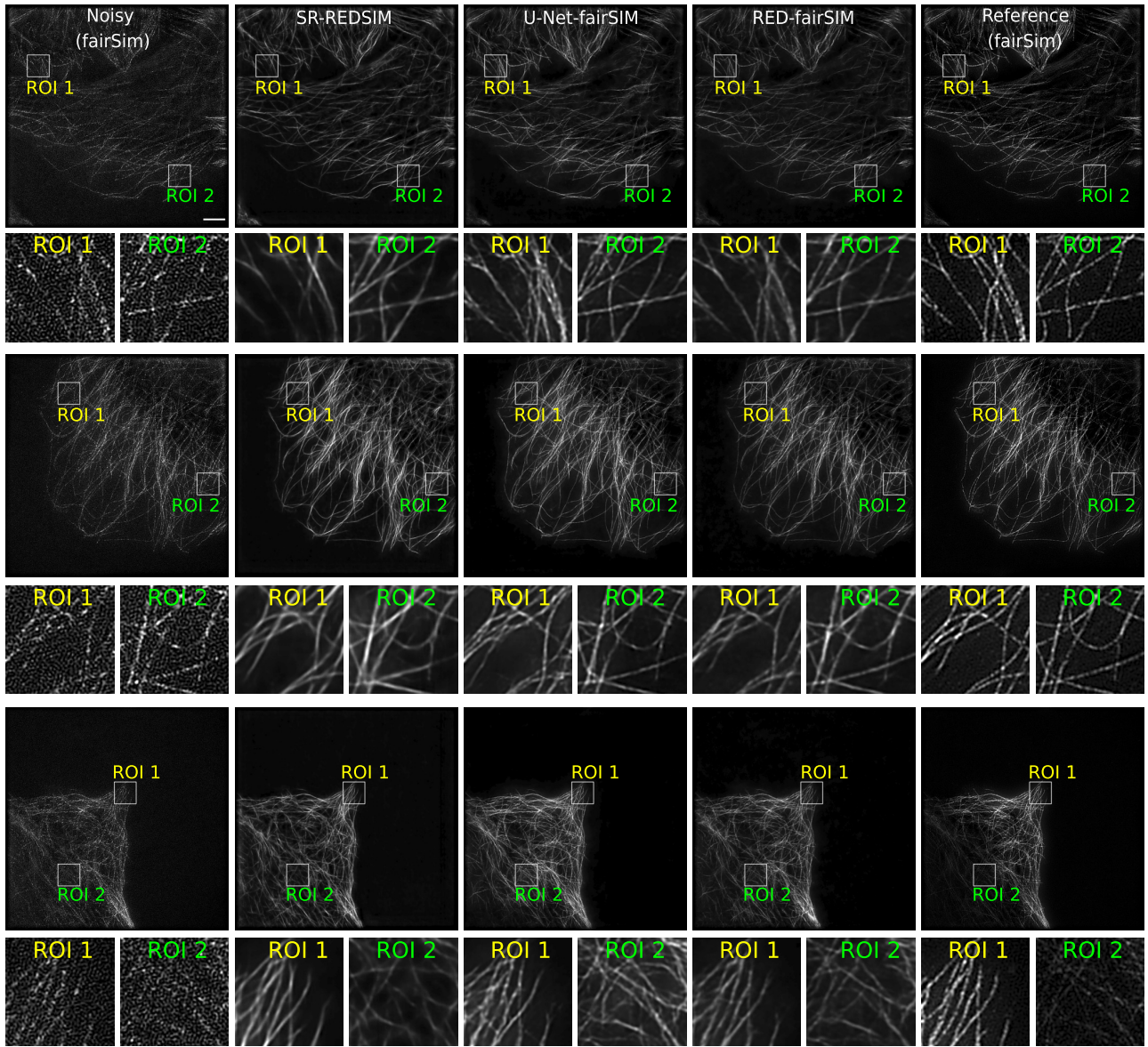
Fig. 3. Super-resolution SIM (SR-SIM) images of three different test samples (U2OS osteosarcoma cells, tubulin cytoskeleton labeled with anti-tubulin-Alexa488) at high-level noise (time stamps 175–199; noise level 4). Each column represents a different reconstruction approach: fairSIM (first column), SR-REDSIM (second column), U-Net-fairSIM (third column), and RED-fairSIM (fourth column). The fifth column depicts the reconstructed reference images which were generated by applying fairSIM image reconstruction to high SNR image data at noise level 0 (lowest noise level; i.e., timestamp 0). All reconstructed SR-SIM images have 1024 × 1024 pixels. The first, third, and fifth rows correspond to the full-size SR-SIM images, whereas the second, fourth, and sixth rows depict magnified ROIs of the white squares (bounding boxes) indicated in the full-size images. The extracted ROIs of size 100 × 100 pixels were upsampled to 300 × 300 pixels using bicubic interpolation for illustration purposes. Scale bar: 4 µm.

Table 1.   Mean PSNR and SSIM Values along with the Standard Deviation for All Methods Calculated on 500 Test Images with Respect to the Reference (Highest SNR) Images

	Mean PSNR (STD) and SSIM (STD) Values at Different Noise Levels							
	Noise level 1		Noise level 2		Noise level 3		Noise level 4	
	PSNR(STD)	SSIM(STD)	PSNR(STD)	SSIM(STD)	PSNR(STD)	SSIM(STD)	PSNR(STD)	SSIM(STD)
fairSIM	28.84(2.47)	0.69(0.10)	25.83(1.70)	0.45(0.09)	24.12(1.64)	0.32(0.07)	23.61(1.54)	0.29(0.07)
SR-REDSIM	26.73(1.93)	0.69(0.07)	26.65(1.55)	0.64(0.07)	26.66(1.73)	0.64(0.07)	26.62(1.79)	0.69(0.09)
U-Net-fairSIM	27.44(1.71)	0.75(0.08)	27.23(1.76)	0.69(0.09)	26.85(1.70)	0.65(0.09)	26.80(1.65)	0.68(0.10)
RED-fairSIM	28.75(1.86)	0.80(0.07)	28.67(1.99)	0.75(0.08)	28.18(2.06)	0.70(0.09)	27.97(2.01)	0.71(0.09)

image from noisy raw SIM images. In the RED-fairSIM method, fairSIM is first used to reconstruct the SIM samples of noisy raw SIM images and then RED-Net is used to denoise the output of fairSIM and to generate high-quality super-resolution images. During the reconstruction process, the stack of 15 noisy images (three angles, five phases) of size 512×512 pixels is again propagated into the fairSIM reconstruction algorithm that then generates a single noisy reconstructed image of size 1024×1024 pixels. The noise and artifacts found in these images do not follow a typical distribution (e.g., Poisson or Gaussian), but have a distinct form that comes from the frequency-based reconstruction algorithm. This single noisy reconstructed image is further passed into the RED-Net architecture to achieve the final result, which can be seen in Fig. 3 (column 4). The complete pipeline of the RED-fairSIM method can be seen in Fig. 1(b) and the architecture of RED-Net is shown in Fig. 2(b). The parameters that were used to generate the SIM reconstructed samples from the raw SIM images are explained in Section 2.E.1. RED-Net was trained in a supervised way where the input–output pairs contain the noisy and reference reconstructed images.

The performance of this method on the unseen test samples is the best among our experiments with respect to PSNR and SSIM values, as shown in Table 1, as well as visually. The ROIs in Fig. 3 show clearly that the output images generated by RED-fairSIM are of high quality with fine details and smooth lines. They are superior compared to the noisy fairSIM, the reference fairSIM, and the SR-REDSIM. Even the artifacts introduced by fairSIM in the reference images are completely removed by RED-fairSIM. It might appear as if the contrast in the denoised SR-REDSIM and RED-fairSIM images in Fig. 3 is weaker when compared to the reference image. This, however, is not the case. The images shown in Fig. 3 are original image data as produced by the various denoising or reconstruction methods without further image processing (i.e. contrast adjustments). Part of the apparent higher contrast in the reference image can also be attributed to graininess of the image, which is caused by uneven antibody staining of the tubulin filaments (a typical phenomenon of antibody staining). It should be noted that the DL-based denoising methods also (at least partly) remove this unevenness.

Furthermore, preliminary tests of RED-fairSIM and SR-REDSIM concerning their ability to generalize to different SIM imaging conditions were carried out, as shown in Fig. 4. As before, U2OS cells were stained for microtubuli, but a dark-red dye with illumination shifted to 642 nm was used, which subsequently also shifts the spatial frequencies of the illumination pattern. The RED-fairSIM approach is able to denoise these images and remove SIM reconstruction artifacts, while the SR-REDSIM approach creates heavy ghosting artifacts. This is unsurprising, as in the case of SR-REDSIM, all specific properties of the SIM pattern (spatial frequencies, orientation, phases) are learned by the network. In the RED-fairSIM approach, parameters specific to the SIM pattern are absorbed by the classic, frequency-domain-based reconstruction, and only reconstruction artifacts are carried into the network. Those artifacts might still depend on the SIM imaging parameters, so
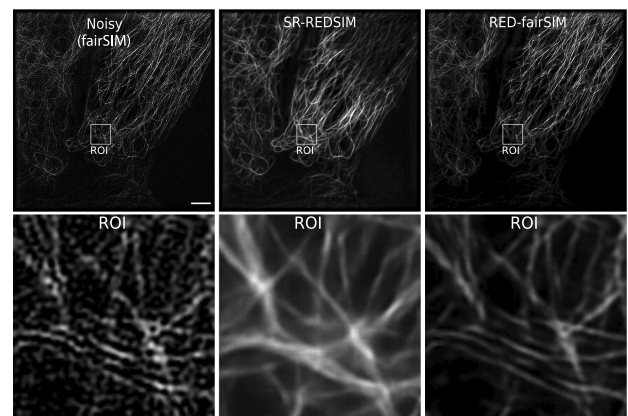
Fig. 4. These SR-SIM images show the difference between the output of the SR-REDSIM and RED-fairSIM methods when applied to imaging conditions that the underlying network was not trained for. To evaluate the generalization capabilities of these methods, we again collected tubulin structure (on U2OS cells), but with a different excitation wavelength. Here, the cell is illuminated by light with a wavelength of 642 nm instead of 488 nm (the latter used for the images in the training set). The different wavelength also changes the spatial frequency of the SIM patterns. This cell structure with unseen illumination properties is then propagated through the pretrained models of both SR-REDSIM and RED-fairSIM. The resulting SR-SIM image shows that RED-fairSIM is more robust against changed microscope settings than SR-REDSIM. Scale bar: 4 μm.

further cross-checks should be carried out. As an initial result, RED-fairSIM seems to generalize well to different SIM pattern settings.

1. U-Net-fairSIM

In the pipeline proposed for RED-fairSIM, we replaced the RED-Net by U-Net and analyzed the resulting super-resolution images. We named this approach as U-Net-fairSIM. Figure 3 (column 3) shows that U-Net-fairSIM also produces better results as compared to the noisy, reference, and SR-REDSIM images. However, it does not surpass RED-fairSIM. Similarly, the U-Net-fairSIM approach outperforms all other counterparts except for RED-fairSIM concerning the PSNR and SSIM values in Table 1. Comparing RED-fairSIM and U-Net-fairSIM directly, as in Fig. 5, the cell structures reconstructed by RED-fairSIM are smoother. Furthermore, they are more faithful when taking the reference as "gold standard" into account. For these reasons, we have focused the presentation in this paper on RED-fairSIM.

2. Data Augmentation

During all of the reported experiments, we did not carry out any preprocessing on the input or output images such as downsampling or cropping, as discussed in Section 2.C. Nonetheless, we tested whether or not image augmentation yields any significant improvement in the results. For this purpose, each image was rotated by an angle of 180° and added in this form to the training set. This increased the amount of training data from 2025 to 4050 images. The mean PSNR and SSIM values after training the methods with the augmentation approach are reported in Table 2. While the image augmentation did not

Table 2. Mean PSNR and SSIM Values along with Standard Deviations (STD) for Noise Level 4 after Applying Data Augmentation (Computed on 500 Test Images)

	PSNR(STD)	SSIM(STD)
SR-REDSIM	26.37(2.07)	0.66(0.10)
U-Net-fairSIM	28.05(2.10)	0.71(0.09)
RED-fairSIM	28.09(2.04)	0.72(0.09)

provide a noticeable advantage overall and did not change the performance-wise ordering of the proposed methods, and instead doubled the training time and increased the preprocessing effort, we decided to focus this paper on the results without image augmentation.

C. Alternative Approaches

1. preRED-fairSIM

Similarly, we also tried to generate high-quality super-resolution SR-SIM images by another method called preRED-fairSIM. The pipeline of preRED-fairSIM is shown in Fig. 6. However, preRED-fairSIM failed to deliver usable results at the end. In preRED-fairSIM, each noisy SIM image from a different phase and orientation is denoised separately and then the whole stack of all 15 denoised images is propagated into the fairSIM algorithm to reconstruct a final super-resolution image. In the preRED-fairSIM approach, we trained a 30-layer RED-Net for one selected phase and orientation and then performed transfer learning [33] (which implies in our scenario no retraining and no changes in the network weights) and fine tuning [34] (which implies adaptation
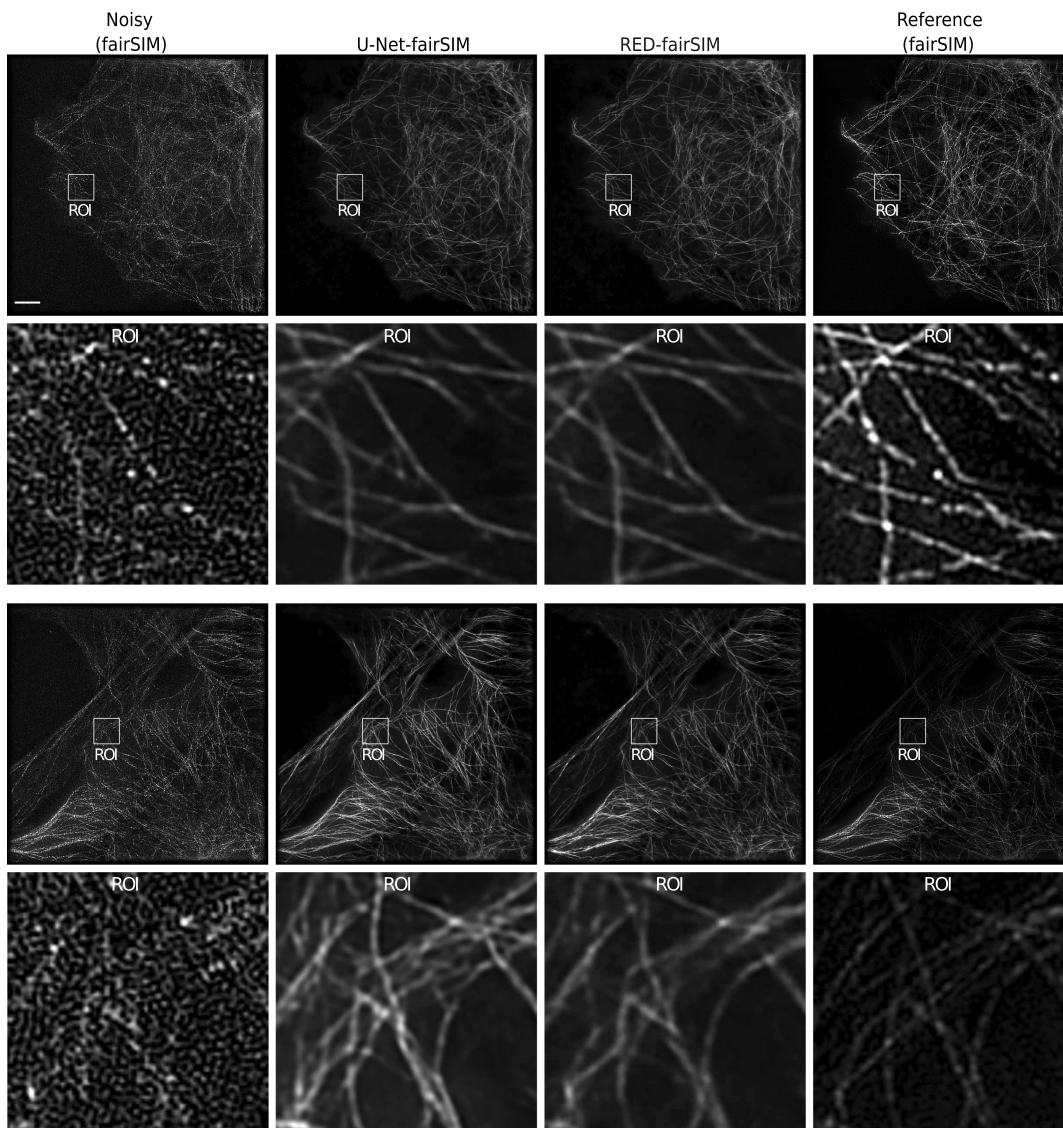
Fig. 5. Reconstruction of SR-SIM images of two different test samples with the fairSIM, U-Net-fairSIM, and RED-fairSIM methods. Each column represents the results of the corresponding method. The first and third rows show the resulting SR-SIM images, whereas the second and fourth rows contain the extracted enlarged ROIs from the full-size images in the rows directly above. The cell structures reconstructed by RED-fairSIM are smoother compared to the U-Net-fairSIM and fairSIM (reference) cell structures. Furthermore, they are more faithful than U-Net-fairSIM when taking the reference as the "gold standard" into account. Scale bar: 4 μm.
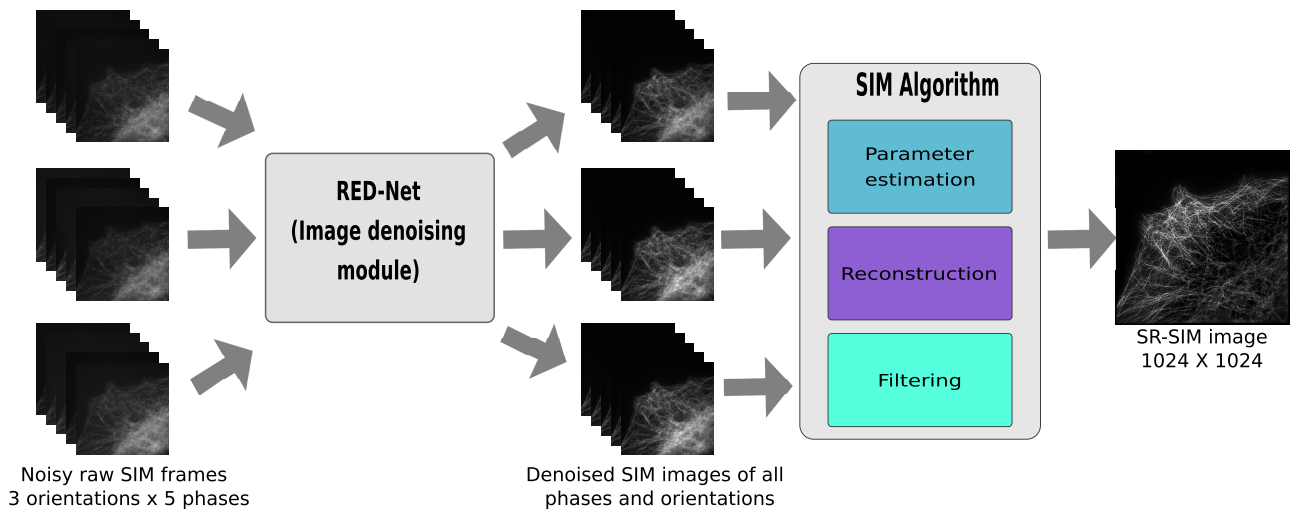
Fig. 6. The complete pipeline of the preRED-fairSIM method. In this pipeline, the raw SIM images [512 × 512 (width × height)] of all phases and orientations are denoised separately with the RED-Net architecture. The complete architecture of RED-Net is shown in Fig. 2. The denoised SIM images of each phase and orientation are then propagated into the fairSIM software in the form of a stack (15 frames) to reconstruct the super-resolution SIM image.

of a subset of the network weights) to other phases and orientations.

The results of transfer learning and fine-tuning are quite promising in terms of the achieved SSIM and PSNR values. The model trained with data of phase 0 and orientation 0 has a mean PSNR = 33.22 dB and a mean SSIM = 0.90 on the test data. Averaged over all other combinations of phase and orientation, transfer learning yields a PSNR = 31.03 dB and an SSIM = 0.89 during testing. After fine-tuning the first and last five layers of the pretrained model, the mean values over all combinations of phase and orientation amount to a PSNR = 33.32 dB and an SSIM = 0.89. These numbers suggest that the denoising of the raw SIM images works well for all orientation phases.

The empirical results of this approach on the image level are shown in Fig. 7 (with fine-tuning applied). They also prove that the raw SIM images are well denoised in the first step of this method. However, fairSIM fails to reconstruct super-resolution images of sufficient quality from the denoised raw images. The resulting reconstructed images contain some additional new artifacts. These artifacts can likely be traced to higher harmonics introduced by the RED-Net in the preRED denoising step, which become very clear in the Fourier power spectrum of the denoised images (see Fig. 7), and then appear similarly as artifacts in the Fourier spectrum of the fully reconstructed image. This is to be expected, as the fairSIM method works in the frequency domain, and highly relies on the precise phases, orientations, and harmonics of the SIM pattern, which the denoising step obviously breaks.

2. Hessian SIM
Hessian SIM [10] is a conventional, frequency-space-based filtering approach, tailored specifically to reduce noise in the band-limited signal of methods such as structured illumination microscopy. Compared to the other methods presented here and compared against, Hessian SIM aims to reduce noise by taking into account both the spatial and temporal frequency distribution in the signal. Thus, while the spatial filtering it offers can be applied to single images, it is most successful when applied to time-lapse data with high temporal sampling. This biases the comparison somewhat in favor of Hessian SIM, because applying the algorithm to the time-lapse data as needed for full performance allows it to see more data than the other algorithms, which can only be applied to a single image. Despite using time-lapse data, the Hessian SIM algorithm does not provide good results for noise level 4, as shown in column 4 of Fig. 8, compared to Red-fairSIM and SR-REDSIM. However, the performance of Hessian SIM was reasonable for the data of lower noise levels.

3. BM3D
BM3D is a conventional state-of-the-art image denoising method from the field of computer vision [35]. During this work, we also used BM3D to denoise the noisy super-resolution images reconstructed by fairSIM. BM3D is able to remove the noise successfully from reconstructed noisy SIM samples, as shown in column 3 of Fig. 8, but fails to recover lost information.

D. SIM Reconstruction at Varying Noise Levels
The raw SIM data in this study was collected as a time-lapse of a fixed sample undergoing photobleaching. Thus, data was collected at different noise levels, which can be assembled into five noise level groups, with noise level 0 representing the lowest and noise level 4 representing the highest level of noise found in the data. As we have previously discussed, the models were trained with input from the highest noise level. Therefore, we investigated whether these pretrained models will also be useful for the SIM images of other lower noise levels. To verify this, we considered only the two best methods from our work: SR-REDSIM and RED-fairSIM are used to evaluate the raw SIM images at different levels. No fine-tuning [34] or transfer learning [33] was performed on these pretrained models.
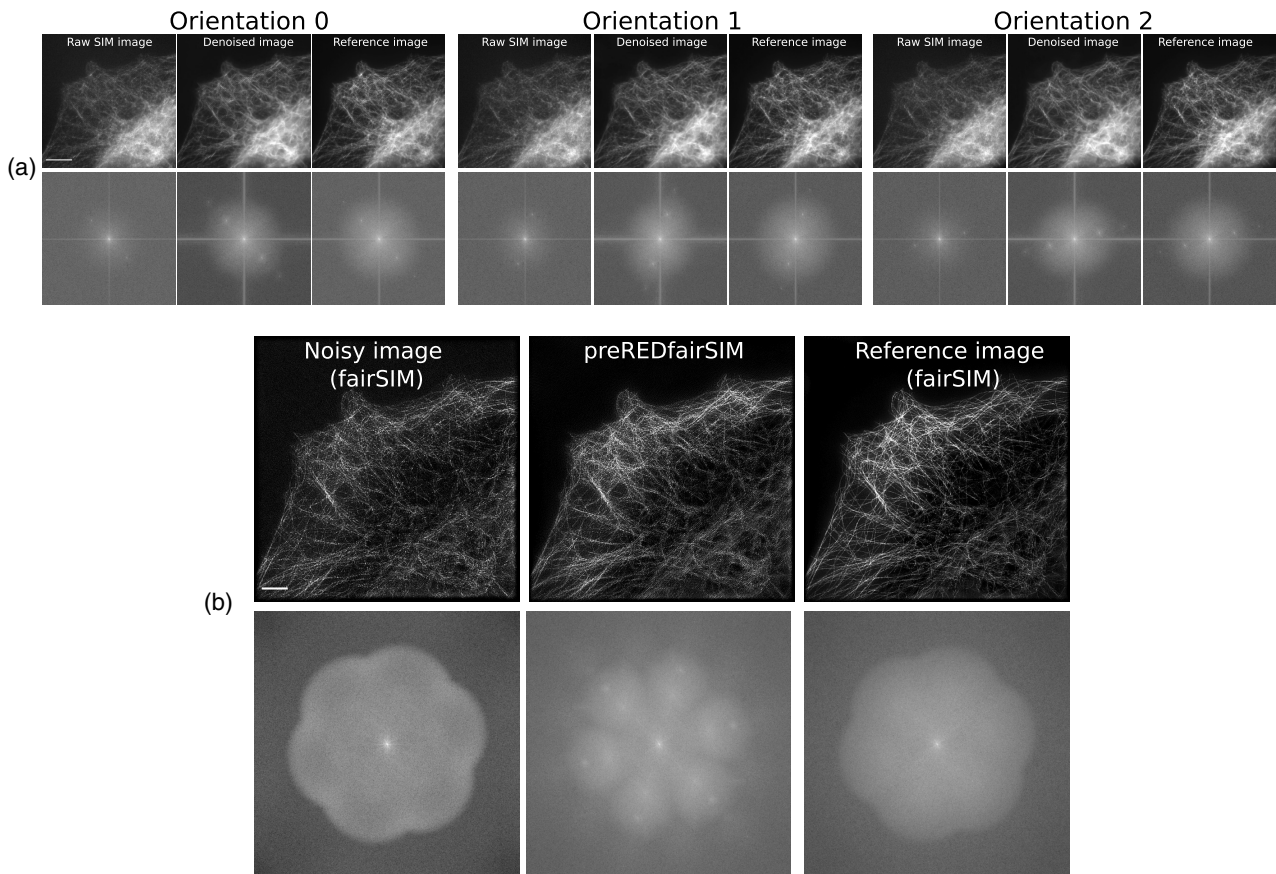
Fig. 7. Results for preRED-fairSIM. (a) Three blocks of images where each block consists of six images. The first block depicts the images from phase 0 and orientation 0, the second block from phase 1 and orientation 1, and the third block from phase 2 and orientation 2. The left image in the first row of each block represents a noisy raw SIM image from noise level 4. The second image in each block is the denoised version, whereas the reference image (rightmost in each block) is the ground truth. The Fourier spectra of the images are shown below each image. The dimensions of each image in these blocks are 512×512 (width × height). Scale bar: 8 μm. (b) Three images reconstructed by fairSIM along with their Fourier spectrum directly below. The noisy image (left) is reconstructed using 15 noisy raw SIM images and the parameter fit summary is: resolution improvement is $(x:1.90, y:1.90, z:1.90)$, and the modulation estimation is $(x:0.310, y:0.341, z:0.332)$ with the assessment as "weak." Similarly, the preRED-fairSIM image (middle) is generated using 15 denoised SIM images and the parameter fit summary is: modulation estimation is $(x:0.310, y:0.341, z:0.332)$ and the assessment "weak"; however, there is no improvement in the resolution. The reference image (right) is reconstructed from the raw SIM images with the highest SNR and the parameter fit summary is: resolution improvement is $(x:1.90, y:1.90, z:1.90)$, and the modulation estimation is $(x:0.558, y:0.580, z:0.578)$ with an assessment of "usable." The Fourier spectrum of preRED-fairSIM shows additional artifacts (white spots) that do not exist in the Fourier spectrum of the reference or the noisy output. Scale bar: 4 μm.

Figure 9 shows the results of this attempt. In this figure, one specific sample was captured at different noise levels. If we further examine the ROIs of all the super-resolution images, it can be seen that both methods show high-quality, super-resolved images at all of the five different noise levels with a slight degradation toward higher noise levels regarding smoothness and clarity of the cell structures. Furthermore, it is again noticeable that the results of RED-fairSIM are overall visually more appealing compared to the other methods.

In addition to visual inspection, quantitative results are given in Table 1, which contains the mean PSNR and SSIM values of 500 test inputs from each noise level. Here, U-Net-fairSIM is included in the comparison. Both RED-fairSIM and U-Net-fairSIM show a gradual decrease in PSNR and SSIM values from noise level 1 (weak noise) to noise level 4 (strong noise). SR-REDSIM performs similarly, but with

noticeably smaller PSNR and SSIM values. The results of fairSIM without denoising deteriorate quickly when moving to higher noise levels. The most important takeaway from Table 1 and Fig. 9 is that the networks—although trained for a specific high noise level—generalize well to conditions with a better SNR.

The reconstruction of the lower noise level 0 through SR-REDSIM and RED-fairSIM highlights a second use case. At this noise level, the SNR of the raw frames is high enough to provide the reference data sets, which, as discussed before, are of high quality, but still feature some SIM reconstruction artifacts. Those artifacts are successfully removed by both SR-REDSIM and RED-fairSIM. A reasonable assumption is that, like noise, reconstruction artifacts are random enough in nature, so they are not picked up by the network during training, and thus cannot be reproduced. This effect is well
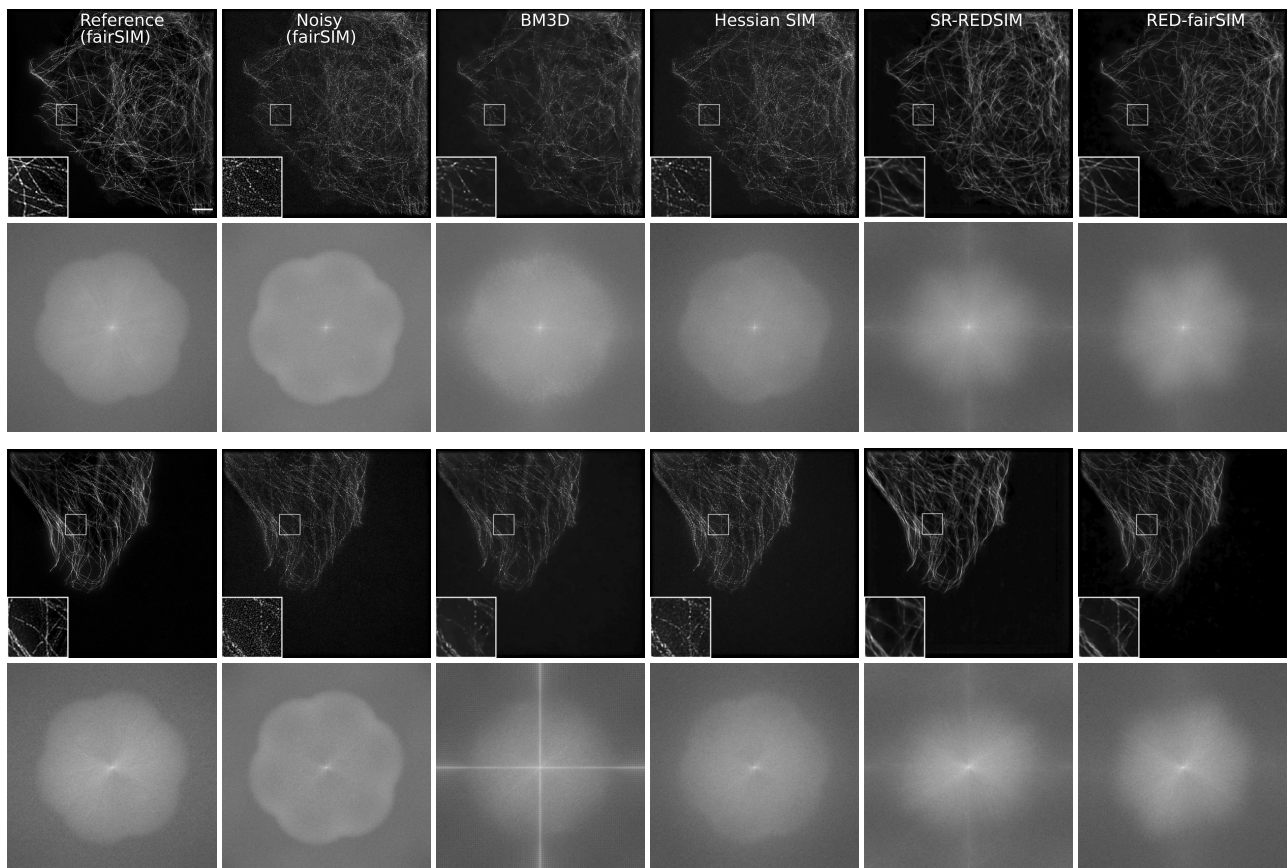
Fig. 8. The reconstructed SR-SIM images of two test samples with different methods. The Fourier spectrum of each SR-SIM image is shown directly below. Each image contains an enlarged ROI at the lower-left bottom. The analysis of ROI of all the methods clearly shows that the results of RED-fairSIM (sixth column) are smoother and more faithful compared to all other methods. Similarly, the Fourier spectra of the RED-fairSIM do not show any additional artifacts in the Fourier space. The SR-SIM images and ROI of SR-REDSIM (fifth column) also show good results; however, there are some artifacts in the high-frequency region of the Fourier spectrum. The ROIs of Hessian SIM (fourth column) do not outperform RED-fairSIM and SR-REDSIM, but do show better results than BM3D (third column). BM3D produces a suppressed cell structure in both of the resultant images. Furthermore, the Fourier spectrum of the BM3D result for the second test sample shows artifacts in both low and high-frequency regions. Scale bar: 4 μm.

known in applications such as Noise2Noise [36], where the inability of a neural network to learn random (noisy) data is explicitly used for denoising.

4. DISCUSSION

The results of this study provide sufficient evidence that SR-REDSIM and RED-fairSIM can both be employed to denoise and reconstruct high-quality SR-SIM images. In contrast, preRED-fairSIM in its current form is not suitable for this purpose because the output of RED-Net, although successfully denoised, contains additional artifacts noticeable in Fourier space, which spoil the performance of the subsequent classical computational SIM reconstruction. We also investigated the robustness of the successful methods (SR-REDSIM and RED-fairSIM) and showed that high-quality reconstruction of SIM samples is possibly irrespective of the noise level in the raw SIM images. The SR-REDSIM and RED-fairSIM methods outperform their counterparts, as shown in Figs. 3 and 4. Furthermore, these approaches are useful even in the absence

of clean ground-truth data, as we have shown especially for RED-fairSIM where the reference data used for training contains many reconstruction artifacts. We have also shown in column of Fig. 9 that the proposed methods SR-REDSIM and RED-fairSIM can be used to remove the reconstruction artifacts from the reference image after training, so even if high SNR data can be acquired easily, SR-REDSIM and RED-fairSIM still offer an improvement over the classical reconstruction approaches.

A recent study [22] used cycle-consistent generative adversarial networks (CycleGANs) [37] to reconstruct SR-SIM images by using three to nine clean raw SIM images. A CycleGAN contains two generators and two discriminators with multiple losses that are trained in a competitive process. Therefore, CycleGANs are generally very difficult to train. Furthermore, the authors did not address the challenge of denoising. Christensen *et al.* [21] trained deep neural networks by using synthetic data instead of real microscope SIM images to reconstruct SR-SIM images. Although the synthetic data used in their studies for training is unrelated to real micro-
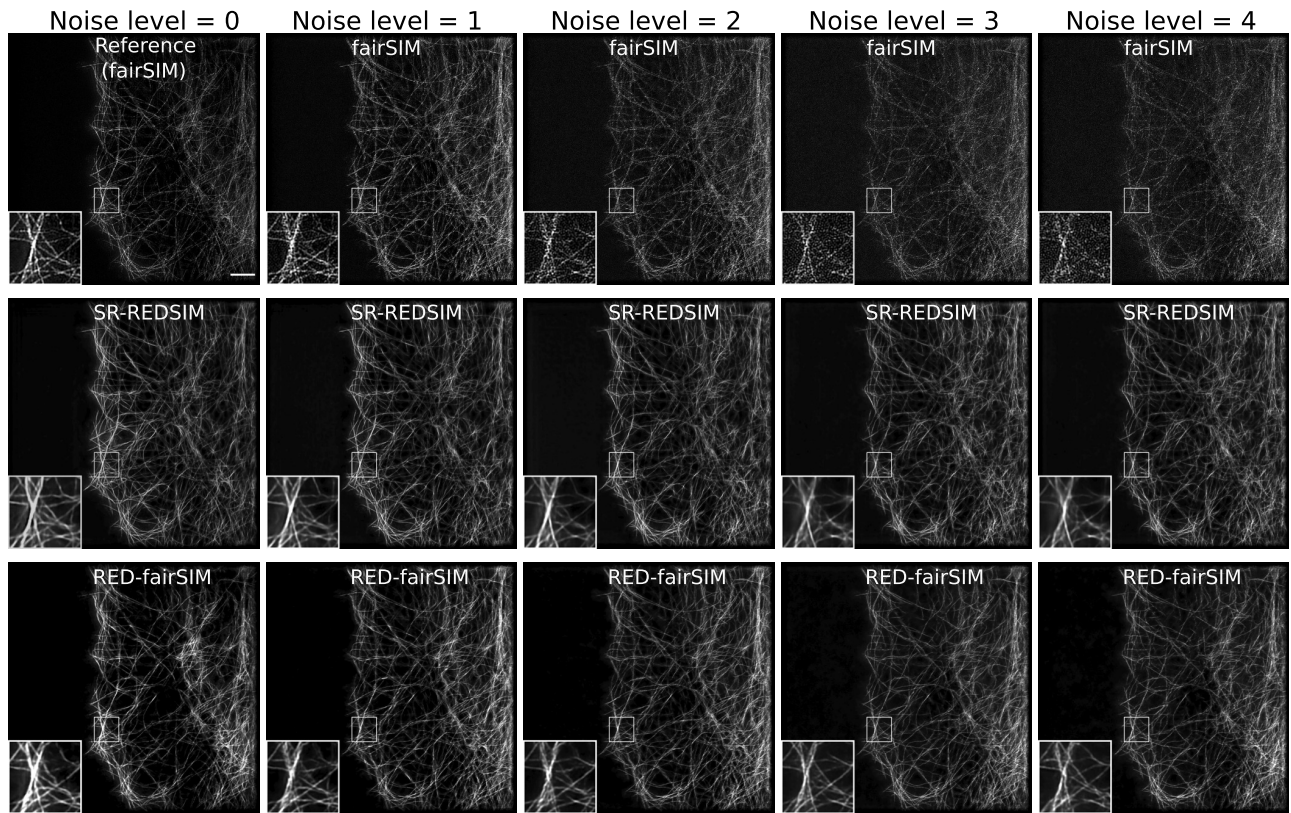
Fig. 9. Reconstructed SR-SIM images at different noise levels with the SR-REDSIM and RED-fairSIM methods for a single field of view. Noise level 0 represents the reference image at timestamp 0, noise level 1 comprises the images from timestamps 25–50, noise level 2 from timestamps 75–100, noise level 3 from timestamps 125–150, and noise level 4 from timestamps 175–200. Each full image contains an enlarged ROI in the bottom left. The images reconstructed by fairSIM in the first row show a significant degradation in quality as the noise level increases. In contrast, the results produced by SR-REDSIM and RED-fairSIM in the second and third columns are far less, depending on the noise level. Scale bar: 4 μm.

scopes, they were successful in generating output comparable to computational tools like fairSIM. However, they did not use real noisy microscope data to test the denoising performance of their networks, and their approach was also not completely successful in the case of (simulated) high-level noise. Jin *et al.* [20] used multiple concatenated U-Nets to reconstruct SR-SIM images by using three to 15 raw SIM images. They trained their models on cropped and resized SIM samples and manually discarded tiles with only background information. These preprocessing steps are time-consuming, and the training of two adjacent U-Net models is also computationally expensive.

Our proposed methods use raw SIM images in their original size, which does not involve any major preprocessing steps. The amount of training data used, about 100 fields of view for training and test data together, is also small enough that specific training, capturing both a given instrument and a specific biological structure of interest, should often be feasible. While SR-REDSIM has similarities to other proposed end-to-end deep learning approaches for SIM [20–22], to the best of our knowledge, RED-fairSIM is a completely novel deep learning approach for SIM which is, as our data shows, superior to SR-REDSIM.

While both SR-REDSIM and RED-fairSIM provide high-quality reconstruction, an obvious difference between them is

their ability to generalize to different SIM imaging settings. As an initial test, we varied the spatial frequency of the SIM pattern [using a 642 nm (instead of 488 nm) excitation light], which commonly happens when designing experiments and choosing dyes, as shown in Fig. 4. We then performed reconstruction with RED-fairSIM and SR-REDSIM, both trained on the original 488 nm data. Here, the RED-fairSIM approach, where the change in spatial frequency of the pattern is absorbed by the classic reconstruction step, still works very well in suppressing noise and SIM artifacts. SR-REDSIM, on the other hand, where the SIM pattern has been learned by the network, created heavy ghosting artifacts. While further validation and cross-testing are needed, this suggests that RED-fairSIM should be able to generalize to different SIM microscopes, excitation wavelengths, and probably illumination types (three-beam, two-beam, TIRF-SIM), while SR-REDSIM would require retraining whenever larger changes in these parameters occur.

Besides visual impression, the quantitative measures chosen for our comparisons are PSNR and SSIM. As a microscopy technique, obviously spatial resolution estimates would present another desirable parameter. Both Fourier ring correlation (FRC) [38,39] and image decorrelation analysis (IDA) [40] are typically chosen for this task, because they offer a quantitative resolution estimate that is not dependent on manual

measurement of single structural features, which easily introduces bias. However, much care is needed to correctly apply these methods, as they are based on strict assumptions about the input data, regarding the statistically independent and spatially uncorrelated distribution of noise. Even slight correlation, introduced, for example, by camera readout electronics [41], will yield an overestimation of spatial resolution. We have performed FRC analysis and IDA on the datasets presented here, and while some results seem reasonable, others are clearly unphysical. As we cannot assume that the denoised output fulfills the assumptions of FRC or IDA, we would argue against using unmodified FRC or IDA to estimate resolution in these images. We believe further research into robust resolution estimation that works more independently of the data generation process is needed.

We also compared our set of deep-learning based approaches to classical denoising algorithms, with BM3D as a general approach and Hessian SIM as a noise filter tailored to ideally time-lapsed SIM data. Both of the classical algorithms do not reach the denoising performance of the deep learning methods. However, these algorithms do not require any prior knowledge, in the form of training data, of the sample. Thus, in contrast to deep learning methods, they can be applied to arbitrary, unknown structures. They can also serve as a cross-check, if concerns arise that the deep-learning based approaches generate artifacts stemming from their training data.

5. CONCLUSION

In this work, we presented two different methods, SR-REDSIM and RED-fairSIM, to reconstruct super-resolution SIM images from raw SIM images with low SNR. We demonstrated that these methods are robust against different noise intensities and do not need any retraining or fine-tuning even if the SNR is varied between the training and application. However, the generalization ability of RED-fairSIM under different SIM imaging conditions (i.e., changed microscope settings) is superior compared to SR-REDSIM. This shows that the combination of fairSIM for reconstruction and RED-Net for denoising is more promising than an end-to-end deep learning approach like SR-REDSIM. Both methods are particularly useful for SIM images with a low SNR since the traditional reconstruction algorithms cannot denoise and generate reconstruction artifacts in the SR-SIM images. Both of our proposed methods can remove these reconstruction artifacts. The overall results also show that our methods outperform other classical denoising methods like BM3D to denoise the noisy SR-SIM images. Furthermore, the proposed methods can potentially be used in the future to handle live-cell SIM imaging data as well as the reconstruction and denoising of SIM images with low SNR from other biological structures.

Disclosures. The authors declare no conflicts of interest.

REFERENCES

1. L. Schermelleh, A. Ferrand, T. Huser, C. Eggeling, M. Sauer, O. Biehlmaier, and G. P. Drummen, "Super-resolution microscopy demystified," Nat. Cell Biol. **21**, 72–84 (2019).
2. J. Demmerle, C. Innocent, A. J. North, G. Ball, M. Müller, E. Miron, A. Matsuda, I. M. Dobbie, Y. Markaki, and L. Schermelleh, "Strategic and practical guidelines for successful structured illumination microscopy," Nat. Protoc. **12**, 988–1010 (2017).
3. R. Heintzmann and T. Huser, "Super-resolution structured illumination microscopy," Chem. Rev. **117**, 13890–13908 (2017).
4. M. G. Gustafsson, "Surpassing the lateral resolution limit by a factor of two using structured illumination microscopy," J. Microsc. **198**, 82–87 (2000).
5. L. M. Hirvonen, K. Wicker, O. Mandula, and R. Heintzmann, "Structured illumination microscopy of a living cell," Eur. Biophys. J. **38**, 807–812 (2009).
6. P. Kner, B. B. Chhun, E. R. Griffis, L. Winoto, and M. G. Gustafsson, "Super-resolution video microscopy of live cells by structured illumination," Nat. Methods **6**, 339–342 (2009).
7. L. Shao, P. Kner, E. H. Rego, and M. G. Gustafsson, "Super-resolution 3D microscopy of live whole cells using structured illumination," Nat. Methods **8**, 1044–1046 (2011).
8. L. Gao, L. Shao, C. D. Higgins, J. S. Poulton, M. Peifer, M. W. Davidson, X. Wu, B. Goldstein, and E. Betzig, "Noninvasive imaging beyond the diffraction limit of 3D dynamics in thickly fluorescent specimens," Cell **151**, 1370–1385 (2012).
9. R. Fiolka, L. Shao, E. H. Rego, M. W. Davidson, and M. G. Gustafsson, "Time-lapse two-color 3D imaging of live cells with doubled resolution using structured illumination," Proc. Natl. Acad. Sci. USA **109**, 5311–5315 (2012).
10. X. Huang, J. Fan, L. Li, H. Liu, R. Wu, Y. Wu, L. Wei, H. Mao, A. Lal, P. Xi, L. Tang, Y. Zhang, Y. Liu, S. Tan, and L. Chen, "Fast, long-term, super-resolution imaging with Hessian structured illumination microscopy," Nat. Biotechnol. **36**, 451–459 (2018).
11. A. Markwirth, M. Lachetta, V. Mönkemöller, R. Heintzmann, W. Hübner, T. Huser, and M. Müller, "Video-rate multi-color structured illumination microscopy with simultaneous real-time reconstruction," Nat. Commun. **10**, 4315 (2019).
12. A. Sandmeyer, M. Lachetta, H. Sandmeyer, W. Hübner, T. Huser, and M. Müller, "DMD-based super-resolution structured illumination microscopy visualizes live cell dynamics at high speed and low cost," bioRxiv 797670 (2019).
13. M. G. Gustafsson, L. Shao, P. M. Carlton, C. R. Wang, I. N. Golubovskaya, W. Z. Cande, D. A. Agard, and J. W. Sedat, "Three-dimensional resolution doubling in wide-field fluorescence microscopy by structured illumination," Biophys. J. **94**, 4957–4970 (2008).

14. M. Müller, V. Mönkemöller, S. Hennig, W. Hübner, and T. Huser, "Open-source image reconstruction of super-resolution structured illumination microscopy data in ImageJ," Nat. Commun. **7**, 10980 (2016).

15. A. Lal, C. Shan, and P. Xi, "Structured illumination microscopy image reconstruction algorithm," IEEE J. Sel. Top. Quantum Electron. **22**, 50–63 (2016).

16. P. Křžek, T. Lukeš, M. Ovesný, K. Fliegel, and G. M. Hagen, "SIMToolbox: a MATLAB toolbox for structured illumination fluorescence microscopy," Bioinformatics **32**, 318–320 (2016).

17. K. Wicker, O. Mandula, G. Best, R. Fiolka, and R. Heintzmann, "Phase optimisation for structured illumination microscopy," Opt. Express **21**, 2032–2049 (2013).

18. J. Fan, X. Huang, L. Li, S. Tan, and L. Chen, "A protocol for structured illumination microscopy with minimal reconstruction artifacts," Biophys. Rep. **5**, 80–90 (2019).

19. D. P. Hoffman and E. Betzig, "Tiled reconstruction improves structured illumination microscopy," bioRxiv 895318 (2020).

20. L. Jin, B. Liu, F. Zhao, S. Hahn, B. Dong, R. Song, T. C. Elston, Y. Xu, and K. M. Hahn, "Deep learning enables structured illumination microscopy with low light levels and enhanced speed," Nat. Commun. **11**, 1934 (2020).

21. C. N. Christensen, E. N. Ward, P. Lio, and C. F. Kaminski, "ML-SIM: a deep neural network for reconstruction of structured illumination microscopy images," arXiv:2003.11064 (2020).

22. C. Ling, C. Zhang, M. Wang, F. Meng, L. Du, and X. Yuan, "Fast structured illumination microscopy via deep learning," Photon. Res. **8**, 1350–1359 (2020).

23. M. Weigert, U. Schmidt, T. Boothe, A. Müller, A. Dibrov, A. Jain, B. Wilhelm, D. Schmidt, C. Broaddus, S. Culley, M. Rocha-Martins, F. Segovia-Miranda, C. Norden, R. Henriques, M. Zerial, M. Solimena, J. Rink, P. Tomancak, L. Royer, F. Jug, and E. W. Myers, "Content-aware image restoration: pushing the limits of fluorescence microscopy," Nat. Methods **15**, 1090–1097 (2018).

24. X. Mao, C. Shen, and Y.-B. Yang, "Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections," in *Advances in Neural Information Processing Systems* (2016), pp. 2802–2810.

25. D. Sage, H. Kirshner, T. Pengo, N. Stuurman, J. Min, S. Manley, and M. Unser, "Quantitative evaluation of software packages for single-molecule localization microscopy," Nat. Methods **12**, 717–724 (2015).

26. T. Novák, T. Gajdos, J. Sinkó, G. Szabó, and M. Erdélyi, "TestSTORM: versatile simulator software for multimodal super-resolution localization fluorescence microscopy," Sci. Rep. **7**, 951 (2017).

27. B. Lim, S. Son, H. Kim, S. Nah, and K. Lee, "Enhanced deep residual networks for single image super-resolution," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* (2017), pp. 1132–1140.

28. Y. Zhang, K. Li, K. Li, L. Wang, B. Zhong, and Y. Fu, "Image super-resolution using very deep residual channel attention networks," in *European Conference on Computer Vision (ECCV)* (2018), pp. 286–301.

29. C. Karras, M. Smedh, R. Förster, H. Deschout, J. Fernandez-Rodriguez, and R. Heintzmann, "Successful optimization of reconstruction parameters in structured illumination microscopy–a practical guide," Opt. Commun. **436**, 69–75 (2019).

30. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," arXiv:1505.04597 (2015).

31. J. F. Abascal, S. Bussod, N. Ducros, S. Si-Mohamed, P. Douek, C. Chappard, and F. Peyrin, "A residual U-Net network with image prior for 3D image denoising," HAL hal-02500664 (2020).

32. A. Hore and D. Ziou, "Image quality metrics: PSNR vs. SSIM," in *20th International Conference on Pattern Recognition* (2010), pp. 2366–2369.

33. L. Torrey and J. Shavlik, "Transfer learning," in *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques* (IGI Global, 2010), pp. 242–264.

34. J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," in *56th Annual Meeting of the Association for Computational Linguistics* (2018), pp. 328–339.

35. K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," IEEE Trans. Image Process. **16**, 2080–2095 (2007).

36. J. Lehtinen, J. Munkberg, J. Hasselgren, S. Laine, T. Karras, M. Aittala, and T. Aila, "Noise2Noise: learning image restoration without clean data," in *35th International Conference on Machine Learning* (2018), pp. 2965–2974.

37. J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *IEEE International Conference on Computer Vision* (2017), pp. 2223–2232.

38. M. Van Heel and M. Schatz, "Fourier shell correlation threshold criteria," J. Struct. Biol. **151**, 250–262 (2005).

39. R. P. Nieuwenhuizen, K. A. Lidke, M. Bates, D. L. Puig, D. Grünwald, S. Stallinga, and B. Rieger, "Measuring image resolution in optical nanoscopy," Nat. Methods **10**, 557–562 (2013).

40. A. C. Descloux, K. S. Grussmayer, and A. Radenovic, "Parameter-free image resolution estimation based on decorrelation analysis," Nat. Methods **16**, 918–924 (2019).

41. R. Van den Eynde, A. Sandmeyer, W. Vandenberg, S. Duwé, W. Hübner, T. Huser, P. Dedecker, and M. Müller, "Quantitative comparison of camera technologies for cost-effective super-resolution optical fluctuation imaging (SOFI)," J. Phys. Photon. **1**, 044001 (2019).