LASER & PHOTONICS REVIEWS
www.lpr-journal.org

Deep Learning-Enabled Pixel-Super-Resolved Quantitative Phase Microscopy from Single-Shot Aliased Intensity Measurement

Jie Zhou, Yanbo Jin, Linpeng Lu, Shun Zhou, Habib Ullah, Jiasong Sun, Qian Chen, Ran Ye, Jiaji Li,* and Chao Zuo**

A new technique of deep learning-based pixel-super-resolved quantitative phase microscopy (DL-SRQPI) is proposed, achieving rapid wide-field high-resolution and high-throughput quantitative phase imaging (QPI) from single-shot low-resolution intensity measurement. By training a neural network with sufficiently paired low-resolution intensity and high-resolution phase data, the network is empowered with the capability to robustly reconstruct high-quality phase information from a single frame of an aliased intensity image. As a graphics processing units-accelerated computational method with minimal data requirement, DL-SRQPI is well-suited for live-cell imaging and accomplishes high-throughput long-term dynamic phase reconstruction. The effectiveness and feasibility of DL-SRQPI have been significantly demonstrated by comparing it with other traditional and learning-based phase retrieval methods. The proposed method has been successfully implemented into the quantitative phase reconstruction of biological samples under bright-field microscopes, overcoming pixel aliasing and improving the spatial-bandwidth product significantly. The generalization ability of DL-SRQPI is illustrated by phase reconstruction of Henrietta Lacks cells at various defocus distances and illumination patterns, and its high-throughput anti-aliased phase imaging performance is further experimentally validated. Given its capability of achieving pixel super-resolved QPI from single-shot intensity measurement over conventional bright-field microscope hardware, the proposed approach is expected to be widely adopted in life science and biomedical workflows.

1. Introduction

Optical microscopy has undergone continuous development since its invention in the 17th century and has gradually become an essential tool for visualizing cellular and subcellular features of biological samples, driven by the increasing demand for biomedical research.[1] However, generating sufficient contrast in most biological samples is challenging due to their low absorption or weak-scattering characteristic.[1,2] To obtain their precise and detailed phase information, extensive research has been conducted for decades. Fluorescence microscopy is one of the most far-reaching developments for weak absorption object visualization. It labels the specimen with fluorescent molecules to provide targeted morphological and biochemical information. With the emergence of new fluorescent molecular probes and novel optical imaging techniques, advanced super-resolution fluorescence microscopy further enables super-resolution subcellular detail observation at the nano-scale well beyond the diffraction limit, such as structured illumination microscopy

J. Zhou, Y. Jin, L. Lu, S. Zhou, H. Ullah, J. Sun, Q. Chen, R. Ye, J. Li, C. Zuo
Smart Computational Imaging (SCI) Laboratory
Nanjing University of Science and Technology
Nanjing, Jiangsu Province 210094, China
E-mail: ran.ye@njnu.edu.cn; jiajili@njust.edu.cn; zuochao@njust.edu.cn

J. Zhou, Y. Jin, L. Lu, S. Zhou, H. Ullah, J. Sun, Q. Chen, J. Li, C. Zuo
Smart Computational Imaging Research Institute (SCIRI) of Nanjing University of Science and Technology
Nanjing, Jiangsu Province 210019, China

J. Zhou, Y. Jin, L. Lu, S. Zhou, H. Ullah, J. Sun, Q. Chen, J. Li, C. Zuo
Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense
Nanjing University of Science and Technology
Nanjing, Jiangsu Province 210094, China

R. Ye
School of Computer and Electronic Information
Nanjing Normal University
Nanjing, Jiangsu Province 210023, China

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/lpor.202300488

DOI: 10.1002/lpor.202300488

(SIM),[3] stimulated emission depletion microscopy (STED),[4,5] photo-activated localization microscopy (PALM), and stochastic optical reconstruction microscopy (STORM).[6] However, the utilization of exogenous agents may introduce photo-toxicity and photo-bleaching issues, which hinder the long-term imaging of living cells. Furthermore, the use of fluorescent dyes and proteins as bio-markers inevitably limits certain non-fluorescent applications where biological samples cannot be easily tagged with fluorescent markers.[7,8]

In recent years, the technique of computational microscopy, including interferometric[9–11] and non-interferometric[12–14] manners for both quantitative phase imaging (QPI)[15–19] and 3D refractive index (3D RI),[20,21] has been proved to be an invaluable tool regarding its distinctive capability to quantify the phase delay of unlabeled biological specimens in a non-destructive way. As two representative QPI approaches, transport of intensity equation (TIE)[22] and Fourier ptychographic microscopy (FPM)[23] have gained wide attention in the application of biomedicine. With a simple optical implementation of an off-the-shelf bright-field microscope, the phase distribution of specimen can be simply reconstructed by TIE using intensity measurements at multiple axially displaced planes. Nevertheless, the achievable imaging resolution of TIE is restricted to the incoherent diffraction limit under partially coherent illumination, and the spatial bandwidth product (SBP) of TIE is fundamentally restrained by the optical system, resulting in a trade-off between imaging resolution and field-of-view (FOV).[24,25] FPM is a recently developed computational imaging technique that could circumvent the imaging resolution-FOV trade-off and improve the throughput of the imaging system.[23,26] FPM maintains high imaging resolution and wide FOV simultaneously by stitching together a series of variously illuminated low-resolution but large-FOV intensity images in Fourier space. However, FPM requires a large amount of data redundancy, which leads to a cumbersome data acquisition process. Additionally, the iterative strategy used by FPM limits its recovery efficiency, preventing its application in high-speed cell imaging.

On the other hand, high-throughput QPI faces another major obstacle posed by pixel-aliasing.[27] In optical systems, detectors are used to collect intensity information and are typically designed with large pixel sizes to accommodate high photosensitivity and large FOVs for high-throughput imaging. However, large pixel sizes may lead to inadequate sampling or digitization of the transmitted intensity, resulting in low pixel resolution and even leading to the infamous pixel-aliasing/undersampling problem. Although deploying magnification camera adapters or using image sensors with smaller pixel sizes could mitigate the pixel-aliasing problem, it comes at the cost of the FOV. Therefore, this trade-off between pixel resolution and FOV leads to sub-optimal use of SBP of the imaging system. Several QPI techniques with anti-aliased ability have been proposed. For instance, the pixel-aliasing in differential phase contrast (DPC) could be alleviated by the iterative de-multiplexing algorithm. However, its efficacy is still restrained by the elaborate illumination scheme, the requirement of multiple intensity images and the iterative strategy.

Benefiting from the accelerating development in computer science and technology, coupled with exponential growth in processing power, the past few years have witnessed rapid progress in deep learning, where high-dimensional representations can be learned directly from captured data based on neural networks. With its unique data-driven methodology, deep learning has solved many tasks in computer vision and computer-aided diagnosis with unprecedented performance.[28] In the field of computational microscopy, deep learning has led to rapid growth in algorithms and methods for solving various ill-posed inverse problems, transcending the limitations of traditional microscopy.[29] For example, deep learning enables super-resolution imaging and reveals microscopic biological details with higher precision.[30] It has also been proved that a conventional microscope aided by deep learning could even enable the observation of nano-scale subcellular details well beyond the diffraction limit, reaching the image resolution of STED.[31] Deep learning also realizes cross-modality imaging of biomedical samples, such as the digital staining technique that accommodates the generation of quantitative phase images for virtual histological staining, therefore circumventing the procedure of laborious and time-consuming sample staining.[32] Thanks to its powerful non-linear ability, deep learning has also been utilized to enhance the phase information acquisition capability of conventional computational imaging techniques by constructing precise mapping relationships between intensity and phase distributions. With a well-trained neural network, FPM can reduce the number of required intensity images from hundreds to five, eliminating the tedious image acquisition process and time-consuming iterations, while maintaining the quality of the reconstructed phase images.[29,33] However, these methods still require multiple input images for phase recovery. The end-to-end capability of deep learning implies that the data redundancy requirement for phase retrieval can be further minimized to single frame. Based on Gerchberg and Saxton (GS)'s iterative phase retrieval algorithm, a deep neural network has successfully obtained accurate amplitude and phase information from a single coaxial hologram amidst the interference of twin images and object artifacts.[34] Another deep learning-assisted method achieves TIE-based phase retrieval from a single intensity image.[35] Nevertheless, though the data efficiency has been improved, the image resolution is still limited, and phase recovery for a large population of cells remains to be investigated. Besides, the problem of pixel-aliasing requires further exploration. Consequently, high-throughput quantitative phase imaging with both wide FOV and pixel super-resolution from single-frame intensity image in bright-field optical implementation has not been proposed yet.

In this work, we present a novel quantitative phase imaging technique, termed deep learning-based single-frame super-resolution quantitative phase imaging (DL-SRQPI). Our method combines deep learning with quantitative phase imaging and achieves high-throughput, high-accuracy phase retrieval in a computational manner without any additional hardware design. After proper training, a neural network identifies the mapping relationship between low-resolution intensity image and high-resolution phase image, with which DL-SRQPI alleviates the pixel-aliasing problem and improves the space-bandwidth product since the inherent large FOV of the low-resolution intensity image is exploited. DL-SRQPI maximizes the data efficiency by reducing the intensity image redundancy requirement to only one frame, and the phase reconstruction speed is greatly accelerated by utilizing the graphics processing unit (GPU). The

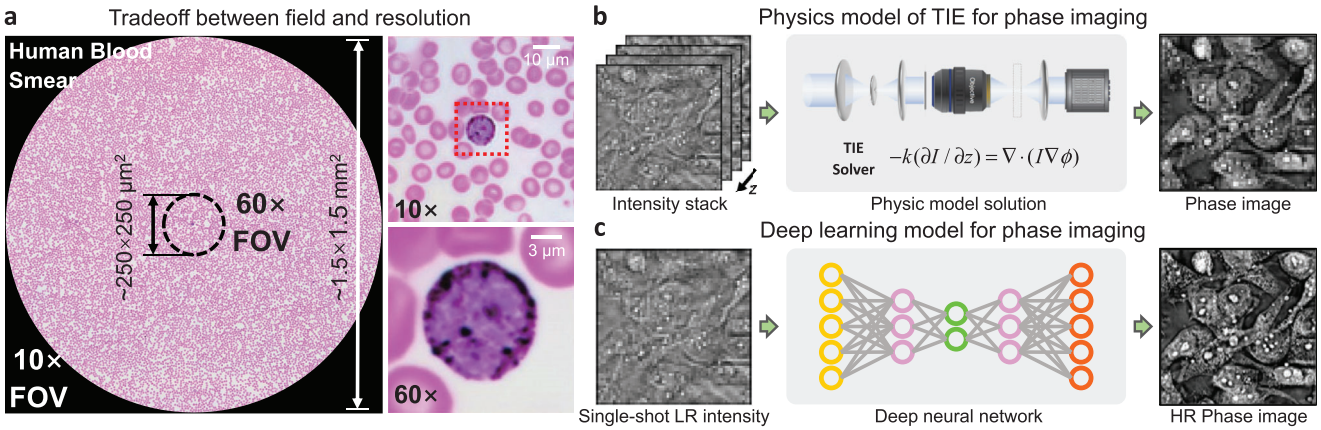


Figure 1. Illustration of the trade-off between field of view and resolution, and comparison of TIE and DL-SRQPI phase retrieval methods. a) Comparison of resolution and field of view of human blood smear microscopic images under 10× and 60× objectives. b) Standard TIE phase retrieval workflow using an axial defocus intensity image stack as input to solve the TIE equation and obtain phase images. c) DL-SRQPI phase retrieval workflow using a single-frame defocused intensity image as input of a well-trained neural network, and outputs a super-resolved phase image.

effectiveness and feasibility of DL-SRQPI has been illustrated by the comparison with other traditional or network-aided phase retrieval methods, and the robustness of DL-SRQPI is also proved by the phase reconstruction of intensity images at various defocus distances and illumination conditions. To demonstrate its strong capability, we use DL-SRQPI to rapidly convert hundreds of frames of 512 × 512 pixels simulated intensity images into the corresponding 2048 × 2048 pixels phase images with high accuracy. We further validate the capability of DL-SRQPI with experimentally acquired intensity images. For a 647 × 490 pixels intensity image obtained by an off-the-shelf bright-field optical microscope, DL-SRQPI precisely retrieves its phase result at a resolution of 2588 × 1960 pixels while maintaining the original FOV, revealing abundant subcellular details that are once embedded in the aliased pixels. With the large-SBP phase reconstruction capability of DL-SRQPI, we provide long-term high-throughput time-lapse videos of Henrietta Lacks (HeLa) cells undergoing division. These superior performances indicate that the proposed DL-SRQPI is a promising tool for achieving high-throughput dynamic quantitative phase imaging of biological cells.

2. Principle and Methods

2.1. High-Throughput QPI via Single-Shot Intensity Measurement

High-throughput microscopy permits access to high-throughput quantitative analysis for multiple events in a large population of cells.[36,37] However, the achievable SBP of conventional microscopy is fundamentally limited by the optical system, leading to an inevitable trade-off between FOV and imaging resolution. This limitation can be intuitively illustrated by **Figure 1**a. A commercial objective lens with low magnitude (UPlanSApo 10×, 0.4 NA, Olympus) allows the observation of tens of thousands of red blood cells across the FOV of ≈2.25 mm^2, but the spatial resolution is insufficient to distinguish detailed structures. In contrast, an alternative objective lens with higher magnitude (UPlanSApo 60×, 1.35 NA, Olympus) enables analysis for high-resolution cellular structures and details, such as the sharp boundaries of red blood cells and the white blood cells' internal particles. However, compared with the large FOV of the 10× objective lens, the achievable FOV of the 60× objective lens shrinks to ≈0.06 mm^2, where only hundreds of cells could be observed. Hence, it is difficult to take into account large FOV and high resolution simultaneously in conventional microscopic imaging systems.[38]

To decouple FOV and resolution from each other in a microscope, considerable research has been conducted, such as imaging stitching,[39] synthetic aperture microscopy,[10,40] lensless on-chip microscopy,[41–43] and FPM,[23,26,44–46] achieving high-throughput microscopic imaging with spatial-domain or frequency-domain methods. Image stitching is a simple and widely used approach that mitigates the trade-off between FOV and resolution by scanning the field with a high numerical aperture (NA) objective lens and then stitching the high-resolution segments in the spatial domain. However, the cost of an image stitching system is usually expensive due to the pricey high-NA objective lens and the high-precision electric scanners used. Besides, the necessary mechanical scanning, refocusing, and registration procedures also induce extra computation, resulting in a restriction of space-bandwidth-time production. In contrast, as mentioned above, FPM is a novel QPI technique that uses a low-NA objective lens to take advantage of its innate large FOV and stitches together images in the frequency domain. By varying the illumination angle, FPM shifts different high spatial frequency components of the object spectrum into the passband of the low-NA objective lens, and realizes high-throughput phase imaging using a low-cost system. Nevertheless, the basic strategy of the above-mentioned techniques is to trade numerous data measurements for high system throughput. This data reliance often requires sophisticated optical setups or elaborate illumination schemes, leading to time-consuming data acquisition and severe storage burden problems. On the other side, TIE is a well-established deterministic QPI approach that simply utilizes intensity measurements at multiple axially displaced planes to obtain the axial intensity derivative and reconstruct the quantitative phase (Figure 1b).[22] Thanks to its Köhler illumination compatibility within an off-the-shelf bright-field microscope, TIE eliminates the need for elaborate illumination schemes and optical

setups. Additionally, TIE recovers phase in a non-iterative manner with a requirement of only a few intensity measurements, which improves the data efficiency, reduces the storage burden and brings higher imaging speed. Nevertheless, despite its high efficiency, the throughput of phase reconstruction is still fundamentally constrained by the optical system, since TIE is always limited by BF illumination, and the maximum attainable imaging resolution is restrained to the incoherent diffraction limit when matched annular illumination is used.[24,25,47] Consequently, a computational QPI technique for wide-field high-resolution and high-throughput phase reconstruction from single-shot intensity measurement is yet to be developed.

The proposed DL-SRQPI has the ability to retrieve high-resolution phase images from low-resolution intensity image using a well-trained neural network (Figure 1c). Thanks to the unique end-to-end mapping mechanism and powerful high-dimensional feature extraction capability of deep learning, DL-SRQPI minimizes the data acquisition requirement to a single intensity measurement. With the help of GPU, this hybrid approach of deep learning and QPI addresses the above-mentioned data reliance limitation, improving the data efficiency and speeding up the phase imaging simultaneously. Meanwhile, DL-SRQPI provides a significant improvement in SBP and realizes high-throughput QPI by enhancing the image resolution without sacrificing the FOV, alleviating the resolution-FOV trade-off in a computational manner. Besides, DL-SRQPI allows for simple and straightforward implementation on a conventional bright-field microscope at a low cost, giving the possibility for its wide application.

2.2. Image Preprocessing and Dataset Construction

As an approach based on the end-to-end supervised-learning strategy, DL-SRQPI gains its capability from the training datasets consisting of well-paired low-resolution (LR) intensities and high-resolution (HR) phases.[28] We constructed simulated and experimental datasets separately to train DL-SRQPI progressively.

In the simulated dataset, we used numerical propagation methods, including angular spectral propagation and Abbe superposition, to accurately and efficiently generate low-resolution intensity images at different defocus distances and illumination patterns from ground truth high-resolution phases. The propagated complex field utilizing the angular spectrum method, which models the propagation of a wave field by using an analytic formula, can be calculated by Equation (1)

$$U(\mathbf{x},z)=\mathscr{F}^{-1}\left\{\hat{U}(u_x,u_y,0)\exp\left[j\frac{2\pi}{\lambda}z\sqrt{1-(\lambda u_x)^2-(\lambda u_y)^2}\right]\right\} \tag{1}$$

where $\mathbf{x}$ represents the 2D spatial coordinate (x, y) in the real space, the scalar coherent field $U(\mathbf{x},0)$ (assuming $z=0$) is decomposed into the coherent superposition of the angular spectrum (plane wave) components $\hat{U}(u_x,u_y,0)=\mathscr{F}\{U(\mathbf{x},0)\}$, and $\exp[j\frac{2\pi}{\lambda}z\sqrt{1-(\lambda u_x)^2-(\lambda u_y)^2}]$ is a phase delay factor, which is also known as the angular spectrum transfer function. Then, the intensity $I(\mathbf{x})$ at propagation distance z can be calculated by multiplying the complex field $U(\mathbf{x},z)$ and its conjugate $U^*(\mathbf{x},z)$.[48,49]

For the generation of intensities of partially coherent field, we utilized Abbe's superposition method, which describes the formation model of intensity images under different illumination conditions. The Abbe's method could be described by Equation (2)

$$I(\mathbf{x})=\int S(\mathbf{u})I_{\mathbf{u}}(\mathbf{x})\mathrm{d}\mathbf{u} \tag{2}$$

where $S(\mathbf{u})$ is the Fourier transform of the source intensity distribution, $I_{\mathbf{u}}(\mathbf{x})$ is the coherent partial image arising from the point of the incoherent source. Equation (2) implies that a partially coherent intensity image can be represented as an incoherent superposition of all intensities $I_{\mathbf{u}}(\mathbf{x})$ generated by all light source points at the condenser aperture plane.[49] By this means, we can generate intensity images under various illumination conditions.

The simulated datasets consist of FPM phase images as labels and their calculated LR intensity images as inputs. The label phase images are from our previously published paper.[50] The accurately matched intensity images were generated from the ground truth phase images by the above-mentioned numerical propagation methods. With the angular spectrum method, we digitally back-propagated a ground truth phase image to 13 intensity images at various defocus distances, within a range of z = (+1 μm, +13 μm) with $\triangle z$ = 1 μm increments. With the Abbe's superposition method, we also generated defocused intensity images (z = +4 μm) under different coherent parameters within a range of S = [0, 0.4] with S = 0.l increments. These generated intensity images were then corrupted by Gaussian noise with a standard deviation of 0.01 to simulate the noise effect. Subsequently, we downsampled the intensity images to simulate the increase of pixel size and introduce the artificial pixel aliasing.[27] After the operation of 4× pixel binning, the pixel resolution of the intensity images was reduced from 2048 × 2048 pixels to 512 × 512 pixels, while the FOV remained unchanged. So far, we had obtained 17 LR intensity images on 13 defocus distances and four coherent parameters. With these images, we constructed 17 simulated datasets, each comprising the original phase image and one intensity image on a specific defocus distance and a coherence parameter generated from the phase. To reduce the memory demand of the computer and speed up the process of network training, we divided the full-field LR intensity images and the corresponding HR phase images into paired image patches, and the patches without valid information were excluded via edge detection algorithm. For each simulated dataset, the full-field intensity image and phase image were segmented into 64 × 64 pixels and 256 × 256 pixels image patches, respectively. With further augmentation by mirroring and rotation, each dataset eventually contains 1480 LR intensity and HR phase image pairs. Out of these images, 1300 pairs were randomly selected to be used as the training dataset, while 80 pairs were used as the validation dataset for validating the network performance and selecting the optimal model, and the remaining 100 pairs formed the testing dataset to blindly quantify the average performance of the final network. To ensure fairness across networks, though an individual dataset was randomly divided, the division of each dataset was identical.

The experimental dataset consists of real experimental intensity images, which are from our previous work.[51] The differences in exposure time during intensity image acquisition and the inhomogeneous light absorption of sample areas resulted in non-uniform field brightness, which had a negative impact on the accuracy of phase retrieval. This influence of spatial variability of the light intensity can be corrected with the algorithm $C_{ij} = R_{ij}\frac{\overline{S}}{\overline{R}}$, where C is the corrected image, R is the original image, $\overline{R}$ is the average gray level of the image R, $\overline{S}$ is the average gray level of all 270 intensity images, and the subscripts i and j indicate that the correction is performed on the ith and jth pixel of the image. The phase images can be recovered by TIE, which is given by Equation (3)

$$-k\frac{\partial I(\mathbf{x})}{\partial z} = \nabla \cdot [I(\mathbf{x})\nabla\phi(\mathbf{x})] \tag{3}$$

where ∇ denotes the 2D gradient operator with respect to x and y, and $k = 2\pi/\lambda$ is the wave number. Then the phase $\phi(\mathbf{x})$ can be extracted by solving the equation. The left hand of TIE is the spatial derivative of intensity at the in-focus plane along the z-axis.[52] The right hand of TIE is a second-order elliptic partial differential equation, and we treat it as a Poisson equation, ideally, which can be easily solved with fast-Fourier transform (FFT).[53]

We corrected the field brightness of the experimentally acquired 135 intensity image stacks and then reconstructed their corresponding phase images via TIE. To introduce noticeable pixel aliasing effect, we downsampled the resolution of the intensity images from 2588 × 1960 pixels to 647 × 490 pixels by performing 4× pixel binning. This can be regarded as an 4× enlargement of the pixel size, resulting in a significant loss of detailed information. The TIE-retrieved phase images and the low-resolution intensity images were used as input and labels for the experimental dataset. Same as the simulated datasets, we constructed the experimental dataset by extracting paired 128 × 128 pixels and 512 × 512 pixels image patches from a pair of experimentally captured full-field LR intensity image and its HR TIE phase image, and augmented the dataset to 433 pairs by rotating and mirroring. 400 pairs were randomly selected as the training dataset, eight pairs were selected as the validation dataset, and the remaining 25 pairs as the testing dataset.

2.3. Network Architecture and Training

DL-SRQPI adopts U-Net (**Figure 2**a) as the neural network to achieve high-speed high-throughput QPI. U-net is a remarkable CNN-based network with excellent performance in biomedical image processing.[35,54,55] In DL-SRQPI, U-Net transforms previously fuzzy inferior intensity images with pixel aliasing to clear, superior, alias-free phase images. As shown in Figure 2a, the network consists of an interpolation operation at the input, an encoder branch for feature extraction, and a decoder branch for feature reconstruction, with skip connections combining the features from two branches. The interpolation operation is performed to align the resolution of the intensity image and the phase image, so that the network can identify the complex mapping relationship of super-resolution phase retrieval. The encoder branch consists of four identical downsampling modules, each including a 2 × 2 max pooling layer and two convolutional layers with a 3 × 3 kernel and stride of 2. The skip connection path connects the extracted features of each stage of the encoder branch to the corresponding feature layer of the decoder branch. The decoder branch consists of four identical upsampling modules, each consisting of an upsampling convolutional layer concatenating with the corresponding feature map in the encoder branch by skip connection, and two convolutional layers with a 3 × 3 kernel. All convolutional layers are followed by a batch normalization module (BN) and a rectified linear unit (ReLU) to achieve faster training and enhance the nonlinear ability. During training (Figure 2b), the mean-square-error (MSE) between the output phase image and the label phase image was calculated as the loss function, and was back-propagated to the network for optimization. For an image with a size of $M \times N$ pixels, this loss function over a mini-batch at size of K is calculated by Equation (4)

$$\text{Loss}(\Theta) = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{M\times N}\sum_{m=1}^{M}\sum_{n=1}^{N}\| Y^{\Theta}_{m,n,k} - Y^{\text{GT}}_{m,n,k}\|^2 \tag{4}$$

where k is the kth image patch among the mini-batch, $Y^{\Theta}_{m,n,k}$ denotes the mth and nth pixel of network output phase image, and $Y^{\text{GT}}_{m,n,k}$ denotes the mth and nth pixel of the training labels (i.e., ground truth). The network's parameter space (e.g., kernels, biases, and weights) is defined by Θ and its output is given by $Y^{\Theta} = F(X_{\text{input}};\Theta)$, where F defines the deep neural network's operation on the network input intensity X_{input}. The adaptive moment estimation (ADAM) optimization algorithm with a learning rate of 0.01 is utilized to minimize the MSE and tune the network parameters. After sufficient training, the network has established the mapping relationship between the LR intensity image and the HR phase image. It is worth mentioning that due to the translation invariance of the convolutional neural network, the network can output full-field phase images from full-field intensity images, despite only patches being used during training. The metric used to measure the accuracy of phase retrieval is given by the structural similarity index (SSIM), which comprehensively evaluates luminance ($l(x,y)$), contrast ($c(x,y)$), and image structure ($s(x,y)$), to quantify the similarity of the ideal phase and the retrieved phase. It can be calculated by Equation (5)

$$\text{SSIM}(x,y) = [l(x,y)]^{\alpha}\cdot[c(x,y)]^{\beta}\cdot[s(x,y)]^{\gamma} \tag{5}$$

where

$$\begin{cases} l(x,y) = \dfrac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \\ c(x,y) = \dfrac{2\sigma_x\sigma_y + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \\ s(x,y) = \dfrac{\sigma_{x,y} + c_3}{\sigma_x\sigma_y + c_3} \end{cases} \tag{6}$$

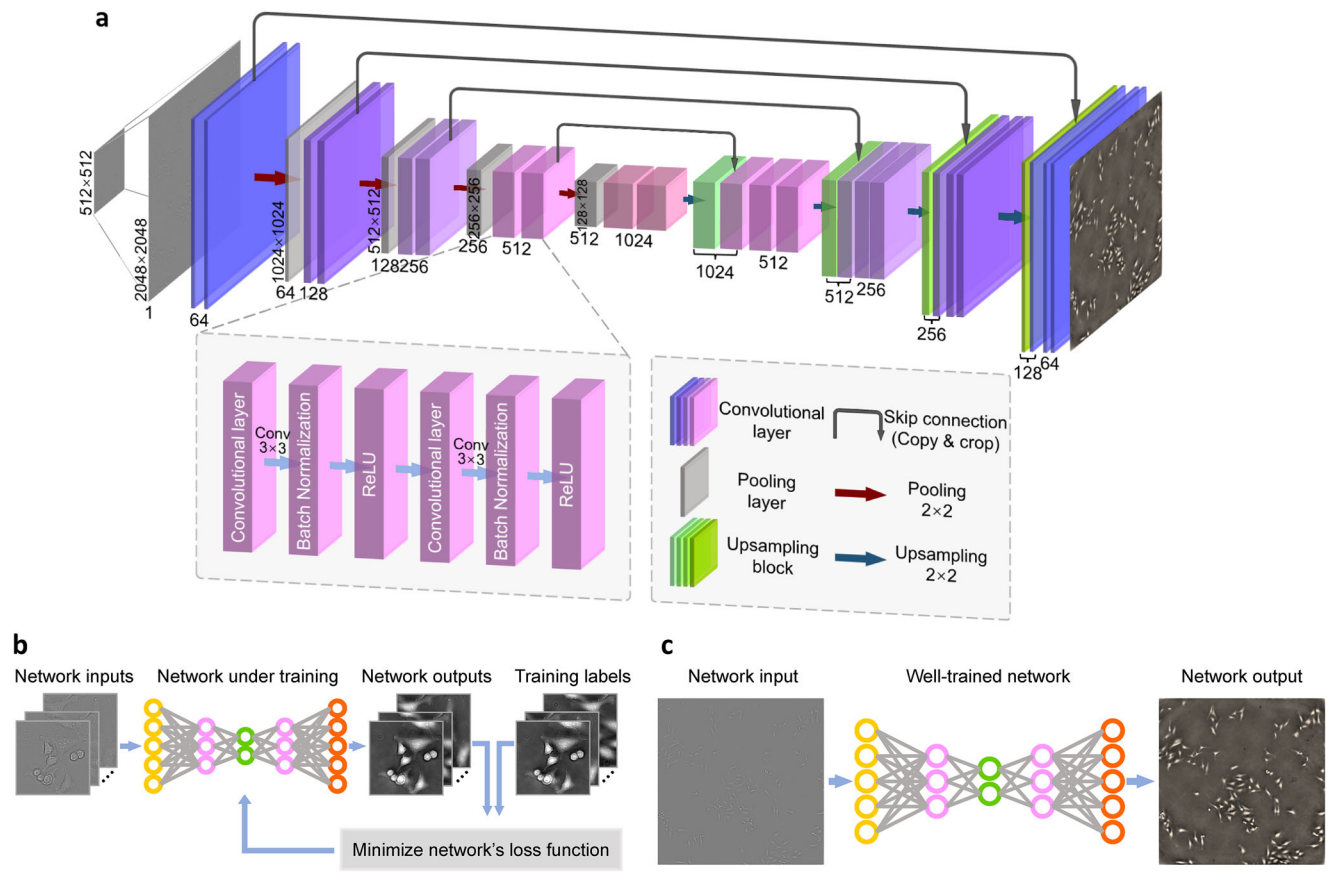


Figure 2. Network architecture and schematics of network training and testing. a) The U-Net structure is depicted, where each block represents a multi-channel feature map. The number of channels is indicated at the bottom of each block, while the size is denoted in the lower left corner. b) The network training workflow involves utilizing low-resolution intensity image patches as inputs and corresponding high-resolution phase image patches as training labels. The network optimizes its parameters by minimizing the loss function (MSE) between the network outputs and the training labels. c) During network testing, a full-FOV low-resolution intensity image is presented as the input, and the well-trained network generates a full-FOV high-resolution phase image as the output.

where μ_x, μ_y, σ_x, σ_y, and σ_{xy} are the local means, standard deviations, and mutual covariances of the images x and y. When $\alpha = \beta = \gamma = 1$ and $c_3 = c_2/2$, the SSIM index simplifies to

$$\mathrm{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{x,y} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{7}$$

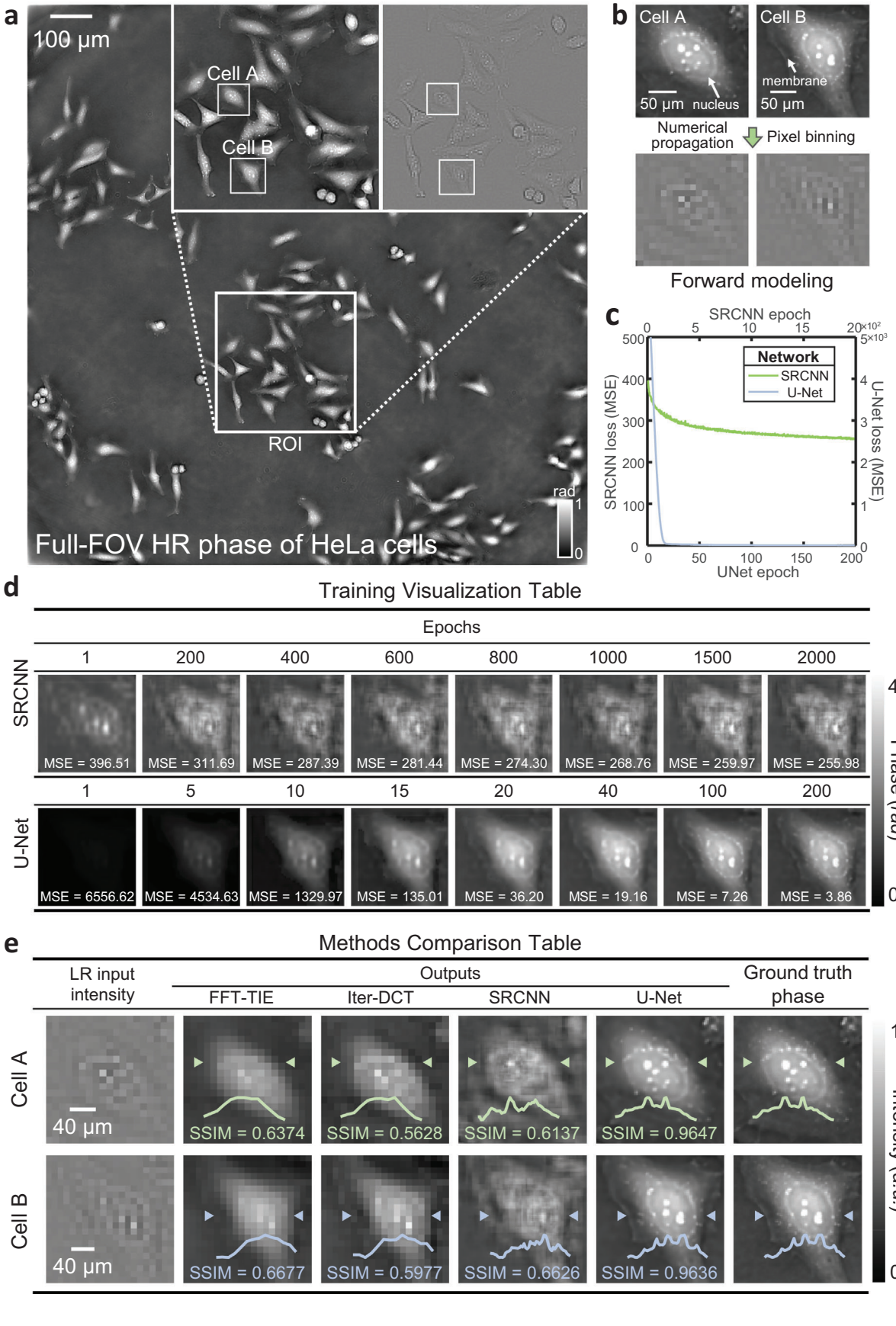
The SSIM index varies between 0 and 1, where 1 can be achieved if predicted and ground truth images are identical to each other. The fixed network gains the capability to blindly output full-field high-resolution phase images at a high reconstruction speed (Figure 2c), providing a great enhancement of space-bandwidth-time product for the QPI system.

3. Results

3.1. Benchmarking of DL-SRQPI

To illustrate the applicability of our proposed method, we compared DL-SRQPI with two traditional TIE-based methods (FFT-TIE,[22] iterative DCT[56]) and a classic super-resolution neural network SRCNN.[57] The full-field FPM phase image was used as the ground truth (**Figure 3**a), with high quality and abundant detail information. As mentioned in Section 2.2, the LR defocused intensity image used for phase reconstruction was generated following the forward model consisting of angular propagation and pixel binning (Figure 3b), and was corrupted by Gaussian noise to simulate the noise effect with a standard deviation of 0.01.

To make a fair comparison between SRCNN and U-Net, both networks are fully trained with the same dataset and loss function (MSELoss). In Figure 3c, we show how the MSE values of two networks decrease. SRCNN has a low initial loss value of ≈400, but it decreases slowly and converges to ≈250 after 2000 epochs of training. In contrast, though the initial loss value of U-Net is as high as ≈6500, the loss function sharply decreases to less than 200 within 15 epochs, and eventually converges to ≈4, exhibiting much stronger mapping ability. More intuitively, the outputs of two networks during the training process are shown in Figure 3d. During training of SRCNN, the output images show little improvement. The poor response at low-spatial frequencies leads to severe noise and artifacts in the output, resulting in difficulty distinguishing the cell boundary and inner structure. On the other hand, as U-Net training proceeds and the loss function steadily



decreases, the precise phase gradually emerges from the dark background. The phase reconstruction of DL-SRQPI achieves an SSIM index of 0.96 with respect to the ground truth, indicating its high precision. In the output of DL-SRQPI, the optically thick nucleus, the cell membrane, and some cytoplasmic organelles are shown with high contrast and clarity. Compared with other methods, the high-quality reconstructions of DL-SRQPI exhibit abundant subcellular features in high resolution, unraveling the aliased pixels in the input images.

In Figure 3e, we compare the phase retrieval of these methods for two HeLa cells. After 4× pixel binning, the intensity images become fuzzy and inferior. Due to its low contrast, the intensity images cannot provide much cellular information, and are further exacerbated by the severe pixel aliasing problem. The high frequency detail is almost completely lost in the under-resolved intensity image, preventing the observation of detailed subcellular structures and making phase retrieval very challenging. Since the FFT-TIE and Iter-DCT methods do not possess super-resolution capability, their retrieved phases are blurry and the detailed information is still buried in oversized pixels. Only optically thick cellular structures in the phase results can be visualized, such as the nucleus, while the high-frequency details are completely lost. In Figure 3e, we demonstrate the line profiles across the nucleus and the cell membrane of each output. The line profiles of FFT-TIE and Iter-DCT only show the general trend of phase changes and are unable to present the detailed phase variation of the internal cellular structures. This result reveals the gap in resolution between TIE-based and deep learning methods. Notably, two intensity images at different defocus distances are required by FFT-TIE and Iter-DCT due to their data redundancy requirement, while the deep learning methods only require a single frame of intensity image as input for its end-to-end mechanism.

3.2. Generalization Capability Analysis for Axial Defocusing and Illumination Condition

Axial intensity derivative estimation is a key issue in TIE-based QPI methods. TIE requires multiple defocused intensity measurements to achieve phase retrieval. The defocus distance has to be large enough to ensure an adequate SNR,[49] but too large a defocus distance tends to introduce phase blurring effect. This noise-resolution trade-off requires a strict and precise choice of defocus distance for TIE-based phase retrieval. Therefore, DL-SRQPI is expected to possess the capability to recover high-quality phase robustly from a single frame of intensity image at a random defocus distance. To initially analyze the generalization capability of DL-SRQPI for axial defocusing, we used the 13 datasets that contain variously defocused intensity images (z = [+1 μm, +13 μm], $\triangle z$ = 1 μm) to train 13 U-Net networks independently, which are mentioned in Section 2.2. Following the forward model in **Figure 4**a, after adequate training, we blindly fed 13 sets of testing intensity images at different defocus distances into each network, each testing dataset containing 100 identically defocused LR intensity images. Note that the image pairs used in the training and test datasets are completely different. All 13 networks responded rapidly to each set of testing images, and the HR phase outputs are shown in Figure 4b. The line profiles show the phase variation along the same region. Notably, with the decrease of testing defocus distance, the phase results become dimmer; while the testing defocus distance becomes higher, the phase results become brighter and even get overexposed. This can be regarded as proof that the neural network has mastered the mapping relationship between intensity and phase: according to Equation (3), too large an estimate of the defocus distance z results in low phase values of the reconstruction result. To quantify the accuracy and the quality of the phase reconstructions, the average SSIM index for the outputs of each testing dataset was evaluated with respect to the ground truth, and the SSIM indexes are shown in Figure 4c. When the defocus distance of the testing images matches the training images, the phase results are of the highest quality with the average SSIM index reaching nearly 0.96. As the testing defocus distance differs from the training defocus distance, the accuracy and the SSIM indexes of the phase results drop slowly. Figure 4 illustrates that though each network was trained with intensity images at one particular defocus distance, DL-SRQPI still has the ability to reconstruct HR phase images from intensity images at a range of defocus distances with a negligible drop in reconstruction quality.

To maximize the generalization capability of DL-SRQPI for axial defocusing, we used 13 sets of intensity images at various defocus distances to construct a training dataset, each containing 100 identically defocused intensity images. After proper training, DL-SRQPI adapts the variation in defocus distance. We blindly tested the network with the above testing dataset and calculated the average SSIM indexes of the phase outputs from each set of intensities at a certain defocus distance. All these average indexes are higher than 0.9, and the overall average SSIM index reaches nearly 0.95, representing the network is adaptive to axial defocusing. So far, we have demonstrated that DL-SRQPI is robust to the variation of defocus distance, and the proposed defocus adaptive DL-SRQPI could further enhance the generalization ability on axial defocusing.

The spatial coherence of the illumination is also an essential factor for phase retrieval. It can be quantified by a normalized factor $S = NA_{\mathrm{cond}}/NA_{\mathrm{obj}}$ (so-called coherent parameter), where NA_{cond} is the numerical aperture of the condenser lens and NA_{obj} is the numerical aperture of the objective lens. Similar to the noise-resolution trade-off in the choice of defocus distance, the spatial coherence also brings a compromise between the contrast and the resolution of phase recovery. Reducing NA_{cond} can effectively improve the phase contrast, but at the same time reduce the imaging resolution of the system. The generalization capability

Figure 3. Methods comparison and training visualization. a) The full-FOV high-resolution ground truth phase image of HeLa cells in vitro, along with the phase of a ROI (region of interest) and its intensity image obtained through numerical propagation. b) The defocused low-resolution intensity images of cell A and cell B were generated by numerical propagation from their high-resolution phase images and subsequently downsampling them through a 4× pixel binning. c) The plot of the Mean Squared Error loss function for SRCNN and U-Net. d) Phase reconstruction images and corresponding loss function values from the outputs of SRCNN and U-Net at different epochs during network training. e) Phase reconstruction results of FFT-TIE, Iter-DCT, SRCNN, and U-Net using the low-resolution intensity images of cell A and cell B as inputs, along with the line profiles of the cells.

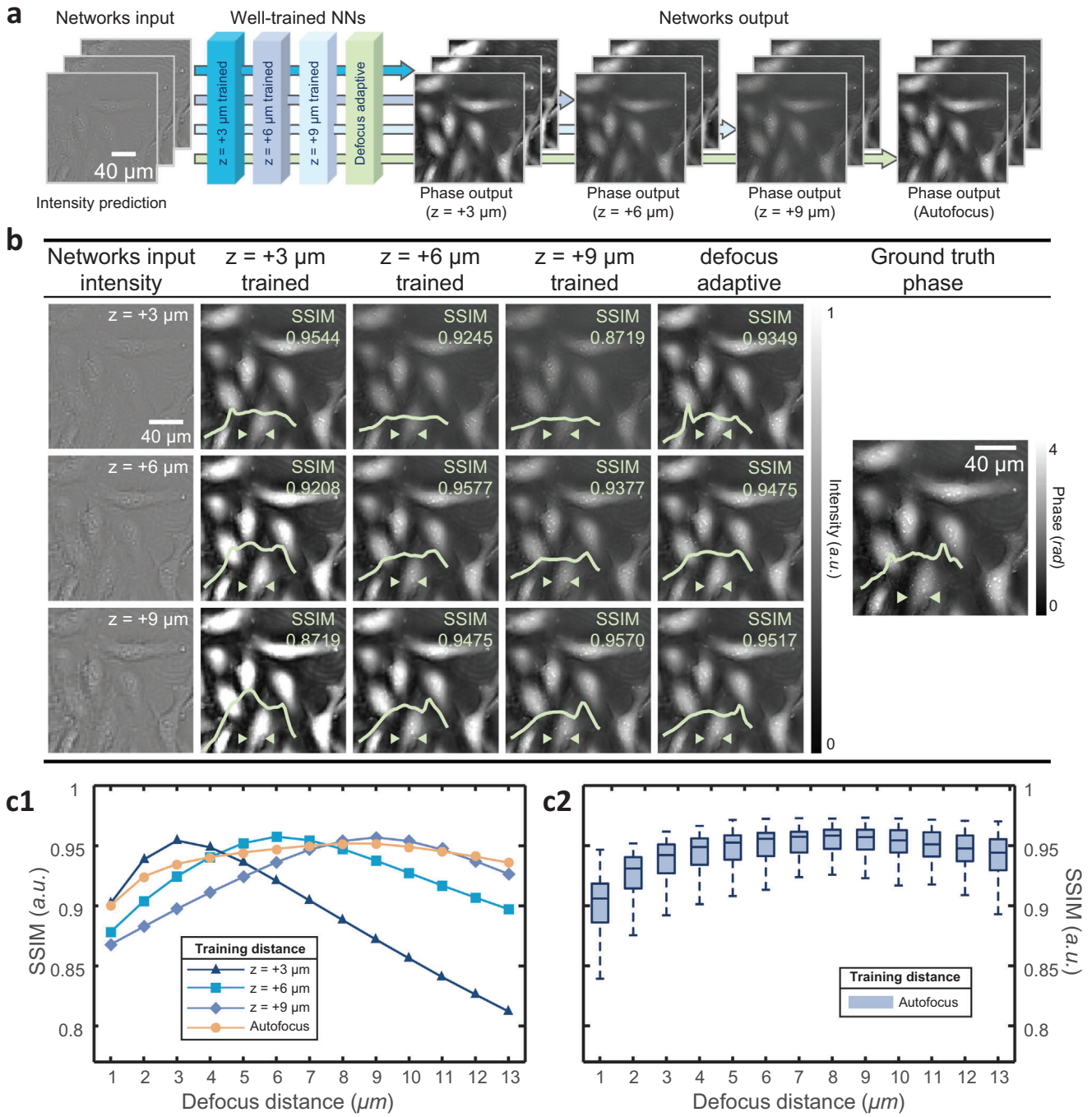


Figure 4. Analysis of axial defocusing generalization capability. a) The forward model of DL-SRQPI for intensity images with different defocus distances. Each neural network is trained with intensity images at a specific defocus distance and tested blindly with intensity images at different defocus distances. b) The phase reconstructions of testing defocused intensity images ($z = +3/+6/+9$ μm), output by three well-trained networks trained on intensity images at defocus distances of $z = +3/+6/+9$ μm), along with their SSIM indexes compared to the ground truth phase image, and their line profiles showing the subcellular features. c1) The average SSIM index curves depicting the similarity between the phase images obtained from intensity images at different defocus distances ($z = +1$ to $+13$ μm) and the ground truth phase, using the defocus adaptive network and three networks trained on three sets of defocused intensity images ($z = +3/+6/+9$ μm). c2) The boxplot of the SSIM index of the phase images obtained from the defocus adaptive network using 13 groups of testing datasets, and the ground truth phase images.

of DL-SRQPI on coherent parameters has been verified. We used the five datasets mentioned in Section 2.2 to train five U-Net networks independently, each containing intensity images at a certain parameter (S = [0, 0.4], $\triangle$S = 0.1). After training, each well-trained network was blindly tested with five sets of testing images at different coherent parameters, each set containing 100 LR intensity images (**Figure 5**a). Note that the image pairs used in the training and test datasets are completely different. The results are shown in Figure 5b and the average SSIM indexes of each set of outputs were evaluated with respect to the ground truth (Figure 5c). The phase results and the line profiles in Figure 5b show that when the coherent parameter of the testing intensity images matches the coherent parameter of the training intensity images, the quality of the phase results reaches the highest, with an average SSIM index of 0.95, which shows the high similarity between ground truth and network output images. As the coherent parameter of the testing image gets lower, the phase imaging contrast reduces, bringing a blurry effect on the internal structures and edges of cells, causing a significant loss of high-frequency detail information. With the increase of testing coherent parameter, the phase imaging contrast becomes excessively high, resulting in a sharpening-like effect in the phase reconstruction, which added difficulty to the cell morphology analysis. The box plot of the SSIM indices for each set of outputs from the $S = 0.4$ trained network shows a small variance in the accuracy of the outputs, indicating the stability of the DL-SRQPI. By far, the robustness of the phase retrieval framework of DL-SRQPI is critically estimated through the above testing experiments.

3.3. Quantitative and Generalizable Characterization of DL-SRQPI

We conducted an experiment to demonstrate the quantitative property of DL-SRQPI using a microlens array as the test subject. For the establishment of the ground truth high-resolution phase image, the TIE algorithm was utilized to reconstruct the accurate phase image of the microlens array. Subsequently, we performed 4× pixel binning on the original defocused intensity image to generate a low-resolution intensity image. This process enlarged the pixel size to 8.8 μm and decreased the pixel resolution of the defocused intensity image from 1280 × 960 pixels to 320 × 240 pixels. With the LR defocused intensity image and ground truth HR phase image, we created a dataset comprising high-resolution phase and low-resolution intensity image pairs of the microlens array using the same method described in Section 2.2. We trained a U-Net network and blindly fed it with a new defocused low-resolution intensity image of the microlens (**Figure 6** a2). The network rapidly generated a high-resolution phase reconstruction (Figure 6 a3), and the heights of the microlens were deduced from the output phase image (Figure 6 a4). In Figure 6 a5, we compare the line profiles of the DL-SRQPI output phase, the ground truth high-resolution phase, and the low-resolution phase for a single lens. It is evident that the three profiles almost completely overlap. As seen in the zoomed area, the discrete pixels of LR phase exhibit a noticeable stair-step pattern due to the pixel binning, indicating a large pixel size and a low resolution. Conversely, the DL-SRQPI achieves a high resolution comparable to the ground truth HR phase. Figure 6 a6 presents the difference between the ground truth phase and the output phase of DL-SRQPI. The difference between the output phase and the ground truth phase is only 1 to 2 gray levels, which is equivalent to 0 to 0.2 μm in height, showing the high accuracy of DL-SRQPI. These results confirm the precise quantitative characteristics of DL-SRQPI.

To further demonstrate the generalizability of DL-SRQPI, we utilized experimentally acquired differently defocused intensity images of HeLa cells. Following the same method described in Section 2.2, we constructed three experimental datasets for HeLa cells, each consisting of HR phase images and LR intensity images captured at different defocus distances ($z = +3/+6/+9$ μm). Subsequently, we trained three separate U-Net networks using these datasets. Figure 6 b1–b3 displays three LR intensity images of a single HeLa cell captured at the corresponding defocus distances of $z = +3/+6/+9$ μm. These LR intensity images were input into their respective networks, which rapidly generated four-fold pixel super-resolved phase images (Figure 6 b4–b6). Importantly, the output phase images obtained from the LR intensity images at different defocus distances all exhibited high SSIM values when compared to the ground truth (Figure 6 b7). This experimental evidence clearly demonstrates the strong generalization capability of DL-SRQPI.

3.4. Evaluation on the Simulation Dataset Using DL-SRQPI

We verified the feasibility of DL-SRQPI to achieve rapid high-throughput single-shot QPI using the simulation dataset of in vitro HeLa cells. **Figure 7** shows the full FOV phase prediction for the simulated dataset of HeLa cells in vitro. The simulated input LR intensity image has a wide FOV of ≈1.77 mm^2, matching the FOV size of the objective lens with 10× magnification. The input intensity image has a resolution of 512 × 512-pixel with an effective pixel size of 2.6 μm. The high-throughput phase reconstruction is displayed in Figure 7a, which shows that our DL-SRQPI is able to achieve a four fold enhancement in the pixel resolution from 512 × 512 pixels to 2048 × 2048 pixels while maintaining the large FOV size of ≈1.77 mm^2, as the effective pixel size improves to 0.65 μm. The comparison between LR intensity and the predicted HR phase of two ROIs is shown in Figure 7a. Compared to the low-contrast aliased intensities, the recovered phases display improved overall contrast of cell organelles and highlight high-spatial-frequency subcellular details.

Additionally, thanks to the non-invasive and nontoxic properties of DL-SRQPI, it can also serve as a practical tool for visualizing the morphological dynamics of living HeLa cells. We used DL-SRQPI to reconstruct phase imaging videos with large SBP (Videos S1 and S2, Supporting Information). As shown in Figure 7b, cells A and B are enlarged to present different typical mitosis phases and the morphological evolution of cells during the mitotic cycle. The subcellular features of both cells, such as cytoplasmic vesicles and pseudopodium, and their sub-pixel organelle motions, such as plasmid migration, are demonstrated in the video. Since each HR phase image was reconstructed within only 0.3 s, which can be further accelerated with higher performance hardware, all the retracting, extending, reorganizing, migrating, and maturing processes of cells could be recovered accurately avoiding motion blur, which lays the foundation for the

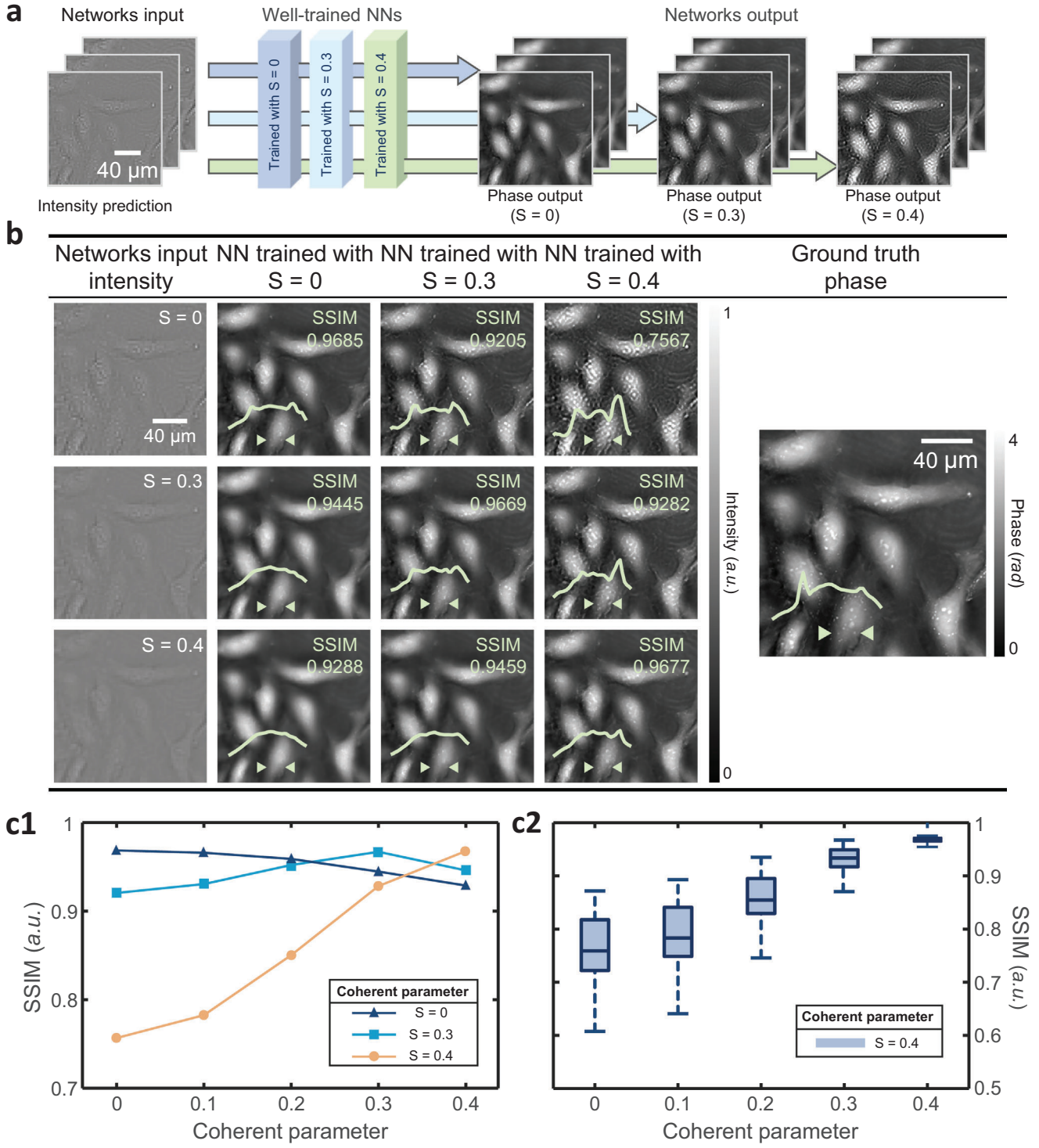


Figure 5. Analysis of generalization capability for illumination condition. a) The forward model of DL-SRQPI for intensity images with various coherent parameters. Each network is trained with intensity images at a specific coherent parameter and tested blindly with intensity images at different coherent parameters. b) The phase reconstructions of testing intensity images ($S = 0/0.3/0.4$), output by three well-trained networks trained on intensity images at illumination condition of $S = 0/0.3/0.4$, along with their SSIM indexes compared to the ground truth phase image, and their line profiles showing the subcellular features. c1) The average SSIM index curves depicting the similarity between the phase images obtained from intensity images at different coherent parameter ($S = 0/0.3/0.4$) and the ground truth phase, using three networks trained on three sets of intensity images at the same coherent parameter. c2) The boxplot of the SSIM index of the phase images obtained from the well-trained network ($S = 0.4$) using five groups of testing datasets, and the ground truth phase images.

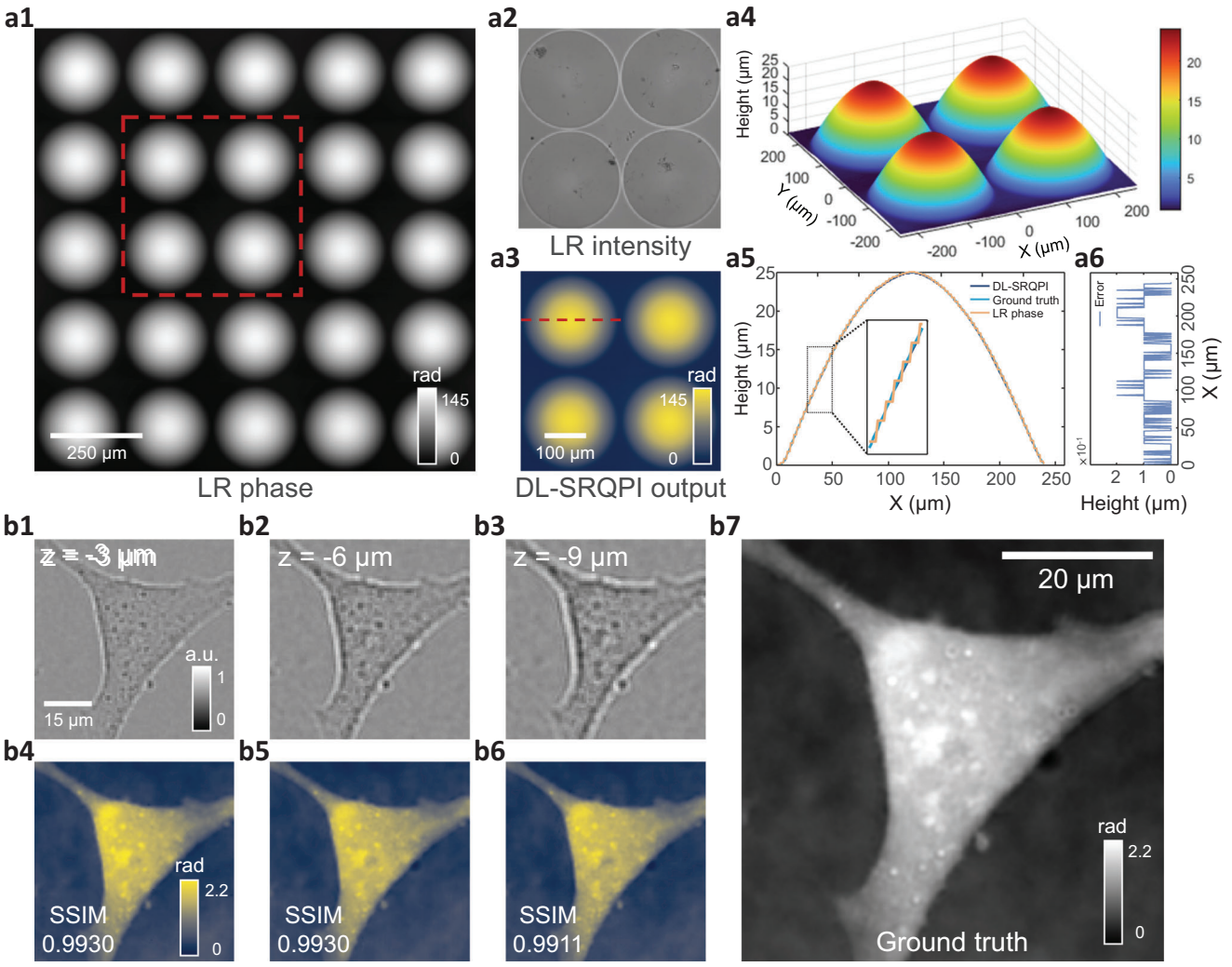


Figure 6. Validation of quantitative and generalization capabilities of experimental data based on a microlens array and a defocused HeLa cell. a1) The low-resolution phase image of the microlens array retrieved from low-resolution intensity images via TIE. a2) The low-resolution defocused intensity image of the microlens array. a3) The high-resolution phase reconstruction result output by DL-SRQPI. a4) The 3D topography of the microlens array. a5) Comparison of DL-SRQPI output, ground truth HR phase and LR phases. a6) The error of DL-SRQPI output, that is, the difference between ground truth phase and DL-SRQPI output phase. b1–b3) The input defocused LR intensity images of a HeLa cell ($z = +3/+6/+9$ μm]). b4–b6) DL-SRQPI reconstructed HR phase images from intensity images of (b1–b3). b7) The ground truth HR phase image of the HeLa cell.

practical application of DL-SRQPI in the fields of cytomorphology, cytokinetics, and cytogenetics.

3.5. Experimental Phase Imaging of HeLa Cells In Vitro Using DL-SRQPI

After validating the feasibility and the generalization capability of DL-SRQPI on simulated data, we further demonstrate the full-field high-resolution phase recovery from experimentally acquired intensity images of HeLa cells in vitro using DL-SRQPI. The HeLa cell image used to test the network was acquired at different culturing time point from the HeLa cell image used to construct the training dataset. As mentioned in Section 2.2, the ground truth phase images were retrieved based on TIE using the experimentally acquired high-resolution intensity images, and the low-resolution intensity images were used as inputs of DL-SRQPI and TIE. As can be seen in **Figure 8**a,b, the low-resolution intensity images suffer from severe pixel aliasing problem, where the high-frequency components almost completely disappear in the extremely low contrast and oversized pixels. With the low-resolution intensity image stack as input, TIE could only provide phase image with the same low resolution (Figure 8c), creating a great obstacle to the observation of the internal structure of cells. In contrast, DL-SRQPI has a minimum requirement and possesses pixel super-resolution capability. With a single frame of low-resolution intensity as the only input, DL-SRQPI rapidly generates the high-resolution phase image with pixel resolution of 2588×1960 pixels (Figure 8d) in 0.2 s, achieving fourfold pixel super-resolution. As shown in three enlarged ROIs in Figure 8e, DL-SRQPI enables the precise observation of plentiful subcellular features of HeLa cells, such as nucleus with large phase,

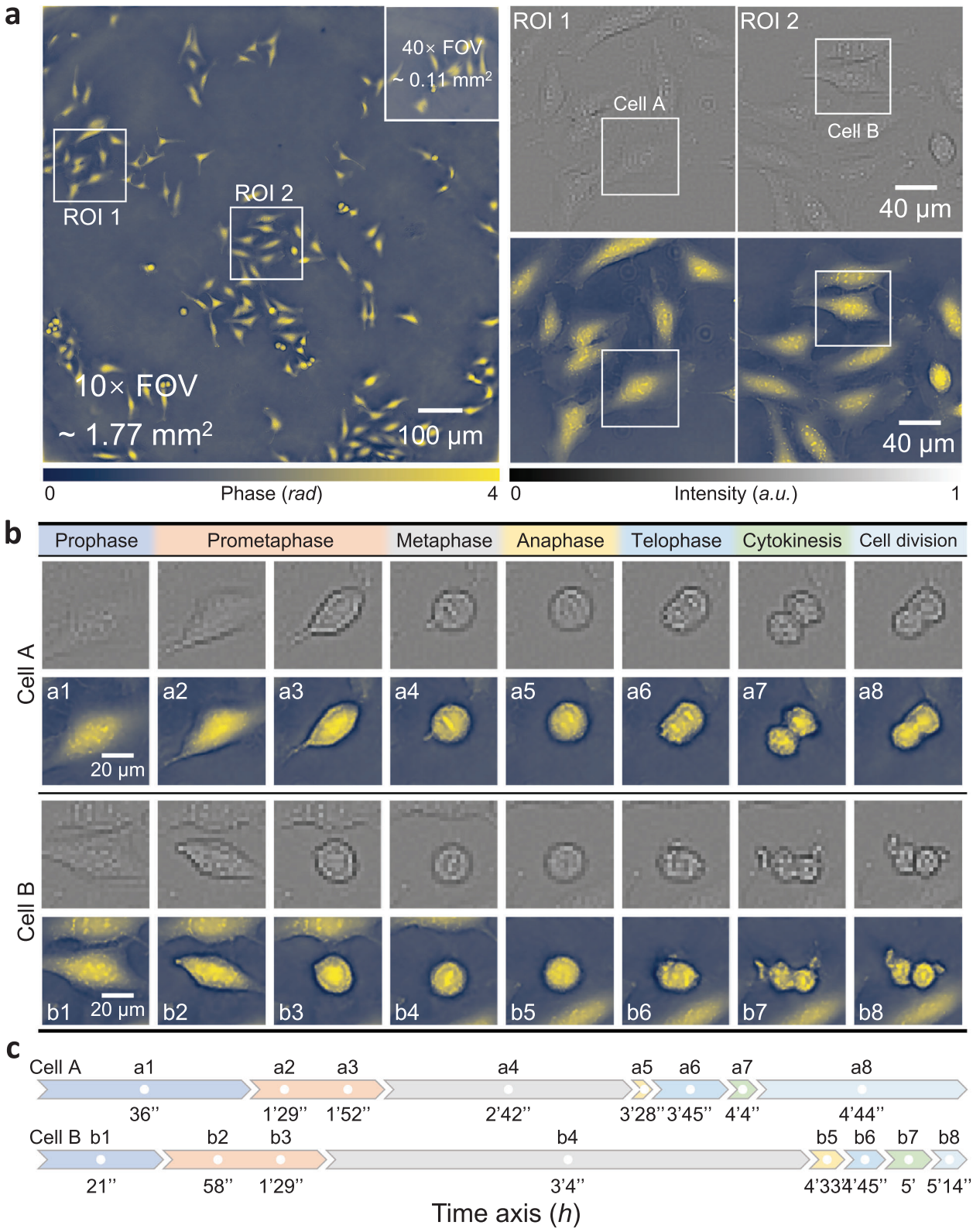


Figure 7. Time-lapse full-FOV high-resolution phase reconstruction of unstained HeLa cells undergoing division using DL-SRQPI. a) The full-FOV high-resolution phase reconstruction of the low-resolution intensity image by DL-SRQPI, and the comparison of low-resolution intensity images and high-resolution phase images of two ROIs. b) Sample frames of the DL-SRQPI reconstructed video (Video S2, Supporting Information) for cells A and B across 5 h, showing their different stages of cell division. c) The timeline of cell division of cells A and B within 5 h judging from DL-SRQPI reconstructed video. Different colors corresponding to (b) represent different cell division stages.

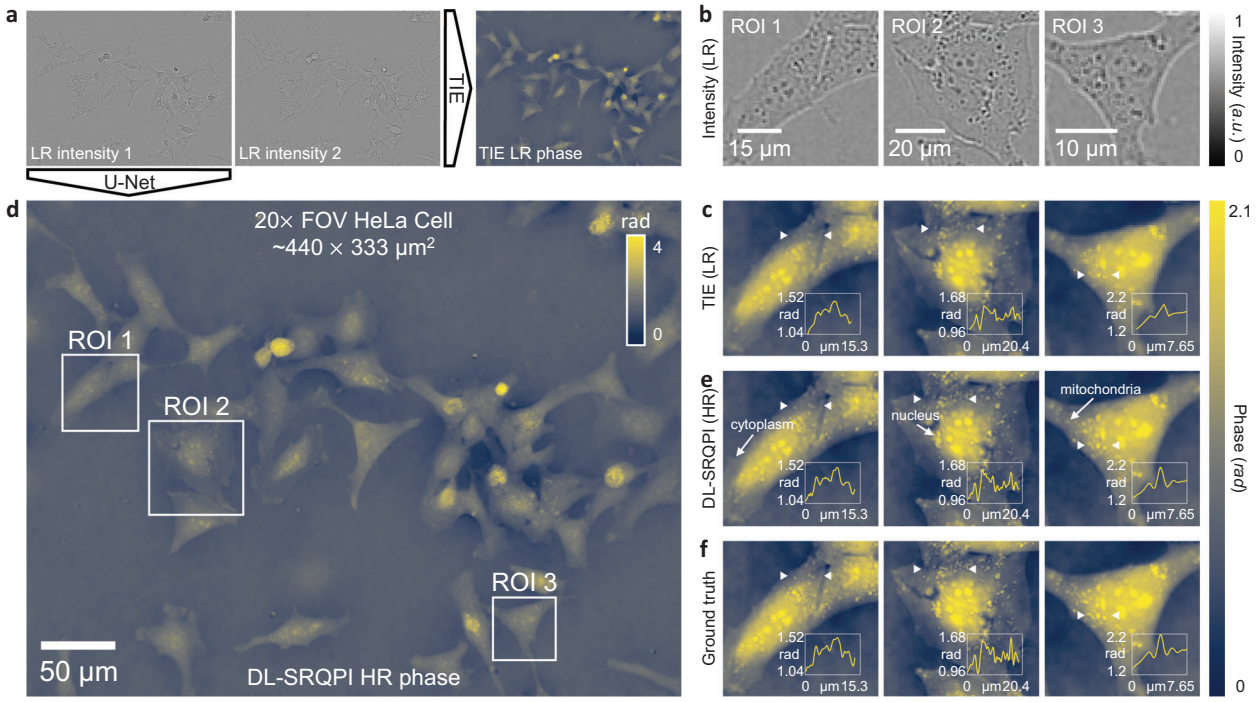


Figure 8. The experimental result of DL-SRQPI to reconstruct the full-FOV high-resolution phase image of unstained HeLa cells. a) The full-FOV low-resolution defocused intensity images, and the low-resolution phase image reconstructed by TIE. b) The low-resolution defocused intensity images of three ROIs. c) The low-resolution phase images of three ROIs reconstructed by TIE. d) The full-FOV high-resolution phase image reconstructed by DL-SRQPI from a single frame of low-resolution intensity image. e) The high-resolution phase images of the three ROIs reconstruted by DL-SRQPI. f) The ground truth phase images of the three ROIs.

mitochondria in transport, and cytoplasm in high contrast. The accuracy of the DL-SRQPI can be quantified by the SSIM index. The SSIM index of the full-field output phase reaches 0.9948, and the SSIM indexes of ROI 1, ROI 2, and ROI 3 are 0.9933, 0.9898, and 0.9920, which show extremely high similarity between the network outputs and the ground truth phases (Figure 8f). Compared with the TIE method that requires two intensity images at different defocus distances, the network reduces the demand for intensity images to only one, lightening the data burden as well as avoiding the phase error caused by the inaccurate estimation of axial intensity derivative. DL-SRQPI improves the resolution without sacrificing FOV, enhancing the throughput of the system effectively. Video S3, Supporting Information, shows the experimental time-lapse full-field phase reconstruction results of DL-SRQPI. These experimental results validate the efficacy and promptness of the DL-SRQPI utilization within a bright-field optical microscopy system.

4. Conclusion

In this study, we have introduced DL-SRQPI, a novel deep learning-based technique for quantitative phase microscopy with pixel super-resolution capability. DL-SRQPI enables full-FOV high-resolution phase imaging of unlabeled specimens using only a single frame of low-resolution intensity image as input, as validated by experimental data. The robust generalization capability of DL-SRQPI has been demonstrated using datasets of simulated and experimental intensity images with varying defocus distances and coherent parameters, which highlights the versatility and adaptability of our approach in handling different imaging conditions. The quantitative property of DL-SRQPI has also been validated by a microlens array. Furthermore, our method exhibits notable advantages in terms of fast speed and high-throughput, as evidenced by successful phase reconstruction of unstained biological cells and dynamic phase observation of HeLa cells. These results demonstrate the feasibility of DL-SRQPI for video-rate living cell phase imaging, opening up new possibilities for real-time cellular dynamics analysis. Importantly, DL-SRQPI significantly reduces the need for intensity data redundancy compared to conventional QPI approaches, thereby mitigating the trade-off between FOV and resolution. This advancement enhances the SBP of fundamental bright-field system equipment, facilitating more efficient and cost-effective imaging capabilities. Overall, DL-SRQPI showcases its potential as a powerful QPI technique with wide-ranging applications in high-throughput microscopy. It holds promise for various fields such as drug discovery, cellular phenotype characterization, and the identification of disease mechanisms.[58] Future investigations will delve into understanding the dependencies of DL-SRQPI on specific cell types and culture conditions, as well as addressing the impact of various experimental configurations on phase retrieval.

5. Experimental Section

Sample Preparation: To prepare biological material, the HeLa cells were cultivated in a glass bottom Petri dish (35 mm, MatTek) with L-glutamine Dulbecco's modified Eagles medium (Gibco, American)

supplemented with 10% Nu-serum (Corning, American), 10% fetal calf serum (Gibco, Australia), and 1% vitamin mix (100×) (Lonza, Cologne, Germany). The cells were cultured in a stage-mounted climate chamber (Tokai Hit INUF-IX3W, Japan) for stabilization of temperature at 37 °C and CO_2 gas at 5%. The medium was changed every other day and cells were passed with trypsin upon reaching 80% confluency. In preparation for cell division imaging, cells were washed once with phosphate-buffered saline and detached with either accutase or trypsin.

Experimental Configuration: For experimental measurements of HeLa cells, the intensity images used for TIE phase retrieval were obtained with an optical system equipped with an inverted microscope (IX71, Olympus, Japan), utilizing a halogen white-light source with a green interference filter (central wavelength $\lambda = 550$ nm, 45-nm bandwidth) for illumination. The microscope was equipped with a 5-megapixel charge-coupled device (CCD) camera (UC50, Olympus, Germany) with a pixel resolution of 2588 × 1960 and a pixel pitch of 3.4 µm. The microscope also included an electrically tunable lens (EL-C-10-30-VISLD, Optotune AG, Switzerland) module that was synchronized with the camera at different focal distances along the *z*-axis and controlled by software via a standard USB cable. The image stack was acquired via plan semiapochromat objective (LUCPlanFLN 40×, Olympus, Half magnification, NA 0.6) in an 8-bit grayscale range.

The microlens array (SUSS MicroOptics pitch 240ROC 297 µm) was imaged using an inverted bright-field microscope (Olympus IX71), and in-focus and out-of-focus intensity images were captured by axially translating the camera. The camera used here had a pixel size of 2.2 µm (The Imaging Source DMK 72BUC02, 1280 × 960 resolution), and the illumination was set to a central wavelength of 550 nm.

Implementation Details: The deep neural network was implemented using Pytorch 1.10.2 based on Python 3.9.7. The network training and testing were performed on a workstation with Intel(R) Core (TM) i9-10900K CPU (3.70 GHz) and 32 GB of RAM, running a Windows 10 operating system (Microsoft) using NVIDIA GeForce RTX 3090 GPU. With a batch size of eight, the training process of each dataset took ≈1 h for 200 epochs. The network took ≈0.3 s to reconstruct a full-field 2048 × 2048 pixels phase image from the 512 × 512 pixels simulated intensity image, and took ≈0.2 s to reconstruct a full-field 2588 × 1960 pixels phase image from the 647 × 490 pixels experimentally acquired intensity image.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (62105151, 62175109, U21B2033, 62227818), the Leading Technology of Jiangsu Basic Research Plan (BK20192003), the Youth Foundation of Jiangsu Province (BK20210338), the Biomedical Competition Foundation of Jiangsu Province (BE2022847), the Key National Industrial Technology Cooperation Foundation of Jiangsu Province (BZ2022039), Fundamental Research Funds for the Central Universities (30920032101), the Open Research Fund of Jiangsu Key Laboratory of Spectral Imaging and Intelligent Sense (JSGP202105, JSGP202201), and the Postgraduate Research Practice Innovation Program of Jiangsu Province (122104010479).

Conflict of Interest

The authors declare no conflict of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Keywords

deep learning, high-throughput microscopy, phase retrieval, quantitative phase imaging, super-resolution

Received: May 29, 2023
Revised: July 30, 2023
Published online: September 17, 2023

[1] J. Mertz, *Introduction to Optical Microscopy*, Cambridge University Press, Cambridge **2019**.
[2] G. Popescu, *Quantitative Phase Imaging of Cells and Tissues*, McGraw-Hill Education, New York **2011**.
[3] M. G. Gustafsson, *J. Microsc.* **2000**, *198*, 82.
[4] S. W. Hell, J. Wichmann, *Opt. Lett.* **1994**, *19*, 780.
[5] P. Gao, B. Prunsche, L. Zhou, K. Nienhaus, G. U. Nienhaus, *Nat. Photonics* **2017**, *11*, 163.
[6] E. Betzig, G. H. Patterson, R. Sougrat, O. W. Lindwasser, S. Olenych, J. S. Bonifacino, M. W. Davidson, J. Lippincott-Schwartz, H. F. Hess, *Science* **2006**, *313*, 1642.
[7] R. H. Webb, *Rep. Prog. Phys.* **1996**, *59*, 427.
[8] D. J. Stephens, V. J. Allan, *Science* **2003**, *300*, 82.
[9] Y. Cotte, F. Toy, P. Jourdain, N. Pavillon, D. Boss, P. Magistretti, P. Marquet, C. Depeursinge, *Nat. Photonics* **2013**, *7*, 113.
[10] P. Gao, C. Yuan, *Light Adv. Manuf.* **2022**, *3*, 105.
[11] Z. Huang, P. Memmolo, P. Ferraro, L. Cao, *PhotoniX* **2022**, *3*, 3.
[12] J. Li, A. Matlock, Y. Li, Q. Chen, C. Zuo, L. Tian, *Adv. Photonics* **2019**, *1*, 066004.
[13] J. Li, N. Zhou, J. Sun, S. Zhou, Z. Bai, L. Lu, Q. Chen, C. Zuo, *Light Sci. Appl.* **2022**, *11*, 154.
[14] S. Zhou, J. Li, J. Sun, N. Zhou, H. Ullah, Z. Bai, Q. Chen, C. Zuo, *Optica* **2022**, *9*, 1362.
[15] V. Mico, Z. Zalevsky, J. García, *Opt. Commun.* **2008**, *281*, 4273.
[16] Y. Park, C. Depeursinge, G. Popescu, *Nat. Photonics* **2018**, *12*, 578.
[17] P. Gao, B. Yao, I. Harder, N. Lindlein, F. J. Torcal-Milla, *Opt. Lett.* **2011**, *36*, 4305.
[18] Y. Fan, J. Li, L. Lu, J. Sun, Y. Hu, J. Zhang, Z. Li, Q. Shen, B. Wang, R. Zhang, et al., *PhotoniX* **2021**, *2*, 19.
[19] D. Dong, X. Huang, L. Li, H. Mao, Y. Mo, G. Zhang, Z. Zhang, J. Shen, W. Liu, Z. Wu, *Light Sci. Appl.* **2020**, *9*, 11.
[20] W. Choi, C. Fang-Yen, K. Badizadegan, S. Oh, N. Lue, R. R. Dasari, M. S. Feld, *Nat. Methods* **2007**, *4*, 717.
[21] Y. Sung, W. Choi, C. Fang-Yen, K. Badizadegan, R. R. Dasari, M. S. Feld, *Opt. Express* **2009**, *17*, 266.
[22] M. R. Teague, *J. Opt. Soc. Am.* **1983**, *73*, 1434.
[23] G. Zheng, R. Horstmeyer, C. Yang, *Nat. Photonics* **2013**, *7*, 739.
[24] C. Zuo, J. Sun, J. Li, J. Zhang, A. Asundi, Q. Chen, *Sci. Rep.* **2017**, *7*, 7654.
[25] J. Park, D. J. Brady, G. Zheng, L. Tian, L. Gao, *Adv. Photonics* **2021**, *3*, 044001.
[26] G. Zheng, C. Shen, S. Jiang, P. Song, C. Yang, *Nat. Rev. Phys.* **2021**, *3*, 207.
[27] J. Sun, Q. Chen, Y. Zhang, C. Zuo, *Opt. Express* **2016**, *24*, 15765.
[28] Y. LeCun, Y. Bengio, G. Hinton, *Nature* **2015**, *521*, 436.
[29] Y. Li, Y. Xue, L. Tian, *Optica* **2018**, *5*, 1181.
[30] Y. Rivenson, Z. Göröcs, H. Günaydin, Y. Zhang, H. Wang, A. Ozcan, *Optica* **2017**, *4*, 1437.
[31] H. Wang, Y. Rivenson, Y. Jin, Z. Wei, R. Gao, H. Günaydın, L. A. Bentolila, C. Kural, A. Ozcan, *Nat. Methods* **2019**, *16*, 103.
[32] Y. Rivenson, H. Wang, Z. Wei, K. de Haan, Y. Zhang, Y. Wu, H. Günaydın, J. E. Zuckerman, T. Chong, A. E. Sisk, *Nat. Biomed. Eng.* **2019**, *3*, 466.

[33] T. Nguyen, Y. Xue, Y. Li, L. Tian, G. Nehmetallah, *Opt. Express* **2018**, *26*, 26470.
[34] Y. Rivenson, Y. Zhang, H. Günaydın, D. Teng, A. Ozcan, *Light Sci. Appl.* **2018**, *7*, 17141.
[35] K. Wang, J. Di, Y. Li, Z. Ren, Q. Kemao, J. Zhao, *Opt. Lasers Eng.* **2020**, *134*, 106233.
[36] C. Zheng, D. Jin, Y. He, H. Lin, J. Hu, Z. Yaqoob, P. T. So, R. Zhou, *Adv. Photonics* **2020**, *2*, 065002.
[37] X. Chang, L. Bian, J. Zhang, *eLight* **2021**, *1*, 4.
[38] J. Sun, C. Zuo, L. Zhang, Q. Chen, *Sci. Rep.* **2017**, *7*, 7654.
[39] P. Stępień, D. Korbuszewski, M. Kujawińska, *ETRI J.* **2019**, *41*, 73.
[40] Y. Fan, J. Sun, Y. Shu, Z. Zhang, G. Zheng, W. Chen, J. Zhang, K. Gui, K. Wang, Q. Chen, C. Zuo, *Laser Photonics Rev.* **2023**, *17*, 2200201.
[41] W. Bishara, T.-W. Su, A. F. Coskun, A. Ozcan, *Opt. Express* **2010**, *18*, 11181.
[42] S. Jiang, J. Zhu, P. Song, C. Guo, Z. Bian, R. Wang, Y. Huang, S. Wang, H. Zhang, G. Zheng, *Lab Chip* **2020**, *20*, 1058.
[43] X. Chang, S. Jiang, Y. Hu, G. Zheng, L. Bian, *ACS Photonics* **2023**.
[44] S. Jiang, C. Guo, P. Song, N. Zhou, Z. Bian, J. Zhu, R. Wang, P. Dong, Z. Zhang, J. Liao, *ACS Photonics* **2021**, *8*, 3261.
[45] C. Zuo, J. Sun, Q. Chen, *Opt. Express* **2016**, *24*, 20724.
[46] Y. Shu, J. Sun, J. Lyu, Y. Fan, N. Zhou, R. Ye, G. Zheng, Q. Chen, C. Zuo, *PhotoniX* **2022**, *3*, 24.
[47] L. Lu, J. Li, Y. Shu, J. Sun, J. Zhou, E. Y. Lam, Q. Chen, C. Zuo, *Adv. Photonics* **2022**, *4*, 056002.
[48] M. Born, E. Wolf, H. Haubold, *Astron. Nachr.* **1980**, *301*, 257.
[49] C. Zuo, J. Li, J. Sun, Y. Fan, J. Zhang, L. Lu, R. Zhang, B. Wang, L. Huang, Q. Chen, *Opt. Lasers Eng.* **2020**, *135*, 106187.
[50] J. Sun, C. Zuo, J. Zhang, Y. Fan, Q. Chen, *Sci. Rep.* **2018**, *8*, 7669.
[51] J. Li, Q. Chen, J. Sun, J. Zhang, C. Zuo, *J. Biomed. Opt.* **2016**, *21*, 126003.
[52] L. Waller, L. Tian, G. Barbastathis, *Opt. Express* **2010**, *18*, 12552.
[53] D. Paganin, K. A. Nugent, *Phys. Rev. Lett.* **1998**, *80*, 2586.
[54] T. Falk, D. Mai, R. Bensch, Ö. Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, *Nat. Methods* **2019**, *16*, 67.
[55] X. Yang, L. Huang, Y. Luo, Y. Wu, H. Wang, Y. Rivenson, A. Ozcan, *ACS Photonics* **2021**, *8*, 2174.
[56] C. Zuo, Q. Chen, A. Asundi, *Opt. Express* **2014**, *22*, 9220.
[57] C. Dong, C. C. Loy, K. He, X. Tang, *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 295.
[58] Z. Chen, M. Segev, *eLight* **2021**, *1*, 2.