Optics and Lasers in Engineering 159 (2022) 107196

Contents lists available at ScienceDirect

Optics and Lasers in Engineering

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

Intensity and phase imaging through scattering media via deep despeckle complex neural networks

Shuai Liu[a,1], Peng Li[a,1], Hao Sha[b], Jiuyang Dong[a], Yue Huang[a], Yanjing Zhao[d], Xuri Yao[e], Qin Peng[d], Xiu Li[a], Xing Lin[c,f,*], Yongbing Zhang[b,*]

[a] Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China
[b] School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen, Guangdong 518055, China
[c] Department of Electronic Engineering, Tsinghua University, Beijing 100084, China
[d] Institute of Systems and Physical Biology, Shenzhen Bay Laboratory, Shenzhen, Guangdong 518055, China
[e] Center for Quantum Technology Research, School of Physics, Beijing Institute of Technology, Beijing 100081, China
[f] Beijing National Research Center for Information Science and Technology, Tsinghua University, Beijing 100084, China

ARTICLE INFO

Keywords:
Despeckle
Quantitative phase imaging
DespeckleNet

ABSTRACT

The existence of a scattering medium causes the degeneration of intensity and phase information, especially in biological imaging. The present techniques to address this challenge only focus on the reconstruction of intensity information, yet few attempts have tried to recover the phase information. We propose a method to simultaneously predict both intensity and phase information from a speckle image employing a deep despeckle complex neural network (DespeckleNet). By combining the advantages of both the complex network and the generative adversarial network framework, our method enables the high contrast single-shot imaging of complicated biological samples through scattering media without labeling. Various experiments demonstrate the superior reconstruction and generalization performance of our method under multiple types of biological samples with different scattering media. We also provide the real-time observation of living cellular activities without any contaminations or damages to the cells. Our method offers simple yet effective imaging through scattering media and paves the way for real-time unlabeled biological imaging.

1. Introduction

As one of the most challenging and practical research topics, imaging through scattering media [1,2] receives more attention recently in many fields, such as cloud tomography [3], underwater imaging [4], and biomedical imaging [5]. In particular, the tissues and cells of most organisms exhibit heterogeneous refractive indices, which causes the relative phase of the laser to be randomly scrambled and thus generates speckle [6]. This limits the resolution of imaging as well as the depth of observation. Many techniques have made great achievements in imaging through scattering media over the years, including optical coherence tomography (OCT) [7], wavefront shaping [8–11], optical transmission matrix [12–18], and others. However, the spatial distribution of the speckle is a complicated function of the microscopic arrangement of the scattering media and the wavefront of the incident field. And it is difficult to provide a wide field of view (FOV) due to the memory effect. OCT and wavefront shaping techniques demand a sophisticated optical design and hardware, which is hard to deploy in practice. In recent years, deep learning (DL) has shown favorable achievements in the field of imaging through scattering media [19–28]. Li built a "one to all" model based on the UNet backbone, which can learn statistical information about similar scattering media with different microstructures and extract statistical invariance of the speckle [23]. However, previous work using deep learning for de-speckle has only reconstructed binary images or simple handwritten digits, rather than complex grayscale samples in real biological microscope scenes.

Meanwhile, Researchers have also developed the reconstruction of phase information through scattering media [29]. Phase information reflects the cellular structures and plays an important role in the imaging of transparent samples. Conventional bright-field microscopy needs to stain the sample or use fluorescent labels to show the structure [30], which may affect the normal vital movement of cells [31,32]. In contrast, phase imaging is a label-free microscopic imaging method that enables the imaging of transparent samples by attaining phase information of different components of cells. A well-known phase imaging technique is phase-contrast microscopy proposed by Zernike. Phase-contrast microscopy uses interference between the scattering and non-

* Correspondence authors.
E-mail addresses: lin-x@tsinghua.edu.cn (X. Lin), ybzhang08@hit.edu.cn (Y. Zhang).
[1] These authors equally contributed to this work

https://doi.org/10.1016/j.optlaseng.2022.107196
Received 13 April 2022; Received in revised form 5 July 2022; Accepted 16 July 2022
Available online 28 July 2022

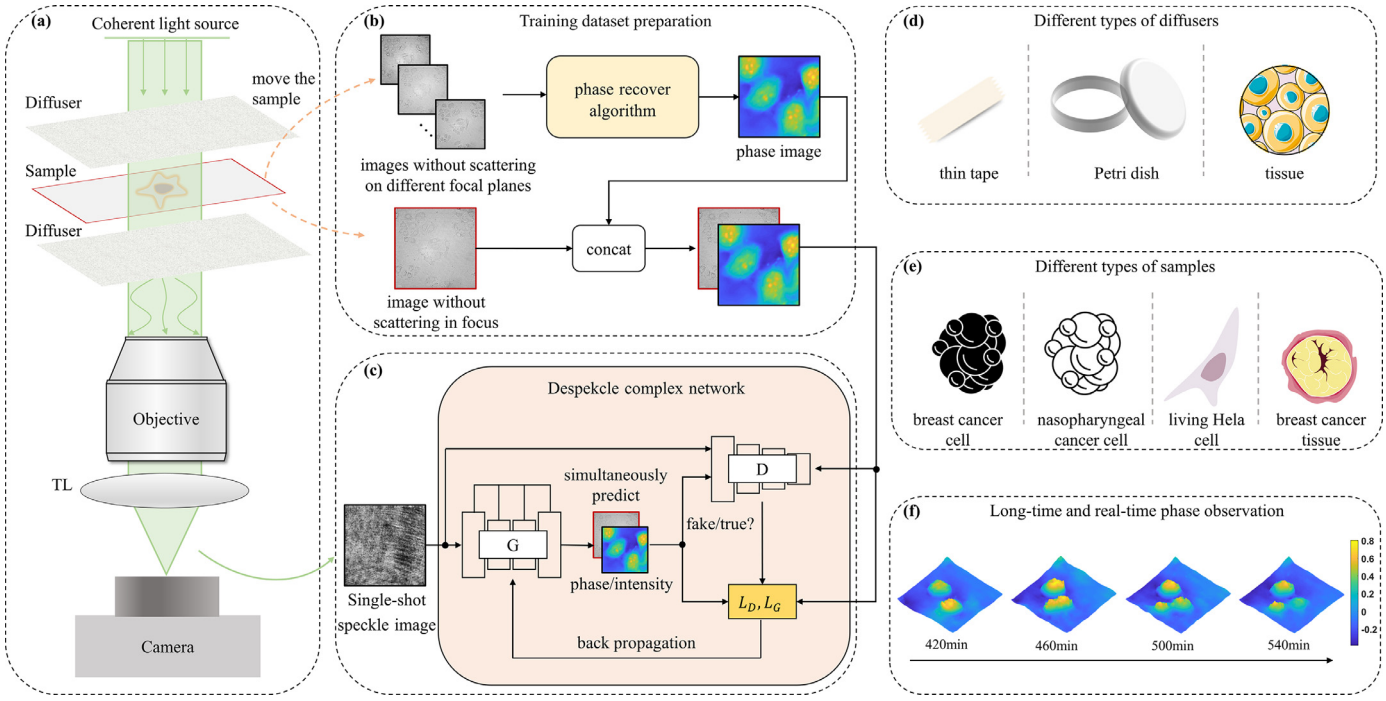

Fig. 1. Workflow overview. (a) Schematic view of the experimental setup. The laser illuminates the sample through diffusers, and the speckle information is recorded by the camera. (b) The intensity images of the sample without scattering on different focal planes are used to generate phase images employing the TIE algorithm. The in-focus intensity image without scattering concatenates the phase image as the ground truth of the proposed network. (c) The overall structure of the DespeckleNet. The network is capable of simultaneously recovering both intensity and phase information from one speckle image. (d) Multiple kinds of diffusers, including thin tape, Petri dish, and tissue, are used to verify the performance of our method. (e) Multiple kinds of diffusers, including breast cancer cells, nasopharyngeal cancer cells, living Hela cells, and breast cancer tissue, are used to demonstrate the generality of our method. (f) A long-time and real-time phase observation of living Hela cells. The unit of phase is radian (rad).

scattering light waves to convert the optical phase into high contrast image [33]. Nomarski further invented differential interference contrast (DIC) microscopy based on phase-contrast microscopy, which can display a three-dimensional projection of the structure [34]. Both of these methods belong to qualitative phase imaging techniques. Quantitative phase imaging techniques have made great progress in the biomedical field recently, including the transport of intensity equation (TIE) [35,36], digital holographic imaging [37,38], and tomographic phase microscopy [39–41]. The digital off-axis holography is the most widely used due to its high-quality phase reconstruction capability [42,43]. But it requires the introduction of an additional reference light path and does not achieve high space bandwidth utilization. The crosstalk between the real image, zero-order and twin images may lead to blurring and artifacts when the space-bandwidth production required by off-axis digital holographic is not satisfied [44,45]. Among these methods, the TIE technique can be easily deployed to conventional microscopy due to the desirable property of simple acquisition manner, no reference beam, and no phase unwrapping. Furthermore, it also suits both coherent and partially coherent illumination [46–48]. However, TIE usually requires a series of images captured at different focal depths, which extends the acquisition time and limits real-time observation of the dynamic process. Although deep learning has made achievements in phase recovery [49,50], all of these methods cannot be directly applied to the simultaneous reconstruction of intensity and phase in scattering problems.

Previous studies indicate that there exists a functional relationship between the speckled image and the original sample [13], and phase information can also be inferred from the intensity distribution [51]. In this paper, we developed a deep despeckle method to simultaneously predict the intensity and phase information of biological samples through scattering media via a complex neural network (DespeckleNet). A speckled image, generated by inserting a diffuser in the optical path, is fed into our network and both intensity and phase are predicted as output. The network is devised based on the generative adversarial network (GAN) framework [52–54], and we obtain a resilient speckle decorrelation for the intensity and phase with a wide range of statistical variations. So we can make high-quality object predictions with a completely different set of diffusers of the same class. Different from previous deep learning-based despeckle works, we observe the biological sample using a microscope rather than a simulated object displayed by the spatial light modulator (SLM) [23,24], which is more difficult for the network to reconstruct. We carry out experiments on both cells and tissue to demonstrate the superior performance of our method. In addition, we also observe the real-time division process of unlabeled living Hela cells using quantitative phase imaging. Our solution is a single-shot imaging method and can be applied in a variety of microscope systems without providing an additional sophisticated optical structure for despeckle or phase reconstruction. Compared to typical holography - off-axis holography, our solution can effectively eliminate the speckle and is not limited by the space-bandwidth production of detector.

2. Methods

2.1. Overview of the workflow

The overall experimental setup is depicted in Fig. 1(a). The coherent light source is modulated through the diffuser to illuminate the biological sample. Due to the optical roughness of the diffuser surface, the image captured by the camera is mixed with the granular pattern. Based on captured speckle image, both intensity and phase information can be simultaneously recovered by our DespeckleNet, whose flowchart is shown in Figs. 1(b) and 1(c). To build the dataset for network training and testing, we replace the laser with an incoherent light source to obtain a series of in-focus and out-of-focus sample images without scattering. The in-focus image without scattering is the ground truth of

intensity. The ground truth of the phase component is obtained by TIE, which takes images from different focal planes as input to extract the phase information (the detail please see supplement TIE algorithm). To reduce the computational complexity, each speckle image and the corresponding ground truth are cropped into lots of smaller-sized patches and then fed into our model. The speckle pattern of each patch is different even they come from the same diffuser. Our model learns the statistical invariance of the diffuser and can output high-quality reconstructions even for the speckle patterns that have not been seen during training. Experimentally, we verify the network performance with three kinds of diffusers, including thin tape, Petri dish, and tissue, as shown in Fig. 1(d). We also use four different biological samples, containing breast cancer cells, nasopharyngeal cancer cells, living Hela cells, and breast cancer tissue, to demonstrate the generality of our method, as shown in Fig. 1(e). In Fig. 1(f), we show real-time quantitative phase observation of living Hela cells without labeling or staining over a long time, which proves that our method can achieve single-shot unlabeled imaging through scattering media. Note that the units of all the reconstructed phases in the text are radians (rad).

2.2. The architecture of the DespeckleNet

The architecture of the DespeckleNet is shown in Fig. 2. Our network realizes an inverse process of the physical system. It inputs a speckle intensity image and outputs high-quality predictions of both phase and intensity of the sample, which corresponds to the components of the coherent transfer function. In our implementation, the DespeckleNet is designed following the generative adversarial network (GAN) framework. The GAN is composed of a generator, as shown in Figs. 2(a)-(d), and a discriminator, as shown in Fig. 2(e). The generator generates the intensity and phase, and the discriminator judges whether the outputs are real or fake. Specifically, we first extract the features of the input image through the feature extractor, as indicated in Fig. 2(a). Next, we add a translate module to cope with the influence of speckle by stacking real-valued and complex residual blocks sequentially, as indicated in Fig. 2(b). The operation of the modified complex residual block is shown in Fig. 2(f), which decouples the intensity and phase information from extracted features. Compared with the real-valued convolution, the complex convolution is more consistent with the optical complex field propagation model. Then we employ two decoders to reconstruct the intensity and phase images, as shown in Figs. 2(c) and 2(d). The two decoders have the same structures except that the intensity branch has an additional tanh activation function in the last layer to normalize the output. For the discriminator, we adopt the PatchGAN to perform discrimination patch-wisely, which is beneficial for the reconstruction of details, as shown in Fig. 2(e).

Generative model. The backbone of our generative model is UNet, which was first proposed for medical image segmentation [55]. The generator contains four components, which are feature extractor, translate module, intensity, and phase reconstruction module, as shown in Fig. 2. The feature extractor consists of five consecutive convolution blocks. Each convolution block contains a convolution layer with 4×4 kernel size, batch normalization layer, and leaky ReLU activation layer. After each convolution block, we double the number of channels and set the convolution stride to 2, which can downscale the feature map and expand the receptive field. Since the formation of speckle results from the nonlinear weighted summation of different scattering mode microstructures within a certain receptive field, expanding the receptive field is beneficial for the despeckle task. Our proposed feature translate module adopts five residual convolution blocks and four modified complex residual blocks rather than the common real-valued networks. Because light propagates in complex form, our module is more applicable in the simultaneous recovery of both intensity and phase processes. The experiment results support this conclusion. Compared to conventional complex network [56,57], our complex residual blocks in the translate module are used for feature transformation with the same size, so the pooling layer is removed. The residual structure in the translate module can speed up the convergence rate.

Traditionally, given a complex-valued convolutional filter $W=W_r+iW_j$ and a complex input $h=h_r+ih_i$, where subscript r represents the real part and i represents the imaginary part, the complex convolution (CConv) is shown as:

$$W * h = \left(W_r * h_r - W_i * h_i\right) + i\left(W_i * h_r + W_r * h_i\right), \tag{1}$$

where $*$ denotes real-valued convolution operation. In our modified complex convolution, the input channels are divided into two groups, standing for real and imaginary features, respectively. The convolution results of these two groups are then fused by 1×1 convolution rather than directly adding the real and the imaginary part. Next, we use a universal activation function, Leaky ReLU, to realize nonlinear mapping. The complex residual block (*CRB*) in our feature translate module is defined as

$$CRB(h) = Leaky\ ReLU\left(conv_{1*1}\left(Leaky\ ReLU\left(conv_{3*3,group=2}(h)\right)\right)\right) + h. \tag{2}$$

The feature map in the translate module has small lateral dimensions (16×16) and a large number of channels (512), which can encode sufficient information beneficial to the following task. Next, these feature maps go through the intensity reconstruction module, which contains five consecutive transposed convolution layers (kernel size=4×4, stride=2) followed by a *tanh* activation layer to normalize the output from -1 to 1. The information across different spatial scales is tunneled through the down-up paths by skip connections to preserve high-frequency information. Similar operations except the *tanh* activation are applied to the phase reconstruction to keep consistent with the absolute value of ground truth.

Discriminative model. The conventional discriminative model consists of a series of convolution layers, a global average pooling layer, and several fully connected layers. The last fully connected layer is followed by a sigmoid activation function whose output is the probability of an input image being real or fake. The global average pooling layer limits the discriminator to focus only on the overall content while ignoring the fine-grained local features of the output. In contrast, the PatchGAN architecture can pay more attention to the structure in local image patches. Here, we build our discriminative model based on PatchGAN, as this method showed excellent results for image translation task [58]. The PatchGAN maps the input image into a $N \times N$ patch matrix M via convolution operation, and the value of M_{ij} represents the probability of being real for the corresponding patch. The final output of the discriminator is the average of all the elements within matrix M.

Loss function. We design our loss function as the combination of the adversarial loss with L1 loss as the regularization terms. To ensure the convergence, we alternatively train the discriminator and generator. The loss function of the discriminator is formulated as:

$$\min_D L(D;G) = -log\,D(y) - log[1 - D(G(x))], \tag{3}$$

where G represents a fixed generator, and D represents the discriminator to be updated, $y=concat(y_I, y_P)$ is the concatenated ground truth of intensity and phase. The loss function of the generator is formulated as:

$$\min_G L(G;D) = -log(D(G(x))) + \alpha\left(L_1(G(x), y)\right), \tag{4}$$

where α is the regularization hyper parameters, $L_1(G(x), y)$ represents the L1 loss between $G(x)$ and y. The symbols D, G and y have the same meaning as those in Eq. (3). While the adversarial loss guides the generative model to map the speckle images into despeckle ones, the regularization terms can speed up the convergence of our network and smooth the training loss. Experimental results reveal that good performance can be achieved when α is set between 0.3 and 0.4.

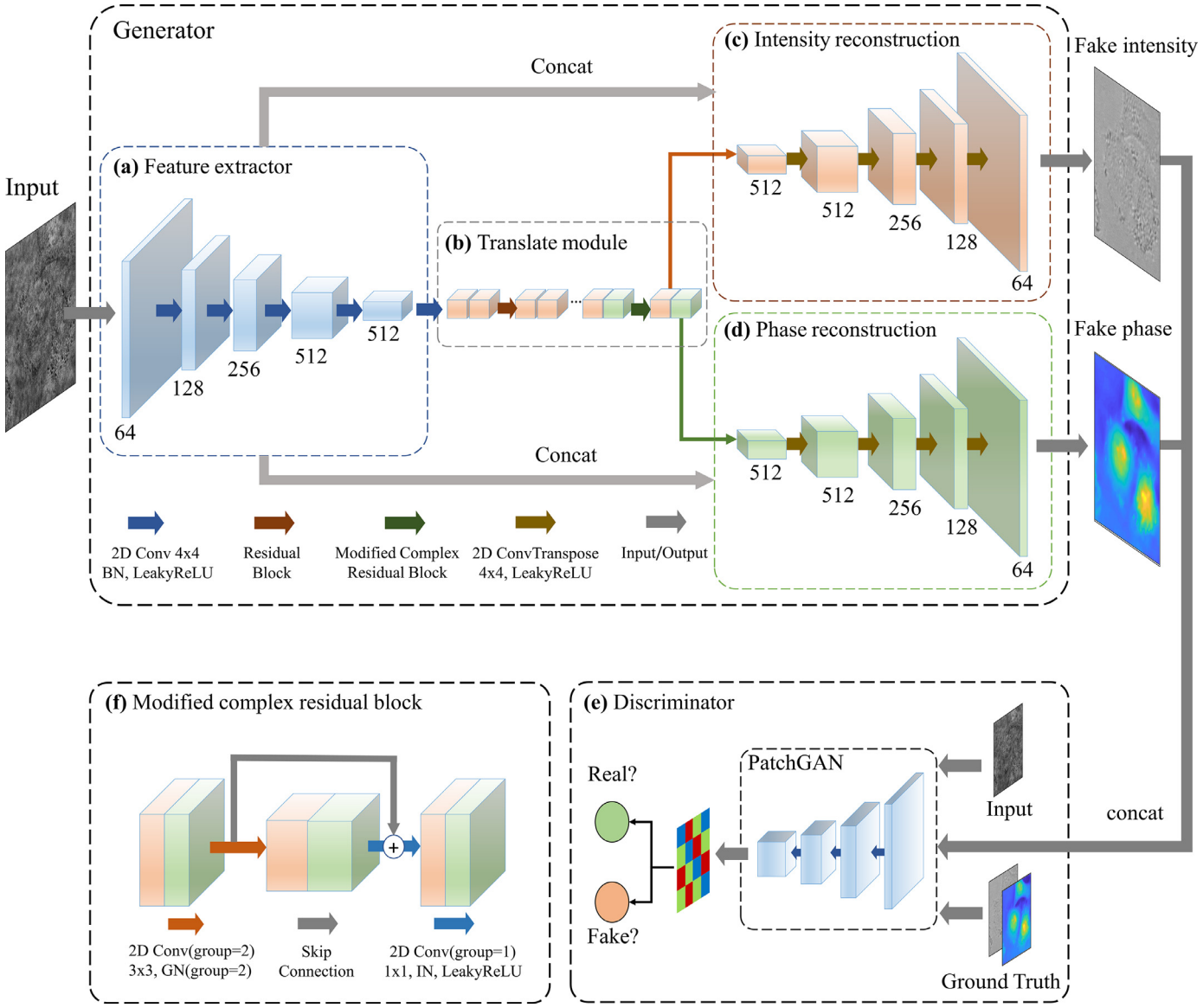

Fig. 2. The architecture of the DespeckleNet. (a) The feature extractor learns the feature representation of the speckle image from the input. (b) The extracted features are refined by the translate module, composed of real-valued and complex residual blocks, for subsequent reconstruction and prediction. (c)-(d) The intensity reconstruction module and phase prediction module use the refined features to predict the intensity and phase image, respectively. (e) The PatchGAN discriminates whether the intensity-phase pair is the ground truth or the pseudo generated by the generative model. (f) The detailed structure of modified complex residual block in (b). It contains group convolution, group normalization, 1 × 1 convolution (channel fusion), instance normalization, and LeakyReLU layer.

2.3. Metrics

Following the numerical evaluation methods proposed in the previous papers, the mean structural similarity (SSIM) [59], learned perceptual image patch similarity (LPIPS) [60] and normalized root mean square error (NRMSE) are used to evaluate the results of the despeckle and phase prediction. The values of the SSIM and NRMSE metrics are between 0 and 1. The higher SSIM and lower NRMSE values represent a high-quality output result. NRMSE is a general objective evaluation index in measuring image quality, while SSIM assures a high degree of structural consistency between output and ground truth. The SSIM is defined as

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_1)}, \tag{5}$$

where μ_x, μ_y are the averages of x, y, $\sigma_x{}^2$, $\sigma_y{}^2$ are the variances of x, y, σ_{xy} is the covariance of x and y; and c_1, c_2 are the variables used to stabilize the division with a small denominator. We also use LPIPS, a learning-based perceptual similarity metric, to measure the distance in feature space, which is more consistent with human perception. The LPIPS is defined as

$$d(x,\hat{x}) = \sum_{l=1}^{L} \frac{1}{H_l W_l} \sum_{h=1,w=1}^{H_l,W_l} \| \omega_l \odot (y_{hw}^l - \hat{y}_{hw}^l) \|_2^2, \tag{6}$$

where x, $\hat{x}$ are the intensity (or phase) ground truth and output from the DespeckleNet, and $y^l, \hat{y}^l \in \mathbb{R}^{Hl \times Wl \times C}$ are the corresponding lth layer features extracted from pre-trained AlexNet, H_l, W_l, C_l are the size of feature map at lth layer, $\omega^l \in \mathbb{R}^{C_l}$ is real-valued scaling weight, and $\odot$ represents the Hardamard product operator. The number of feature layers L is set to 5. The lower LPIPS value means that the images are more similar to the ground truth in the aspect of human perception.

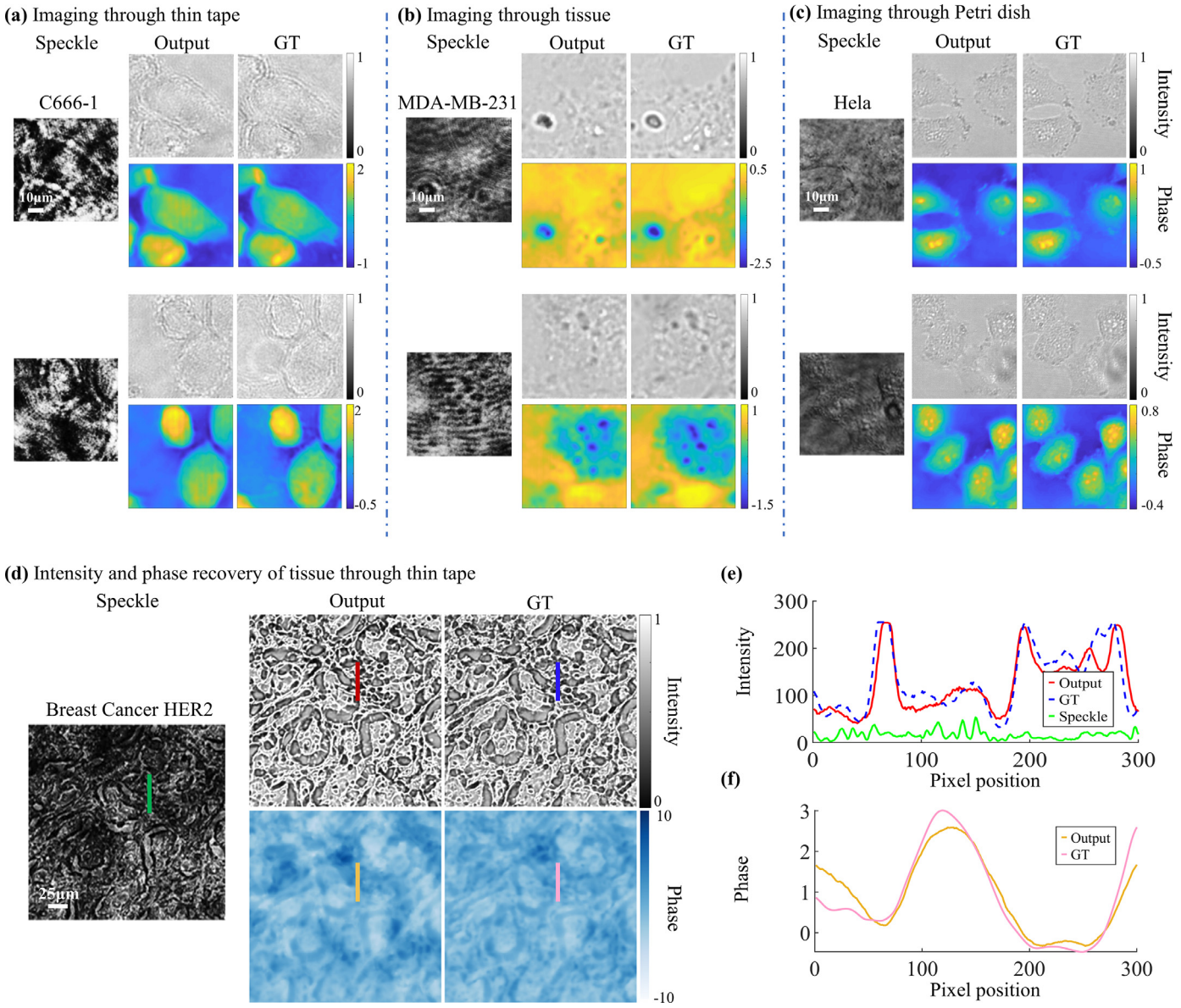

Fig. 3. The reconstructions via the DespeckleNet. Recovered intensity and phase of (a) nasopharyngeal carcinoma cells (C666-1) through thin tapes, (b) breast cancer cells (MDA-MB-231) through unstained tissue, (c) Hela cells through Petri dish, and (d) breast cancer tissue with HER2 through thin tapes, respectively. (e) Intensity and (f) phase profile along the corresponding line trace in (d). The speckle patterns used for testing are not previously seen by the network. Output: the output of DespeckleNet, GT: Ground truth.

2.4. Data acquisition

In our experiments, we image different types of cells and tissues through different scattering media including thin tape, tissue, and Petri dish. To obtain the ground truth for network training, we collect ±2 μm, ±4 μm out-of-focus images, and an in-focus image. The in-focus image is used as the ground truth of the intensity image, and the ground truth of the phase image is calculated by the TIE algorithm using four out-of-focus and one in-focus intensity image. The input of the network is the speckle image, which is obtained by illuminating the biological sample through the diffuser using the laser. For each sample under a specific diffuser, we collect 20 sets of images, which are then randomly cropped into about 2,000 512 × 512 patches for training, and about 200 patches for testing. And we will re-collect 2000 patches as a training dataset and 200 patches as a testing dataset in the same way, when replacing the sample or diffuser types for experiments. The images in the testing dataset are not used for training, which ensures that our network never sees the speckle patterns before. Therefore, the datasets have the following two characteristics. One is that each patch corresponds to a different scattering pattern, which can test the robustness to the scattering modes. Another is that our datasets consist of different samples, which can test the generalization of the sample types.

We design two scattering modes for data acquisition. In mode one, the living Hela cells are imaged through a Petri dish, and the laser goes through the diffuser (the plastic shell of the Petri dish) twice. In mode two, the breast cancer cells, nasopharyngeal carcinoma cells, and breast cancer tissue with HER2 are imaged through thin tape and tissue section, respectively, and the laser go through the diffuser (thin tape and tissue section) once. A 40 × /1.25NA objective is used to magnify the Hela cells in mode one. A 40 × /0.65NA objective is employed for the cell

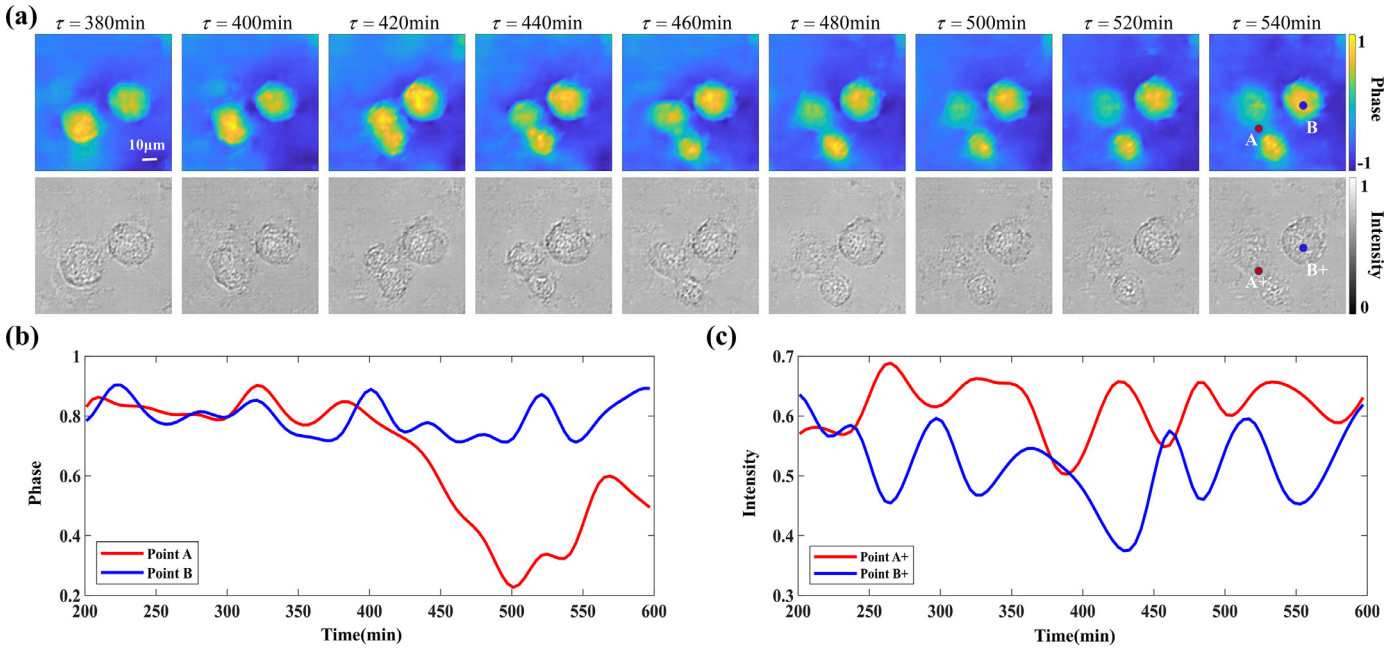

Fig. 4. Morphological changes of living Hela cell during the division process. (a) Phase, intensity, and speckle profiles at different timestamps (see Visualization 1). (b) Phase variations with time of points A and B, indicated by red curve and blue curve, respectively. (c) Intensity variations with time. The red and blue solid curves represent the variations for points A+ and B+ in recovered intensity images.

samples and a 20 × /0.5NA objective is employed for the tissue samples in mode two. The detailed description of the experimental setup can be found in Figs. S1 and S2.

For living cell imaging samples, 4×10^4 Hela cells were seeded on Φ15 mm glass-bottom dish (NEST 801002) pre-treated with 20 μg/mL human fibronectin (Coring 356008) at 37 ˚C for 1 hour, treated in DMEM (GIBCO 10569010) supplemented with 10% FBS (GIBCO 10099-141C) and 100 U/mL Penicillin-Streptomycin (GIBCO 15140122) and returned to CO2 incubator overnight. Next day, the culture medium was replaced with no phenol red DMEM (GIBCO) supplemented with 10% FBS (GIBCO) and 100 U/mL Penicillin-Streptomycin (GIBCO). The samples were placed at a 37 ˚C chamber supplemented with 5% CO2 for imaging.

3. Result

3.1. Multiple types of samples imaging through different scattering media

To verify the generalization and robustness of our method, we experimentally image cell samples through different scattering media, including thin tape, tissue section, and Petri dish. In addition, we test tissue samples to further demonstrate the superior property in terms of different sample types. Notably, all the samples preserve the original structure without fluorescence labeling or staining. These biological samples are illuminated by the laser through the diffuser, and on the other side, a camera captures the raw image mixed with the speckle pattern. The detailed setup is shown in Fig. S1. The captured raw images, whose speckle patterns have not been seen by the network, are fed into our trained pixel to pixel network. Since the unstained cells are transparent, the captured raw images have poor contrast, which leads to a difficult solution to the pixel-level prediction task in bioimaging. Nevertheless, our network outputs the simultaneously recovered intensity and phase images with high quality, compared with the ground truth obtained under a non-scattering scenario.

We evaluate the performance of our DespeckleNet over nasopharyngeal carcinoma cells (C666-1) and breast cancer cells (MDA-MB-231). The first sample is imaged through thin tapes to keep consistent with previous methods [61–63]. Furthermore, we replaced the thin tape with a breast cancer tissue section (~4 μm thick) as a diffuser for the second sample, which is more in line with real bioimaging. Representative examples of the speckle and prediction pairs are shown in Figs. 3(a) and 3(b). The speckle prevents our normal observation of the cells. We input the speckle into our network and the outputs are clear images of the cell's intensity and phase, which are very close to the ground truth. We can see that the intensity images of cells without labeling have low contrast and it is difficult to recognize the boundary of cells against the background. In contrast, the corresponding cell phase images have high contrast and we can observe the difference between the morphological features of the cells. Our results show that our network performs well not only under the thin tape but also in the real biological tissue scattering media, providing a potential solution for unlabeled deep tissue laser imaging. Next, we test the performance of the DespeckleNet over living Hela cells through a Petri dish. Imaging of living cells often requires the cells to be placed in a Petri dish, which is equivalent to a diffuser. Different from other experimental configurations, the Petri dish consists of the upper and lower surfaces, and the laser will go through the scattering media twice in the whole optical path. The results are shown in Fig. 3(c), where the outputs of the network are still satisfactory even after scattering twice. Our method enables high quality coherent illumination imaging of unlabeled living cells.We also carry out the experiment over breast cancer tissue with HER2 through a thin tape to prove the robustness of our network with different types of samples. As shown in Fig. 3(d), the pathological tissue structure is more complicated and the texture has more details, which inevitably imposes a great challenge for high-quality reconstruction. On the other hand, the observation of tissue usually requires a wide FOV, so we capture the raw image with 2048 × 2048 pixels (FOV ~250 × 250 μm^2). The conventional scattering imaging technique has a small imaging FOV due to the limitation of the optical memory effect, and our method can break through the limitation. In Fig. 3(d), the outputs of the network remain high fidelity compared to the ground truth, even though the tissue speckle images are blurred and coarse. To quantitively analyze the performance of our network, the intensity and phase profile along the corresponding line trace are plotted in Figs. 3(e) and 3(f), respectively. The red solid line (network output) almost coincides with the blue dashed one (ground truth).

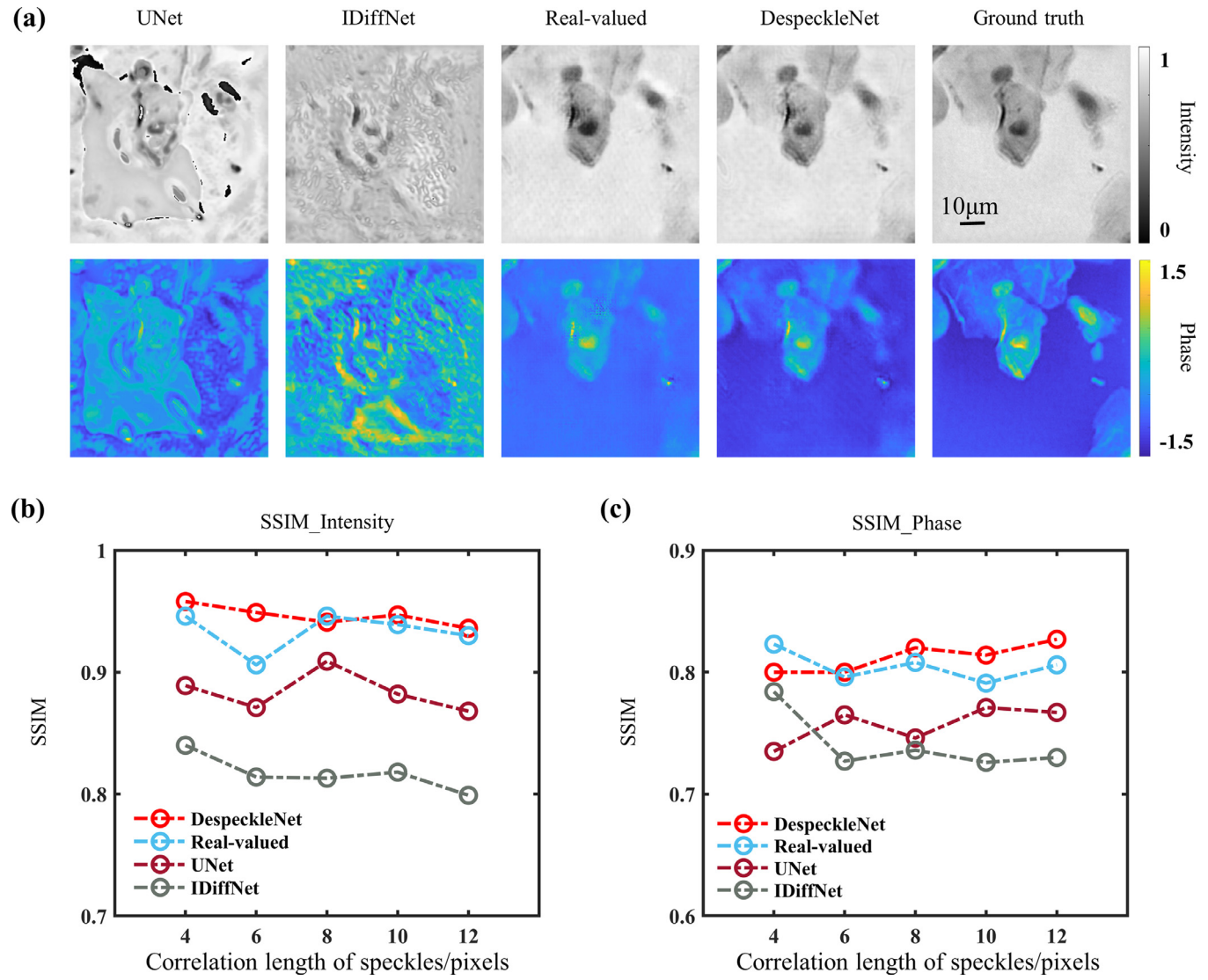

Fig. 5. Simulation results over different coherence scales. (a) The recovered results of different networks including Unet, IDiffNet, real-valued network, and DespeckleNet. The mean SSIM of (b) intensity and (c) phase based on different correlation lengths of speckles [66].

There is no visual correlation between the green solid line (speckle) and the blue dashed one, which means the intensity distribution of the input is different from the ground truth. Besides, we can draw a similar conclusion in terms of phase profile in Fig. 3(f).

3.2. *The real-time intensity and phase imaging of living Hela cells*

Observing living cells is one of the most challenging and meaningful experiments in biological microscopic imaging, which just falls into the category of the application of our method. One significant strength of our network is the ability to achieve high contrast and long-time observation of living cells without labeling, which preserves the vital movement of the cells to the maximum extent. The other advantage of our method is the real-time observation based on a single-shot speckle image. Here, we provide an example of real-time imaging of living Hela cells. For the data collection, we used a Spinning Disk Confocal microscope (Dragonfly 200, Andor) with a 40 × /1.25NA objective, and the sample is illuminated by an extra laser (405nm, 50mW). The other experimental configuration is the same as those in Fig. 3(c), as shown in Fig. S1. We take a video of the living Hela cell (see Visualization 1), where the whole division process in the vital movement can be observed. The duration of the video covers 800 minutes with an interval of 20 minutes.

We extract 9 sequential representative frames from Visualization 1 ranging from 380 min to 540 min, as shown in Fig. 4(a). The complete division process of Hela cells is observed within the interval of 160 minutes. There are only two cells at 380 min, and the left cell begins to divide at 420 min. The division of the left cell is finished while the right cell remains unchanged at 540 min. The optical morphology of the cell changes during the process of division. In Fig. 4(a), the morphology of Hela cells in phase information changes dramatically during the process of division, and there is a clear difference between the two divided daughter cells. Otherwise, the intensity images reconstructed by the network are ambiguous in terms of the morphological characteristics of the cells.

The temporal variations of both point A and point B in the phase image of Fig. 4(a) are plotted in Fig. 4(b). The red curve shows a decreasing trend within 400 min to 500 min, reflecting the phase change of the region where the cell divides. Meanwhile, the blue curve exhibits a stable trend with small fluctuations in phase along the temporal axis, indicating there is no dividing process at point B. In addition, Fig. 4(c) shows the intensity variations of points A+, B+ in the corresponding recovered intensity image along the temporal axis. Although points A and

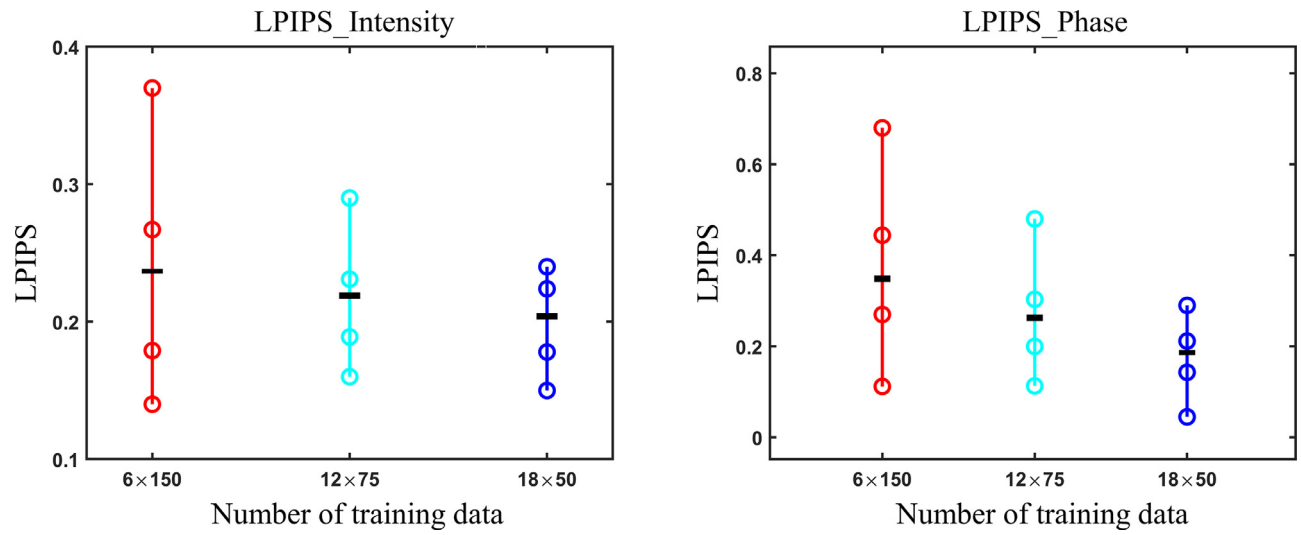

Fig. 6. The performance of our networks with different training data distributions. We compare the performance of our networks trained on Hela data with different training data distributions (patches cropped from 6, 12, and 18 sets of speckle images) by LPIPS. Each circle from top to bottom represents the maximum, upper quarter, lower quarter, and the minimum of the intensity and phase LPIPS, respectively. The mean LPIPS is marked by black horizontal bars. The results show that the mean and variance of the LPIPS metrics become better as the diversity of data distribution increases.

A+ locate the same position in the phase image and intensity image, the red solid curve in Fig. 4(c) has no obvious upward or downward tendency during the division process. It reveals that it is difficult to distinguish the cell division process by only using the recovered intensity image. Overall, it is more accurate to identify the dividing state of the living Hela cells employing phase information. In particular, there is an obvious phase difference between A and B in Fig. 4(b) after the left cells finished the division process at 540 min. However, the intensity difference between the two points at this moment is slight due to the high transparency of the cells. This further reveals the benefits of simultaneously recovering both intensity and phase in our method.

4. Discussion

In coherent bioimaging, the coherent transfer function is a complex-valued function describing the response of intensity and phase. Considering this optical property during imaging, we design the DespeckleNet. In contrast to the traditional real-valued network, our method utilizes the complex convolution block. We first extract features using a series of general convolution layers, then the features are further divided into two groups for intensity and phase reconstruction separately. Our network stacks 5 common convolution blocks and 4 complex convolution blocks in the translate module, as shown in Fig. 2. More details about the complex convolution block can be seen in Method. To evaluate the performance of our model, we compare it with the real-valued network model and two previous deep learning despeckle networks (UNet [23] and IDiffNet [24]) on the simulated dataset with five scatter level images. The real-valued network has the same structure as the complex network by replacing the complex convolution block with the traditional real-valued one. Particularly, the parameter sizes of the complex and real-valued networks are the same.

When a coherent signal consists of a large number of complex phasor components, the speckle would be generated by the summation of independent phases. In particular, the phase and amplitude of the laser distribute randomly after passing through a diffuser, and the superposition of these complex components results in the “random walk” [64]. In addition, the high coherence of the laser causes coherent superposition in a certain optical plane to form a speckle. The complex wavefront $\boldsymbol{A}$ is given by

$$\boldsymbol{A} = Ae^{j\theta} = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} a_n e^{j\varphi_n}, \tag{7}$$

where A is the magnitude of the resultant phasor component summation, θ is the phase of the resultant phasor component summation, N denotes the number of phasor components in the random walk, a_n and φ_n represent the nth magnitude and phase of the phasor vector, respectively.

To fairly compare the performance of networks, we use the open-source dataset [65] to generate paired simulated speckle images for training different despeckle networks. The corresponding code for speckle simulation can refer to Ref 4646. The simulation results are shown in Fig. 5(a), where the DespeckleNet has better output compared to other networks. In the result of the UNet and IDiffNet, the contour of the recovery is not distinct, contaminated by the artifacts. The real-valued network outperforms those two networks, especially in the intensity recovery. However, there is still an obvious difference between the phase output and the ground truth. On the contrary, the reconstruction of the DespeckleNet is the closest to the ground truth in both intensity and phase, which demonstrates the superior advantages of our method over other competitive ones. It should be noted that the real-valued network is consistent with DespeckleNet in terms of the number of parameters. Quantitively, we provide the SSIM indices of these four networks, as shown in Figs. 5(b) and 5(c). Our DespeckleNet achieves the highest score and is robust even when the correlation length of speckles increases.

As shown in Fig. 6, we evaluate the impact of training datasets with different distributions on network performance using the LPIPS metric, which is closer to human perception in visual similarity judgments. We acquire multiple sets of data under different experimental conditions (including light intensity, light incidence angle, etc.) to construct our training datasets, and the number of sets reflects the data diversity in training datasets. We constructed three training sets for comparison, cropped from 6, 12, and 18 sets of speckle images, respectively. For the sake of fairness, the total number of images is the same in different training datasets. Through quantitative analysis, the training dataset with the largest number of sets has the lowest mean of LPIPS, which demonstrates that the diversity of training sets is conducive to improving the generalization ability of the network.

5. Conclusion

We achieved biological imaging through scattering, where simultaneous phase and intensity imaging through scattering media are real-

ized via a deep despeckle complex neural network. By fully integrating the GAN network and complex neural network, our network can obtain high contrast and real-time observation of unlabeled living Hela cells, which preserves the vital movement of the cells to the maximum extent. And our solution breaks the limitation of the memory effect to achieve biological tissue imaging through scattering media with a wide field of view. Notably, our network achieved simultaneous prediction of grayscale intensity and quantitative phase that can generalize to the various biological samples through unseen diffuser patterns. We believe our work provides a new method for imaging through scattering media by using artificial intelligence technology.

Funding

This work was supported in part by the National Natural Science Foundation of China (61922048 & 62031023), in part by the Shenzhen Science and Technology Project (JCYJ20200109142808034), in part by the Guangdong Special Support (2019TX05×187), and in part by the Guangdong Provincial Key Laboratory of Novel Security Intelligence Technologies (2022B1212010005).

Data availability

Due to the large size of the complete data, we have uploaded a partial dataset of cells at https://drive.google.com/drive/folders/1jF5DLO8Ug0hElA0rOrNcWN_RJUkaqjCp?usp=sharing

Code availability

Our code of DespeckleNet is available at https://github.com/hitsh95/DespeckleNet.git.

Declaration of Competing Interest

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

CRediT authorship contribution statement

Shuai Liu: Conceptualization, Methodology, Software, Visualization, Writing – original draft. **Peng Li:** Methodology, Software, Visualization, Writing – review & editing. **Hao Sha:** Software, Visualization, Writing – review & editing. **Jiuyang Dong:** Software. **Yue Huang:** Methodology, Writing – review & editing. **Yanjing Zhao:** Methodology, Writing – review & editing. **Xuri Yao:** Writing – review & editing. **Qin Peng:** Supervision. **Xiu Li:** Supervision. **Xing Lin:** Writing – review & editing, Supervision. **Yongbing Zhang:** Writing – review & editing, Supervision, Funding acquisition.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.optlaseng.2022.107196.

References

[1] Tatarski VI. Wave propagation in a turbulent medium. Courier Dover Publications; 2016.
[2] Ishimaru A. Wave propagation and scattering in random media. New York: Academic press; 1978.
[3] Holodovsky V, Schechner YY, Levin A, Levis A, Aides A. In-Situ multi-view multi--scattering stochastic tomography. In: IEEE International Conference on Computational Photography (ICCP). Evanston, IL: Ieee; 2016. p. 22–33. NW Univ.
[4] Sheinin M, Schechner YY. The next best underwater view. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA: Ieee; 2016. p. 3764–73. doi:10.1109/cvpr.2016.409.
[5] Gibson AP, Hebden JC, Arridge SR. Recent advances in diffuse optical imaging. Phys Med Biol 2005;50:R1–R43. doi:10.1088/0031-9155/50/4/r01.
[6] Goodman JW. Statistical optics. John Wiley & Sons; 2015.
[7] Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, Hee MR, Flotte T, Gregory K, Puliafito CA, Fujimoto JG. Optical coherence tomography. Science 1991;254:1178–81. doi:10.1126/science.1957169.
[8] Yaqoob Z, Psaltis D, Feld MS, Yang CH. Optical phase conjugation for turbidity suppression in biological samples. Nat Photonics 2008;2:110–15. doi:10.1038/nphoton.2007.297.
[9] Vellekoop IM, Lagendijk A, Mosk AP. Exploiting disorder for perfect focusing. Nat Photonics 2010;4:320–2. doi:10.1038/nphoton.2010.3.
[10] Katz O, Small E, Guan YF, Silberberg Y. Noninvasive nonlinear focusing and imaging through strongly scattering turbid layers. Optica 2014;1:170–4. doi:10.1364/optica.1.000170.
[11] Horstmeyer R, Ruan H, Yang C. Guidestar-assisted wavefront-shaping methods for focusing light into biological tissue. Nat Photonics 2015;9:563–71. doi:10.1038/nphoton.2015.140.
[12] Kim M, Choi W, Choi Y, Yoon C, Choi W. Transmission matrix of a scattering medium and its applications in biophotonics. Opt Express 2015;23:12648–68. doi:10.1364/OE.23.012648.
[13] Popoff S, Lerosey G, Fink M, Boccara AC, Gigan S. Image transmission through an opaque material. Nat Commun 2010;1:81. doi:10.1038/ncomms1078.
[14] Liutkus A, Martina D, Popoff S, Chardon G, Katz O, Lerosey G, Gigan S, Daudet L, Carron I. Imaging with nature: compressive imaging using a multiply scattering medium. Sci Rep 2014;4. doi:10.1038/srep05552.
[15] Dremeau A, Liutkus A, Martina D, Katz O, Schulke C, Krzakala F, Gigan S, Daudet L. Reference-less measurement of the transmission matrix of a highly scattering material using a DMD and phase retrieval techniques. Opt Express 2015;23:11898–911. doi:10.1364/oe.23.011898.
[16] Nakamura T, Horisaki R, Tanida J. Compact wide-field-of-view imager with a designed disordered medium. Opt Rev 2015;22:19–24. doi:10.1007/s10043-015-0060-2.
[17] Bertolotti J, van Putten EG, Blum C, Lagendijk A, Vos WL, Mosk AP. Non-invasive imaging through opaque scattering layers. Nature 2012;491:232–4. doi:10.1038/nature11578.
[18] Katz O, Heidmann P, Fink M, Gigan S. Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations. Nat Photonics 2014;8:784–90. doi:10.1038/nphoton.2014.189.
[19] Horisaki R, Takagi R, Tanida J. Learning-based imaging through scattering media. Opt Express 2016;24:13738–43. doi:10.1364/OE.24.013738.
[20] Lyu M, Wang H, Li G, Situ G. Exploit imaging through opaque wall via deep learning; 2017. arXiv preprint arXiv:1708.07881.
[21] Horisaki R, Takagi R, Tanida J. Learning-based focusing through scattering media. Appl Opt 2017;56:4358–62. doi:10.1364/ao.56.004358.
[22] Li J, Mengu D, Yardimci NT, Luo Y, Li X, Veli M, Rivenson Y, Jarrahi M, Ozcan A. Spectrally encoded single-pixel machine vision using diffractive networks. Sci Adv 2021;7:eabd7690. doi:10.1126/sciadv.abd7690.
[23] Li Y, Xue Y, Tian L. Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media. Optica 2018;5. doi:10.1364/optica.5.001181.
[24] Li S, Deng M, Lee J, Sinha A, Barbastathis G. Imaging through glass diffusers using densely connected convolutional networks. Optica 2018;5. doi:10.1364/optica.5.000803.
[25] Sun YW, Shi JH, Sun L, Fan JP, Zeng GH. Image reconstruction through dynamic scattering media based on deep learning. Opt Express 2019;27:16032–46. doi:10.1364/oe.27.016032.
[26] Satat G, Tancik M, Gupta O, Heshmat B, Raskar R. Object classification through scattering media with deep learning on time resolved measurement. Opt Express 2017;25:17466–79. doi:10.1364/OE.25.017466.
[27] Rahmani B, Loterie D, Konstantinou G, Psaltis D, Moser C. Multimode optical fiber transmission with a deep learning network. Light Sci Appl 2018;7. doi:10.1038/s41377-018-0074-1.
[28] Tegin U, Rahmani B, Kakkava E, Borhani N, Moser C, Psaltis D. Controlling spatiotemporal nonlinearities in multimode fibers with deep neural networks. Apl Photonics 2020;5. doi:10.1063/1.5138131.
[29] Peng T, Li RZ, Min JW, Dan D, Zhou ML, Yu XH, Zhang CM, Bai C, Yao BL. Quantitative phase retrieval through scattering medium via compressive sensing. IEEE Photonics J 2022;14. doi:10.1109/jphot.2021.3136509.
[30] Diaspro A. Optical fluorescence microscopy: From the spectral to the nano dimension. Springer Science & Business Media; 2010.
[31] Campagnola PJ, Wei MD, Lewis A, Loew LM. High-resolution nonlinear optical imaging of live cells by second harmonic generation. Biophys J 1999;77:3341–9. doi:10.1016/s0006-3495(99)77165-1.
[32] Dixit R, Cyr R. Cell damage and reactive oxygen species production induced by fluorescence microscopy: effect on mitosis and guidelines for non-invasive fluorescence microscopy. Plant J 2003;36:280–90. doi:10.1046/j.1365-313X.2003.01868.x.
[33] Zernike F. How I Discovered Phase Contrast. Science 1955;121:345–9. doi:10.1126/science.121.3141.345.
[34] Allen RD, David GB, Nomarski G. The zeiss-Nomarski differential interference equipment for transmitted-light microscopy. Zeitschrift fur wissenschaftliche Mikroskopie und mikroskopische Technik 1969;69:193–221.
[35] Paganin D, Nugent KA. Noninterferometric phase imaging with partially coherent light. Phys Rev Lett 1998;80:2586–9. doi:10.1103/PhysRevLett.80.2586.
[36] Zuo C, Li JJ, Sun JS, Fan Y, Zhang JL, Lu LP, Zhang RN, Wang BW, Huang L, Chen Q. Transport of intensity equation: a tutorial. Opt Lasers Eng 2020;135. doi:10.1016/j.optlaseng.2020.106187.
[37] Gabor D. A new microscopic principle. Nature 1948;161:777–8. doi:10.1038/161777a0.

[38] Marquet P, Rappaz B, Magistretti PJ, Cuche E, Emery Y, Colomb T, Depeursinge C. Digital holographic microscopy: a noninvasive contrast imaging technique allowing quantitative visualization of living cells with subwavelength axial accuracy. Opt Lett 2005;30:468–70. doi:10.1364/ol.30.000468.
[39] Choi W, Fang-Yen C, Badizadegan K, Oh S, Lue N, Dasari RR, Feld MS. Tomographic phase microscopy. Nat Methods 2007;4:717–19. doi:10.1038/nmeth1078.
[40] Cotte Y, Toy F, Jourdain P, Pavillon N, Boss D, Magistretti P, Marquet P, Depeursinge C. Marker-free phase nanoscopy. Nat Photonics 2013;7:113–17. doi:10.1038/nphoton.2012.329.
[41] Charriere F, Marian A, Montfort F, Kuehn J, Colomb T, Cuche E, Marquet P, Depeursinge C. Cell refractive index tomography by digital holographic microscopy. Opt Lett 2006;31:178–80. doi:10.1364/ol.31.000178.
[42] Cuche E, Marquet P, Depeursinge C. Spatial filtering for zero-order and twin-image elimination in digital off-axis holography. Appl Opt 2000;39:4070–5. doi:10.1364/ao.39.004070.
[43] Zhang WH, Cao LC, Brady DJ, Zhang H, Cang J, Zhang H, Jin GF. Twin-image-free holography: a compressive sensing approach. Phys Rev Lett 2018;121. doi:10.1103/PhysRevLett.121.093902.
[44] Liu YW, Jiang ZQ, Wang YB, Sun QY, Chen H. Single-frame reconstruction for improvement of off-axis digital holographic imaging based on image interpolation. Opt Lett 2020;45:6623–6. doi:10.1364/ol.405578.
[45] He SY, Pan XC, Liu C, Zhu JQ. Further improvements to iterative off-axis digital holography. Opt Express 2021;29:18831–44. doi:10.1364/oe.425150.
[46] Teague MR. Deterministic phase retrieval: a Green's function solution. J Opt Soc Amer B Opt Phys 1983;73:1434–41. doi:10.1364/josa.73.001434.
[47] Zuo C, Chen Q, Qu WJ, Asundi A. High-speed transport-of-intensity phase microscopy with an electrically tunable lens. Opt Express 2013;21:24060–75. doi:10.1364/oe.21.024060.
[48] Ampem-Lassen E, Huntington ST, Dragomir NM, Nugent KA, Roberts A. Refractive index profiling of axially symmetric optical fibers: a new technique. Opt Express 2005;13:3277–82. doi:10.1364/opex.13.003277.
[49] Rivenson Y, Zhang YB, Gnaydin H, Teng D, Ozcan A. Phase recovery and holographic image reconstruction using deep learning in neural networks. Light Sci Appl 2018;7. doi:10.1038/lsa.2017.141.
[50] Xue YJ, Cheng SY, Li YZ, Tian L. Reliable deep-learning-based phase imaging with uncertainty quantification. Optica 2019;6:618–29. doi:10.1364/optica.6.000618.
[51] Goy A, Arthur K, Li S, Barbastathis G. Low photon count phase retrieval using deep learning. Phys Rev Lett 2018;121. doi:10.1103/PhysRevLett.121.243902.
[52] Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. Generative adversarial networks. Commun ACM 2020;63:139–44. doi:10.1145/3422622.
[53] Zhu L, Chen YS, Ghamisi P, Benediktsson JA. Generative adversarial networks for hyperspectral image classification. IEEE Trans Geosci Remote Sens 2018;56:5046–63. doi:10.1109/tgrs.2018.2805286.
[54] Wang HD, Rivenson Y, Jin YY, Wei ZS, Gao R, Gunaydin H, Bentolila LA, Kural C, Ozcan A. Deep learning enables cross-modality super-resolution in fluorescence microscopy. Nat Methods 2019;16:103 -+. doi:10.1038/s41592-018-0239-0.
[55] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. Springer; 2015. p. 234–41.
[56] Trabelsi C, Bilaniuk O, Zhang Y, Serdyuk D, Subramanian S, Santos JF, Mehri S, Rostamzadeh N, Bengio Y, Pal CJ. Deep Complex Networks; 2017.
[57] Zhang ZM, Wang HP, Xu F, Jin YQ. Complex-valued convolutional neural network and its application in polarimetric SAR image classification. IEEE Trans Geosci Remote Sens 2017;55:7177–88. doi:10.1109/tgrs.2017.2743222.
[58] Isola P, Zhu JY, Zhou TH, Efros AA. Image-to-image translation with conditional adversarial networks. In: 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI: Ieee; 2017. p. 5967–76. doi:10.1109/cvpr.2017.632.
[59] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process 2004;13:600–12. doi:10.1109/tip.2003.819861.
[60] Zhang R, Isola P, Efros AA, Shechtman E, Wang O. The unreasonable effectiveness of deep features as a perceptual metric. In: 31st IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, UT: Ieee; 2018. p. 586–95. doi:10.1109/cvpr.2018.00068.
[61] Zhang H, Bian ZC, Jiang SW, Liu J, Song PM, Zheng GA. Field-portable quantitative lensless microscopy based on translated speckle illumination and sub-sampled ptychographic phase retrieval. Opt Lett 2019;44:1976–9. doi:10.1364/ol.44.001976.
[62] Guo K, Zhang ZB, Jiang SW, Liao J, Zhong JG, Eldar YC, Zheng GA. 13-fold resolution gain through turbid layer via translated unknown speckle illumination. Biomed Opt Express 2018;9:260–74. doi:10.1364/boe.9.000260.
[63] Lu LP, Sun JS, Zhang JL, Fan Y, Chen Q, Zuo C. Quantitative phase imaging camera with a weak diffuser. Front Phys 2019;7. doi:10.3389/fphy.2019.00077.
[64] Goodman JW. Speckle phenomena in optics: theory and applications. Roberts and Company Publishers; 2007.
[65] National Cancer Institute GDC Data Portal, https://portal.gdc.cancer.gov.
[66] Gatti A, Magatti D, Ferri F. Three-dimensional coherence of light speckles: theory. Phys Rev A 2008;78. doi:10.1103/PhysRevA.78.063806.