Optica Applicata, Vol. LVI, No. 1, 2026
DOI: 10.37190/oa/213503

Unpaired self-focusing quantitative phase imaging based on cycle-GAN

Huaying Wang[1,*], Guoqing Xu[1], Shuo Wang[2], Qiaofen Zhu[2,3], Yanhui Huang[4]

[1] School of Information Engineering, Zhengzhou Shengda University, Zhengzhou, Henan 451191, China

[2] School of Mathematics & Physics, Hebei University of Engineering, Handan, Hebei 056038, China

[3] Hebei Computational Optical Imaging and Photoelectric Detection Technology Innovation Center, Hebei University of Engineering, Handan, Hebei 056038, China

[4] Division of Personnel and Party Affairs, Handan Vocational College of Science and Technology, Handan, Hebei 056046, China

[*]Corresponding author: pbxsyingzi@126.com

In order to solve the problem that traditional microscopes cannot directly obtain the focus phase map from defocusing micrographs under incoherent illumination, a non-paired self-focusing quantitative phase imaging method based on Cycle-GAN is proposed. It only needs to obtain the bright field image of the cell by using an ordinary optical bright field microscope, and obtain the focused quantitative phase image by using the optimal weight model trained by the Cycle-GAN network. This method can directly reconstruct the phase image from an incoherent intensity map while achieving self-focusing. The quantitative analysis of the network-output phase images and holographically captured intensity images proves that the Cycle-GAN method can realize the self-focusing and quantitative phase image reconstruction of biological samples.

Keywords: phase imaging, deep learning, autofocus, Cycle-GAN.

1. Introduction

Cell tissue is an important part of the organism. By analyzing the cell structure, we can understand the morphology and change process of the cell and study its internal structural characteristics. In order to obtain quantitative phase information of cell tissues, many quantitative phase imaging techniques have been proposed [1-4]. Some examples show that deep learning provides new ideas for the development of quantitative phase microscopy. Some examples show that deep learning becomes a powerful tool for solving inverse problems in various areas of optical imaging, such as mark dyeing [5-6], improving image quality [7-8], quantitative phase imaging [9-10], phase unrolling [11-13], and self-focusing [14-15]. Among them, the sample data set is a key step in the process of high-quality imaging. Most networks need to pair data sets, that is, they only work if they match one by one. Phase-GAN (generative adversarial network) uses unpaired data sets to solve the phase problem in real time [16], while patch-GAN uses unpaired and few images for training. Its special point is that it does not need the real distribution of objects. High-precision reconstructed images can be obtained by using unpaired training and label data [17-18].

In order to obtain the quantitative phase information of self-focusing, it is necessary to use special optical means to convert the intensity information of defocusing into focused phase information, but the existing auto-focusing and phase imaging methods are independent techniques. Currently, no microscope can directly obtain the focusing phase from the defocus intensity map under incoherent illumination. Such cross-mode imaging requires high requirements, and images of the same content that need to be captured in different modes may show little of the same information [19-20]. In this paper, based on Cycle-GAN, unpaired cross-mode self-focusing phase imaging can be carried out. This model converts the image in the source mode to the image in the target mode, and there is a mapping relationship between the intensity and phase information of the sample. Once the intensity image and quantitative phase image of the sample are available as datasets, the deep learning network can learn this mapping relationship. The experimental results show that the proposed method can obtain the corresponding focusing phase information from the intensity image, which proves that Cycle-GAN can realize the reconstruction of focusing and quantitative phase images of biological samples. In a sense, this method combines the advantages of phase imaging and incoherent bright field imaging modes, and combines the advantages of two different fields together to bridge the domain gap between different imaging modes.

2. Experimental principles

2.1. Preparation data set

Quantitative phase imaging (QPI) is an optical imaging technique used to study unlabeled and non-invasive samples [21]. QPI has unique features. On the one hand, some optical configurations can retrieve phase information from a single image acquisition, which facilitates long-term study of living cells. This method is only affected by the exposure speed limit of image sensor and the phase reconstruction algorithm, and has good real -time performance. On the other hand, QPI technology is almost non-phototoxic, does not affect the metabolism of cells, and avoids lengthy sample preparation. The whole process of this technology can be roughly divided into two parts; First, obtain one or more intensity images; Secondly, the acquired images are processed numerically to extract quantitative phase information.

Focusing on these two processes, this paper realizes cross-mode self-focusing phase imaging based on deep learning method. The training data set consists of defo-

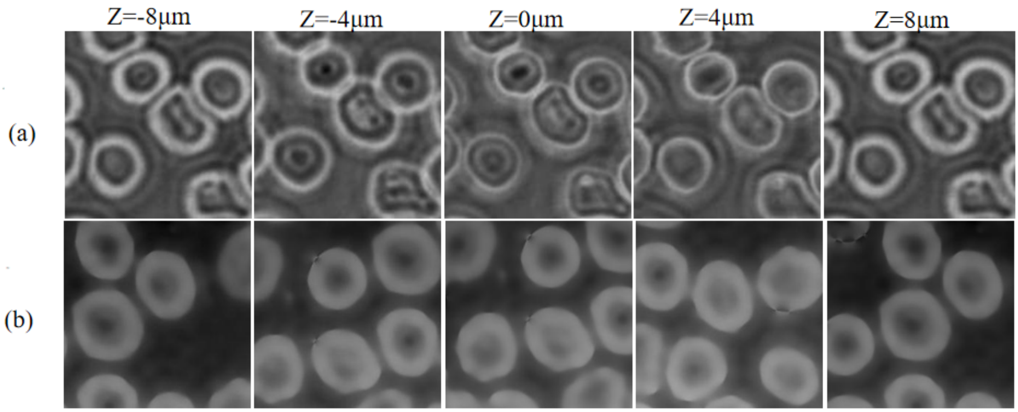

Fig. 1. Part of the training dataset. (a) Defocus micrograph of red blood cells. (b) Phase diagram of red blood cells.

cusing micrographs obtained by white light microscope at different distances [22] and phase graphs reconstructed by hologram obtained by Mach–Zehnder interference system. An important feature of Cycle-GAN is that the source domain image and the target domain image do not need to match one by one to achieve cross-mode. Data set preparation is mainly divided into three parts: First, the acquisition of intensity maps under incoherent illumination. As shown in Fig. 1(a), by manually adjusting the focal length to 0 μm, ±4 μm and ±8 μm, 500 microscopic images at different focal lengths were obtained. Second, hologram acquisition under coherent illumination. The main steps of hologram recording and reconstruction are as follows: digital hologram is obtained by Mach–Zehnder interference system, hologram is reconstructed by angle spectrum reconstruction algorithm [23], and phase distribution of red blood cells is obtained, as shown in Fig. 1(b). Third, the data set is expanded by translation, rotation and other methods. Red blood cells were selected as the sample, and the network input-output image size was set to 256×256, using 5000 intensity plots and phase plots at different defocusing distances as neural network data sets, 90% as the training set and 10% as the test set. Figure 1 shows a partial training data set.

2.2. Network structure

The special feature of Cycle-GAN is that it is composed of two Gans, including two generators and two discriminators [24-25]. Compared with other neural networks, the biggest advantage of Cycle-GAN is that there is no need for one-to-one matching between the input image and the target image in the network, which means that the target image can be any image of the same style. Cycle-GAN greatly reduces the difficulty of data acquisition, and provides the possibility for image transformation that cannot obtain one-to-one correspondence to the target image. The working principle of the Cycle-GAN network is shown in Fig. 2, starting from the samples in the *X* domain,

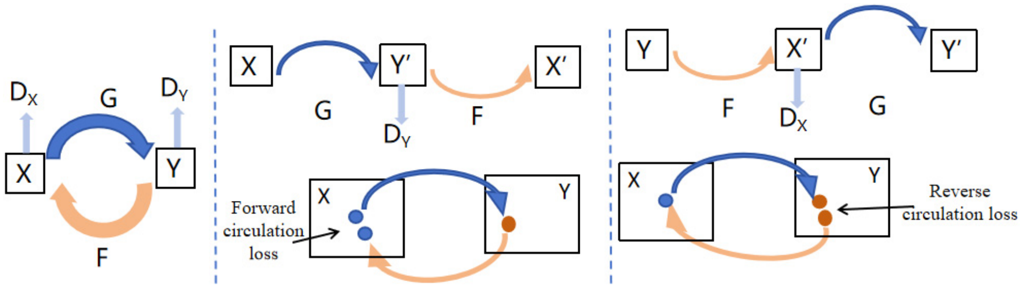

Fig. 2. Schematic of Cycle-GAN [22].

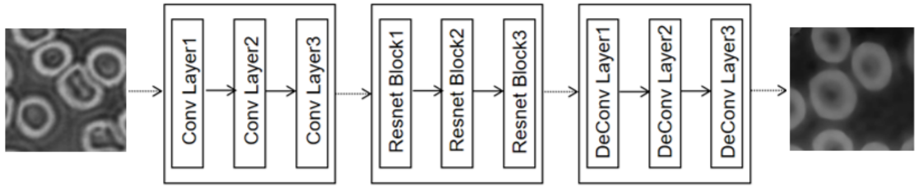

Fig. 3. The network structure of the generator in Cycle-GAN.

they are transformed into pseudo-samples Y' in the Y domain through generator G, and then the pseudo-samples in the Y domain are reconstructed back into the X domain by generator F. Ultimately, by calculating the difference between the original X and the reconstructed X, the forward loop consistency loss is obtained. Similarly, for the reverse loop consistency loss of Y and Y', achieve the dual-domain conversion of $X \rightarrow Y \rightarrow X$.

2.2.1. Generator network architecture

The generator network of Cycle-GAN is mainly composed of encoders, converters and decoders. The specific process is shown in Fig. 3. The image input generator network first extracts the image feature information through the way of subsampling through the encoder; After the feature map enters the converter, the converter is composed of residual blocks, whose function is to delete useless features and retain useful information; Finally, it enters the decoder, which is composed of three deconvolution, and restores the obtained feature information into the corresponding input image by upsampling. The detailed structure inside the generator is shown in Table 1.

2.2.2. Discriminator network architecture

The discriminator of Cycle-GAN uses Patch GAN of 70×70 pixels to determine whether the local Patch is true. Using the method of image cutting to distinguish, the amount of training data can be effectively reduced, and the network parameters can be correspondingly reduced, and the network training becomes easier. The network structure of Patch GAN is shown in Fig. 4, which consists of five convolutional layers. The role

Table 1. A detailed structure of the generator internals in Cycle-GAN.

Module	Detailed description			
Encoder	Convolution kernel size	Connection mode	Number of filters	Step size
	7×7	Conv, BN, ReLu	64	1
	5×5	Conv, BN, ReLu	128	1
	3×3	Conv, BN, ReLu	256	2
Convertor	Module		Rate of increase	
	Resblock1		32	
	…		…	
	Resblock9		32	
Decoder	Convolution kernel size	Connection mode	Number of filters	Step size
	3×3	Deconv, BN, ReLu	128	1/2
	3×3	Deconv, BN, ReLu	64	1/2
	7×7	Deconv, ReLu	3	1

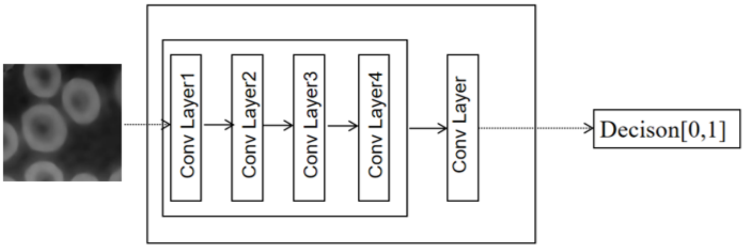

Fig. 4. The network structure of Patch GAN.

of the first four convolutional layers is to segment the image and distinguish each region. The segmentation region is 70×70 pixels. The function of the last convolution layer is to average the results obtained from the first four convolution layers, and this value is used as the discriminant result.

2.3. Experimental process

In this paper, the overall experimental flow of cross-mode auto-focusing phase imaging based on the Cycle-GAN network is shown in Fig. 5. Cycle-GAN is actually an X → Y one-way GAN plus a Y → X one-way GAN. Therefore, the network consists of two generators (F and G) and two discriminators (DX and DY). In Fig. 5, the dark blue dots represent fake samples similar to phase images generated by generator F after learning the original defocusing intensity map → phase image map; The dark red dots represent the fake samples similar to the defocusing intensity map generated by the generator G after learning the phase map → defocusing intensity map; The light blue dots represent the fake samples generated by fake_B after generator F corresponding to the phase images in the target domain Y; The light red dots represent the virtual sam-

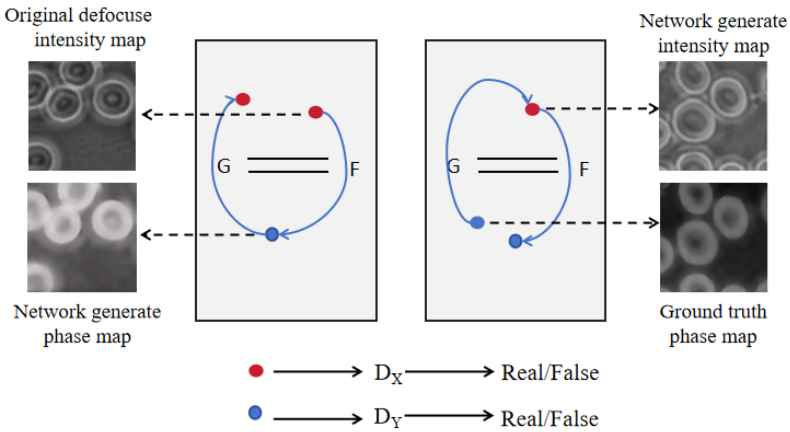

Fig. 5. Structure diagram of Network training overall.

ples generated by the generator G that correspond to the original defocusing intensity image in the source domain X. The discriminator DX and DY are used to judge whether the input defocusing intensity and phase plots are real samples.

In the neural network training process, Adam was used as the optimizer to optimize the weight. Due to the adaptation of video memory and the retention of sample uniqueness, the batch setting is set to 1 to meet the adversarial training requirements of Cycle-GAN. The dynamic balance between the generator and the discriminator is extremely unstable. Selecting a low learning rate can prevent optimization oscillations. It is set to 0.0001, and the loss value tends to stabilize after 40 cycles. Table 2 describes the experimental environment.

Table 2. The experimental environment configuration.

Experimental environment	Environment configuration
Operating system	Ubuntu 18.04
framework	Pytorch 10.2
Computer language	Python 3.7
Graphics card type	GeForce GTX2060
Video memory	8 GB
Running memory	32 GB
Processor	Intel core i7-9700CPU
Processor frequency	3 GHz

3. Experimental results and analysis

3.1. Network test results

Figure 6 shows the test results of different regions at different defocusing distances, where input is the microscopic image with the defocusing distance of 0 μm, ±4 μm and

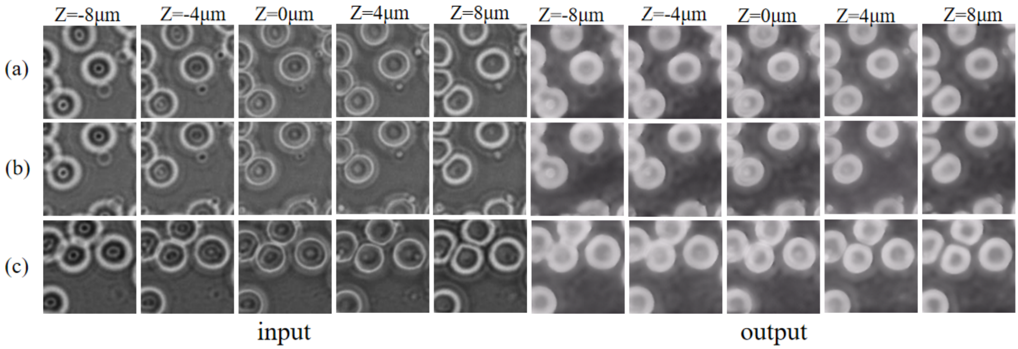

Fig. 6. Test results of different defocus distances. (a)(b)(c)Defocus micrograph of different regions of red blood cells.

±8 μm input from the network, and output is the focusing phase image output from the network.

When the defocusing micrograph of the test set is input into Cycle-GAN, the defocusing distance ranges from –8 μm to 8 μm, the trained neural network model can quickly output the quantitative phase diagram of the sample, achieve cross-modal conversion, and eliminate the interference of background noise to a certain extent. From the perspective of vision, microscopic images at different defocusing distances can obtain their corresponding focusing phase maps through this network. In the case of pixel overlap caused by defocus, although there is a slight error, it also shows good matching performance. Figures 7(a), (b) and (c), respectively, show the 3D phase distribution corresponding to the output results of Fig. 6. From the 3D phase images of blood cells in different regions, it can be seen that the neural network has the ability to learn between different modes. The feasibility of this method can be preliminarily proved.

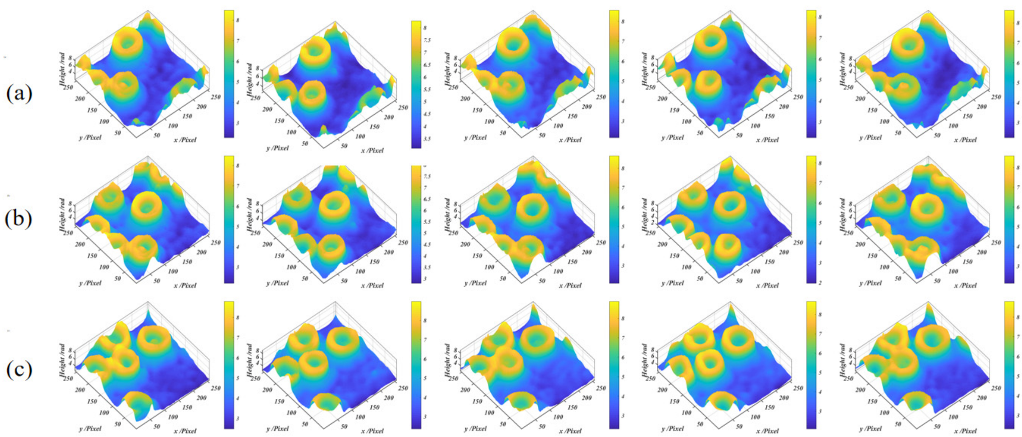

Fig. 7. Phase distribution of network output.

3.2. Hologram strength verification

The unpaired phase recovery method based on Cycle-GAN proposed in this paper is cross-modal image conversion, so the SSIM (structural similarity index measure) value and PSNR (peak signal-to-noise ratio) value between the network output value and the real value cannot be evaluated quantitatively. Further, we obtained the matched data set through the digital holographic optical path, used Cycle-GAN to train the unmatched data set, and then quantitatively measured the evaluation index between the network output value and the real value to prove the effectiveness of the method.

The verification experiment still adopts red blood cells as the data set. The specific process is as follows: The experimental system adopts Mach–Zehnder off-axis digital holographic optical path, the light source adopts He-Ne laser with wavelength of 633 nm, and the light source is divided into reference light and object light through the beam splitting prism. The hologram formed by the interference of reference light and object light is received by CCD, and the off-axis digital hologram is obtained. The reference light is blocked, and the defocusing intensity image corresponding to the object is obtained by moving CCD with 3D translation platform, which is used as the input data set of neural network. The hologram is reconstructed by Angle spectrum method [23], and the real phase distribution of the sample is obtained, which is used as the control group of the neural network. The experimental training data are shown in Fig. 8: (a) the erythrocyte strength diagram, and (b) the corresponding phase diagram. When preparing the data set, label images that are not in the corresponding order can be prepared, which increases the fault tolerance rate in the operation process. After the training of the neural network is completed, the random input intensity map can quickly and accurately give the real phase distribution of the phase object.

The image size is set to 256×256. 1000 intensity plots and phase plots at different defocusing distances were used as the neural network data set, 90% as the training set and 10% as the test set.

When the intensity diagram of the test set is input into the trained Cycle-GAN network, the network can quickly output the quantitative phase diagram of the sample, and

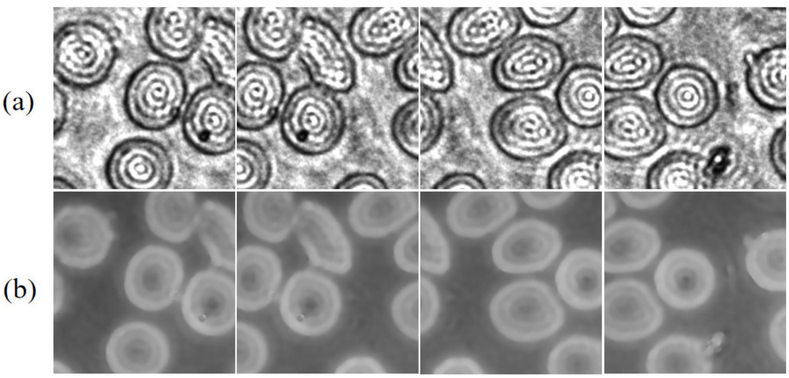

Fig. 8. Part of the training dataset. (a) Intensity images; (b) phase images.

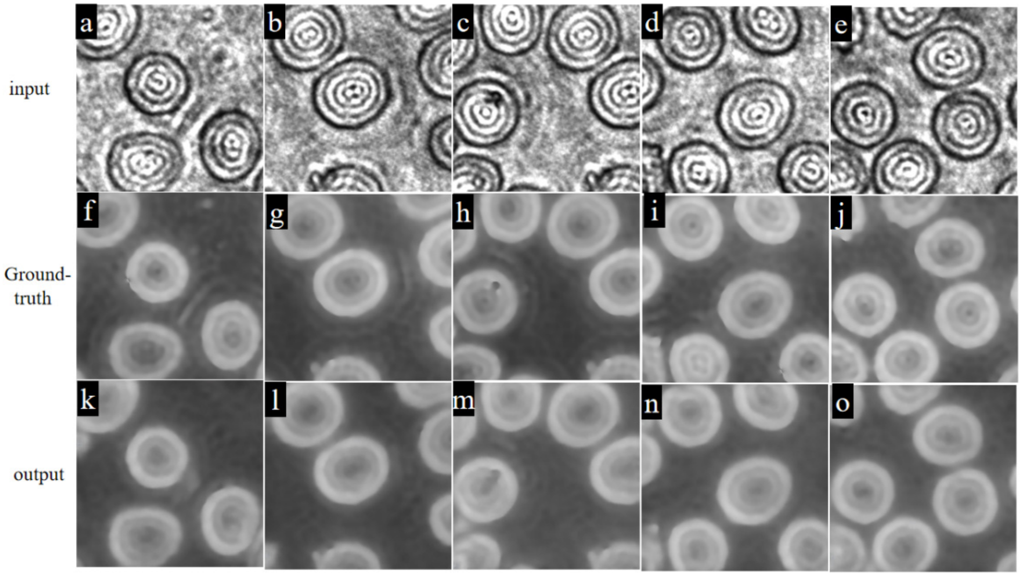

Fig. 9. Test results of different defocus distances. (a–e) Intensity images for testing, (f–j) ground truth, and (k–o) network output.

at the same time, the interference of background noise can be suppressed to a certain extent. Figure 9 shows part of the test results of the intensity map. From the two-dimensional plane, it can be seen that the output result of the neural network can effectively restore the phase information of red blood cells, and the output result is very close to the real value, with the average SSIM value above 0.93 and the average PSNR value above 23.

Here, the SSIM value is used to calculate the similarity index between the reconstructed phase image and the real phase image. The formula is:

$$\mathrm{SSIM} = \frac{(2\boldsymbol{\mu}_x\boldsymbol{\mu}_y + c_1)(2\boldsymbol{\sigma} + c_2)}{(\boldsymbol{\mu}_x^2 + \boldsymbol{\mu}_y^2 + c_1)(\boldsymbol{\sigma}_x^2 + \boldsymbol{\sigma}_y^2 + c_2)} \tag{1}$$

where μ_x and μ_y are respectively the average values of x and y, $\boldsymbol{\sigma}_x^2$ and $\boldsymbol{\sigma}_y^2$ are respectively the variances of x and y, σ is the covariance of y, c_1 and c_2 are constants that maintain stability. The larger the SSIM value, the better the quality of the generated image.

PSNR is the ratio of the target peak power to the background noise power, used to measure the difference between two images. The formula is:

$$\mathrm{PSNR} = 10\,\mathrm{lg}_{10}\left(\frac{(2^n - 1)^2}{\mathrm{MSE}}\right) \tag{2}$$

$$\mathrm{MSE} = \frac{1}{mn}\sum_{m}\sum_{n}\left[U_0(i,j) - U(i,j)\right]^2 \tag{3}$$

U represents the original image, U_0 represents the generated image, and $m \times n$ is the resolution of the image. The larger the PSNR value, the better the image quality.

To verify the advantages of the method proposed in this paper over other methods in reconstruction, the comparison with the TIE traditional method and the U-Net paired dataset is shown in Fig. 10. It can be seen that tilt compensation is not carried out in the TIE method. The reconstruction result of the U-Net method has too much background noise and severe speckle noise. The method proposed in this paper can effectively perform error compensation and be clearly distinguished from the background interference environment. This experimental result further proves the effectiveness and advantages of the phase recovery method based on unpaired datasets. This method is simple to operate, does not require background images, and has a certain attenuation effect on noise.

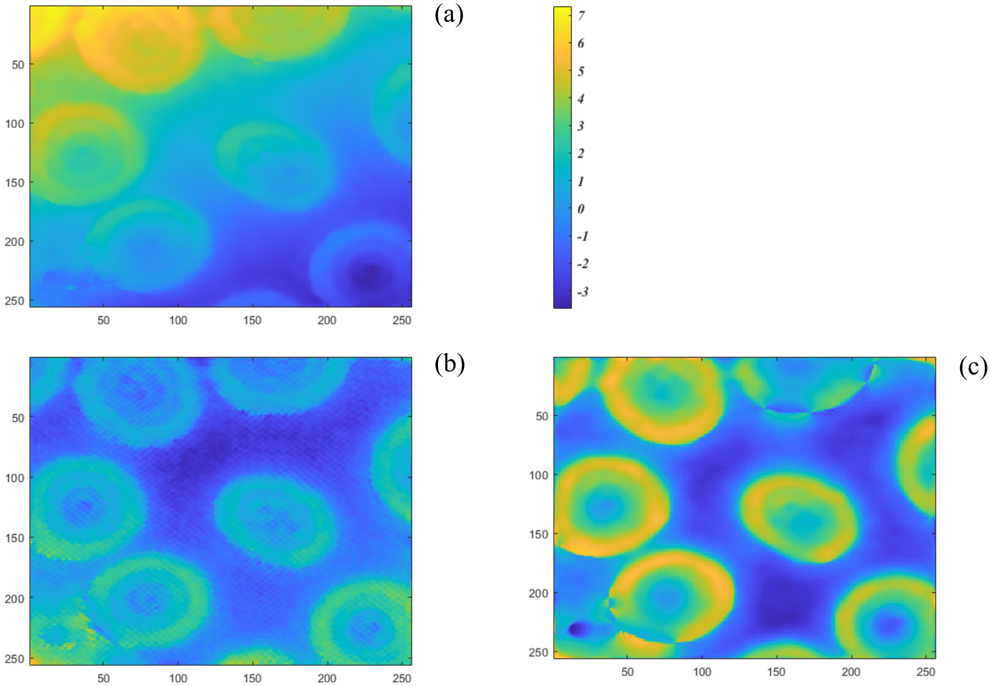

Fig. 10. Comparison of experimental results. (a) TIE method, (b) U-Net paired dataset method, and (c) Cycle-GAN.

4. Conclusion

This paper proposes a cross-modal self-focusing quantitative phase imaging method based on deep learning technology, which can carry out unsupervised learning and complete training by using mismatched data sets. Experimental results show that the trained neural network can quickly output quantitative phase graphs of samples. The focus phase diagram of the microimages at different defocusing distances can be obtained by neural network, which realizes cross-mode conversion and eliminates the interference of background noise to some extent. This method combines the advantages of digital holography and incoherent bright field imaging, and further provides the possibility for quantitative phase imaging of living cells under incoherent illumination.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No.62175059).

Conflict of 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.

References

[1] TODA K., TAMAMITSU M., NAGASHIMA Y., HORISAKI R., IDEGUCHI T., *Molecular contrast on phase-contrast microscope*, Scientific Reports, **9**(1), 2019: 9957. https://doi.org/10.1038/s41598-019-46383-6

[2] CHENG H., ZHANG H.Y., LU W., ZHANG Q.B., HU Z., *An enhanced multimode phase imaging method based on the transport of intensity equation*, Journal of Biophotonics, **17**(8), 2024: e202400137. https://doi.org/10.1002/jbio.202400137

[3] BIAN Y.X., JIANG Y.N., HUANG Y.R., YANG X.F., DENG W.J., SHEN H., SHEN R.B., KUANG C.F., *Smartphone phase contrast microscope with a singlet lens and deep learning*, Optics & Laser Technology, **139**, 2021: 106900. https://doi.org/10.1016/j.optlastec.2020.106900

[4] FU S., *Research on phase recognition and analysis method based on quantitative phase imaging and deep learning*[D]. Jiangsu: Jiangsu University, 2020.

[5] RIVENSON Y., LIU T.R., WEI Z.S., ZHANG Y.B., DE HAAN K., OZCAN A., *PhaseStain: The digital staining of label-free quantitative phase microscopy images using deep learning*, Light: Science & Applications, **8**, 2019: 23. https://doi.org/10.1038/s41377-019-0129-y

[6] CHEN Q.D., LIU X., DU L.L., SONG B., SUN X.B., SHEN Z.Y., *NIR-to-RGB image colorization based on a conditional GAN*, Applied Optics, **64**(11), 2024: 2968-2978. https://doi.org/10.1364/AO.557734

[7] LIU Z.F., YANG Y.Z., FANG Y., *et al.*, *Deep learning-based virtual phase contrast imaging method*, Acta Optica, 2021(022): 041.

[8] BABA M., BLINDER D., WANG F., BIRNBAUM T., SCHELKENS P., ITO T., SHIMOBABA T., *Improving image quality of a phase-added stereogram using deep learning*, Applied Optics, **63**(28), 2024: G1-G8. https://doi.org/10.1364/AO.521087

[9] DARDIKMAN-YOFFE G., ROITSHTAIN D., MIRSKY S.K., TURKO N.A., HABAZA M., SHAKED N.T., *PhUn-Net: Ready-to-use neural network for unwrapping quantitative phase images of biological cells*, Biomedical Optics Express, **11**(2), 2020: 1107-1121. https://doi.org/10.1364/BOE.379533

[10] XU X.Q., XIE M., JI Y., WANG Y.W., *Rapid quantitative phase imaging using deep learning for phase object with refractive index variation*, Journal of Modern Optics, **68**(6), 2021: 327-338. https://doi.org/10.1080/09500340.2021.1896815

[11] ZHANG T., JIANG S.W., ZHAO Z.X., DIXIT K., ZHOU X.F., HOU J., ZHANG Y.B., YAN C.G., *Rapid and robust two-dimensional phase unwrapping via deep learning*, Optics Express, **27**(16), 2019: 23173-23185. https://doi.org/10.1364/OE.27.023173

[12] QIAN J.M., FENG S.J., TAO T.Y., HU Y., LI Y.X., CHEN Q., ZUO C., *Deep-learning-enabled geometric constraints and phase unwrapping for single-shot absolute 3D shape measurement*, APL Photonics, **5**(4), 2020: 046105. https://doi.org/10.1063/5.0003217

[13] WANG K.Q., LI Y., KEMAO Q., DI J.L., ZHAO J.L., *One-step robust deep learning phase unwrapping*, Optics Express, **27**(10), 2019: 15100-15115. https://doi.org/10.1364/OE.27.015100

[14] LUO Y., HUANG L., RIVENSON Y., OZCAN A., *Single-shot autofocusing of microscopy images using deep learning*, ACS Photonics, **8**(2), 2021: 625-638. https://doi.org/10.1021/acsphotonics.0c01774

[15] YANG L., WANG H.Y., DONG Z., *et al.*, *Application of autofocus phase contrast technique based on improved U-Net in cell imaging*, Chinese Laser, **49**(15), 2022: 133-142.

[16] ZHANG Y., NOACK M.A., VAGOVIC P., FEZZAA K., GARCIA-MORENO F., RITSCHEL T., VILLANUEVA -PEREZ P., *PhaseGAN: A deep-learning phase-retrieval approach for unpaired datasets*, Optics Express, **29**(13), 2021: 19593-19604. https://doi.org/10.1364/OE.423222

[17] TRIEU Q., NEHMETALLAH G., *Deep learning based coherence holography reconstruction of 3D objects*, Applied Optics, **63**(7), 2024: B1-B15. https://doi.org/10.1364/AO.503034

[18] YIN D., GU Z., ZHANG Y., GU F., NIE S., MA J., YUAN C., *Digital holographic reconstruction based on deep learning framework with unpaired data*, IEEE Photonics Journal, **12**(2), 2020: 3900312. https://doi.org/10.1109/JPHOT.2019.2961137

[19] RIVENSON Y., ZHANG Y.B., GÜNAYDIN H., TENG D., OZCAN A., *Phase recovery and holographic image reconstruction using deep learning in neural networks*, Light: Science & Applications, **7**, 2018: 17141. https://doi.org/10.1038/lsa.2017.141

[20] CACACE T., BIANCO V., FERRARO P., *Quantitative phase imaging trends in biomedical applications*, Optics and Lasers in Engineering, **135**, 2020: 106188. https://doi.org/10.1016/j.optlaseng.2020.106188

[21] HE X., *Study on imaging technique of bright field, dark field and phase contrast multimode microscopy*[D], Chengdu: University of Electronic Science and Technology of China, 2017.

[22] ZHU J.Y., PARK T., ISOLA P., *et al.*, *Unpaired image-to-image translation using cycle-consistent adversarial networks*, 2017 IEEE International Conference on Computer Vision (ICCV), October 22-29, 2017, Venice, Italy. New York: IEEE Press, 2017: 2242-2251.

[23] MATSUSHIMA K., SHIMOBABA T., *Band-limited angular spectrum method for numerical simulation of free-space propagation in far and near fields*, Optics Express, **17**(22), 2009: 19662-19673. https://doi.org/10.1364/OE.17.019662

[24] REN Y., NIE M.Z., LI S.C., LI C.K., *Single image de-raining via improved generative adversarial nets*, Sensors, **20**(6), 2020: 1591. https://doi.org/10.3390/s20061591

[25] GOODFELLOW I., POUGET-ABADIE J., MIRZA M., XU B., WARDE-FARLEY D., OZAIR S., COURVILLE A., BENGIO Y., *Generative adversarial networks*, Communications of the ACM, **63**(11), 2020: 139-144.

Received June 26, 2025
in revise form September 5, 2025