Neural network model combined with pupil recovery for Fourier ptychographic microscopy

MINGLU SUN,[1,2] **XIONG CHEN,**[1,2] **YOUQIANG ZHU,**[1,2] **DAYU LI,**[1,3] **QUANQUAN MU,**[1,2,4] **AND LI XUAN**[1,2]

[1]*State Key Laboratory of Applied Optics, Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China*
[2]*Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China*
[3]*lidayu@ciomp.ac.cn*
[4]*muquanquan@ciomp.ac.cn*

Abstract: Fourier ptychographic microscopy (FPM) is a recently developed imaging approach aiming at circumventing the limitation of the space-bandwidth product (SBP) and acquiring a complex image with both wide field and high resolution. So far, in many algorithms that have been proposed to solve the FPM reconstruction problem, the pupil function is set to be a fixed value such as the coherent transfer function (CTF) of the system. However, the pupil aberration of the optical components in an FPM imaging system can significantly degrade the quality of the reconstruction results. In this paper, we build a trainable network (FINN-P) which combines the pupil recovery with the forward imaging process of FPM based on TensorFlow. Both the spectrum of the sample and pupil function are treated as the two-dimensional (2D) learnable weights of layers. Therefore, the complex object information and pupil function can be obtained simultaneously by minimizing the loss function in the training process. Simulated datasets are used to verify the effectiveness of pupil recovery, and experiments on the open source measured dataset demonstrate that our method can achieve better reconstruction results even in the presence of a large aberration. In addition, the recovered pupil function can be used as a good estimate before further analysis of the system optical transmission capability.

1. Introduction

Recently, a data-driven image reconstruction technology, called Fourier Ptychographic microscopy (FPM), has been proposed [1–3] which provides a coherent imaging approach for achieving a wide-field and high-resolution result by circumventing optical space-bandwidth product (SBP). This method integrates the theories of phase retrieval [4] and aperture synthesis [5] to recover both the intensity and phase information. The FPM approach is derived from a similar imaging technique called ptychography, which is a lensless imaging technique [6–8]. Both the techniques could acquire a high-resolution complex object image based on a series of low-resolution images. However, unlike ptychography, FPM adopts a microscopic system which can transform the spectral information received in ptychography into the spatial information [1]. In addition, FPM replaces the fixed light source with an LED array. As a result, the sample could be illuminated by plane waves from multiple angles. In this way, the higher-frequency information that would normally surpass the system bandwidth could be received and detected by the image sensor.

Since FPM was first proposed in 2013, the system structure and reconstruction algorithm have been modified several times to further improve the performances of the technology. In order to minimize the negative impact of background noise during the reconstruction process, several methods have been proposed [9–13]. To increase the reconstruction speed and robustness of FPM, some optimization methods have been developed such as nonlinear optimization algorithm [14], Gauss-Newton method [15,16], Wirtinger flow optimization [17], convex relaxation method

[18]. By selectively updating the most informative sub-regions [19,20], the reconstruction speed of FPM could be further improved. In addition, the algorithms for diminishing the influences of system aberrations, like the positional misalignment of the LED array [21–23] and the aberrations of the optical lens [24–26] have been proposed. Meanwhile, the concept of FPM was also adopted in lensless imaging [27] and by combining with the multi-layer modeling approach, Tian *et al*. developed a multi-slice FPM which can obtain the three-dimensional (3D) information of the sample [28].

In addition to the approaches mentioned above, recently an increasing number of algorithms based on deep convolutional neural network (DCNN) have been published to solve image reconstruction problems such as single image super-resolution [29–33] and phase retrieval [34,35]. The traditional DCNN is a machine learning technique that learns the mathematical mapping between the input and output. A DCNN model usually needs to be trained with the existing training dataset and the loss function is calculated by the residual between the output of the model and the label images in the training dataset. By using the gradient descent method to minimize the loss function, the model could be updated and learn the mapping process more correctly. Due to the excellent ability of DCNN to learn the non-linear relationship between the input and output, these works produce great results. Because the purpose of FPM is to solve the nonlinear ill-posed inverse problem of synthesizing a high-resolution complex image from low-resolution images [1,4,14], it is naturally to introduce the idea of DCNN into FPM by learning the mapping process from several low-resolution images to a high-resolution target [35–39]. However, there still exist some issues in these methods. First, this end-to-end DCNN framework relies on a large dataset to learn the rules of the underlying inverse process while lacks the constrains that could represent the actual physical process [38]. In this case, it is usually necessary to train the dataset for hundreds of times to acquire acceptable results [37,38], which makes the training process of the network very time-consuming. Second, after the training is completed, the network can only work properly for a specialized FPM system. If any system parameters change, the network needs to be retrained to match the new parameters, which makes the reconstruction algorithm lack of universality and stability. Jiang *et al*. recently proposed a novel FPM algorithm which utilizes a deep learning library called TensorFlow [40]. They first establish a forward imaging model of FPM and then use back-propagation to train the model. Because this method introduces the physical process of the actual FPM system and is still essentially an iterative-based algorithm, it does not have the same issues as the DCNN-based methods do. However, none of these methods takes into account the influences of the system pupil aberrations.

In this paper, we propose a novel FPM reconstruction model termed forward imaging neural network with pupil recovery (FINN-P), which can simultaneously reconstruct an aberration-free high-resolution complex image of the sample and the pupil function of the system. Similar to Jiang's method, we build a TensorFlow-based trainable network to solve the FPM problem, and the high-resolution results can be obtained by minimizing the loss between the network's output and measured low-resolution images. However, due to the more effective model workflow and optimization methods employed, our algorithm can achieve better reconstruction results and the imaging wavefront aberrations could be estimated through the phase of the recovered pupil function [24].

This paper is structured as follows: in Section 2, we present the overall structure of the FPM system and describe the forward imaging model mathematically. in Section 3, we explain our reconstruction network and verify the effectiveness of pupil recovery through simulation. in Section 4, we demonstrate that our method can improve the reconstruction quality and compare the performances with several other reconstruction algorithms such as GS [1], AS [9], and Jiang's method [40]. Finally, Section 6 describes the summaries and discussions as well as our prospect for future works.

2. The principle of FPM

Before we introduce the details of our algorithm, it is worthwhile to briefly review the structure of the FPM system and the imaging process. A traditional FPM system comprises five parts as shown in Fig. 1. A thin specimen is located far enough away from an LED array so that the incident waves which irradiate the specimen from different angles can be approximately considered as plane waves. An optical 4f system consisting of an objective lens and a tube lens is placed behind the specimen. In order to obtain the spectra of the sample at the back focal plane of the objective lens, the specimen need to be placed at the front focal plane of the objective lens. Eventually, an image sensor placed behind the tube lens is utilized to detect the light waves propagating through the system.
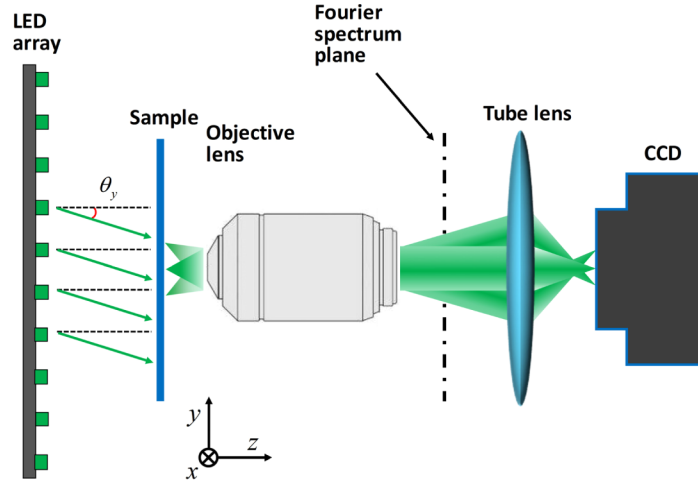

Fig. 1. Schematic of a typical Fourier ptychographic microscopy.

Since the image sensor (CCD) can only capture the intensity of the outgoing waves, for the mth LED, the intensity image recorded by the sensor can be formulated as:

$$I_m(\boldsymbol{r}) = |\mathcal{F}^{-1}\{\mathcal{F}\{o(\boldsymbol{r}) \cdot exp(i2\pi \cdot \boldsymbol{u}_m \boldsymbol{r})\} \cdot P(\boldsymbol{u})\}|^2, \tag{1}$$

where $\boldsymbol{r} = (x, y)$ represents the 2D coordination in the spatial domain, "·" symbol donates element-wise multiplication and $\boldsymbol{u} = (f_x, f_y)$ is the 2D coordination in the Fourier domain, $o(\boldsymbol{r})$ refers to the transmission function of the sample, $\mathcal{F}$ and $\mathcal{F}^{-1}$ represent the Fourier transform and inverse Fourier transform respectively, and $\boldsymbol{u}_m$ represents the wave vector shift corresponding to the illumination angle of the mth LED, which can be expressed as:

$$\boldsymbol{u}_m = (\frac{sin\theta_x^{(m)}}{\lambda}, \frac{sin\theta_y^{(m)}}{\lambda}), \tag{2}$$

where $(\theta_x^{(m)}, \theta_y^{(m)})$ represent the incident angles for the mth LED, and λ represents the illumination wavelength in the air. Therefore, the imaging formation process can be rewritten as:

$$I_m(\boldsymbol{r}) = |\mathcal{F}^{-1}\{O(\boldsymbol{u} - \boldsymbol{u}_m) \cdot P(\boldsymbol{u})\}|^2, \tag{3}$$

where $O(\boldsymbol{u} - \boldsymbol{u}_m)$ represents the sample spectrum corresponding to the mth LED, and $P(\boldsymbol{u})$ represents the pupil function of the imaging system, which is considered as the coherent transfer

function (CTF) of the system in many traditional algorithms and can be expressed as:

$$P(f_x, f_y) = \text{CTF} = \begin{cases} 1, & \text{if } (f_x^2 + f_y^2) \leq \left(\frac{NA}{\lambda}\right)^2 \\ 0, & \text{otherwise} \end{cases}, \tag{4}$$

where NA is the numerical aperture of the objective lens. According to Eqs. (3) and (4), when light passes through the system, the field is low-pass filtered by the pupil function $P(\pmb{u})$, and when the LEDs which provide different wave vectors $\pmb{u}_m$ are sequentially activated, we can obtain a series of low-resolution images containing information from different sub-regions of the spectrum.

3. FPM reconstruction model with pupil recovery

Based on the imaging model discussed above, we propose a new algorithm called FINN-P by reproducing the forward imaging process through a network similar to Jiang's method [40]. However, Jiang's method uses a previously fixed pupil function, such inaccuracy of the pupil function will degrade the reconstruction quality when the system suffers from severe aberrations. To solve this problem, we set both the Fourier spectrum of the sample and the pupil function as trainable multiplication layers, unlike the usual deep learning method, this algorithm does not require a pre-training process. Instead, we use the characteristic of gradient back-propagation to optimize the layers by alternately training one and keeping the other one fixed during the reconstruction process. The overall workflow of FINN-P is shown in Fig. 2.

First, the algorithm needs to make an initial guess on the sample spectrum $O_e(\pmb{u})$, which is the Fourier transform of an up-sampled low-resolution image under the normal incidence condition, and the pupil function is set as the circular low-pass filter $P(\pmb{u})$ defined in Eq. (4). Second, we obtain the estimated low-resolution complex field $E_{e,m}(\pmb{r})$ from the initial spectrum similar to Eq. 3.

$$E_{e,m}(\pmb{r}) = \mathcal{F}^{-1}\{O_e(\pmb{u} - \pmb{u}_m) \cdot P(\pmb{u})\}. \tag{5}$$

Since our forward imaging network is built based on the TensorFlow library, in which the trainable parameters need to be real, we separate the spectrum of the sample and the pupil function into real and imaginary parts, (O_r, O_j) and (P_r, P_j) respectively, such that:

$$O_e(\pmb{u}) = O_r(\pmb{u}) + iO_j(\pmb{u})$$
$$P(\pmb{u}) = P_r(\pmb{u}) + iP_j(\pmb{u}) \tag{6}$$

In this case, there are four layers in the entire network that need to be trained and by substituting Eq. (6) into Eq. (5), the simulated low-resolution complex field can be expressed as:

$$E_{e,m}(\pmb{r}) = \mathcal{F}^{-1}\{[O_r(\pmb{u} - \pmb{u}_m) \cdot P_r(\pmb{u}) - O_j(\pmb{u} - \pmb{u}_m) \cdot P_j(\pmb{u})]$$
$$+ i[O_r(\pmb{u} - \pmb{u}_m) \cdot P_j(\pmb{u}) + O_j(\pmb{u} - \pmb{u}_m) \cdot P_r(\pmb{u})]\} \tag{7}$$

In order to suppress the negative effects caused by the intensity fluctuation of different LED elements [26], after obtaining $E_{e,m}(\pmb{r})$, an intensity correction process is imposed using the measured intensity $\hat{I}_m(\pmb{r})$.

$$\tilde{E}_{e,m}(\pmb{r}) = \frac{\sum_{pix} \sqrt{\hat{I}_m(\pmb{r})}}{\sum_{pix} E_{e,m}(\pmb{r})} \cdot E_{e,m}(\pmb{r}), \tag{8}$$

where the operator $\sum_{pix}$ means the summation of all the pixel values in the matrix. Moreover, the loss function is set as L_2-norm so that we still can obtain good reconstruction results under a
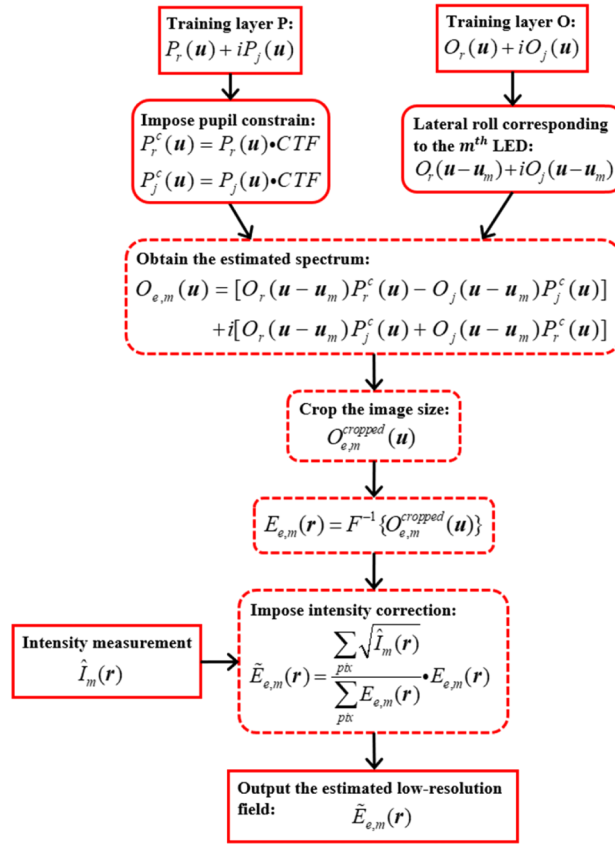
Fig. 2. Overall workflow of FINN-P method

fixed learning rate.

$$loss = \sum_{m=1}^{M} \sum_{pix} ||\tilde{E}_{e,m}(\boldsymbol{r})| - \sqrt{\hat{I}_m(\boldsymbol{r})}|^2, \tag{9}$$

where M represents the number of employed LEDs. To minimize the loss, we apply stochastic gradient descent with Nesterov acceleration (NAG) as the optimizer instead of Adaptive Moment Estimation (Adam) used in Jiang's method. This optimizer updates the current parameters by calculating the gradient information of the lookahead position and combining it with the momentum [41,42]. Therefore, it could get a faster convergence speed and better result than the traditional first-order gradient-based optimizers. Meanwhile, Nesterov-accelerated Adaptive Moment Estimation (Nadam) is employed as the optimizer to train the pupil function, which uses the estimation of the first and second moments of the gradient to dynamically adjust the learning rate of each epoch. Moreover, as the improved version of Adam, Nadam combines Adam with NAG to have better estimates in both the learning rate and direction of the gradient [42,43].

It is worth noting that since the recovered high-resolution image has different size from measured image. After the simulated high-resolution spectrum is multiplied by the pupil function, it needs to be cropped to fit the corresponding measured size. In addition, during the reconstruction process of the pupil function, the pixel values outside the circular function which should always be zero could become non-zero due to the noise in the measured images. Therefore, we propose a pupil constraint by multiplying the CTF of the system. Only after all

captured intensity images are used to train the network, a single epoch of the training process is completed. Then the entire training process needs to be repeated for dozens of epochs to ensure the convergence towards the final sample spectrum and pupil function.

4. Performance in simulation

To quantitatively verify the effectiveness of our model in improving the reconstruction quality through pupil recovery, a simulated FPM dataset is trained on our networks with and without pupil recovery, denoted as FINN-P, FINN respectively. The overall workflow of FINN is similar to the one mentioned in Section 3, except that the pupil function remains unchanged after being fed into the network.

In the simulation, two images with a size of 512×512 pixels are employed as the high-resolution amplitude and phase images respectively, the magnification of the simulated system is set to $8\times$ and the NA is set to 0.2, the sample is placed $67.5\,\mu m$ behind a 7×7 LED array and the distance between the adjacent LEDs is $4\,mm$. Therefore, the synthetic NA_{syn} could be up to 0.44. In order to simulate the aberrations and the situation of incoherent imaging that occurs in the system, we set the amplitude of pupil function as the incoherent optical transfer function (OTF) and the phase as non-zero during the process of obtaining the low-resolution images as shown in Figs. 3(a4) and 3(a5). OTF is the normalized autocorrelation of CTF which could be formulated as follows:

$$OTF(\boldsymbol{u}) = \frac{CTF(\boldsymbol{u}) \star CTF(\boldsymbol{u})}{\sum_{pix} |CTF(\boldsymbol{u})|^2}, \tag{10}$$

where '$\star$' means the autocorrelation operation. The same dataset is also employed to the AS algorithm and Jiang's method for comparison. In all algorithms, the image corresponding to the center LED is up-sampled and Fourier transformed to act as the initial guess of the spectrum. The reconstruction results are shown in Fig. 3. Instead of using the original high-resolution images as the ground truth, we obtain the ground-truth images by passing the original high-resolution spectrum through a synthetic low-pass circular filter with the radius of $2\pi(NA_{syn})/\lambda$. The ground-truth of the amplitude, phase and spectrum are shown in Figs. 3(a1)–3(a3).

All the four methods use the circular low-pass filter mentioned in Eq. (4) as the initial guess of the pupil function. To ensure the convergence of all algorithms, FINN and FINN-P are trained for 40 epochs, Jiang's method is trained for 200 epochs, the AS algorithm stops automatically after 21 steps since it adopts an adaptive step-size strategy [9]. However due to the imprecise pupil function used in the AS algorithm, Jiang's method and FINN, the reconstruction results are subject to significant crosstalk between the amplitude and phase, which seriously degrades the reconstruction quality as shown in Figs. 3(b1)–3(b3), Figs. 3(c1)–3(c3) and Figs. 3(d1)–3(d3). Whereas, the FINN-P method can correctly reconstruct the amplitude and phase information of the sample with only a small amount of crosstalk which is similar to the ground truth (Figs. 3(e1)–3(e3)).

The recovered pupil functions are shown in Figs. 3(e4) and 3(e5), the similarity between the recovered pupil function and ground truth is further evaluated by using the mean square error (MSE) and the structural similarity index (SSIM). The result is listed in Table 1. It can be seen that the recovered pupil functions are very close to the ground truth in both amplitude and phase. Therefore, the FINN-P method could be used to evaluate the optical transmission capability of the system.

Table 1. The comparison between the recovered pupil function and ground truth

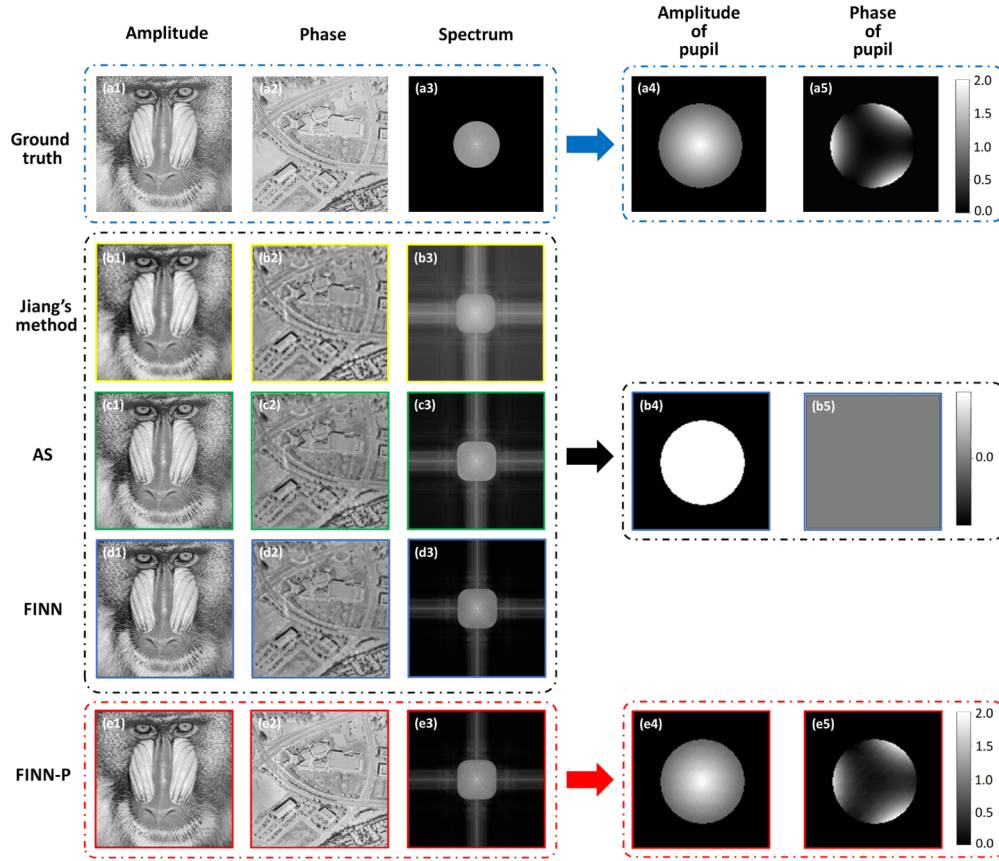|	MSE ($\times e^{-3}$)	SSIM
Amplitude of pupil	0.346	0.993
Phase of pupil	2.808	0.937

Fig. 3. Reconstruction results of Jiang's method, AS, FINN and FINN-P using the simulated dataset. (a1-a2) The ground-truth amplitude (Baboon) and phase (Aerial) for comparison. (a3) The truncated spectrum based on the synthetic NA. (a4-a5) The actual pupil function added into the imaging system. (b1-b3) The reconstruction amplitude, phase and spectrum using Jiang's method. (c1-c3) The reconstruction amplitude, phase and spectrum using the AS method. (d1-d3) The reconstruction amplitude, phase and spectrum using FINN. (b4-b5) the amplitude and phase of the circular low-pass filter defined in Eq. (4), which is also the initial guess of the pupil function in all the four methods. (e1-e3) The reconstruction results using FINN-P. (e4-e5) The reconstructed amplitude and phase of the pupil function through FINN-P, showing a similar distribution as Figs. 3(a4) and 3(a5).

In addition, due to the use of an LED array which provides a periodic sample pattern, there is a slightly periodic pattern in the recovered pupil function [44,45]. However, this pattern could be eliminated by using a non-periodic LED array [45], which will not be further discussed in this paper.

In addition to using MSE as the evaluation indicator, the reconstruction results are also quantitatively evaluated by calculating the normalized mean square error (NMSE) between the reconstruction spectrum and the ground truth spectrum [24,46]:

$$
\text{NMSE} = \frac{\sum_{pix} \left| \hat{O}(\boldsymbol{u}) - \frac{\sum_{pix} \hat{O}(\boldsymbol{u}) O_e^*(\boldsymbol{u})}{\sum_{pix} (|O_e(\boldsymbol{u})|^2)} O_e(\boldsymbol{u}) \right|^2}{\sum_{pix} |\hat{O}(\boldsymbol{u})|^2}, \tag{11}
$$

where $\hat{O}(\boldsymbol{u})$ represents the ground truth spectrum of the sample. The NMSE is calculated over the overlapping area of 128×128 pixels at the center of two spectra. NMSEs at different epochs are plotted in Fig. 4.
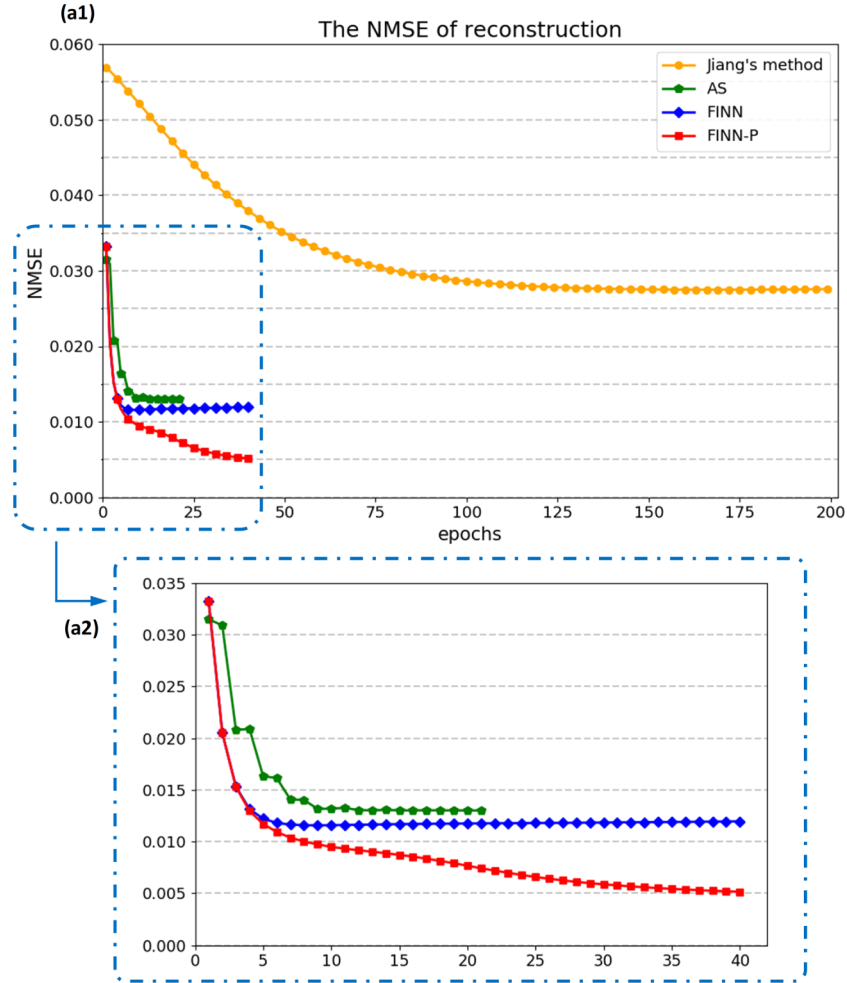

Fig. 4. The NMSE between the reconstruction results of different methods and the ground truth at different epochs. (a1) To ensure the convergence of different algorithms, FINN and FINN-P are trained for 40 epochs, the AS algorithm stops automatically after 21 steps, Jiang's method is trained for 200 epochs. (a2) the locally enlarged image of Fig. 4(a1).

It can be seen from Fig. 4(a1) that, for Jiang's method, the NMSE converges at the slowest rate and is still much larger than the other three methods even after 200 epochs of training. In order to better demonstrate the comparison between the reconstructed results of AS, FINN and FINN-P, the sub-region framed in Fig. 4(a1) is enlarged and shown in Fig. 4(a2). Compared with the AS algorithm, Both the FINN and FINN-P have faster convergence rate and end up with a smaller NMSE. Meanwhile, because of the inaccurate estimation of the pupil function, the NMSE of the FINN stops dropping after the first 8 epochs and slightly increase during the subsequent training process. Whereas, due to the embedded pupil recovery procedure in FINN-P, the impacts of the pupil inaccuracy will gradually decrease as the increase of the training epoch, which leads to a

continuous convergence process and better reconstructed result. The NMSE could reach about $5 \times e^{-3}$ after 40 epochs.

We further compare our results of MSE and SSIM with GS, AS and Jiang's method, the GS algorithm is iterated for 40 times, those two evaluation indicators are calculated between the reconstructed amplitude and the ground truth. The comparison is listed in Table 2. Benefiting from the pupil recovery procedure, we can see that FINN-P presents the best results in NMSE, MSE and SSIM than the other four methods.

Table 2. The comparison of reconstruction results

	NMSE of spectrum ($\times e^{-3}$)	MSE^a ($\times e^{-3}$)	SSIM
GS	42.47	23.87	0.904
Jiang's method^b	28.86	3.75	0.953
AS^c	13.09	2.09	0.979
FINN	11.96	1.35	0.985
FINN-P	5.15	1.05	0.991

^{a}MSE is calculated after normalizing the reconstruction amplitude into 0~1.
^{b}Jiang's method is modified from the open source code [40].
^{c}The AS algorithm is implemented with Python

5. Performance in experience

In this section, we implement FINN-P in experimental data. First, we use an open source USAF dataset provided by Zuo *et al* [9] and compare the reconstruction result with GS, AS, and Jiang's method. The hardware setup of the FPM system consists of a 2× objective lens with an NA of 0.1, a CCD sensor with a pixel size of $6.5\,\mu m$ and a 21×21 programmable LED matrix with a lateral distance between two adjacent LEDs of 2.5 mm, and the sample which is placed $87.5\,\mu m$ behind the LED. The initial guess of the high-resolution spectrum is the Fourier transform of the up-sampled image captured under the illumination of the center LED. The reconstruction results are shown in Fig. 5.
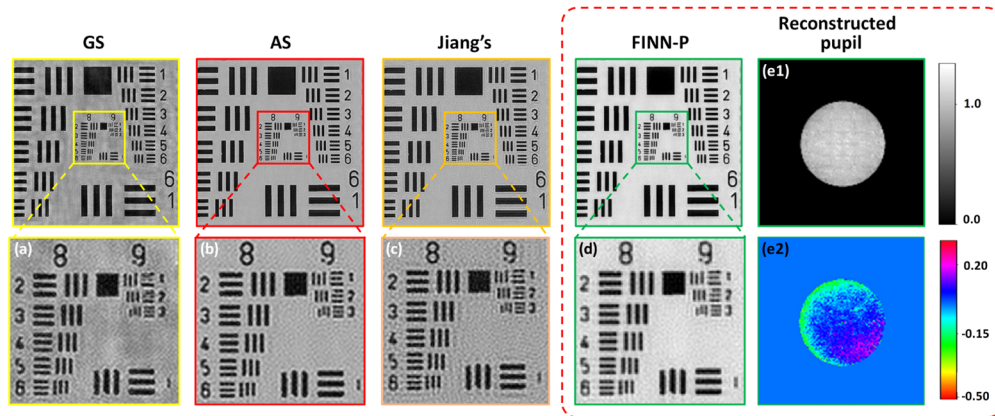

Fig. 5. The comparison of reconstruction results using the USAF dataset. (a-d) The reconstruction results by GS, AS, Jiang's and FINN-P respectively. (e1). The amplitude of the pupil function recovered through FINN-P. (e2) The phase of the pupil function recovered through FINN-P.

Due to the sufficient number of low-resolution images and overlap in the Fourier domain [47], all four algorithms can reconstruct the line pairs in Group 9, Element 3 (0.775μm). However, GS

method has the worst reconstruction result with much more background noise than other methods, and some convergence errors appears in Group 9, Element 1 and 2 (Fig. 5(a)). The reconstruction result of Jiang's method suffers from obvious ring artifacts around Group 8, Element 1 (Fig. 5(c)). The AS algorithm proposed by Zuo *et al*. can reduce the influences of the background noise to a certain extent and present better reconstruction quality than GS and Jiang's method, but there is still serious background noise (Fig. 5(b)). In contrast, FINN-P jointly optimizes the sample spectrum and pupil function, improves the reconstruction quality and obtains a better result with minimal background noise (Fig. 5(d)). The amplitude and phase of the pupil function recovered by FINN-P are shown in Figs. 5(e1) and 5(e2) respectively, from the pupil function we can see that the optical transfer function of the system is very similar to CTF and there is only a slight pupil aberration in the system, which is why some algorithms that do not consider the pupil aberrations, such as AS, can still get acceptable reconstruction results.

In addition to using USAF dataset, the FINN-P algorithm is also implemented on an open source dataset (stained Human Bone Osteosarcoma Epithelial U2OS sample) provided by Tian *et al* [16]. The parameters of the FPM system are the same as those used in the previous simulation. The performance of our method under a slight aberration is first analyzed by choosing an area with a size of 256×256 pixels ($204~\mu m \times 204~\mu m$) at the center of the entire field of view (FOV). The same dataset is also employed to the AS algorithm, and the comparison of both methods are shown in Fig. 6. In order to better illustrate the effectiveness of our method in improving the reconstruction quality, the same small portion of the reconstruction amplitude from both algorithms is zoomed in and shown in Figs. 6(a3) and 6(c3). From the recovered pupil phase which is shown in Fig. 6(c4), it can be seen that the aberration at this location is so small that both methods can obtain fine results without obvious convergence errors. However, Fig. 6(a3) shows that some details are not well recognized because the AS method ignores the aberration effects. In contrast, FINN-P can successfully separate the influences of the aberration from the
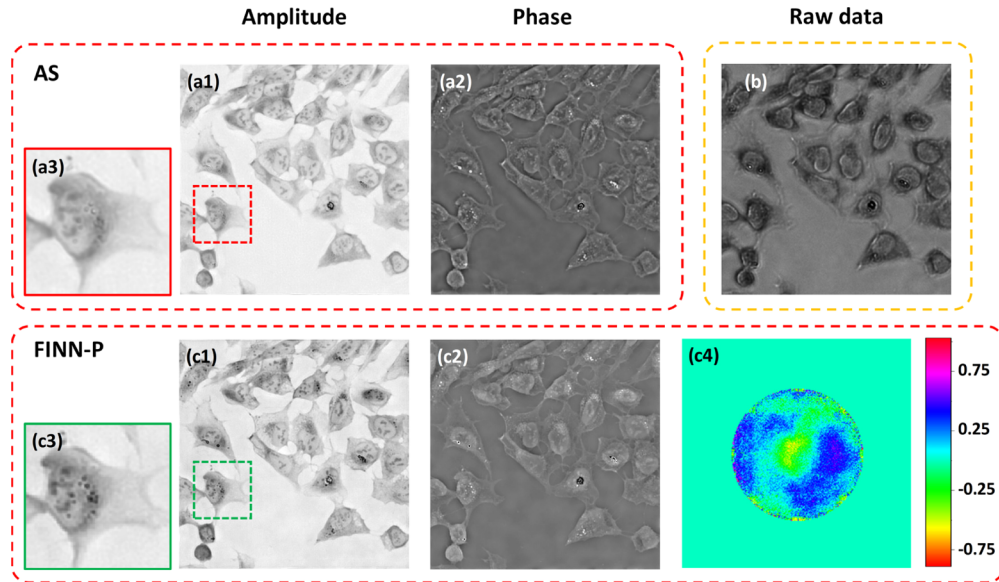

Fig. 6. The comparison of reconstruction results using the open source dataset (U2OS). The reconstruction region located at the center of the FOV. (a1-a3) The reconstruction results by the AS method. (b) The low-resolution image captured under the illumination of the center LED. (c1-c3) The reconstruction results by FINN-P. (c4) The recovered pupil phase by FINN-P.

input dataset through the pupil recovery procedure, which results in a higher-quality result. As shown in Fig. 6(c3), more details such as the organelles inside the cell can be identified, and the phase information of the sample obtained by our method exhibits a higher contrast and smoother background than AS.

We further analyze the region in the upper-right corner of the FOV, in which the aberration is no longer non-negligible. Figure 7(c4) shows the pupil phase recovered by FINN-P. The amplitude and phase images of the same sub-region reconstructed by AS and FINN-P are shown in Figs. 7(a1) and 7(a2), Figs. 7(c1) and 7(c2) respectively. It can be seen from the local enlarged
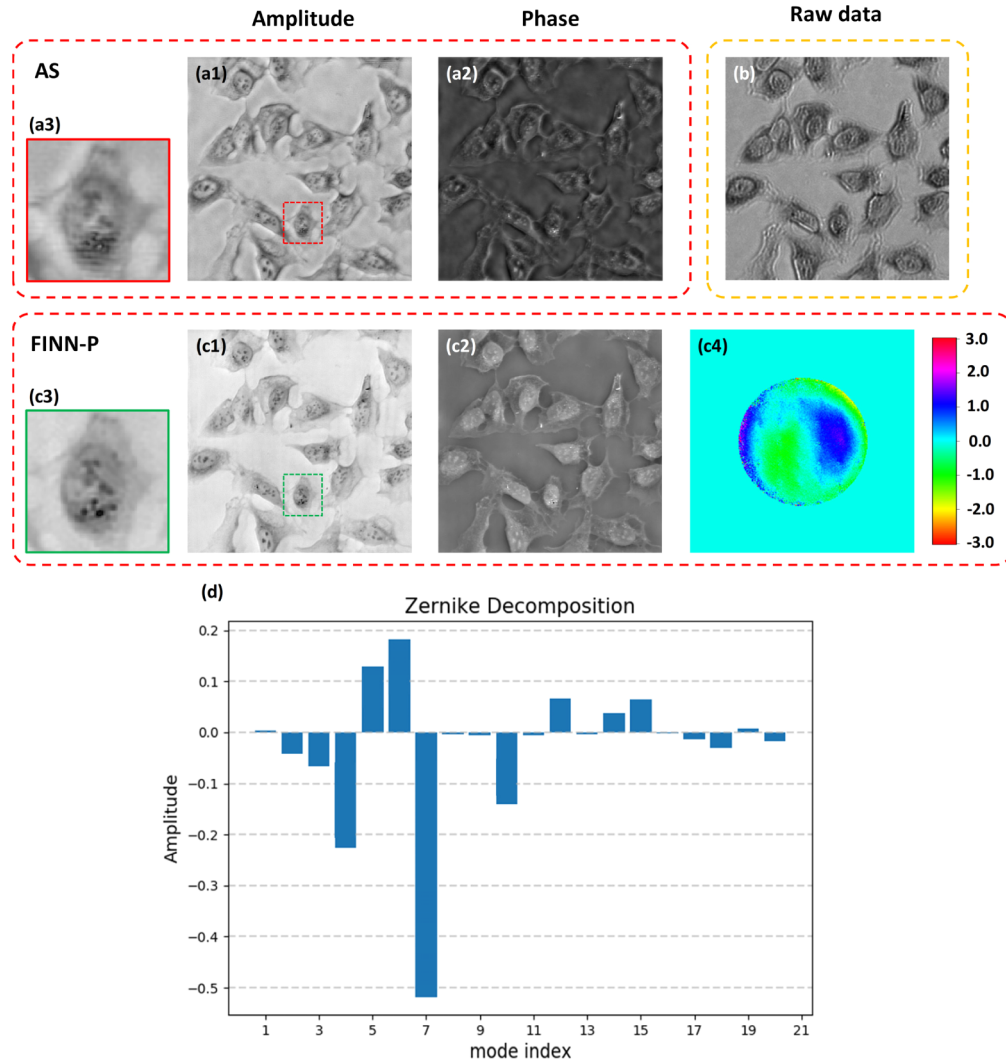

Fig. 7. The comparison of reconstruction results using the open source dataset (U2OS). The reconstruction region located in the upper-right corner with a non-negligible pupil aberration. (a1-a3) The reconstruction amplitude and phase by the AS method. (b) The low-resolution image captured under the illumination of the center LED. (c1-c3) The reconstruction results using FINN-P method. (c4) The reconstruction phase of pupil function which could represent the wavefront aberration of the system. (d) The Zernike decomposition of the reconstruction pupil phase.

images shown in Figs. 7(a3) and 7(c3) that due to a very significant pupil aberration, the AS method is unable to converge to a clear result and generates some grating fringe errors. The detailed structure in Fig. 7(a3) is difficult to distinguish, and the reconstruction phase shown in Fig. 7(a2) contains a lot of background fluctuation. Whereas, even with such a large aberration, our method still can obtain clear results in amplitude and phase with more recognizable details and smoother background (Fig. 7(c1)–7(c3)).

In addition, a Zernike decomposition of the pupil phase component is performed to better demonstrate the aberration. The coefficients of the first 20 Zernike polynomials are decomposed and shown in Fig. 7(d). The three main Zernike components of the aberration are mode 7 which represents the tilt in the y-direction, mode 6 which represents the tilt in the x-direction, mode 4 which represents the astigmatism at 90°. In this way, a good estimate can be obtained before further analysis of the system wavefront aberration.

6. Conclusion and discussion

In this paper, based on the machine-learning platform TensorFlow, a Fourier ptychographic forward imaging network embedded with pupil recovery (FINN-P) is proposed and the effectiveness in reducing aberration effects on both simulated and measured datasets is also demonstrated. In FINN-P, both the spectrum of the sample and pupil function are separated into the real and imaginary parts respectively, and all four parts are treated as trainable layers. By using the imaging network with a more effective workflow and different optimizers, FINN-P can achieve much better reconstruction results than Jiang's method which is also based on neural network. Due to the co-optimization of the sample's spectrum and pupil function, FINN-P can obtain clearer results even if there is a significant aberration and the recovered pupil function can be used as a good initial estimate when further analysis of the system optical transmission capability is required. Meanwhile, the reconstruction speed of FINN-P can be further improved by training the proposed network in the neural engine or tensor processing unit (TPU) as mentioned in [40].

Due to the periodic LED array used in the FPM system, there is a slight periodic pattern in the recovered pupil function, which will become an obstacle to further application of our method. However, this issue could be solved by using a different illumination source, such as a non-periodic LED array, which will be our future work. In addition, in order to improve the reconstruction speed, we could combine FINN-P with a DCNN by adopting the structure of generative adversarial network (GAN) [32], in which FINN-P can be treated as a discriminator and the DCNN acts as a generator. By training the entire network using a loss function such as cross-entropy or L_2-norm, we could obtain a well-trained generator which could quickly map the low-resolution images into a high-resolution result and also reduce the influences of the pupil aberration.

References

1. G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," Nat. Photonics **7**(9), 739–745 (2013).

2. V. Mico, Z. Zalevsky, P. García-Martínez, and J. García, "Synthetic aperture superresolution with multiple off-axis holograms," J. Opt. Soc. Am. A **23**(12), 3162–3170 (2006).

3. K. Guo, S. Dong, and G. Zheng, "Fourier Ptychography for Brightfield, Phase, Darkfield, Reflective, Multi-Slice, and Fluorescence Imaging," IEEE J. Sel. Top. Quantum Electron. **22**(4), 77–88 (2016).

4. X. Ou, R. Horstmeyer, C. Yang, and G. Zheng, "Quantitative phase imaging via Fourier ptychographic microscopy," Opt. Lett. **38**(22), 4845–4848 (2013).

5. J. Di, J. Zhao, H. Jiang, P. Zhang, Q. Fan, and W. Sun, "High resolution digital holographic microscopy with a wide field of view based on a synthetic aperture technique and use of linear CCD scanning," Appl. Opt. **47**(30), 5654–5659 (2008).

6. H. M. L. Faulkner and J. M. Rodenburg, "Movable Aperture Lensless Transmission Microscopy: A Novel Phase Retrieval Algorithm," Phys. Rev. Lett. **93**(2), 023903 (2004).

7. J. M. Rodenburg, A. C. Hurst, A. G. Cullis, B. R. Dobson, F. Pfeiffer, O. Bunk, C. David, K. Jefimovs, and I. Johnson, "Hard-X-ray lensless imaging of extended objects," Phys. Rev. Lett. **98**(3), 034801 (2007).

8. J. M. Rodenburg, "Ptychography and related diffractive imaging methods," Adv. Imaging Electron Phys. **150**, 87–184 (2008).

9. C. Zuo, J. Sun, and Q. Chen, "Adaptive step-size strategy for noise-robust Fourier ptychographic microscopy," Opt. Express **24**(18), 20724–20744 (2016).

10. Y. Fan, J. Sun, Q. Chen, M. Wang, and C. Zuo, "Adaptive denoising method for Fourier ptychographic microscopy," Opt. Commun. **404**, 23–31 (2017).

11. Y. Zhang, P. Song, and Q. Dai, "Fourier ptychographic microscopy using a generalized Anscombe transform approximation of the mixed Poisson-Gaussian likelihood," Opt. Express **25**(1), 168–179 (2017).

12. L. Hou, H. Wang, M. Sticker, L. Stoppe, J. Wang, and M. Xu, "Adaptive background interference removal for Fourier ptychographic microscopy," Appl. Opt. **57**(7), 1575–1580 (2018).

13. S. Chen, T. Xu, J. Zhang, X. Wang, and Y. Zhang, "Optimized Denoising Method for Fourier Ptychographic Microscopy Based on Wirtinger Flow," IEEE Photonics J. **11**(1), 1–14 (2019).

14. Y. Zhang, W. Jiang, and Q. Dai, "Nonlinear optimization approach for Fourier ptychographic microscopy," Opt. Express **23**(26), 33822 (2015).

15. L. H. Yeh, J. Dong, J. Zhong, L. Tian, M. Chen, G. Tang, M. Soltanolkotabi, and L. Waller, "Experimental robustness of Fourier ptychography phase retrieval algorithms," Opt. Express **23**(26), 33214–33240 (2015).

16. L. Tian, X. Li, K. Ramchandran, and L. Waller, "Multiplexed coded illumination for Fourier Ptychography with an LED array microscope," Biomed. Opt. Express **5**(7), 2376–2389 (2014).

17. L. Bian, J. Suo, G. Zheng, K. Guo, F. Chen, and Q. Dai, "Fourier ptychographic reconstruction using Wirtinger flow optimization," Opt. Express **23**(4), 4856–4866 (2015).

18. R. Horstmeyer, R. Y. Chen, X. Ou, B. Ames, J. A. Tropp, and C. Yang, "Solving ptychography with a convex relaxation," New J. Phys. **17**(5), 053044 (2015).

19. L. Bian, J. Suo, G. Situ, G. Zheng, F. Chen, and Q. Dai, "Content adaptive illumination for Fourier ptychography," Opt. Lett. **39**(23), 6648 (2014).

20. Y. Zhang, W. Jiang, L. Tian, L. Waller, and Q. Dai, "Self-learning based Fourier ptychographic microscopy," Opt. Express **23**(14), 18471–18486 (2015).

21. A. Pan, Y. Zhang, T. Zhao, Z. Wang, D. Dan, M. Lei, and B. Yao, "System calibration method for Fourier ptychographic microscopy," J. Biomed. Opt. **22**(09), 1–11 (2017).

22. J. Sun, Q. Chen, Y. Zhang, and C. Zuo, "Efficient positional misalignment correction method for Fourier ptychographic microscopy," Biomed. Opt. Express **7**(4), 1336–1350 (2016).

23. A. Zhou, W. Wang, N. Chen, E. Y. Lam, B. Lee, and G. Situ, "Fast and robust misalignment correction of Fourier ptychographic microscopy for full field of view reconstruction," Opt. Express **26**(18), 23661–23674 (2018).

24. X. Ou, G. Zheng, and C. Yang, "Embedded pupil function recovery for Fourier ptychographic microscopy," Opt. Express **22**(5), 4960–4972 (2014).

25. M. Zhang, L. Zhang, D. Yang, H. Liu, and Y. Liang, "Symmetrical illumination based extending depth of field in Fourier ptychographic microscopy," Opt. Express **27**(3), 3583–3597 (2019).

26. Z. Bian, S. Dong, and G. Zheng, "Adaptive system correction for robust Fourier ptychographic imaging," Opt. Express **21**(26), 32400–32410 (2013).

27. W. Luo, A. Greenbaum, Y. Zhang, and A. Ozcan, "Synthetic aperture-based on-chip microscopy," Light: Sci. Appl. **4**(3), e261 (2015).

28. L. Tian and L. Waller, "3D intensity and phase imaging from light field measurements in an LED array microscope," Optica **2**(2), 104 (2015).

29. C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a Deep Convolutional Network for Image Super-Resolution," IEEE Trans. Pattern Anal. Machine Intell. **38**(2), 295–307 (2016).

30. J. Kim, J. K. Lee, and K. M. Lee, "Deeply-Recursive Convolutional Network for Image Super-Resolution," arXiv preprint arXiv: 1511.04491v2 (2015).

31. J. Kim, J. K. Lee, and K. M. Lee, "Accurate Image Super-Resolution Using Very Deep Convolutional Networks," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2016), pp. 1646–1654.

32. I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative Adversarial Networks," arXiv preprint arXiv: 1406.2661v1 (2014).

33. C. Ledig, L. Theis, F. Huszar, J. Caballero, A. Cunningham, A. Acosta, A. Aitken, A. T. ejani, J. Totz, Z. Wang, and W. Shi, "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE, 2017), pp. 105–114.
34. Y. Rivenson, Y. Zhang, H. Gunaydin, D. Teng, and A. Ozcan, "Phase recovery and holographic image reconstruction using deep learning in neural networks," Light: Sci. Appl. **7**(2), 17141 (2018).
35. L. Boominathan, M. Mainiparambil, H. Gupta, R. Baburaj, and K. Mitra, "Phase retrieval for Fourier Ptychography under varying amount of measurements," arXiv preprint arXiv: 1805.03593 (2018).
36. Y. F. Cheng, M. Strachan, Z. Weiss, M. Deb, D. Carone, and V. Ganapati, "Illumination pattern design with deep learning for single-shot fourier ptychographic microscopy," arXiv preprint arXiv: 1810.03481 (2018).
37. A. Kappeler, S. Ghosh, J. Holloway, O. Cossairt, and A. Katsaggelos, "Ptychnet: CNN based fourier ptychography," *IEEE International Conference on Image Processing (ICIP)* (IEEE, 2017), pp. 1712–1716.
38. T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, "Deep learning approach for Fourier ptychography microscopy," Opt. Express **26**(20), 26470 (2018).
39. F. Shamshad, F. Abbas, and A. Ahmed, "Deep Ptych: Subsampled Fourier Ptychography using Generative Priors," arXiv preprint arXiv: 1812.11065 (2018).
40. S. Jiang, K. Guo, J. Liao, and G. Zheng, "Solving Fourier ptychographic imaging problems via neural network modeling and TensorFlow," Biomed. Opt. Express **9**(7), 3306 (2018).
41. Y. U. Nesterov, "A method for unconstrained convex minimization problem with convergence rate o(1/k2)," Doklady AN SSSR **269**, 543–547 (1983).
42. S. Ruber, "An overview of gradient descent optimization algorithms," arXiv preprint arXiv: 1609.04747v2 (2016).
43. T. Dozat, "Incorporating Nesterov Momentum into Adam," ICLR Workshop **1**, 2013–2016 (2016).
44. P. Thibault, M. Dierolf, O. Bunk, A. Menzel, and F. Pfeiffer, "Probe retrieval in ptychographic coherent diffractive imaging," Ultramicroscopy **109**(4), 338–343 (2009).
45. G. Kaikai, D. Siyuan, and N. Pariksheet, "Optimization of sampling pattern and the design of Fourier ptychographic illuminator," Opt. Express **23**(5), 6171 (2015).
46. J. R. Fienup, "Invariant error metrics for image reconstruction," Appl. Opt. **36**(32), 8352–8357 (1997).
47. J. Sun, Q. Chen, Y. Zhang, and C. Zuo, "Sampling criteria for Fourier ptychographic microscopy in object space and frequency space," Opt. Express **24**(14), 15765 (2016).