


Fourier Ptychography Reconstruction Based on Reweighted Amplitude Flow With Regularization by Denoising and Deep Decoder

Baopeng Li , Caiwen Ma, Okan K. Ersoy, *Fellow, IEEE*, Zhibin Pan, Wansha Wen, Zhonghan Sun, and Wei Gao

Abstract—Fourier ptychography (FP) is a computational imaging technique with the advantage that it can obtain large field-of-view (FOV) and high-resolution (HR) imaging. We propose an algorithm for Fourier ptychography based on reweighted amplitude flow (RAF) with regularization by denoising (RED) and deep decoder (DD), which is an untrained deep generative model. The proposed method includes two loops, using reweighted amplitude flow with regularization by denoising as an inner loop for phase retrieval and deep decoder for further denoising as an outer loop in the Fourier ptychography recovery system. The proposed method does not need any training dataset, just adds a little computer time during the image recovery process. The proposed method has no bias due to training images, which is different from other deep learning methods. The experimental results show that the proposed method can improve the reconstruction quality in both PSNR and SSIM.

Index Terms—Coherent imaging, imaging systems, technologies for computing.

I. INTRODUCTION

FOURIER ptychography is a new imaging technique that bypasses the resolution limit of the employed optics. In particular, it transforms the general challenge of high throughput, high-resolution imaging from one that is coupled to the physical limitations of the optics to one that is solvable through computation [1], [2], [3], [4]. FP integrates phase retrieval and

aperture synthesis together to recover both amplitude and phase information.

Most FP research are focusing on microscopy applications [2], [5], [6], [7], [8]. A light-emitting diode (LED) array is used to provide different angle illumination, and intensity images of the sample are recorded and used to recover a high-resolution sample image. Recently, researchers consider using the FP method to obtain long-distance, subdiffraction-limited imaging instead of building a larger lens, which is usually expensive and heavy, with bulky optics and mechanics [9]. Holloway et al. [10] proposed using camera arrays coupled with coherent illumination to improve spatial resolution in long-distance images. Synthetic apertures for visible imaging (SAVI) [9] using FP extended the synthetic aperture techniques in the visible light domain to achieve subdiffraction-limited resolution. They also achieved imaging of optically rough objects experimentally using macroscopic FP in a reflection imaging geometry.

The captured low-resolution images always include complicated noise for long-distance FP imaging systems. For example, coherent illumination of rough surfaces produces speckle noise [11]. Some works research the noise modeling and denoising of the FP reconstruction. Zuo et al. [12] proposed an adaptive step-size strategy which improved the stability and robustness of the reconstruction under noisy conditions. Bian et al. [13] proposed a Wirtinger flow optimization for Fourier Ptychographic (WFP), which incorporated phase retrieval and noise relaxation together. WFP tests with the different additive Gaussian noises. Reweighted amplitude flow (RAF) is more robust than the intensity-based algorithms in some applications, for example, in the presence of Poisson noise [14]. Li et al. [11] proposed a new FP algorithm, which combined reweighted amplitude flow and regularization by denoising (RED). This method achieved better image reconstruction, and reduced the effects of laser speckle noise in long-distance experiments. The above studies have been carried out on different noises and have achieved good results to a certain extent. However, there are still some drawbacks of using the above methods. First, the image prior knowledge adopted by these methods are mostly based on human knowledge, so it is difficult to capture the full features of the image [15]. Therefore, the performance may be limited. Next, most of these methods only use the internal information of the input image without use of any external information [15]. So there is room for further improvement.

Manuscript received 17 September 2022; revised 20 November 2022; accepted 14 December 2022. Date of publication 19 December 2022; date of current version 30 December 2022. This work was supported in part by Open Research Fund of State Key Laboratory of Transient Optics and Photonics, under Grant SKLST202114, and in part by High-resolution Earth Observation System of China, under Grant GFZX04014307. (*Corresponding author: Baopeng Li.*)

Baopeng Li is with the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China, also with the Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China, and also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: libaopeng@opt.ac.cn).

Caiwen Ma, Wansha Wen, Zhonghan Sun, and Wei Gao are with the Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Beijing 100045, China (e-mail: cwma@opt.ac.cn; wenwansha19@mails.u.ac.cn; sunzhonghan@opt.ac.cn; gaowei@opt.ac.cn).

Okan K. Ersoy is with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907 USA (e-mail: ersoy@purdue.edu).

Zhibin Pan is with the Faculty of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: zbpan@mail.xjtu.edu.cn).

Digital Object Identifier 10.1109/JPHOT.2022.3230422

In recent years, deep neural networks (DNN) have been successfully used in image processing and computer vision, such as image denoising, image restoration and optical metrology tasks [16], [17]. With the rapid development of DNN, researchers have developed a lot of very powerful denoising architectures [16], such as DnCNN [18], FFDNet [19], VNet [20] and DANet [21]. Deep learning based denoising methods are state-of-the-art image denoising methods. However, it is well-known that deep learning methods need a large training dataset to a certain degree. The more training data are used the better the results are. However, training data are not easy to be obtained in some areas, such as astronomy and remote sensing. It is difficult to acquire adequate an number of clean-noisy image pairs of FP denoising tasks. Moreover, if the testing data misfit the training data, the performance of deep learning's prediction will be affected.

Recently, researchers proposed some new image denoising neural networks which do not involve in any training dataset, which are different from traditional DNNs. The biggest advantage of an untrained network is that researchers can avoid building a large database. Ulyanov et al. [22] first proposed an untrained neural network method, namely Deep Image Prior (DIP), which showed that the structure of a generator network is sufficient to capture image information. Heckel et al. [23] proposed a new untrained neural network method called as Deep Decoder (DD), which is a deep neural network that can generate natural images with very few weight parameters. DD maps a lower-dimensional space to a higher-dimensional space, which is similar to sparse image representation. DD has a better performance compared with state-of-the-art untrained denoising methods, such as DIP and BM3D [24]. Since the untrained neural network is valuable for real data, we consider using a state-of-the-art untrained denoising network deep decoder in the FP reconstruction.

In this paper, we propose an algorithm called RAF-RED-DD, which combines reweighted amplitude flow (RAF), regularization by denoising (RED) and deep decoder (DD). The recovery algorithm effectively combines two kinds of denoising algorithms, in the inner loop based on RED denoising, and in the outer loop based on DD denoising for the amplitude image within the Fourier ptychography recovery process. RAF-RED-DD is a flexible framework for using denoising algorithms as priors for FP reconstruction algorithm. Deep decoder, which is flexibly embedded into the FP reconstruction algorithm, can be considered as a plug-and-play prior [25]. RAF-RED-DD allows denoising methods that are not explicitly formulated as optimization problems to be used. Experiments show that RAF-RED-DD can get a better result than RAF-RED without considering DD. The proposed method is robust in different noise.

The remainder of the paper is organized as follows. In Section II, we first demonstrate the FP imaging model and then show the proposed method. In Section III, we discuss the hyper parameters of the proposed method and show the performance of the proposed method with numerical simulation and real data. Finally, Section IV gives the conclusion and future work.

II. PROPOSED METHOD

In this section, we first introduce the forward Fourier ptychography imaging model, and then introduce the reweighted amplitude flow with regularization by denoising for the FP reconstruction process. Next, we combine the reweighted amplitude flow with regularization by denoising and deep decoder together. Finally, we summarize our proposed reconstruction algorithm.

A. Forward FP Imaging Model

In the forward imaging model of FP, multiple intensity images of the sample are captured under different incident positions. The field emanating from the object is a complex wave $\psi(x, y)$, which passes through a long distance to the imaging system which satisfies the far-field Fraunhofer approximation. The field at the aperture plane of the camera is related to the field at the object by

$$\Psi(u, v) \propto F\{\psi(x, y)\} \quad (1)$$

where F is the 2D Fourier transform. In order to reduce clutter, spatial coordinates (x, y) are represented as \mathbf{x} , and frequency coordinates (u, v) are represented as \mathbf{u} .

Considering the finite diameter of the optical lens, we can only obtain part of information on the camera sensor. Let $P(\mathbf{u})$ represent pupil. For a circular lens, $P(\mathbf{u})$ can be defined as

$$P(\mathbf{u}) = \begin{cases} 1 & \text{if } \|\mathbf{u}\| < d/2 \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where d is the diameter of the optical lens.

Since CCD or CMOS can only measure the intensity images of the optical waves, we can capture low-resolution images at different locations in the Fourier domain, and the captured image is given by

$$I(x, c) = |F^{-1}\{\Psi(\mathbf{u} - \mathbf{c})P(\mathbf{u})\}|^2 \quad (3)$$

where \mathbf{c} is the center of the pupil, F^{-1} is the inverse Fourier transform. Shifting the camera aperture to capture different regions of the Fourier domain can be achieved by shifting the Fourier pattern relative to the camera. Because of the finite extent of the lens aperture, the captured images will not contain all of the frequency content of the propagated field $\Psi(\mathbf{u})$. By recovering the complex field $\Psi(\mathbf{u})$, we can reconstruct a high-resolution image of the object. A forward Fourier ptychography imaging system is shown in Fig. 1.

B. Optimization Framework

The FP measurement process can be defined as $y = |Ax|^2$, where $A \in \mathbb{C}^{m \times n}$ is the relation between high-resolution reconstruction and low-resolution measurements. A includes two linear operations: (1) down-sampling caused by the lens aperture, (2) inverse Fourier transform of the low-resolution spectrum caused by the imaging system. The target of FP reconstruction is to recover the high-resolution spatial spectrum (a plural image z) from a series of low-resolution images (measurements y_i) which are captured in the spatial space.

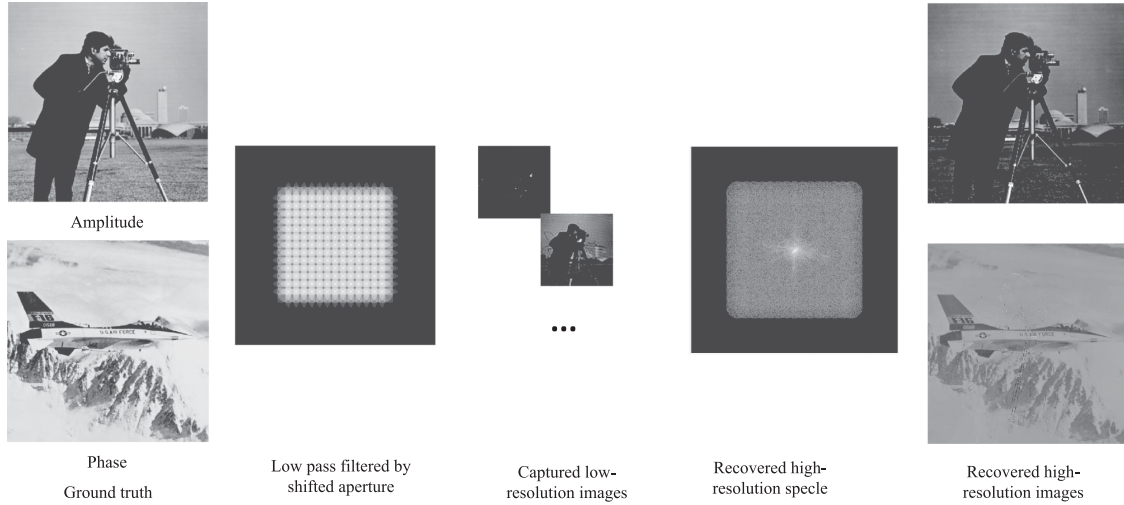


Fig. 1. A forward Fourier ptychography imaging system. Using amplitude and phase images make up the complex object image. After the forward processing, we get a series of low-resolution images. There are $M \times M$ captured low-resolution images, where M is the number of shifted apertures. Using the FP algorithm, we obtain a high-resolution image. In the simulations, the amplitude and phase images are 512×512 pixels, M is 15, and the low-resolution images are 102×102 pixels.

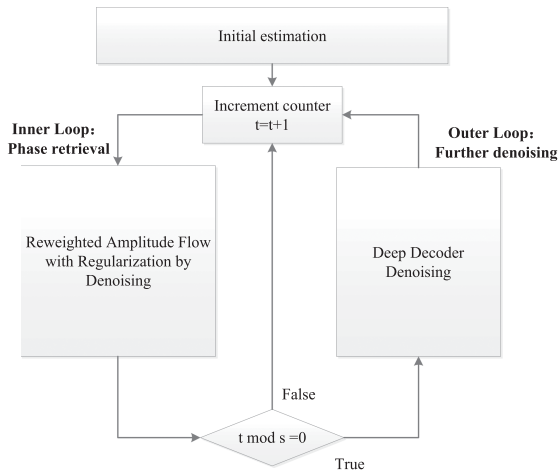


Fig. 2. Block diagram of the proposed algorithm, which includes inner loop (phase retrieval) and outer loop (deep decoder denoising).

Our method combines reweighted amplitude flow with regularization by denoising and deep decoder together. The proposed method includes two parts, an inner loop for phase retrieval using reweighted amplitude flow (RAF) with regularization by denoising (RED) and an outer loop for further denoising based on the deep decoder (DD). The deep decoder is applied every S iterations in the algorithm. The structure of the proposed algorithm is shown in Fig. 2.

1) *Inner Loop (Phase Retrieval)*: Since the reweighted amplitude flow with regularization by denoising for Fourier Ptychography is a flexible and powerful algorithm, we choose it as the inner loop. The optimization framework of FP using reweighted amplitude flow can be expressed as follows:

$$\operatorname{argmin} L(z) = \frac{1}{2m} \sum_{i=1}^m (|A_i z| - \sqrt{y_i})^2 \quad (4)$$

where y_i is a low-resolution intensity measurement, A_i is a linear transform matrix which incorporates the inverse Fourier transform and low-pass filtering of the pupil function, z is eventually the high-resolution spectrum of the object. The root-mean-squared measurements are employed as an initial estimator for FP. We can use the gradient descent method to optimize the above equation. Then, the update formula is:

$$\begin{aligned} z^{t+1} &= z^t - u^t w_i^t \nabla L(z) \\ &= z^t - \frac{u^t}{m} \sum_{i=1}^m w_i^t \left(A_i z - \sqrt{y_i} \frac{A_i z}{|A_i z|} \right) A_i^H \end{aligned} \quad (5)$$

where t is the iteration number, A_i^H is the transposed conjugate matrix of A_i , and w_i is given by

$$w_i = \frac{|A_i z|^2 / y_i}{|A_i z|^2 / y_i + \beta_i} \quad (6)$$

Regularization by denoising (RED) [26] is an efficient approach to solve imaging inverse problems, which can flexibility incorporate a denoiser to regularize imaging. RED is tested and demonstrated with state-of-the-art results in image deblurring and super-resolution applications. The RED framework defines a regularization term $R(z)$ as follows:

$$R_z = \frac{1}{2} z^H [z - f(z)] \quad (7)$$

where $f(z)$ is a denoising engine which can be flexibly chosen as an image denoising filter. Reweighted amplitude flow with RED for FP can be modeled as an energy function as follows:

$$E_z = L_z + R_z \quad (8)$$

where L_z is a data-fidelity term given by (4) and R_z is a regularization term given by (7).

The gradient of (8) is as follows:

$$\nabla E_z = \nabla L_z + \nabla R_z \quad (9)$$

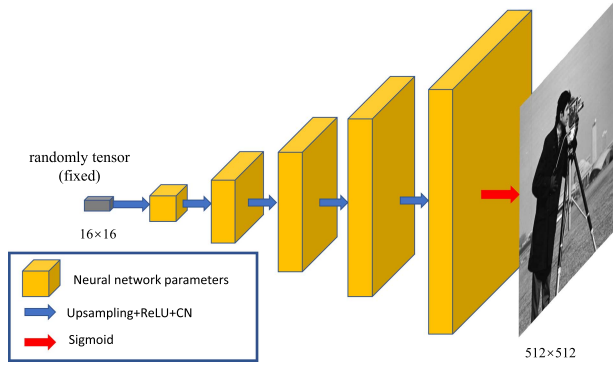


Fig. 3. The network structure of the Deep Decoder (DD). DD includes upsampling, ReLU and channel normalization (CN) in each layer. The output of the network uses Sigmoid activation function. DD does not require training data. DD itself acts as a natural data model [23].

The update formula is

$$\begin{aligned}
 z^{t+1} &= z^t - u^t \nabla E(z) \\
 &= z^t - u^t \left[\frac{1}{m} \sum_{i=1}^m w_i^t \left(A_i z - \sqrt{y_i} \frac{A_i z}{|A_i z|} \right) A_i^H \right. \\
 &\quad \left. + \lambda^t (z - f(z)) \right]
 \end{aligned} \quad (10)$$

More details about RAF and RED can be found in the papers [11], [26], [27], [28]. Using the inner loop, we can get a good Fourier ptychography reconstruction result. However, due to complex FP processing and different noise effects, the FP algorithm can still be further improved.

2) *Outer Loop (Deep Decoder Denoising)*: Deep decoder denoising transforms a randomly chosen and fixed tensor $B_0 \in \mathbb{R}^{n_0}$ consisting of n_0 -dimensional channels to a n_d -dimensional image x . This is done by using upsampling, rectified linear units (ReLU) and normalizing. The $(i+1)$ th layer is given by

$$B_{i+1} = cn(\text{relu}(U_i B_i C_i)), \quad i = 0, \dots, d-1 \quad (11)$$

where $C_i \in \mathbb{R}^{k_i \times k_{i+1}}$ contains the weights of the network, $U_i \in \mathbb{R}^{n_{i+1} \times n_i}$ is a bi-linear upsampling tensor, and $cn(\cdot)$ is a channel normalization operation. The final output of the network is given by

$$x = \text{sigmoid}(B_d C_d), \quad (12)$$

where $C_d \in \mathbb{R}^{k_d \times k_{out}}$. In our simulation, $n_0 = 16 \times 16$ and $n_d = 512 \times 512$. The structure of DD is shown in Fig. 3.

We can estimate an unknown image from the noisy measurement x by minimizing the loss function:

$$L(C) = \|f(DD(C)) - x\|_2^2 \quad (13)$$

where C is the network model parameters. By minimizing (13), we can obtain the parameters \hat{C} . We denoise the noisy image using:

$$\hat{x} = DD(\hat{C}) \quad (14)$$

where \hat{x} is the final reconstruction image.

Algorithm 1: Reweighted Amplitude Flow with Regularization by Denoising and Deep Decoder for Fourier Ptychography (RAF-RED-DD).

Input: Captured low-resolution images $\{y_i\}_{i=1}^m$; sampling matrix $\{A_i\}_{i=1}^m$.
Output: Retrieved signal z (recovered high-resolution spatial spectrum).

1: *Parameters:* maximum number of iterations T ; step size μ ; weighting parameters $\beta = 0.3$; regularization parameter $\lambda = 0.05$; DD epochs $T_D = 2000$; iteration number using DD $S = 50$.

2: *Initialization:* $z^0 = FFT\left(\frac{1}{m} \sum_{i=1}^m \sqrt{y_i}\right)$.

3: *Loop:*

for $t = 1$ to T

if $t = S$

$$x^t = |IFFT(z^t)|, \quad p^t = \angle(IFFT(z^t)),$$

$$x^{t+1} = DD(\hat{C}), \quad \text{where } \hat{C} = \text{argmin} \|f(DD(C)) - x^t\|_2^2$$

$$z^{t+1} = FFT(x^{t+1} \exp(1j p^t))$$

else

$$\begin{aligned}
 z^{t+1} &= z^t - u^t \left[\frac{1}{m} \sum_{i=1}^m w_i^t \left(A_i z - \sqrt{y_i} \frac{A_i z}{|A_i z|} \right) A_i^H \right. \\
 &\quad \left. + \lambda^t (z - f(z)) \right]
 \end{aligned}$$

$$\text{where } w_i = \frac{|A_i z|^2 / y_i}{|A_i z|^2 / y_i + \beta_i}, \text{ for } 1 \leq i \leq m$$

4: end

The updated parameter of the (10) is spectrum z , which is complicated and different from natural image. We notice that DD itself acts as a natural data model [23]. So we translate the spectrum information to the image domain based on inverse Fourier transform. Next, we use DD to improve the reconstructed amplitude image. Then, we get the updated spectrum z based on Fourier transform.

The deep decoder is applied every ‘S’ iterations in the proposed algorithm, allowing adequate time for information transfer between the spatial and Fourier domains in the process of phase retrieval. Another reason is to balance the computational time and computer resources.

3) *RAF-RED-DD*: Our approach for FP utilizes RAF-RED (inner loop) together with DD (outer loop), namely, using RAF-RED to get a reconstruction and using DD to further improve the reconstruction of FP. The overall framework of the proposed method is illustrated in Fig. 2. Based on the above derivations, the imaging reconstruction framework of Fourier ptychography is summarized in Algorithm 1.

III. EXPERIMENTAL RESULTS

In this section, we discuss the hyper parameters of the algorithm. We also discuss the performance of our algorithm with numerical simulation data and real data. The results show that the proposed method is robust in different noises.

A. Numerical Simulation

The forward FP imaging process are simulated. First, we choose a high-resolution complex image as an object. We choose ‘cameraman’ as the input amplitude and ‘jetplane’ as the input phase. The input amplitude and phase images are both 512×512 pixels. Then, based on the FP forward imaging model, a series of low-resolution images are generated using (3). The number

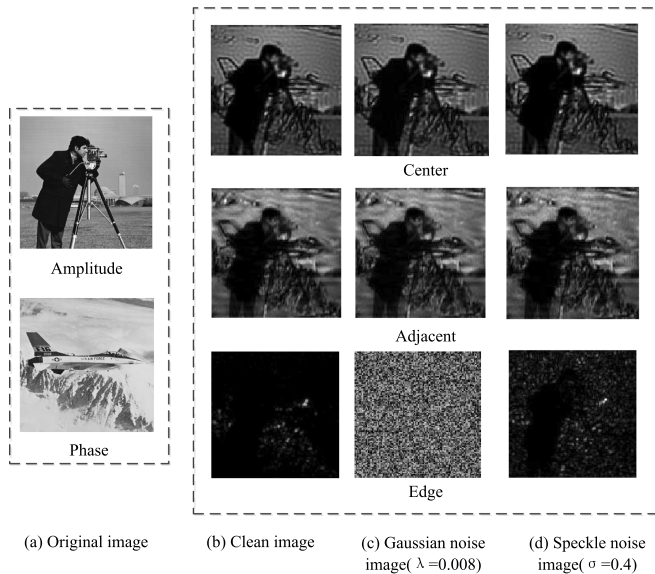


Fig. 4. Gaussian noise ($\lambda = 0.008$) images, speckle noise ($\sigma = 0.4$) images and clean images. Center means that the cameraman image is captured in the center. Adjacent and edge are relative to the center image. All images are normalized to (0-1) for clear display.

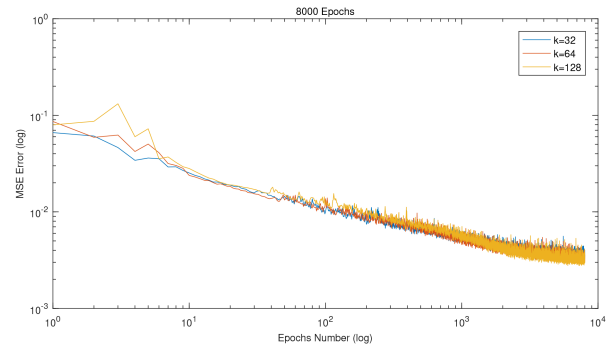
of horizontal and vertical measurements are 15×15 . The total number of captured low-resolution images are 225. The low-resolution images are 102×102 pixels. The overlap ratio is about 65%. Lastly, we add Gaussian noise and speckle noise in the numerical simulation. It should be noted that speckle noise is caused by random phase in the forward imaging processing [29]. Hence, random phase are added to simulate ‘speckle noise’. The image formation model is shown in the left part of Fig. 1. The center clean images, Gaussian noise images and speckle noise images are shown in Fig. 4. For clearer observation of images, all images are normalized to (0–1). We could clearly see the speckle noise from the Fig. 4. We also could see that the Gaussian noise is really strong, especially on the edge of the captured area.

In order to assess the quality of the image reconstructions, the reconstructions of the proposed approach are compared with the true images by using the peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) [30]. For PSNR, larger values are better, and SSIM values closer to 1 (always ranging between 0 and 1) indicate better visual quality.

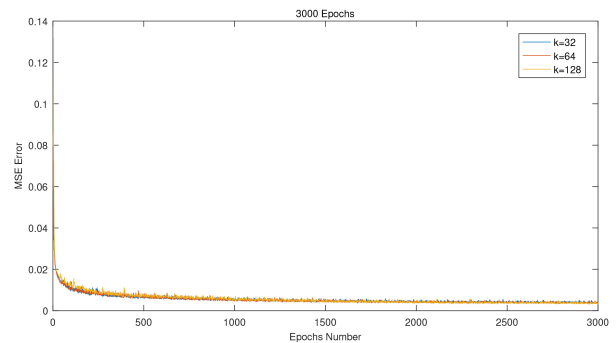
B. Hyper-Parameters of the Proposed Method

1) *The Number of Deep Decoder Layers and Training Epochs*: The number of deep decoder layers K and the number of training epochs are two key hyper-parameters of the deep decoder. K enables trading off amount of noise that is removed versus the representation error by the model. Smaller K may remove more noise, but also increase the error of approximating an image with the deep decoder.

We use mean square error (MSE) to measure the average of the squares of the error of the output of the deep decoder denoising image and the true image. The high-resolution image has Gaussian noise $\lambda = 0.004$ and speckle noise is $\sigma = 0.3$. The



(a) Epochs 8000



(b) Epochs 3000

Fig. 5. MSE of the image deep decoder denoising. The horizontal and vertical axis are log coordinates in the sub-figure (a). We notice that the MSE is not obvious decrease after 2000 epochs. we choose 2000 epochs and $K = 128$.

MSEs of the DD reconstructed result are shown in Fig. 5. We see that about 2000 epochs of DD result in MSE less than 5×10^{-2} , which is small enough. We find that using $K = 128$ gives a better reconstruction result than using $K = 32$ or $K = 64$ after 1500 iterations. We also test other noise level images, and the curves are similar to Fig. 5. Considering the computational time and denoising performance, we choose the number of epochs of DD as 2000 and $K = 128$ in the following experiments.

2) *The Iterations Number of Using Deep Decoder*: In the proposed method, We use DD denoising (every ‘S’ iteration) to improve the RAF-RED reconstruction result. The iteration number of using DD is one key hyper-parameters of the proposed method. We test different iteration numbers, the results are shown in Fig. 6. From Fig. 6, we see that including deep decoder denoising improves accuracy. We notice that decrease the number of ‘S’ will obtain a better result. For example, every 10 iterations using deep decoder is better than other iterations using deep decoder in Fig. 6. However, considering saving computational time, we choose $S = 50$ in the following experiments.

C. Numerical Simulation Results

In the simulations, the standard deviation of Gaussian noise ranges from 0.002 to 0.008 with a 0.002 interval, which is the same as the simulation experiments of WFP method [13]. We add 0.1 to 0.4 times random noise in the phase part to simulate

TABLE I
QUANTITATIVE COMPARISONS OF THE RECONSTRUCTION OF DIFFERENT METHODS WITH THE PROPOSED ALGORITHM AT DIFFERENT GAUSSIAN NOISE LEVELS

Gaussian Noise	$\lambda = 0.002$		$\lambda = 0.004$		$\lambda = 0.006$		$\lambda = 0.008$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
TowardCCA	14.07	0.1982	13.30	0.1333	12.30	0.1067	11.50	0.0927
WFP	21.40	0.5628	21.07	0.5290	20.75	0.4968	20.46	0.4652
RAF-Median	20.33	0.6692	20.75	0.5310	19.54	0.4448	18.70	0.3922
RAF-Median-DD	22.62	0.7794	<u>23.17</u>	0.6834	21.77	0.6074	<u>21.31</u>	0.5345
RAF-BM3D	<u>24.53</u>	<u>0.7933</u>	23.00	<u>0.7081</u>	<u>22.11</u>	<u>0.6615</u>	20.87	<u>0.6117</u>
RAF-BM3D-DD	24.54	0.8076	23.83	0.7390	22.68	0.6793	22.08	0.6482

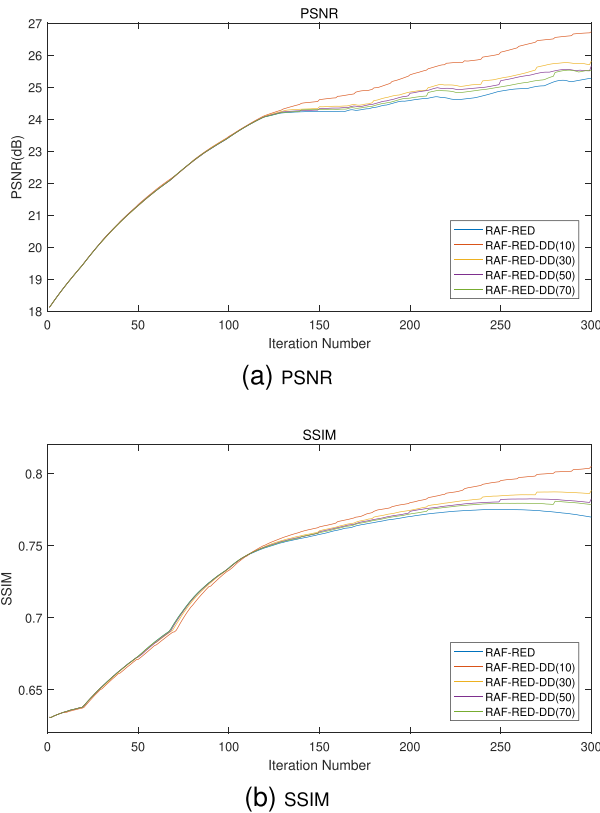


Fig. 6. Quantitative comparison for the reconstruction results of RAF-RED-DD and RAF-RED. RAF-RED-DD gives better results in both PSNR and SSIM. Using small ‘S’ could obtain a better reconstruction result. Considering saving the computation time, we choose $S = 50$ in the following experiments.

speckle noise, which follows the simulation experiments of RAF-RED [11]. The value range of the amplitude is from 0 to 1 and the phase is from $-\pi$ to π .

First, we would like to prove that using deep decoder as an outer loop is useful in the proposed algorithm. We compare our proposed algorithm RAF-RED-DD and RAF-RED (without deep decoder) with the image which Gaussian noise ($\lambda = 0.002$) and speckle noise ($\sigma = 0.1$) are added. From Fig. 6, the results show that RAF-RED-DD yields a better reconstruction than RAF-RED both in PSNR and SSIM. Hence, DD is useful in the proposed algorithm.

We test the RAF-RED-DD algorithm with different Gaussian and speckle noises. We also compare the results with the method of alternating projections method TowardCCA [10], Wirtinger flow optimization for Fourier Ptychography (WFP) [13] and

RAF-RED [11]. In the following experiments, we use median filter and BM3D [24], [31] filter in RED processing. The median filter is simple but always effective, BM3D is a state-of-art image denoiser filter. We notice that BM3D needs a lot of time (about 6 s) to denoise a single image. The proposed algorithm usually needs hundreds of iterations. Using BM3D in every iteration processing, a huge time is then needed for computation.

The experimental results with Gaussian noise are shown in Table I. It can be seen that RAF-RED-DD is better than other methods, and RAF-RED-DD is robust even with high level noise.

The quantitative comparison of the reconstruction results using speckle noise are shown in Table II. It can be seen that RAF-RED-DD obtain the best PSNR results and the best or the second best SSIM results.

We also test with complicated noise situations, which include different levels of Gaussian noise and speckle noise. The reconstruction results are shown in Table III. In summary, we see that RAF-BM3D-DD gives the best reconstruction result with Gaussian noise. Most often RAF-BM3D-DD gives a better reconstruction result than other methods with speckle noise and complicated noise. RAF-BM3D or RAF-Median-DD sometimes gives a better result and RAF-BM3D obtains the second best result with speckle noise and complicated noise.

As shown in Tables I, II, and III, using DD prior as an outer loop in the proposed method, both PSNR and SSIM give a high value. PSNR and SSIM increase by 1.73 dB/0.90 dB (Median/BM3D) and 0.096/0.022 (Median/BM3D) on average with DD and without DD under different noise situations. Hence, the experimental results show that RAF-RED-DD can give a better result than RAF-RED.

A visual representation of the recovered images with complicated noise (Gaussian noise $\lambda = 0.002$ and speckle noise $\sigma = 0.1$) is shown in Fig. 7. The RAF-RED-DD reconstruction is more smooth and better than the other methods.

D. Real Captured Data Experiment Results

To further validate the effectiveness of RAF-RED-DD, we use TowardCCA [10] data for FP reconstruction. The TowardCCA data contains the captured low-resolution images as inputs for FP reconstruction. We compare our proposed method with TowardCCA [10] method using the fingerprint data. The fingerprint data contains 441 (recorded with a grid of 21×21) low-resolution images, with 600×600 pixels. The fingerprint is pressed on a glass slide. The distance of the fingerprint from the camera is about 1.5 m. The reconstructed high-resolution image is $820 \times$

TABLE II
QUANTITATIVE COMPARISONS OF RECONSTRUCTION OF DIFFERENT METHODS WITH THE PROPOSED ALGORITHM AT DIFFERENT SPECKLE NOISE LEVELS

Speckle Noise	$\sigma = 0.1$		$\sigma = 0.2$		$\sigma = 0.3$		$\sigma = 0.4$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
TowardCCA	16.54	0.5937	16.05	0.5262	15.54	0.4759	14.08	0.4226
WFP	21.84	0.5981	21.71	0.5955	21.35	0.5890	20.93	0.5844
RAF-Median	28.02	0.9028	26.22	0.8927	27.10	0.8817	27.76	0.8655
RAF-Median-DD	29.50	0.9017	27.64	0.8895	28.21	0.8837	28.24	0.8656
RAF-BM3D	29.91	0.9130	29.15	0.9133	28.31	0.9055	26.90	0.8766
RAF-BM3D-DD	30.44	0.9067	29.27	0.9023	29.43	0.9130	28.27	0.8747

TABLE III
QUANTITATIVE COMPARISONS OF RECONSTRUCTION OF DIFFERENT METHODS WITH THE PROPOSED ALGORITHM IN COMPLICATED NOISE SITUATIONS

Gaussian Noise Speckle Noise	$\lambda = 0.002$ $\sigma = 0.1$		$\lambda = 0.004$ $\sigma = 0.2$		$\lambda = 0.006$ $\sigma = 0.3$		$\lambda = 0.008$ $\sigma = 0.4$	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
TowardCCA	13.68	0.2036	13.05	0.1337	11.97	0.1074	11.53	0.0874
WFP	21.36	0.5611	20.98	0.5261	20.55	0.4835	20.01	0.4645
RAF-Median	22.33	0.6664	21.42	0.5368	20.49	0.4310	18.74	0.4095
RAF-Median-DD	24.36	0.7715	22.61	0.6908	22.36	0.5978	20.79	0.5601
RAF-BM3D	24.32	0.7776	20.05	0.6614	21.38	0.6527	20.84	0.6415
RAF-BM3D-DD	24.62	0.8074	22.87	0.7448	21.91	0.6674	21.47	0.6628

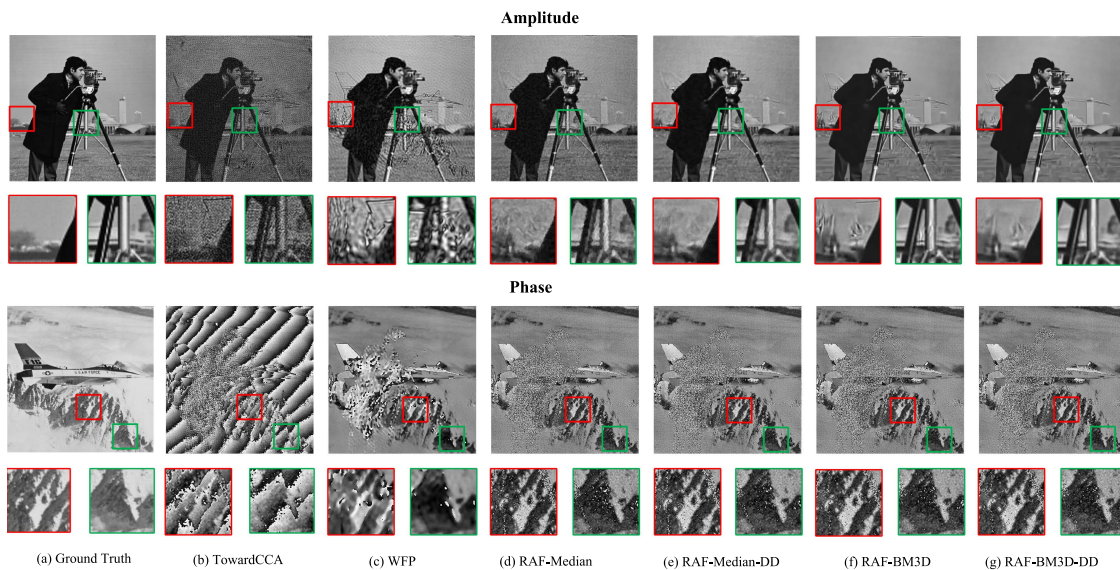


Fig. 7. The recovered amplitude and phase images. From the amplitude reconstruction results, we see that the RAF-BM3D-DD reconstruction is the best. We also notice that RAF-RED-DD is better and smooth than RAF-RED.

820 pixels. In Fig. 8(a), we see that the captured center image is blurred. The reconstructed images are clear and have detail information.

Because the images are real captured data, we could not evaluate them using PSNR or SSIM. We just show their reconstruction results and use visual evaluation. The fingerprint reconstruction results of towardCCA, RAF-BM3D and RAF-BM3D-DD are shown in Fig. 8. We see that the proposed algorithm can obtain a better visual result than the TowardCCA. Compared with Fig. 8(i) and (j), we see that the amplitude result and the phase result of RAF-BM3D-DD contain more detail information than RAF-BM3D, especially in the phase result. We can see the

texture of the fingerprint in the phase result, which is useful for reconstructing details.

E. Discussion and Analysis

Deep learning denoisers (such as DnCNN [18] and FFDNet [19]) are very popular and give good performance. For example, DnCNN achieves state-of-the-art performance on Gaussian denoising tasks. Deep learning based denoisers trained with synthetic clean-noisy image pairs work well with white Gaussian noise. However, real world noises are much more complex than white Gaussian noise, and little work has been done with real

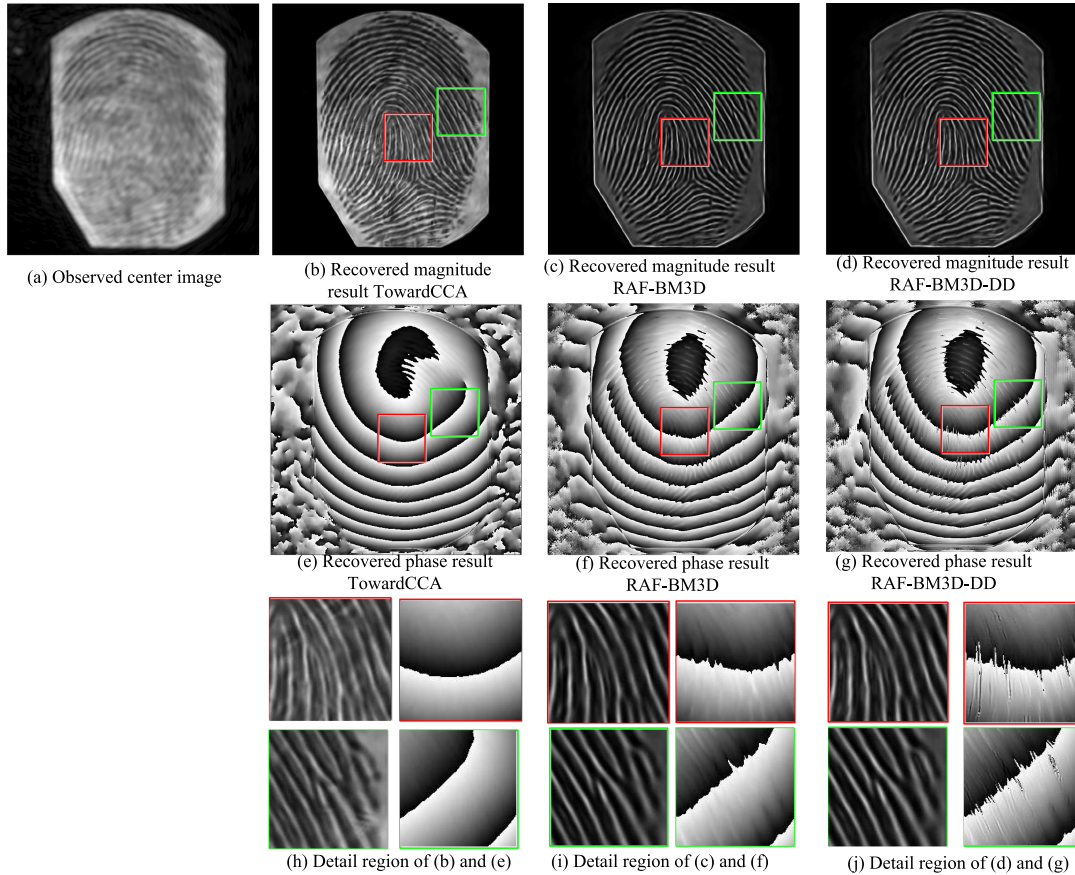


Fig. 8. Fingerprint reconstruction results of TowardCCA, RAF-BM3D and RAF-BM3D-DD. The RAF-BM3D and RAF-BM3D-DD results look more smooth than TowardCCA. The RAF-BM3D-DD result has more details, especially in the phase result.

image denoising, because it is very difficult to capture adequate number of clean-noisy training image pairs. For keeping the proposed algorithm flexible and unbiased with the training images, we do not use any deep learning based denoiser in the RED step.

We test the effect of the number of iterations on the experimental performance. The PSNR and SSIM results of the complicated noise (Gaussian noise $\lambda = 0.002$ and speckle noise $\sigma = 0.1$) are shown in Fig. 9. In the other noise situations, the results are similar. Using more iterations after 300, the improvement is little in our experiments. It should be noted that the number of iterations is a key hyperparameter. To obtain the best results, different iteration numbers should be tested for different experiments.

We also test the proposed method using different sampling measurement numbers with Gaussian noise ($\lambda = 0.002$) and speckle noise ($\sigma = 0.1$). The horizontal and vertical measurement numbers of 13×13 and 17×17 are tested. The results show that with DD could yield a better reconstruction than without DD under different sampling measurement numbers from Table IV.

The TowardCCA and WFP are written in Matlab codes. The other programs are written in python. The TowardCCA and WFP algorithms are tested using Matlab R2020b on a laptop Lenovo W540 with an Intel i7-1075H CPU, 64GM RAM. The RAF-RED and RAF-RED-DD algorithms are tested by Dell T5820, with an Intel i9-10980XE CPU, 32 GB RAM, NVIDIA

TABLE IV
QUANTITATIVE COMPARISONS OF THE RECONSTRUCTION OF DIFFERENT METHODS WITH THE PROPOSED ALGORITHM AT DIFFERENT SAMPLING MEASUREMENTS

Sampling	13×13		15×15		17×17	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
RAF-Median	21.91	0.6508	22.33	0.6664	21.61	0.7013
RAF-Median-DD	23.41	<u>0.7581</u>	<u>24.36</u>	0.7715	24.02	0.7997
RAF-BM3D	23.97	0.7538	24.32	<u>0.7776</u>	<u>24.63</u>	0.8030
RAF-BM3D-DD	24.15	0.8034	24.62	0.8074	24.99	0.8216

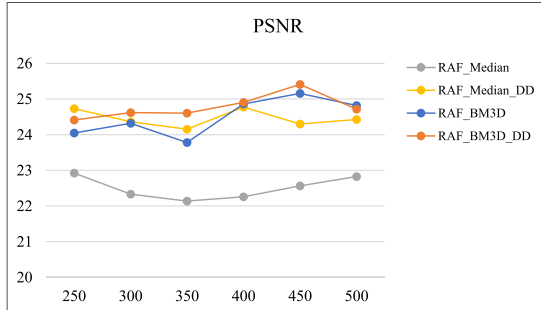
RTX 2080. In RED, we choose median and BM3D filters for RAF-RED and RAF-RED-DD. To reduce computation time, we use median filter at the first 120 iterations of the RAF-BM3D and RAF-BM3D-DD algorithms. The computation time of the above algorithms are shown in Table V.

IV. CONCLUSION

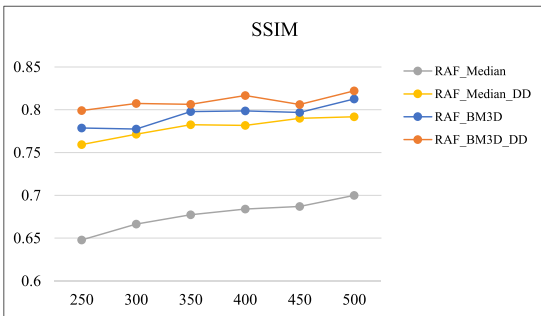
We propose an algorithm for Fourier ptychography based on reweighted amplitude flow with regularization by denoising and deep decoder. The proposed algorithm can iteratively improve the reconstruction results. Considering deep decoder does not need any training images, the proposed algorithm has no bias due to training images. Experiments show that

TABLE V
COMPARISON OF COMPUTATION TIME (S) OF DIFFERENT ALGORITHMS WITH SIMULATION DATA (ITERATION NUMBER: 300)

Methods	TowardCCA	WFP	RAF-Median	RAF-Median-DD	RAF-BM3D	RAF-BM3D-DD
Time(s)	31	112	198	407	1246	1441



(a) PSNR



(b) SSIM

Fig. 9. Quantitative comparison for the reconstruction results of RAF-Median, RAF-Median-DD, RAF-BM3D and RAF-BM3D-DD using different iterations. We choose 300 iterations in our experiments.

the proposed algorithm is a powerful method for improving Fourier ptychography even when the system contains complex noise. The proposed algorithm is flexible since it can easily combine a new image denoiser during regularization by denoising step. Using deep decoder as an outer loop denoiser can also be used in phase retrieval and other image inverse problems.

ACKNOWLEDGMENT

Baopeng Li would like to thank Zhixin Li for her kind and constructive discussion about the Fourier Ptychography methods, which helped to improve the quality of the proposed method.

REFERENCES

- [1] G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," *Nature Photon.*, vol. 7, no. 9, pp. 739–745, 2013.
- [2] G. Zheng, "Breakthroughs in photonics 2013: Fourier ptychographic imaging," *IEEE Photon. J.*, vol. 6, no. 2, pp. 1–7, Apr. 2014.
- [3] L. Tian, X. Li, K. Ramchandran, and L. Waller, "Multiplexed coded illumination for Fourier ptychography with an LED array microscope," *Biomed. Opt. Exp.*, vol. 5, no. 7, pp. 2376–2389, 2014.
- [4] S. Chen, T. Xu, J. Zhang, X. Wang, and Y. Zhang, "Optimized denoising method for Fourier ptychographic microscopy based on wirtinger flow," *IEEE Photon. J.*, vol. 11, no. 1, pp. 1–14, Feb. 2019.
- [5] G. Zheng, X. Ou, R. Horstmeyer, J. Chung, and C. Yang, "Fourier ptychographic microscopy: A gigapixel superscope for biomedicine," *Opt. Photon. News*, vol. 25, no. 4, pp. 26–33, 2014.
- [6] J. Zhang, T. Xu, J. Liu, S. Chen, and X. Wang, "Precise brightfield localization alignment for Fourier ptychographic microscopy," *IEEE Photon. J.*, vol. 10, no. 1, pp. 1–13, Feb. 2018.
- [7] H.-Y. Liu, E. Jonas, L. Tian, J. Zhong, B. Recht, and L. Waller, "3D imaging in volumetric scattering media using phase-space measurements," *Opt. Exp.*, vol. 23, no. 11, pp. 14461–14471, 2015.
- [8] L. Tian and L. Waller, "3D intensity and phase imaging from light field measurements in an LED array microscope," *Optica*, vol. 2, no. 2, pp. 104–111, 2015.
- [9] J. Holloway, Y. Wu, M. K. Sharma, O. Cossairt, and A. Veeraraghavan, "SAVI: Synthetic apertures for long-range, subdiffraction-limited visible imaging using Fourier ptychography," *Sci. Adv.*, vol. 3, no. 4, 2017, Art. no. e1602564.
- [10] J. Holloway et al., "Toward long-distance subdiffraction imaging using coherent camera arrays," *IEEE Trans. Comput. Imag.*, vol. 2, no. 3, pp. 251–265, Sep. 2016.
- [11] Z. Li, D. Wen, Z. Song, G. Liu, W. Zhang, and X. Wei, "Sub-diffraction visible imaging using macroscopic Fourier ptychography and regularization by denoising," *Sensors*, vol. 18, no. 9, 2018, Art. no. 3154.
- [12] C. Zuo, J. Sun, and Q. Chen, "Adaptive step-size strategy for noise-robust Fourier ptychographic microscopy," *Opt. Exp.*, vol. 24, no. 18, pp. 20724–20744, Sep. 2016.
- [13] L. Bian, J. Suo, G. Zheng, K. Guo, F. Chen, and Q. Dai, "Fourier ptychographic reconstruction using Wirtinger flow optimization," *Opt. Exp.*, vol. 23, no. 4, pp. 4856–4866, 2015.
- [14] L.-H. Yeh et al., "Experimental robustness of Fourier ptychography phase retrieval algorithms," *Opt. Exp.*, vol. 23, no. 26, pp. 33214–33240, 2015.
- [15] J. Chen, J. Chen, H. Chao, and Y. Ming, "Image blind denoising with generative adversarial network based noise modeling," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 3155–3164.
- [16] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C.-W. Lin, "Deep learning on image denoising: An overview," *Neural Netw.*, vol. 131, pp. 251–275, 2020.
- [17] C. Zuo et al., "Deep learning in optical metrology: A review," *Light: Sci. Appl.*, vol. 11, no. 4, 2022, Art. no. 54.
- [18] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.
- [19] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Trans. Image Process.*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018.
- [20] Z. Yue, H. Yong, Q. Zhao, D. Meng, and L. Zhang, "Variational denoising network: Toward blind noise modeling and removal," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, vol. 32, pp. 1690–1701.
- [21] Z. Yue, Q. Zhao, L. Zhang, and D. Meng, "Dual adversarial network: Toward real-world noise removal and noise generation," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 41–58.
- [22] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 9446–9454.
- [23] R. Heckel and P. Hand, "Deep decoder: Concise image representations from untrained non-convolutional networks," in *Proc. Int. Conf. Learn. Representations*, 2019, pp. 1–17.
- [24] K. Dabov, A. Foi, V. Katkovich, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.
- [25] S. V. Venkatakrisnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *Proc. Glob. Conf. Signal Inf. Process.*, 2013, pp. 945–948.

- [26] Y. Romano, M. Elad, and P. Milanfar, "The little engine that could: Regularization by denoising (red)," *SIAM J. Imag. Sci.*, vol. 10, no. 4, pp. 1804–1844, 2017.
- [27] G. Wang, G. B. Giannakis, Y. Saad, and J. Chen, "Phase retrieval via reweighted amplitude flow," *IEEE Trans. Signal Process.*, vol. 66, no. 11, pp. 2818–2833, Jun. 2018.
- [28] C. Metzler, P. Schniter, A. Veeraraghavan, and R. Baraniuk, "prDeep: Robust phase retrieval with a flexible deep network," in *Proc. Int. Conf. Mach. Learn.*, 2018, pp. 3501–3510.
- [29] C.-F. Cheng, D.-P. Qi, D.-L. Liu, and S.-Y. Teng, "The computational simulations of the Gaussian correlation random surface and its light-scattering speckle field and the analysis of the intensity probability density," *Acta Physica Sinica*, vol. 48, no. 9, pp. 1635–1643, 1999.
- [30] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [31] Y. Mäkinen, L. Azzari, and A. Foi, "Exact transform-domain noise variance for collaborative filtering of stationary correlated noise," in *Proc. IEEE Int. Conf. Image Process.*, 2019, pp. 185–189.