DEEP PTYCH: SUBSAMPLED FOURIER PTYCHOGRAPHY USING GENERATIVE PRIORS

Fahad Shamshad, Farwa Abbas, Ali Ahmed

Department of Electrical Engineering, Information Technology University, Lahore, Pakistan

ABSTRACT

This paper proposes a novel framework to regularize the highly ill-posed and non-linear Fourier ptychography problem using generative models. We demonstrate experimentally that our proposed algorithm, *Deep Ptych*, outperforms the existing Fourier ptychography techniques, in terms of quality of reconstruction and robustness against noise, using far fewer samples. We further modify the proposed approach to allow the generative model to explore solutions outside the range, leading to improved performance.

Index Terms— Fourier ptychography, phase retrieval, generative models, noise robustness, subsampling

1. INTRODUCTION

Resolution loss in long distance imaging can primarily be attributed to the diffraction blur, that is caused by limited aperture of the imaging system [1]. To mitigate the effects of the diffraction blur, recently an emerging computational imaging technique known as Fourier Ptychography (FP) has shown promising results [2, 3]. The objective of FP is to recover a high-resolution image from multiple diffraction-limited low-resolution images. In this paper, we consider recovering the signal $\boldsymbol{x} \in \mathbb{C}^n$ captured via forward acquisition model of FP, given by:

$$\boldsymbol{y}_\ell = |\mathcal{A}_\ell(\boldsymbol{x})| + \boldsymbol{n}_\ell, \quad \text{for } \ell = 1, 2, ..., L, \tag{1}$$

where $\boldsymbol{y}_\ell \in \mathbb{R}^n$ is diffraction-limited image corresponding to ℓ^{th} camera, $\mathcal{A}_\ell : \mathbb{C}^n \to \mathbb{C}^n$ is the linear operator representing the forward acquisition model, and $\boldsymbol{n}_\ell \in \mathbb{R}^n$ denotes noise perturbation. For ℓ^{th} camera, the linear operator $\mathcal{A}_\ell$ has the form $\mathcal{F}^{-1}\mathcal{P}_\ell \circ \mathcal{F}$, where $\mathcal{F}$ denotes 2D Fourier transform, $\mathcal{P}_\ell$ is a pupil mask that acts as a bandpass filter in the Fourier domain, and $\circ$ represents the Hadamard product (details are provided in subsequent sections). Specifically, FP works by iteratively stitching together a sequence of frequency limited low-resolution images $\boldsymbol{y}_\ell$ in Fourier domain to recover the high-resolution true image $\boldsymbol{x}$. Since optical sensors can measure only the magnitude of the signal [4], phase information is lost during the acquisition process — making the FP problem highly ill-posed.

Traditional approaches to overcome the ill-posedness of FP are iterative phase retrieval algorithms [5]. Generally, phase retrieval algorithms fall into two categories. The *first* approach is to introduce redundancy into the measurement system, where more measurements are taken than the dimension of true signal usually in the form of oversampled Fourier transform [6], short-time Fourier transform [7], structured illuminations [8], etc. The *second* approach is to exploit the structural assumptions on the true signal as a prior, such as sparsity [9] or non-negativity [10]. These prior based phase retrieval approaches have recently attained much significance for the purpose of reducing number of measurements as the acquisition of measurements is usually expensive and time-consuming, especially at high resolutions.

FP recovery algorithms that exploit redundant information perform well provided the set of low-resolution images $\boldsymbol{y}_\ell$ have high overlapping frequency bands (typically 60% overlap) in the Fourier domain [11]. The reconstruction quality of these algorithms quickly degrades with the decrease in overlap. This high overlap requirement for faithful reconstruction of the true image $\boldsymbol{x}$ results in long acquisition times and high computational cost [2].

To reduce the computational cost, algorithms that exploit prior information about true signal have recently been explored for FP. Among them, priors learned from large datasets by utilizing the power of deep neural networks have shown promising results for low overlap case, thus reducing computational cost [12, 13, 14]. Specifically, these deep learning based approaches invert the forward acquisition model of FP via end-to-end training of deep neural networks in a supervised manner [15]. However, even a slight change in the noise level $\boldsymbol{n}_\ell$ or parameters of FP measurement model, such as aperture diameter or overlap would require costly retraining of these models. Apart from reducing overlap, an alternative approach to reduce computational cost for FP, is to design realistic subsampling strategies by exploiting some prior information about the true signal. This alternative has not yet been investigated by the deep learning based FP approaches. Recently, Jagatap et al. [16] explored this alternative by leveraging sparsity prior. The resulting algorithm (CoPRAM - *Compressive Phase Retrieval using Alternative Minimization*) significantly reduces the number of samples required for faithful reconstruction of the true signal $\boldsymbol{x}$ in FP setup. However, it has been shown in [17] that sparsity priors often fail to capture the complex structure that many natural signals exhibit resulting in unrealistic signals also fitting the sparse prior modelling assumption.

Recently, neural network based implicit generative models such as Generative Adversarial Networks (GANs) [18] and Variational Autoencoders (VAEs) [19] have been quite successful in modelling complex data distributions especially that of images. In these models, the generative part (G), learns a mapping from low dimensional latent space $\boldsymbol{z} \in \mathbb{R}^k$ to a high dimensional sample space $G(\boldsymbol{z}) \in \mathbb{R}^n$ where $k \ll n$. During training, these generative models are encouraged to produce samples that resemble with that of training data $\mathcal{X}$. A well-trained generator, given by deterministic function $G : \mathbb{R}^k \to \mathbb{R}^n$ with a distribution P_Z over $\boldsymbol{z}$ (usually random normal or uniform), is therefore capable of generating fake data indistinguishable from the real data it has been trained on. Due to their power to effectively model natural image distributions, generative priors have recently been introduced to solve ill-posed inverse problems like compressed sensing [20], phase retrieval [21, 22], blind deconvolution [23], etc. Notably, these generative prior based approaches, have been shown to improve over sparsity-based approaches, thus advancing the state of the art in several fields [24].

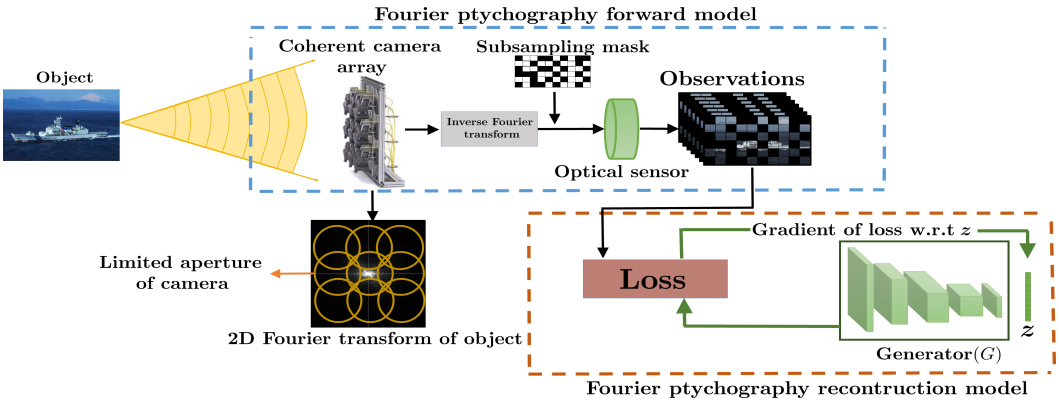


Fig. 1. *Overview of Fourier ptychography forward model and proposed reconstruction algorithm. A coherent camera array captures an image of the object. The bandlimited signal is then focused to an image plane and a subsampling operator is applied. Subsequently, an optical sensor measures the magnitude while discarding the phase of signal. During the reconstruction phase, the generator (G) optimizes a latent code $\boldsymbol{z}$ via gradient descent algorithm to find a corresponding $G(\boldsymbol{z})$ that best explains the observations.*

1.1. Our Contributions

In this paper, we leverage the expressive power of generative models to regularize FP problem. To the best of our knowledge, this is the first work that explores the generative models for FP setup. In particular, we make the following contributions:

1. We combine the expressive power of deep generative models (GANs and VAEs) with non-linear inverse problem of FP for the first time. We show that the resultant problem can be solved effectively using conventional gradient descent algorithm yielding promising results. Moreover, unlike deep learning based approaches [12, 14], our approach does not require expensive retraining in case of slight changes in parameters such as subsampling ratio or noise level.
2. We show experimentally that the proposed approach produces impressive results for low subsampling ratios and high noise.
3. Finally, to further improve the performance of the proposed scheme, especially for the rich image datasets not well learned by the generator, we present a modification in the loss function that allows more flexible reconstructions.

2. FORMULATION

In this section, we describe the forward acquisition model of FP for long-distance imaging following the setup of [2], as shown in Fig. 1. In long-distance imaging, resolution of the image is limited by diffraction of light at the limited aperture of the imaging system. To emulate capturing the scene with a large synthetic aperture, multiple cameras (L) are deployed, usually in a square grid. Such an arrangement is called a coherent camera array. In long-distance imaging, coherent camera array is usually placed in far field of the object ($\boldsymbol{x}$), and the Fraunhofer approximation is satisfied. Under this approximation, illumination field emerging from the object when intercepted by the camera array can be treated as 2D Fourier transform ($\mathcal{F}$) of the image at the object plane [25]. Each camera in the array have a limited aperture denoted by $\mathcal{P}_\ell$ ($\ell = 1, 2, ..., L$) that acts as a bandpass filter covering different parts of Fourier domain image as shown in Fig. 1. The bandlimited signal is then focused to an image plane where this propagation to the image plane can be modelled as an inverse Fourier transform ($\mathcal{F}^{-1}$). After that, a subsampling operator ($\mathcal{M}_\ell(\cdot)$), is applied to effectively reduce the number of measurements. Subsequently, the complex spatial domain image is captured by an optical sensor that measures only the magnitude while discarding the phase information. Mathematically, observation $\boldsymbol{y}_\ell$ for ℓth camera can be modelled as

$$\boldsymbol{y}_\ell = |\mathcal{M}_\ell(\mathcal{A}_\ell(\boldsymbol{x}))| + \boldsymbol{n}_\ell, \tag{2}$$

where $\mathcal{A}_\ell = \mathcal{F}^{-1}\mathcal{P}_\ell \circ \mathcal{F}$ is the measurement model prior to optical sensor acquisition step, $\circ$ denotes the Hadamard product and $\mathcal{M}_\ell$ is the subsampling operator. Subsampling operator when applied to an image, randomly picks a fraction of samples (f) discarding the others [16]. We define the subsampling ratio as the fraction of samples retained by $\mathcal{M}_\ell$ divided by the total number of observed samples i.e.

$$\text{Subsampling Ratio (\%)} = \frac{\text{Fraction of samples retained } (f) \times 100}{\text{Total observed samples } (nL)}.$$

3. DEEP PTYCH

In this section, we formally introduce our proposed approach *Deep Ptych*. We assume that the true image $\boldsymbol{x} \in \mathbb{R}^n$ is a member of some structured class of images, denoted by $\mathcal{X}$. For example, $\mathcal{X}$ may be set of face images or digits. A generative model is trained on a representative sample set from class $\mathcal{X}$. Given low-dimensional input vector $\boldsymbol{z} \in \mathbb{R}^k$, the generator G after training, generates new samples $G(\boldsymbol{z})$ similar to the representative samples of class $\mathcal{X}$. Once trained, the weights of the generator are kept fixed (pretrained generator). To recover the estimate of true image $\hat{\boldsymbol{x}}$ from FP measurements $\boldsymbol{y}_\ell$ for $\ell = 1, 2, ..., L$, in (2), we propose minimizing the following objective function

$$\hat{\boldsymbol{x}} := \underset{\boldsymbol{x} \in \text{Range}(G)}{\text{argmin}} \sum_{\ell=1}^{L} \|\boldsymbol{y}_\ell - |\mathcal{M}_\ell(\mathcal{A}_\ell(x))|\|_2^2, \tag{3}$$

where Range(G) is the set of all the images that can be generated by pretrained G. In other words, we want to find an image $\hat{\boldsymbol{x}}$ that is closest to the observed measurements and lies within the range of the generator. Ideally, the range of the generator must comprise of only the samples drawn from the distribution of the images it has

Algorithm 1 Deep Ptych

Input: $\boldsymbol{y}$, $\mathcal{M}$, $\mathcal{A}$, G, and η
Output: Estimate $\hat{\boldsymbol{x}}$
Initialize:
$\boldsymbol{z}_0 := \mathcal{N}(0, I_k)$
for $t = 1, 2, 3, \dots T$ **do**
$\boldsymbol{z}_{t+1} \leftarrow \boldsymbol{z}_t$ - $\eta \nabla_{\boldsymbol{z}_t} \mathcal{L}(\boldsymbol{z}_t)$.
end for
$\hat{\boldsymbol{x}} \leftarrow G(\boldsymbol{z}_T)$

been trained on, i.e., $\exists\, \boldsymbol{z}$ such that $\boldsymbol{x} = G(\boldsymbol{z})$. It must be noted that constraining the solution $\hat{\boldsymbol{x}}$ to lie inside the generator range will force the solution to be a member of image class $\mathcal{X}$ thus, implicitly reducing the solution ambiguities associated with FP [5]. The minimization program in (3) can be equivalently formulated in the lower dimensional, latent representation $\boldsymbol{z}$ as follows:

$$\hat{\boldsymbol{z}} = \underset{\boldsymbol{z} \in \mathbb{R}^k}{\operatorname{argmin}} \sum_{\ell=1}^{L} \|\boldsymbol{y}_\ell - |\mathcal{M}_\ell(\mathcal{A}_\ell(G(\boldsymbol{z})))|\|_2^2. \tag{4}$$

This optimization program modifies the latent representation vector $\boldsymbol{z}$ such that the generator generates an image $G(\boldsymbol{z})$ that is consistent with (4). Since the latent representation $\boldsymbol{z}$ is much smaller in dimension as compared to the ambient dimension of the true image $\boldsymbol{x}$, it not only reduces the number of unknowns in the FP problem but also allows an efficient implementation of gradients using backpropagation through the generators. Since optimization program in (4) is non-convex and non-linear owing to the modulus operator and non-linear generative model, we resort to the gradient descent algorithm to find an optimal $\hat{\boldsymbol{z}}$. The estimated image $\hat{\boldsymbol{x}}$ is acquired by a forward pass of the $\hat{\boldsymbol{z}}$ through the generator G. Mathematically, $\hat{\boldsymbol{x}} = G(\hat{\boldsymbol{z}})$. The gradient descent scheme, namely, *Deep Ptych* is formally given in Algorithm 1, where the objective function in (4) is denoted by $\mathcal{L}(\boldsymbol{z})$, that is, $\mathcal{L}(\boldsymbol{z}) := \sum_{\ell=1}^{L} \|\boldsymbol{y}_\ell - |\mathcal{M}_\ell(\mathcal{A}_\ell(G(\boldsymbol{z})))|\|_2^2$.

3.1. Beyond the range of generator: *Deep Ptych+*

The shortcoming of the generative network in effectively learning the distribution of the image dataset, especially more complex ones, hampers its ability to reliably reproduce a given new (test) sample from its range [26]. To address this shortcoming, we allow the reconstructed image to deviate a bit from the range of the generator if it improves the measurement loss. To achieve this, we propose solving the following modified version of the optimization program in (4)

$$(\hat{\boldsymbol{x}}, \hat{\boldsymbol{z}}) = \underset{\boldsymbol{z} \in \mathbb{R}^k,\, \boldsymbol{x} \in \mathbb{R}^n}{\operatorname{argmin}} \sum_{\ell=1}^{L} \|\boldsymbol{y}_\ell - |\mathcal{M}_\ell(\mathcal{A}_\ell(\boldsymbol{x}))|\|_2^2 + \lambda \|\boldsymbol{x} - G(\boldsymbol{z})\|_2^2. \tag{5}$$

The first term in the objective favors an image $\boldsymbol{x}$ with smaller measurement loss, while the second term ensures that in doing so such an $\boldsymbol{x}$ should not deviate too far away from the range of the pretrained generative network. The free parameter $\lambda \geq 0$ controls the degree to which we are willing to deviate away from the range of the generator, for example, one might choose a smaller λ when the generator is known to not reliably reproduce the image samples from the dataset.

To minimize (5), we use an alternating gradient descent scheme that takes gradient steps in one of the unknowns $\boldsymbol{x}$ or $\boldsymbol{z}$ while keeping the other fixed. This enables us to land at the local minima $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{z}})$ of the objective in (5). We take $\hat{\boldsymbol{x}}$ to be the reconstructed image, which generally improves over the reconstructions from Deep Pytch algorithm above. We refer to this modified approach as Deep Ptych+.

4. EXPERIMENTAL RESULTS

We now provide simulation experiments to evaluate the performance of *Deep Ptych* and *Deep Ptych+*. To quantitatively evaluate the performance of our algorithm, we use two metrics Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). In all our experiments, we report results on a held out test set, unseen by the generative model during training.

4.1. Datasets and Generative models

We use three datasets (MNIST [27], CelebA (64×64) [28], and high-resolution CelebA (128×128) [29]) and two generative models (VAE and GAN). This provides some evidence that our approach is generalizable to many types of datasets and generative models.

MNIST with VAE: MNIST consists of 28×28 grayscale images of handwritten digits, with 60,000 training and 10,000 test examples. We super-resolve original MNIST dataset to size 56×56 using super-resolution method proposed in [30], that gives promising results without requiring external training data. We train VAE, having same architecture as proposed in [23] for image datasets, on this super-resolved dataset with size of latent dimension set to 50, batch size of 128, λ set to 2×10^{-4}, and learning rate of 5×10^{-4} with Adam optimizer.

CelebA with DCGAN: We use aligned and cropped version of CelebA dataset having more than 200,000 RGB face images of size 64×64×3. Each pixel value is scaled between $[-1, 1]$. Deep convolutional generative adversarial network (DCGAN) is trained on this rescaled dataset with architecture as proposed in [31] having latent dimension size set to 100, batch size of 64, β_1 set to 0.5, λ set to 1.4×10^{-4}, and learning rate of 2×10^{-4} using Adam optimizer.

CelebA (128×128) with Progressive GAN [29]: To show the dependence of *Deep Ptych* on the expressive power of generator, we use a powerful generative model named progressive GAN [29], trained on 128×128 face images. We use the originally proposed architecture having the latent dimension set to 512.

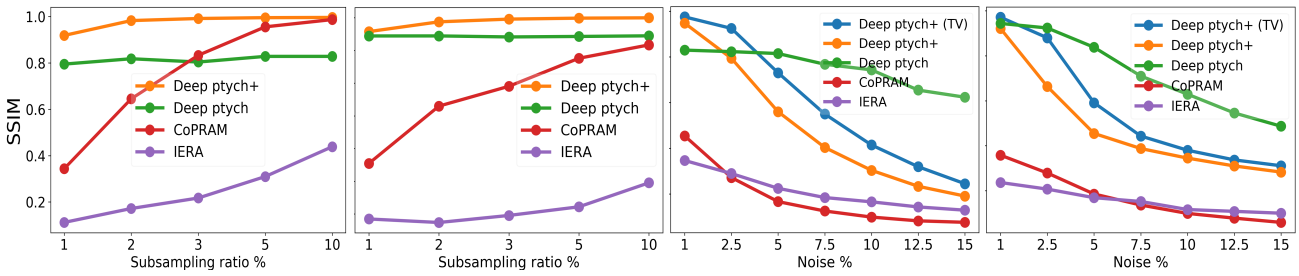
4.2. Baseline methods

We consider IERA (Iterative Error Reduction Algorithm) [2] and CoPRAM [16] as baseline methods. IERA solves FP problem by alternatively enforcing spatial and Fourier domain constraints whereas CoPRAM exploits the underlying sparse structure of true signal for faithful reconstruction at low subsampling ratios. For CoPRAM, we assume sparsity for MNIST and CelebA in spatial and Fourier domains respectively. We use the default algorithmic parameters for IERA and CoPRAM, unless stated otherwise.

4.3. Subsampling results

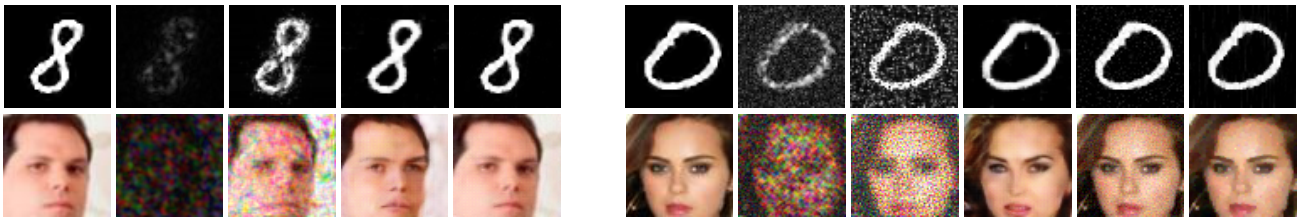
We observed that CoPRAM does not work well with additive noise therefore, to carry out a fair comparison, we perform all subsampling experiments without noise. For MNIST, we run gradient descent algorithm using Adam optimizer with learning rate of 0.05 for 2000 steps. We set aperture diameter to 15 pixels and overlap to 65% with camera array consisting of 81 cameras (square grid of 9×9). We use same parameters for CelebA, except learning rate and aperture diameter, that are set to 0.01 and 16 pixels respectively.

We observe that *Deep Ptych* and *Deep Ptych+* are able to attain much higher SSIM and PSNR values at low subsampling ratios as shown in Fig. 2 and Table 1 respectively. The Qualitative results at 2% subsampling ratio show faithful reconstruction of *Deep*



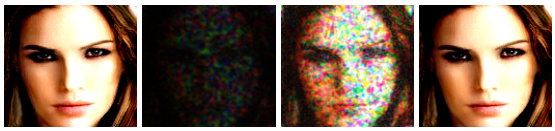
(a) CelebA (b) MNIST (c) CelebA (d) MNIST

Fig. 2. *SSIM plots of proposed approaches against baseline methods for different subsampling ratios and noise levels for CelebA ((a) and (c)) and MNIST ((b) and (d)) respectively. Results are averaged over 10 test images of each dataset.*



(a) Original (b) IERA (c) CoPRAM (d) DP (e) DP+ (a) Original (b) IERA (c) CoPRAM (d) DP (e) DP+ (f) DP+(TV)

Fig. 3. ***(Left)**: Subsampling results for MNIST and CelebA for 2% subsampling ratio. Results of Deep Ptych (DP)(d) and Deep Ptych+ (DP+)(e) are far superior as compared to IERA (b) and CoPRAM (c). (**Right**): Noise results for MNIST and CelebA for 2.5% noise. Results of DP, DP+, and DP+ (Total Variation) as shown in (d), (e), and (f) are visually appealing as compared to IERA (b) and CoPRAM (c).*



(a) Original (b) IERA (c) CoPRAM (d) Deep Ptych

Fig. 4. *Deep Ptych results for 128×128 face images with progressive GAN generator for 1% subsampling ratio and aperture diameter of 30 pixels. Using powerful generator, results are realistic and visually appealing as compared to IERA (b) and CoPRAM (c).*

	MNIST				CelebA			
	Subsampling		Noise		Subsampling		Noise	
	1%	3%	1%	10%	1%	3%	1%	10%
IERA	10.59	11.72	14.15	9.59	5.89	7.29	12.36	11.25
CoPRAM	14.17	19.16	16.63	6.45	13.95	23.85	17.53	9.34
FP+(TV)	-	-	27.10	12.26	-	-	26.41	15.81
DP	24.70	24.80	25.31	**19.27**	22.54	22.79	23.57	**21.21**
DP+	**29.78**	**39.39**	34.83	15.20	**27.99**	**38.72**	32.08	14.00
DP+(TV)	-	-	**35.98**	16.82	-	-	**33.61**	16.46

Table 1. PSNR (dB) results for MNIST and CelebA.

Ptych and *Deep Ptych+* as compared to IERA and CoPRAM results. Since output of *Deep Ptych* is constrained to lie in generator range, increasing subsampling ratio results in its performance saturation. CoPRAM has no such limitation so its performance eventually surpasses *Deep ptych* at higher subsampling ratios. We show in Fig. 4 that for a more expressive generator this range constraint becomes less effective, resulting in visually appealing reconstructions for *Deep Ptych*. However, *Deep Ptych+* outperforms IERA and CoPRAM at all subsampling ratios.

4.4. Noise results

For all experiments in this section a noise of 1%, for image scaled between 0 to 1, translates to Gaussian noise with zero mean and a standard deviation of 0.01. All parameter values are set to same as in subsampling experiments with subsampling ratio set to 10%.

Quantitative results in Fig. 2 and Table 1 manifest that *Deep Ptych* exhibits far superior performance than baseline approaches for high noise levels. Conversely, performance of *Deep Ptych+* is superior to IERA and CoPRAM, but degrades quickly by increasing noise percentage. This is because its output is not constrained to lie in generator range (that spans the set of representative samples of noiseless training data). We find that by adding total variation (TV) regularization (with weighting factor of 10^{-4}) to *Deep Ptych+* loss function, the performance is improved (generally $0.5 - 2$dB gain in PSNR) at all noise levels. We dubbed this modified *Deep Ptych+* as *Deep Ptych+ (TV)*. We also show PSNR results in Table 1 by regularizing FP with TV regularization i.e. $\mathcal{L}(\boldsymbol{x}) = \|\boldsymbol{y}_\ell - |\mathcal{M}_\ell(\mathcal{A}_\ell(\boldsymbol{x}))|\|_2^2 + \|\boldsymbol{x}\|_{TV}$, and dubbed it as FP+(TV). Qualitatively, as depicted in Fig. 3, the proposed approaches outperform baseline methods in terms of quality of reconstruction at 2.5% noise.

5. CONCLUSION

To conclude, we demonstrated the effectiveness of integrating deep generative priors with Fourier ptychography problem. We showed, experimentally, that the proposed approach is effective at low subsampling ratios and is highly robust to noise. We further refined the proposed algorithm to allow the generative model to explore solutions outside the range, leading to improved performance.

6. REFERENCES

[1] Jason Holloway, Yicheng Wu, Manoj K Sharma, Oliver Cossairt, and Ashok Veeraraghavan, "Savi: Synthetic apertures for long-range, subdiffraction-limited visible imaging using fourier ptychography," *Science advances*, vol. 3, no. 4, pp. e1602564, 2017.

[2] Jason Holloway, M Salman Asif, Manoj Kumar Sharma, Nathan Matsuda, Roarke Horstmeyer, Oliver Cossairt, and Ashok Veeraraghavan, "Toward long-distance subdiffraction imaging using coherent camera arrays," *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 251–265, 2016.

[3] Guoan Zheng, Roarke Horstmeyer, and Changhuei Yang, "Wide-field, high-resolution fourier ptychographic microscopy," *Nature photonics*, vol. 7, no. 9, pp. 739, 2013.

[4] Yoav Shechtman, Yonina C Eldar, Oren Cohen, Henry Nicholas Chapman, Jianwei Miao, and Mordechai Segev, "Phase retrieval with application to optical imaging: a contemporary overview," *IEEE signal processing magazine*, vol. 32, no. 3, pp. 87–109, 2015.

[5] Chao Yang, Jianliang Qian, Andre Schirotzek, FRNC Maia, and Stefano Marchesini, "Iterative algorithms for ptychographic phase retrieval," *arXiv preprint arXiv:1105.5628*, 2011.

[6] Tamir Bendory, Robert Beinert, and Yonina C Eldar, "Fourier phase retrieval: Uniqueness and algorithms," in *Compressed Sensing and its Applications*, pp. 55–91. Springer, 2017.

[7] Tamir Bendory, Yonina C Eldar, and Nicolas Boumal, "Non-convex phase retrieval from stft measurements," *IEEE Transactions on Information Theory*, vol. 64, no. 1, pp. 467–484, 2018.

[8] Emmanuel J Candes, Xiaodong Li, and Mahdi Soltanolkotabi, "Phase retrieval from coded diffraction patterns," *Applied and Computational Harmonic Analysis*, vol. 39, no. 2, pp. 277–299, 2015.

[9] Edouard Jean Robert Pauwels, Amir Beck, Yonina C Eldar, and Shoham Sabach, "On fienup methods for sparse phase retrieval," *IEEE Transactions on Signal Processing*, vol. 66, no. 4, pp. 982–991, 2018.

[10] Robert Beinert, "Non-negativity constraints in the one-dimensional discrete-time phase retrieval problem," *Information and Inference: A Journal of the IMA*, vol. 6, no. 2, pp. 213–224, 2017.

[11] Jianliang Qian, Chao Yang, A Schirotzek, F Maia, and S Marchesini, "Efficient algorithms for ptychographic phase retrieval," *Inverse Problems and Applications, Contemp. Math*, vol. 615, pp. 261–280, 2014.

[12] Armin Kappeler, Sushobhan Ghosh, Jason Holloway, Oliver Cossairt, and Aggelos Katsaggelos, "Ptychnet: Cnn based fourier ptychography," in *Image Processing (ICIP), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1712–1716.

[13] Thanh Nguyen, Yujia Xue, Yunzhe Li, Lei Tian, and George Nehmetallah, "Deep learning approach for fourier ptychography microscopy," *Optics Express*, vol. 26, no. 20, pp. 26470–26484, 2018.

[14] Lokesh Boominathan, Mayug Maniparambil, Honey Gupta, Rahul Baburajan, and Kaushik Mitra, "Phase retrieval for fourier ptychography under varying amount of measurements," *arXiv preprint arXiv:1805.03593*, 2018.

[15] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning," *nature*, vol. 521, no. 7553, pp. 436, 2015.

[16] Gauri Jagatap, Zhengyu Chen, Chinmay Hegde, and Namrata Vaswani, "Sub-diffraction imaging using fourier ptychography and structured sparsity," in *Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.(ICASSP)*, 2018.

[17] Christopher A Metzler, Arian Maleki, and Richard G Baraniuk, "Bm3d-prgamp: compressive phase retrieval based on bm3d denoising," in *Multimedia & Expo Workshops (ICMEW), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–2.

[18] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[19] Diederik P Kingma and Max Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.

[20] Ashish Bora, Ajil Jalal, Eric Price, and Alexandros G Dimakis, "Compressed sensing using generative models," *arXiv preprint arXiv:1703.03208*, 2017.

[21] Fahad Shamshad and Ali Ahmed, "Robust compressive phase retrieval via deep generative priors," *arXiv preprint arXiv:1808.05854*, 2018.

[22] Paul Hand, Oscar Leong, and Vladislav Voroninski, "Phase retrieval under generative prior," *arXiv preprint arXiv:1807.04261*, 2018.

[23] Muhammad Asim, Fahad Shamshad, and Ali Ahmed, "Solving bilinear inverse problems using deep generative priors," *arXiv preprint arXiv:1802.04073*, 2018.

[24] Paul Hand and Vladislav Voroninski, "Global guarantees for enforcing deep generative priors by empirical risk," *arXiv preprint arXiv:1705.07576*, 2017.

[25] Max Born and Emil Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*, Elsevier, 2013.

[26] Manik Dhar, Aditya Grover, and Stefano Ermon, "Modeling sparse deviations for compressed sensing using generative models," *arXiv preprint arXiv:1807.01442*, 2018.

[27] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[28] Ziwei Liu, Ping Luo, Xiaogang Wang, and Xiaoou Tang, "Deep learning face attributes in the wild," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 3730–3738.

[29] Tero Karras, Timo Aila, Samuli Laine, and Jaakko Lehtinen, "Progressive growing of gans for improved quality, stability, and variation," *arXiv preprint arXiv:1710.10196*, 2017.

[30] Cristóvão Cruz, Rakesh Mehta, Vladimir Katkovnik, and Karen O Egiazarian, "Single image super-resolution based on wiener filter in similarity domain," *IEEE Transactions on Image Processing*, vol. 27, no. 3, pp. 1376–1389, 2018.

[31] Alec Radford, Luke Metz, and Soumith Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.