DEEP PTYCH: SUBSAMPLED FOURIER PTYCHOGRAPHY USING DEEP GENERATIVE PRIORS

Fahad Shamshad, Farwa Abbas, and Ali Ahmed
Information Technology University, Lahore, Pakistan

Presented by Salman Asif

INFORMATION TECHNOLOGY UNIVERSITY

Abstract

- **Problem**: Faithful recovery of estimate of true signal from subsampled Fourier Ptychography measurements.
- **Novel Solution**: We propose a novel framework to regularize the highly ill-posed Fourier Ptychography problem using Generative Models.
- **Numerics**: We demonstrate experimentally that our proposed algorithm *Deep Ptych* outperforms existing subsampled Fourier Ptychography approaches
  - in terms of quality of reconstruction
  - robustness against noise
- **Deep Ptych+**: We further modify the proposed approach to allow the generative model to explore solutions outside the range, leading to improve performance.

Problem Formulation

- **Aim**: Reliable estimation of image from subsampled Fourier ptychography measurements.
- **Observation Model**: Forward acquisition model of subsampled Fourier Ptychography is

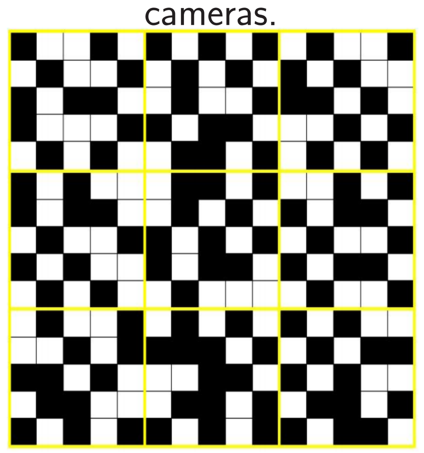
$$y_\ell = |\mathcal{M}_\ell(\mathcal{A}_\ell(x))| + n_\ell, \text{ for } \ell = 1, \dots, L,$$

where $y_\ell \in \mathbb{R}^n$ is subsampled image corresponding to ℓ^{th} camera, $\mathcal{A}_\ell : \mathbb{C}^n \rightarrow \mathbb{C}^n$ is the linear operator representing the forward acquisition model, and $n_\ell \in \mathbb{R}^n$ 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 pupil mask, and $\circ$ represents the Hadamard product.

- 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.

$$\text{Subsampling Ratio} = \frac{\text{Number of samples retained}}{\text{Total observed samples}}.$$

Subsampling mask for 3×3 cameras.



- **Challenges**: Problem is highly ill-posed due to its non-linear and non-convex nature.
- **Naive Approaches**:
  - Introduce redundancy into measurement system but it is expensive and time consuming at high resolutions.
  - Exploit structural assumptions on true signal like sparsity, non-negativity but nature exhibits far richer non-linear structure than sparsity or non-negativity alone [1].

Generative Models

- **Generative Prior**: Learn the structure of class of natural signals like faces or numbers from the training data using Generative Adversarial Networks or Variational Autoencoders.
- In these models, the generative part (G), learns a mapping from low dimensional latent space $z \in \mathbb{R}^k$ to a high dimensional sample space $G(z) \in \mathbb{R}^n$ 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 $G : \mathbb{R}^k \rightarrow \mathbb{R}^n$ with a distribution P_Z over (usually random normal or uniform), is therefore capable of generating fake data indistinguishable from the real data it has been trained on.
- Notably, these generative prior based approaches, have been shown to improve over sparsity-based approaches, thus advancing the state of the art in several image restoration tasks [2].
- **Assumption**: The generator $G(.)$ well approximates the set $\mathcal{X}$.

Deep Ptych

- We aim to solve for $\hat{x}$, given measurements y, forward operator $\mathcal{A}$, and subsampling mask $\mathcal{M}$:

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

where Range(G) is the set of all the images that can be generated by pretrained G.

- The minimization program in (1) can be equivalently formulated in the latent representation z as follows:

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

Deep Ptych

- Overview of Fourier ptychography forward acquisition model and proposed reconstruction algorithm.

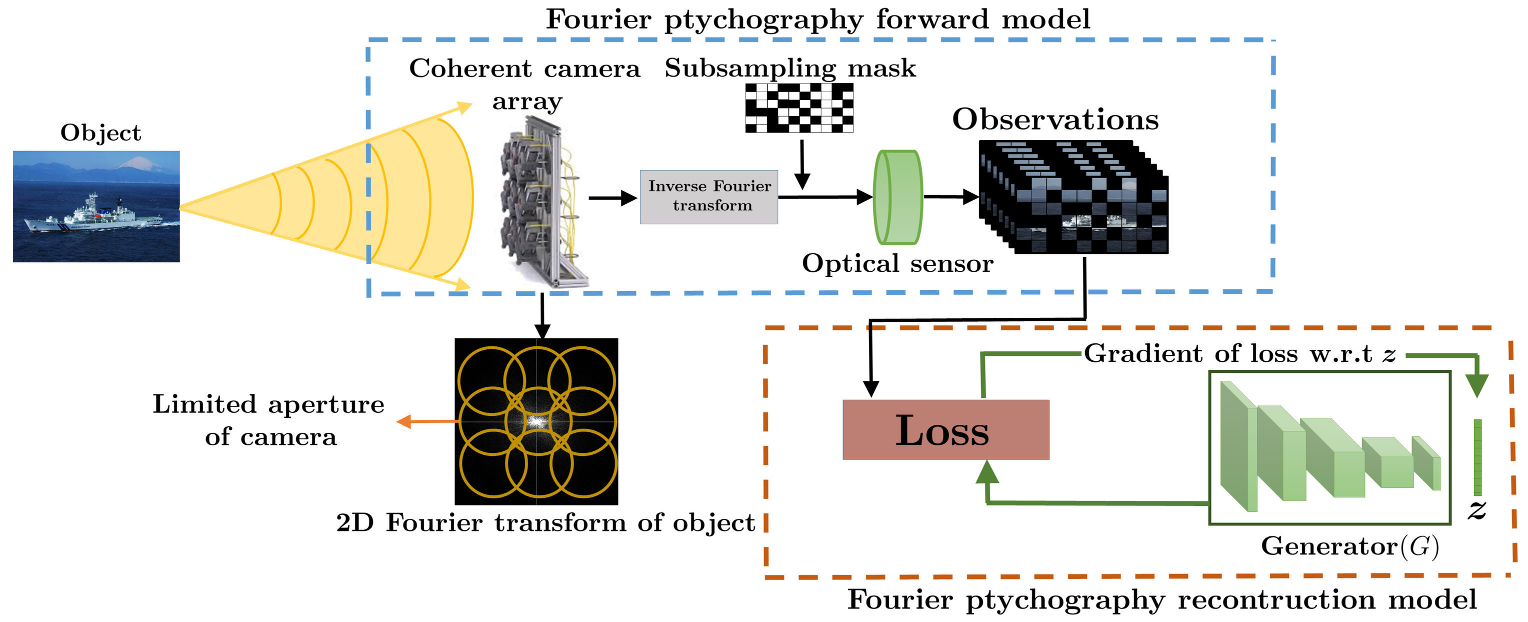


Figure: Overview of Fourier ptychography forward model and proposed reconstruction algorithm. A coherent camera array captures an image of the object. The bandlimited signal is then focused to an image plane and a subsampling operator is applied. Subsequently, an optical sensor measures the magnitude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

Deep Ptych+

- Output of Deep Ptych is constrained to lie in the range of generator.
- 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 (2) and dubbed this approach as **Deep Ptych+**.

$$(\hat{x}, \hat{z}) = \underset{z \in \mathbb{R}^k, x \in \mathbb{R}^n}{\text{argmin}} \sum_{\ell=1}^{L} \| y_\ell - |\mathcal{M}_\ell(\mathcal{A}_\ell(x))| \|_2^2 + \lambda \|x - G(z)\|_2^2. \tag{3}$$

  - First term in the objective favors an image x with smaller measurement loss
  - Second term ensures that x should not deviate too far away from the range of the pretrained generative network $G()$.
  - λ is free parameter
- We find that further adding total variation regularization results in improved results, especially in high noise regime. We dubbed this approach as **Deep Ptych + (TV)**.

$$(\hat{x}, \hat{z}) = \underset{z \in \mathbb{R}^k, x \in \mathbb{R}^n}{\text{argmin}} \sum_{\ell=1}^{L} \| y_\ell - |\mathcal{M}_\ell(\mathcal{A}_\ell(x))| \|_2^2 + \lambda_1 \|x - G(z)\|_2^2 + \lambda_2 \|x\|_{tv}^2. \tag{4}$$

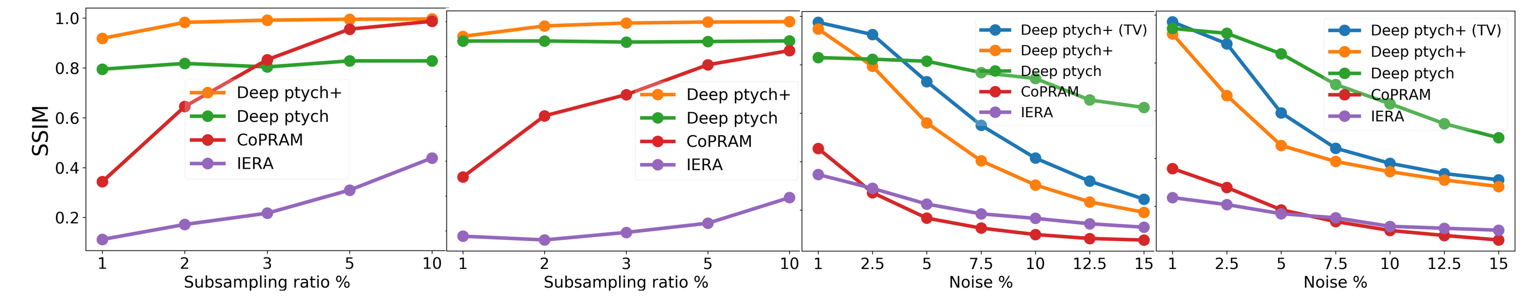
  - λ_1 and λ_2 are hyperparameters.

Experimental Setup

- Generative Models
  - Generative Adversarial Networks
  - Variational Autoencoders
- Datasets
  - MNIST (Super-resolved to 56×56)
  - CelebA (64×64)
  - CelebA HQ (128×128)
- Baseline Methods
  - Iterative Error Reduction Algorithm (IERA) [2]
  - Compressive Phase Retrieval using Alternative Minimization (CoPRAM) [3]
- Quantitative Measures
  - Peak Signal to Noise Ratio (PSNR)
  - Structural Similarity Index Measure (SSIM)
- Adam optimizer with learning rate of 0.05.
- Aperture diameter of each camera in 9×9 camera grid is 15 and 16 pixels for MNIST and CelebA, respectively.
- Overlap between cameras is 65%

Numerical Simulations

- SSIM plots for Subsampling and noise robustness for MNIST and CelebA.



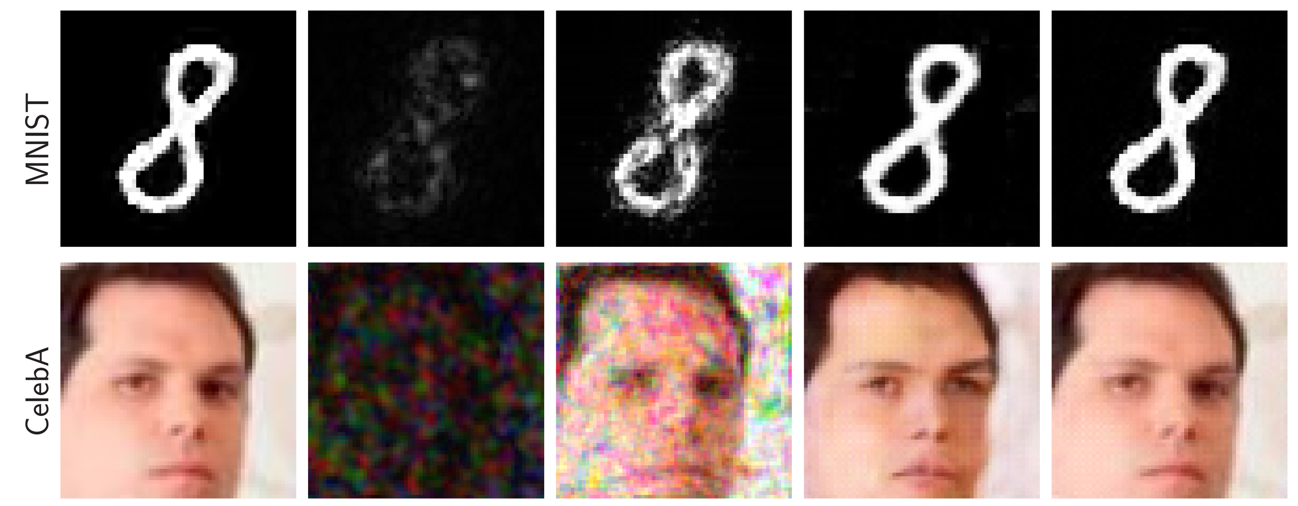
(a) CelebA (b) MNIST (c) CelebA (d) MNIST

- PSNR results

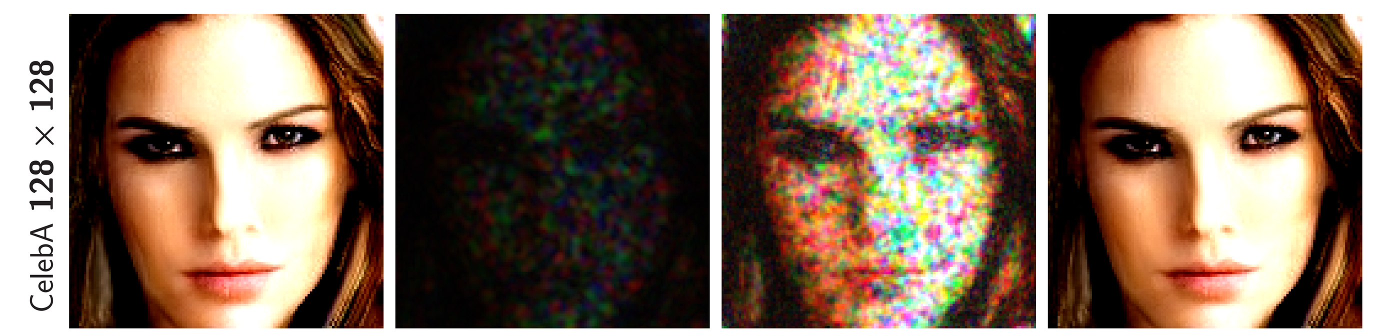
	MNIST				CelebA			
	Subsampling		Noise		Subsampling		Noise	
	1%	3%	1%	10%	1%	3%	1%	10%
IERA	10.59	11.72	14.15	9.59	5.89	7.29	12.36	11.25
CoPRAM	14.17	19.16	16.63	6.45	13.95	23.85	17.53	9.34
FP+(TV)	-	-	27.10	12.26	-	-	26.41	15.81
DP	24.70	24.80	25.31	**19.27**	22.54	22.79	23.57	**21.21**
DP+	**29.78**	**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: PSNR (dB) results for MNIST and CelebA.

- Subsampling Results (2% subsampling ratio, $\lambda = 0.01$ for Deep Ptych+.)

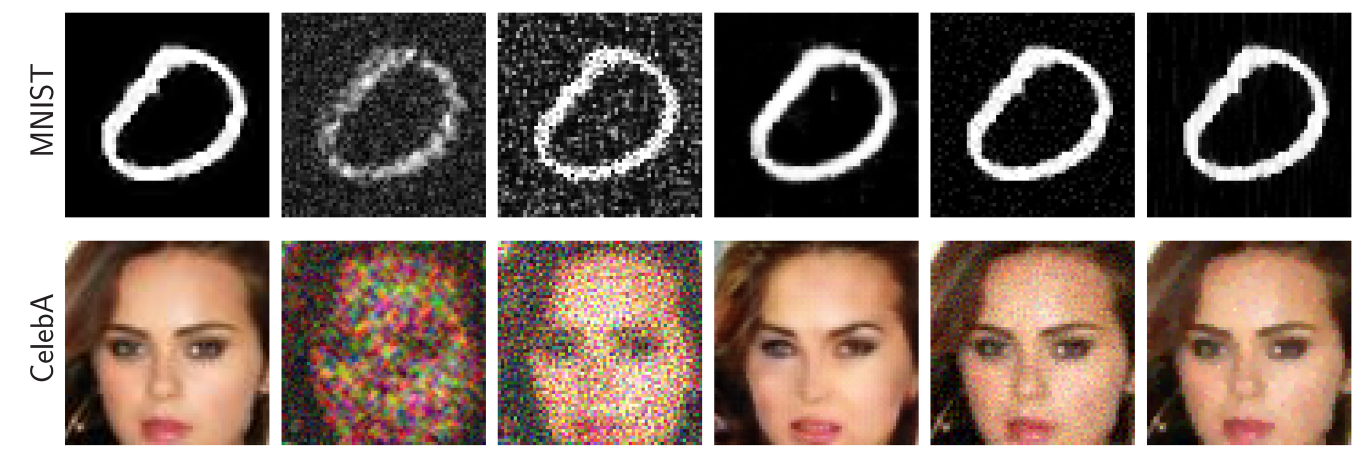


(a) Original (b) IERA (c) CoPRAM (d) DP (e) DP+



(f) Original (g) IERA (h) CoPRAM (i) Deep Ptych

- Noise Results (2.5% noise, $\lambda = 0.01$ for Deep Ptych+, $\lambda_1 = 0.01$ and $\lambda_2 = 10^{-4}$)



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

References

[1] P. Hand and V. Voroninski,"Global guarantees for enforcing deep generative priors by empirical risk minimization" arXiv preprint arXiv:1705.07576, 2017.
[2] M. Asim, F. Shamshad, and A. Ahmed, " Blind image deconvolution using deep generative priors." ArXiv eprints, February 2018.
[3] J. Holloway, M.S. Asif, M.K. Sharma, N Matsuda, R Horstmeyer, O Cossairt, and A Veeraraghavan, "Toward long-distance subdiffraction imaging using coherent camera arrays," IEEE Transactions on Computational Imaging, 2016.
[4] G. Jagatap, Z. Chen, C. Hegde, and N. Vaswani," Sub-diffraction imaging using fourier ptychography and structured sparsity," in Proc. IEEE Int. Conf. Acoust., Speech, and Sig. Proc.(ICASSP), 2018.