

Compressive Ptychography using Deep Image and Generative Priors

Semih Barutcu, Doğa Gürsoy, Aggelos K. Katsaggelos

Abstract—Ptychography is a well-established coherent diffraction imaging technique that enables non-invasive imaging of samples at a nanometer scale. It has been extensively used in various areas such as the defense industry or materials science. One major limitation of ptychography is the long data acquisition time due to mechanical scanning of the sample; therefore, approaches to reduce the scan points are highly desired. However, reconstructions with less number of scan points lead to imaging artifacts and significant distortions, hindering a quantitative evaluation of the results. To address this bottleneck, we propose a generative model combining deep image priors with deep generative priors. The self-training approach optimizes the deep generative neural network to create a solution for a given dataset. We complement our approach with a prior acquired from a previously trained discriminator network to avoid a possible divergence from the desired output caused by the noise in the measurements. We also suggest using the total variation as a complementary before combat artifacts due to measurement noise. We analyze our approach with numerical experiments through different probe overlap percentages and varying noise levels. We also demonstrate improved reconstruction accuracy compared to the state-of-the-art method and discuss the advantages and disadvantages of our approach.

Index Terms—Ptychography, Deep Generative Priors, Deep Image Priors, Image Reconstruction, Self-Training, Phase Retrieval, Generative Adversarial Network, Neural Network Optimization, Total Variation

I. INTRODUCTION

Ptychography is an important computational imaging tool in materials science that enables imaging of large field-of-views at a sub-10 μm spatial resolution [1]. In ptychography, a focused beam of x-ray is scanned across a sample while recording diffraction images in the far-field plane. The images of the sample are obtained after data collection computationally by solving the problem known as phase retrieval [2]. Because ptychography uses computing to form an image instead of a physical lens, it avoids the limitations of lens-based imaging systems because of the low photon efficiency of the lenses at hard x-ray energies or the lens-induced distortions in the acquired images. However, despite its high photon efficiency and imaging resolution, the time to obtain images is still long compared with the conventional x-ray microscopes [3]. The

main reason for its slow operation is the mechanical scanning procedure, which requires raster scanning of a micrometer-sized beam of illumination across the sample. For example, while small samples are routinely imaged with ptychography, 3D imaging or imaging of large samples is still challenging [4]. Improving the data acquisition time can enable a range of capabilities that are critical for materials science studies, including capturing dynamic phenomena under *operando* conditions or imaging larger sample volumes.

Traditionally, the ptychographic inverse problem is solved by iterative methods such as the Difference Map (DM) [5]–[7] and extended Ptychographical Iterative Engine (ePIE) [8], and reconstruction quality is increased by more advanced algorithms [9]. Initial efforts to improve the scanning time have been concentrated on the development of high-precision scanners that can reliably work at high scan velocities [10]–[12]. Alternatively, one can scan fewer data points and use compressive sensing [13], [14] to complement the sparsely acquired diffraction images.

[15]–[17] provide theoretical limits for lowering the amount of data collected by manipulation of Fourier coefficients, usage of lifting techniques, or employing iterative reweighing for phase retrieval problems. [18] utilizes Gabor’s decomposition into elementary signals and [19] lowers the sampling by random removal of detector pixels and diffraction data as examples of the application of compressed sensing to ptychography. However, none of these approaches can produce decent solutions at extreme compression levels, like when the overlapping requirement is completely ignored or when the noise is significantly high due to a dim source or a low exposure time.

To overcome the weaknesses of the traditional ptychographic phase retrieval, deep learning has emerged as a promising alternative or a complementary approach [20]–[27]. Self-optimization approaches such as the deep image prior (DIP) [28] have been effective for phase retrieval [29]–[32]. Its effectiveness is shown also for ptychographic inverse problem in [33]. In addition, supervised approaches such as the deep generative priors (DGPs) have also been applied for reducing imaging artifacts associated with undersampling [34], [35], as well as to overcome issues related to partial-coherence of the source [36]. Successful reconstructions using generative networks are demonstrated in [37] for Fourier ptychography and [38] for x-ray ptychography.

However, both approaches have limitations. On the one hand, the convergence behavior of a DIP approach is unpredictable because it is heavily dependent on the initial state of the network. Also, the improvement expected from an

Semih Barutcu (*Corresponding Author*) is with Department of Electrical and Computer Engineering, Northwestern University in Evanston, IL 60201 USA (semihbarutcu@u.northwestern.edu).

Doğa Gürsoy is with X-ray Science Division, Argonne National Laboratory, Lemont, IL 60439 USA, and also with Department of Electrical Engineering, Northwestern University in Evanston, IL 60201 USA (dgursoy@anl.gov).

Aggelos K. Katsaggelos is with Department of Electrical and Computer Engineering, Northwestern University in Evanston, IL 60201 USA, and is also with Department of Computer Science Engineering, Northwestern University in Evanston, IL 60201 USA (a-katsaggelos@northwestern.edu)

unsupervised method is limited compared to a supervised method. On the other hand, the success of supervised methods such as DGPs is limited by the representation capability of the pre-trained network. Especially when working with sub-optimal training data, DGPs can yield undesired solutions.

In this paper, to address the weaknesses of both supervised and unsupervised methods for iterative reconstruction for x-ray ptychography, we suggest complementing DIPs with DGPs. We suggest that DGPs can be used as a good initialization of DIPs. We show that using a pre-trained generator network on a specific dataset to limit the generator to the target domain would significantly improve the reconstruction quality. The process is done in two major steps: first, a random vector is optimized to get the best reconstruction in the limited representation capability of the network learned from the pre-training. Then, the network weights are optimized for the target reconstruction by expanding the representation capability towards the required result. In addition to combining DIPs and DGPs, we suggest using iterative updates of the weights with a progressive approach for achieving linear convergence. Furthermore, we suggest the addition of priors such as total variation prior for piecewise-constant smoothness and discriminative prior to penalize divergence from the target domain constrained by the pre-trained network, which is particularly useful for avoiding artifacts in the reconstructions from the noisy data. We validate our method with numerical tests simulating a compressive x-ray ptychography experiment. Since the DGP is trained to generate samples in the specified and trained domain, it enables us to complete the missing data in the low overlap x-ray ptychography. Moreover, we demonstrate the effect of regularization on the reconstructed images to evaluate the artifacts. Finally, we argue that good quality reconstructions may be possible with data acquired with no overlap during scanning, enabling a fully flexible data acquisition with no hard restriction.

II. PTYCHOGAN

In this section, we present the ptychographic phase retrieval problem, describe the components that we suggest using in our model, and describe our approach for solving the problem.

A. The Forward Problem

In ptychography, we record the intensity measurements in the far-field plane while scanning the object with a focused probe (Fig. 1). Assume that $x \in \mathbb{C}^{N \times N}$ is the unknown 2D complex object of size $N \times N$, and the values of x to depend on the material and the energy of the x-ray source to be scanned. The acquisition process can be described in multiple steps. The object first interacts with the complex illumination function $P \in \mathbb{C}^{M \times M}$ where $M \leq N$. The resulting wave $\psi \in \mathbb{C}^{M \times M}$ in the sample plane is written as,

$$\psi_i = P_i \odot x \quad (1)$$

where we use i to represent the position of the probe in the sample plane, and \odot denotes the element-wise multiplication. Then the wave goes through far-field diffraction represented by the Fourier transform operator $F \in \mathbb{C}^{M \times M}$. The intensity

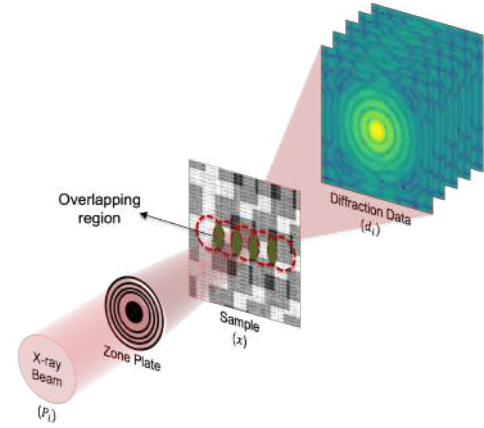


Fig. 1. X-ray ptychography experimental setup and overview of the scanning process. A focused beam of x-rays is scanned across the sample by maintaining an overlap in the illumination of consecutive scan points (shown in red). The intensity of the diffracted wave at each scan point is then recorded with a pixel-area detector.

is recorded at the detector, which is represented by the $|\cdot|^2$ operator.

$$d_i = |F\psi_i|^2 + \eta_i \quad (2)$$

where $d_i \in \mathbb{R}^{M \times M}$ is the intensity image recorded in the detector, and η_i is the additive noise at the scan position i .

B. Ptychographic Inverse Problem

The phase retrieval problem is to recover the phase $\Phi(k)$ of the complex signal $A(k)\exp(j\Phi(k))$ in the detector plane using only the magnitude information $A(k)$. Phase retrieval algorithms aim at iteratively solving for the phase of the signal by imposing some prior knowledge. While some algorithms assume phase-only objects, such as Gerchberg-Saxton Error Reduction algorithm [39], some algorithms assume a finite spatial extent object [40]–[42]. On the other hand, ptychography is created to avoid these problems, where the required constraints come from the overlapping scanning pattern on the object. We can express the phase retrieval minimization problem as:

$$\hat{x} = \arg \min_x \sum_i \|d_i - |F(P_i \odot x)|^2\|_1. \quad (3)$$

where \hat{x} is the solution. The redundant information in the overlapping regions is utilized to recover the phase of the object. In traditional methods such as ePIE [8], the object is updated for each scanning position consecutively, similar to a gradient-descent-based optimization. This allows the algorithm to update the overlapped regions more than once and use the updated information to recover lost data. However, high overlap constraints significantly increase the data acquisition time.

C. Compressive Ptychography with Deep Priors

Compressed sensing is a method using simultaneous acquisition and compression to reduce acquisition time and storage space, increase transmission efficiency, or decrease processing times of a signal. The goal of compressed sensing is to

recover a signal $t \in \mathbb{C}^P$ from measurements $y = Et \in \mathbb{C}^R$ where $E \in \mathbb{C}^{R \times P}$ is the forward model and $P \gg R$. Prior information is necessary to support the solution for recovering the true signal in compressed sensing. We pose the phase retrieval as a compressed sensing problem and focus on combining two state-of-the-art methods: deep generative priors and deep image priors.

1) *Deep Generative Priors*: Generative Adversarial Networks (GANs) [43] are proven to be effective in learning a representation of a group of samples from a training set which results in having the capability of generating a sample example that can be included in the same sample group. That is done by a method called adversarial training. GANs consist of two neural networks, a generator ($G(\cdot)$) which learns to generate an image in the desired domain, and a discriminator ($D(\cdot)$) that trains the generator by scoring the generated samples. The training is done in an adversarial structure where the generator aims at "fooling" the discriminator, and the discriminator aims at distinguishing between acceptable and unacceptable generated samples [44].

If trained successfully, the generator learns a mapping from a low dimensional latent vector $z \in \mathbb{R}^k$ to a higher dimensional object space $G(z) \in \mathbb{R}^{NxN}$ where $NxN \gg k$. The representation capability of the generator network depends on the training process and the complexity of the higher dimensional object space. An ideal generator should span an adequate distribution P_z mapped from a latent vector z .

In this paper, we suggest using deep generative priors to reconstruct the object scanned in the x-ray ptychography. Replacing the phase of the unknown object x in Eq. 3 with the output of the generator $G(\cdot)$, and setting the magnitude of the object to 1, the minimization problem becomes:

$$\hat{z} = \arg \min_z \sum_i \|d_i - |F(P_i \odot e^{jG(z)})|^2\|_1, \quad (4)$$

where,

$$\hat{x} = e^{jG(\hat{z})}. \quad (5)$$

Therefore, the phase retrieval problem becomes solving for the latent vector z that minimizes Eq. 4. To minimize the network's approximation error, the generator should have a significant representation capability, which often requires training on a large dataset of images in the same subset of images with the scanned object. However, this is usually not practical due to the limited availability of open-access datasets for training.

2) *Deep Image Priors*: Deep networks trained on a limited number of samples will have fundamentally limited effectiveness defined by the training data used. Even if a generator network is trained with a considerably large dataset, the representation capability is still not ideal, and the solution is likely beyond the representation capability of the network. A path forward to combat with the limited representation capability of pre-trained networks is to use untrained networks, i.e., deep image priors. With DIPs, the network architecture itself constrains the solution and allows the reconstruction to fit the measurement data. They do so by leveraging the superior low-dimensional image representation capability of

deep generative networks. We propose a method to use deep image priors as a fine-tuning method to get the required reconstructions after the initial reconstruction acquired by the deep generative prior. Based on this approach, Eq. 4 can be modified as:

$$\hat{z}, \hat{\theta} = \arg \min_{z, \theta} \sum_i \|d_i - |F(P_i \odot e^{jG_\theta(z)})|^2\|_1, \quad (6)$$

where,

$$\hat{x} = e^{jG_{\hat{\theta}}(\hat{z})}, \quad (7)$$

and θ represents the weights of the network. Here the network structure of the generator can be used as the prior, and by optimizing the parameters of the network, we can achieve an improved reconstruction of the object.

D. Progressive Adjustment Optimization

Optimization of the weights of a large network for a specific problem is not straightforward. The function is bound to many variables is non-convex, and minimization is challenging. To avoid this problem, we suggest using progressive optimization in which we update the parameters of the generator network starting from the shallowest layer. Then, the number of updated layers is increased progressively until the whole network weights are completely updated.

The progressive adjustment allows the network to focus on different layers in a sequential manner and to improve the reconstruction quality by minimizing the possibility of being stuck at a local minimum. We denote the weight of the shallowest layer of a network by θ_0 , and the total number of layers by L . The process of updating weights can be shown by Algorithm 1.

Algorithm 1: Progressive Optimization

```

1 for  $l$  in  $L$  do
2    $\hat{z}, \hat{\theta}_0, \dots, \hat{\theta}_l$ 
   =  $\arg \min_{z, \theta_0, \dots, \theta_l} \sum_i \|d_i - |F(P_i \odot e^{jG_\theta(z)})|^2\|_1$ 
3 end for
4 return  $e^{jG_{\hat{\theta}_0, \dots, \hat{\theta}_L}(\hat{z})}$ 

```

E. Regularization

Prior information about the target object is crucial for an accurate solution to the phase retrieval problem. Regularization applied to the inverse problem can significantly increase the reconstruction quality depending on the specific problem or enable reconstructions with limited or noisy data. Total variation regularization and discriminative regularization are the regularization types that we apply as part of our method.

1) *Total Variation Regularization*: For many objects in x-ray ptychography piece-wise continuity is a common feature to have as the objects scanned usually consist of artificial structures. The information available in this form can be a valuable prior to reduce artifacts in the reconstruction and to force the regularizer to a better solution. In addition, total variation minimization can leverage the solution for noisy data

significantly by eliminating the sparse artifacts. For the x-ray ptychography minimization we suggest, the total variation distance can be defined as:

$$TV(z, \theta) = \|\nabla(G_\theta(z))\|_1 \quad (8)$$

where ∇x designates the gradient operation on the object. Adding the regularization term to our minimization function in Eq. 6 with the regularization parameter λ_1 controlling the trade-off between the data fidelity and regularization parts of the function, the minimization can be expressed as:

$$\hat{z}, \hat{\theta} = \arg \min_{z, \theta} \sum_i \|d_i - |F(P_i \odot e^{jG_\theta(z)})|^2\|_1 + \lambda_1 \|\nabla(G_\theta(z))\|_1. \quad (9)$$

2) *Discriminative Regularization:* As described in Section II-C1, DGPs are strong priors to drive the reconstruction closer to the true solution. However, the addition of DIPs can drive the solution away from the true solution due to missing information and the added noise in measurements. To avoid divergence from the true solution and getting trapped in an undesired local minimum, we employ a discriminative regularization in our DGP approach.

The network used for the optimization is the trained generator network of a GAN structure, while the discriminator is discarded. However, the discriminator is valuable as it is trained to distinguish if the reconstruction from the generator belongs to the same domain as the objects from the training set. Thus, the loss from the discriminator output can be used to create a regularizer punishing the divergence from the required domain. $D(\cdot)$ describing the discriminator, the discriminator loss (DL) can be written as:

$$DL(z, \theta) = \log(1 - D(G_\theta(z))). \quad (10)$$

The discriminator loss can be added to the minimization function in Eq. 9 with the regularization parameter λ_2 as:

$$\hat{z}, \hat{\theta} = \arg \min_{z, \theta} \sum_i \|d_i - |F(P_i \odot e^{jG_\theta(z)})|^2\|_1 + \lambda_1 \|\nabla(G_\theta(z))\|_1 + \lambda_2 \log(1 - D(G_\theta(z))). \quad (11)$$

F. Bayesian Approach

Photon detection by a detector can be described as a Poisson process. Therefore, the probability of measuring the diffraction image d_i , can be expressed by the likelihood function. With the assumption that the measurements are independent of each other, instead of the minimization in Eq. 3, we can maximize the probability of measuring all data points.

$$\hat{x} = \arg \max_x p(d|x) \quad (12)$$

$$= \arg \max_x \prod_i \frac{e^{-|F(P_i \odot x)|^2} |F(P_i \odot x)|^{2d_i}}{d_i!} \quad (13)$$

Here, we assume that we have a prior knowledge on the object to be reconstructed. So, instead, we can maximize the

posteriori probability $p(x|d)$ which can be calculated by the Bayes' theorem:

$$\begin{aligned} \hat{x} &= \arg \max_x p(x|d) = \arg \max_x \frac{p(d|x)p(x)}{p(d)} \\ &= \arg \max_x p(d|x)p(x) \\ &= \arg \min_x (-\log(p(d|x)) - \log(p(x))) \end{aligned} \quad (14)$$

where,

$$-\log(p(d|x)) = \sum_i (F(P_i \odot x))^2 - 2d_i \log(|F(P_i \odot x)|). \quad (15)$$

As we have previously discussed, we would like to apply piece-wise continuity to the reconstruction to suppress imaging noise. So, we select the object in the Gibbs form, which allows us to apply total variation regularization.

$$p(x) = e^{\lambda_1 \|\nabla(x)\|_1} \quad (16)$$

$$-\log(p(x)) = \lambda_1 \|\nabla(x)\|_1. \quad (17)$$

Combining these information, maximum a posteriori probability (MAP) estimate for the x-ray ptychography problem can be written as the minimization:

$$\hat{x} = \arg \min_x \sum_i (|F(P_i \odot x)|^2 - 2d_i \log|F(P_i \odot x)|) + \lambda_1 \|\nabla(x)\|_1 \quad (18)$$

which can be used to update our minimization equation Eq. 11:

$$\begin{aligned} \hat{z}, \hat{\theta} &= \arg \min_{z, \theta} \sum_i \| |F(P_i \odot e^{jG_\theta(z)})|^2 \\ &\quad - 2d_i \log|F(P_i \odot e^{jG_\theta(z)})| \|_1 \\ &\quad + \lambda_1 \|\nabla(G_\theta(z))\|_1 \\ &\quad + \lambda_2 \log(1 - D(G_\theta(z))). \end{aligned} \quad (19)$$

Algorithm 2: Proposed Solution

Input : $z \in \mathbb{R}^k$, a random vector
 $G_\theta(\cdot) \in \mathbb{R}^{N \times N}$, a generator network
 $D_\omega(\cdot) \in \mathbb{R}^1$, a discriminator network
 I_{train} , a set of images to train the networks
 $d_i \in \mathbb{R}^{M \times M}$, intensity of diffractions at pos. i

Output: $\hat{x} \in \mathbb{R}^{N \times N}$, reconstructed object

```
// Pre-train the networks
1 Train  $G_\theta(\cdot)$  and  $D_\omega(\cdot)$  using  $I_{train}$  // Find an
  initial reconstruction  $x_{init}$ 
2  $\hat{z} = \arg \min_z \sum_i \|d_i - |F(P_i \odot e^{jG_\theta(z)})|^2\|_1$ 
3  $x_{init} = e^{jG_\theta(\hat{z})}$ 
  // Improve reconstruction quality
  // Apply progressive adjustment (Alg. 1)
4  $\hat{z}, \hat{\theta} = \arg \min_{z, \theta} \sum_i \| |F(P_i \odot e^{jG_\theta(z)})|^2$ 
   $- 2d_i \log|F(P_i \odot e^{jG_\theta(z)})| \|_1$ 
   $+ \lambda_1 \|\nabla(G_\theta(\hat{z}))\|_1$ 
   $+ \lambda_2 \log(1 - D_\omega(G_\theta(\hat{z})))$ .
5 return  $\hat{x} = e^{jG_\theta(\hat{z})}$ 
```

Combining the steps described in previous sections, we can solve the minimization problem of undersampled ptychography in Eq. 2 with low or no overlap data. The novel algorithm we propose in this paper is shown in Algorithm 2 and can be summarized as:

- Train generator-discriminator networks with enough data for the network to learn generating samples in the desired output domain.
- Find an initial reconstruction using Eq. 4.
- Improve reconstruction quality using Eq. 19 while applying Algorithm 1 to update the network weights.

III. EXPERIMENTS AND RESULTS

In this section, we describe the numerical experiments to validate the proposed algorithm, show the results, and compare the capabilities of the approach with a well-understood and widely used method.

A. Simulation Setup

To evaluate the effectiveness of the proposed method against the state-of-the-art methods, numerical tests are performed on face images to represent the idea. Instead of designing generator-discriminator network architecture, the networks structures are taken from the StyleGAN2 architecture [45], which is specialized in generating different faces from a random vector. A PyTorch implementation of these networks is used for accurate representation [46].

To utilize the deep generative priors, the networks are trained on a large face dataset called Flickr-Faces-HQ (FFHQ) [47]. The resulting generator network has the capacity to generate different realistic faces from a given random vector, and the discriminator network is capable of classifying the generated images as real or fake. Three images from the test dataset are chosen as the images to be reconstructed,

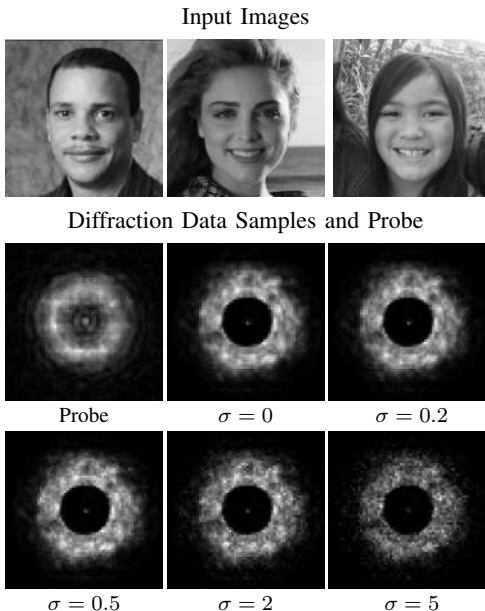


Fig. 2. Sample input images, phase of the probe used, and sample acquired data in for different noise amounts.

which the network has never seen before. To simulate the measurement data, these images are turned into gray-scale images and converted to complex matrices with the size of 256×256 pixels by setting the complex numbers' phase values to the gray-scale objects and the amplitude values to 1. The probe for the simulations is obtained from the reconstruction of a real object using the algorithm proposed in [48], and the diameter of the probe is set to be 32 pixels. With the acquired complex object and the probe, the forward model of x-ray ptychography described by Eqs 1 and 2 is applied with varying step sizes and with varying Poisson noise levels according to the peak intensities. The ground truth images and the acquired data without noise and the highest noise level are seen in Fig. 2. As it can be observed, apart from they are all face images for getting advantage of DGPs, the input images are significantly different from each other, and the noisy data has some missing information.

For demonstration, different step sizes are utilized for various experiments resulting in different overlap percentages. The ptychographic step sizes for the probe movement are chosen to be 8, 16, 24, 32, and 48 pixels, which correspond to 75%, 50%, 25%, 0%, and -50% overlaps. Most of the experiments for comparison are done for 50% overlap cases to show the limitations in a high overlap percentage, but reconstructions for no-overlap situations are done to show the effectiveness of the algorithm. Four different Poisson noise levels are used for comparison, which are denoted by σ values showing their relative intensity compared to the peak intensity values, and these are chosen to be 0.2, 0.5, 2, and 5.

B. Reconstructions and Comparison

To evaluate the robustness of the proposed approach, we show the reconstructions for different face images, different noise levels, various adjustments, and changing priors. We also present methods in the literature to be compared to the proposed approach, present the reconstructions for the simula-

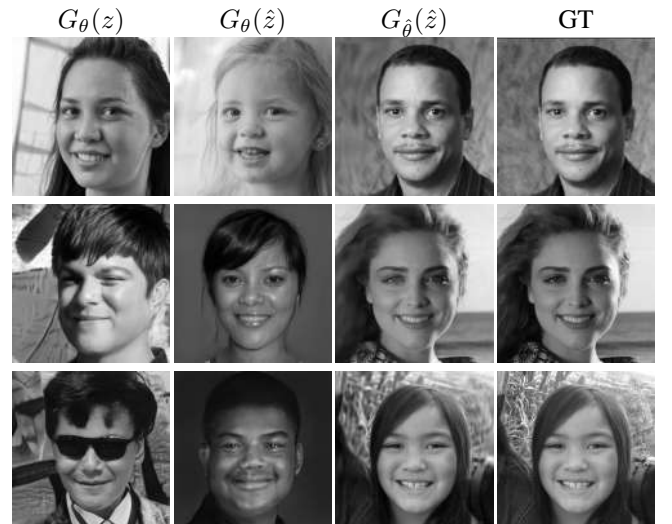


Fig. 3. Initial reconstruction without any adjustments, reconstructions after adjustment of latent vector and ground truth (GT) images for 50% overlap and noise-free cases.

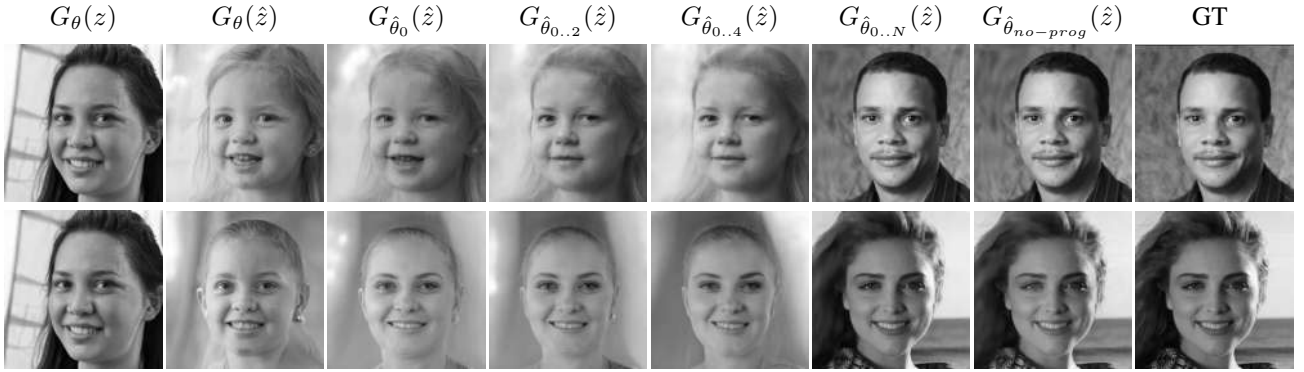


Fig. 4. Initial reconstruction without any adjustments, reconstructions after adjustment of latent vector, and reconstructions after adjusting the parameters of the network with progressive adjustment at different levels, ePIE reconstruction, and ground truth (GT) images for 50% overlap and noise-free cases.

tions for all methods, and discuss the results qualitatively and quantitatively.

The reconstructions are done in PyTorch with the proposed algorithm by using the minimization shown Eq. 19. In each experiment, a different aspect of the algorithm is tested with varying parameters. For experiments optimizing the input latent vector, standard gradient descent with a learning rate between 2×10^{-6} and 2×10^{-5} for 1000 epochs is used to solve the minimization in Eq. 4 by keeping the weights of the network constant. For experiments optimizing the network weights without progressive adjustment, the expression in Eq. 19 is minimized after initial adjustment of the latent vector using Adam optimizer [49] with learning rates between 2×10^{-5} and 2×10^{-4} is used till convergence. In the cases where progressive adjustment is applied, the update of the weights starts from the last layer and consequent layers are unfrozen at each 800th step. At the 5th step, all layers are started to be updated.

In Fig 3, reconstructions for three face images can be seen compared to the ground truth image without TV and discriminator priors. Reconstructions in the second column show that adjustment of the latent vector is not enough for a successful reconstruction. This is due to the limited representation capability of the generator network showing that deep generative priors (DGPs) are not enough for this problem by itself. However, the reconstructions after the adjustment of the network weights are similar to the ground truth images showing the capability of deep image priors (DIPs). On the other hand, although the reconstructions are very good in the facial parts, the background is not as close to the ground truth. That is due to the lack of DGPs in these parts.

We show that the reconstructions with the proposed method are successful. However, we suggested that applying progressive adjustment of the network weights would increase the reconstruction quality. In Fig. 4, the reconstructions for 50% probe overlap with noise-free data is shown. Here, it can be observed that optimizing the layers while freezing some of them gives reconstructions that are far from the ideal solution. Thus, adjustment of all layers is required for successful outputs. However, it is a highly convex minimization problem, and adjustment of layers progressively gives a significantly better starting point for the final step. This can be observed

in the fifth column of Fig. 4, where we can see that the background and the general shape of the heads look similar to the initial steps in the reconstructions; however, the facial expressions become more similar to the required solution. That results in an improved solution compared to the non-progressive reconstruction.

The noise in the diffraction data for x-ray ptychography creates some artifacts in the reconstructions. For the proposed method, these artifacts are affected in two directions. DGPs and DIPs used in the proposed method provide some additional information, leading to a constrained solution avoiding the noise. On the other hand, since the loss function is minimized designed to fit the diffraction images acquired from the reconstructed data to the input data, the network forces the reconstructions to be noisy as well. Thus, we add two different regularizers: discriminator loss and total variation as shown in Eq. 19. Reconstructions are shown for varying λ levels of regularizers for different noise levels in Fig 5. For low noise levels, the algorithm performs well without a need for priors. However, the reconstructions show significant artifacts with the increased noise in diffraction data.

The structural similarity index (SSIM) scores for Fig 5 are given in Table I. If we investigate the effect of regularization in both of them, we can see that the addition of discriminator loss as a regularizer increases the effect of DGPs, forcing the output not to diverge from the face domain. Although it seems like this improves the reconstruction quality, it might cause

TABLE I
STRUCTURAL SIMILARITY INDEX (SSIM) FOR DIFFERENT NOISE LEVELS AND DIFFERENT AMOUNTS OF REGULARIZATION.

| SSIM | Image 1 | | | | Image 2 | | | |
|--------------------|----------------------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | Noise Level (σ^2) | 0.2 | 0.5 | 2 | 5 | 0.2 | 0.5 | 2 |
| DL(0), TV(0) | 0.96 | 0.94 | 0.86 | 0.75 | 0.94 | 0.88 | 0.85 | 0.76 |
| DL(1e-5), TV(3e-4) | 0.97 | 0.95 | 0.86 | 0.77 | 0.95 | 0.95 | 0.86 | 0.75 |
| DL(1e-5), TV(1e-3) | 0.96 | 0.96 | 0.89 | 0.81 | 0.95 | 0.93 | 0.89 | 0.81 |
| DL(1e-5), TV(3e-3) | 0.94 | 0.94 | 0.89 | 0.87 | 0.92 | 0.91 | 0.91 | 0.86 |
| DL(1e-4), TV(3e-4) | 0.96 | 0.95 | 0.86 | 0.75 | 0.97 | 0.93 | 0.87 | 0.77 |
| DL(1e-4), TV(1e-3) | 0.94 | 0.95 | 0.88 | 0.78 | 0.95 | 0.95 | 0.86 | 0.79 |
| DL(1e-4), TV(3e-3) | 0.94 | 0.92 | 0.91 | 0.86 | 0.92 | 0.92 | 0.90 | 0.86 |
| DL(1e-3), TV(3e-4) | 0.92 | 0.90 | 0.90 | 0.81 | 0.93 | 0.93 | 0.88 | 0.78 |
| DL(1e-3), TV(1e-3) | 0.92 | 0.88 | 0.90 | 0.81 | 0.87 | 0.92 | 0.87 | 0.83 |
| DL(1e-3), TV(3e-3) | 0.87 | 0.87 | 0.90 | 0.86 | 0.88 | 0.89 | 0.88 | 0.85 |

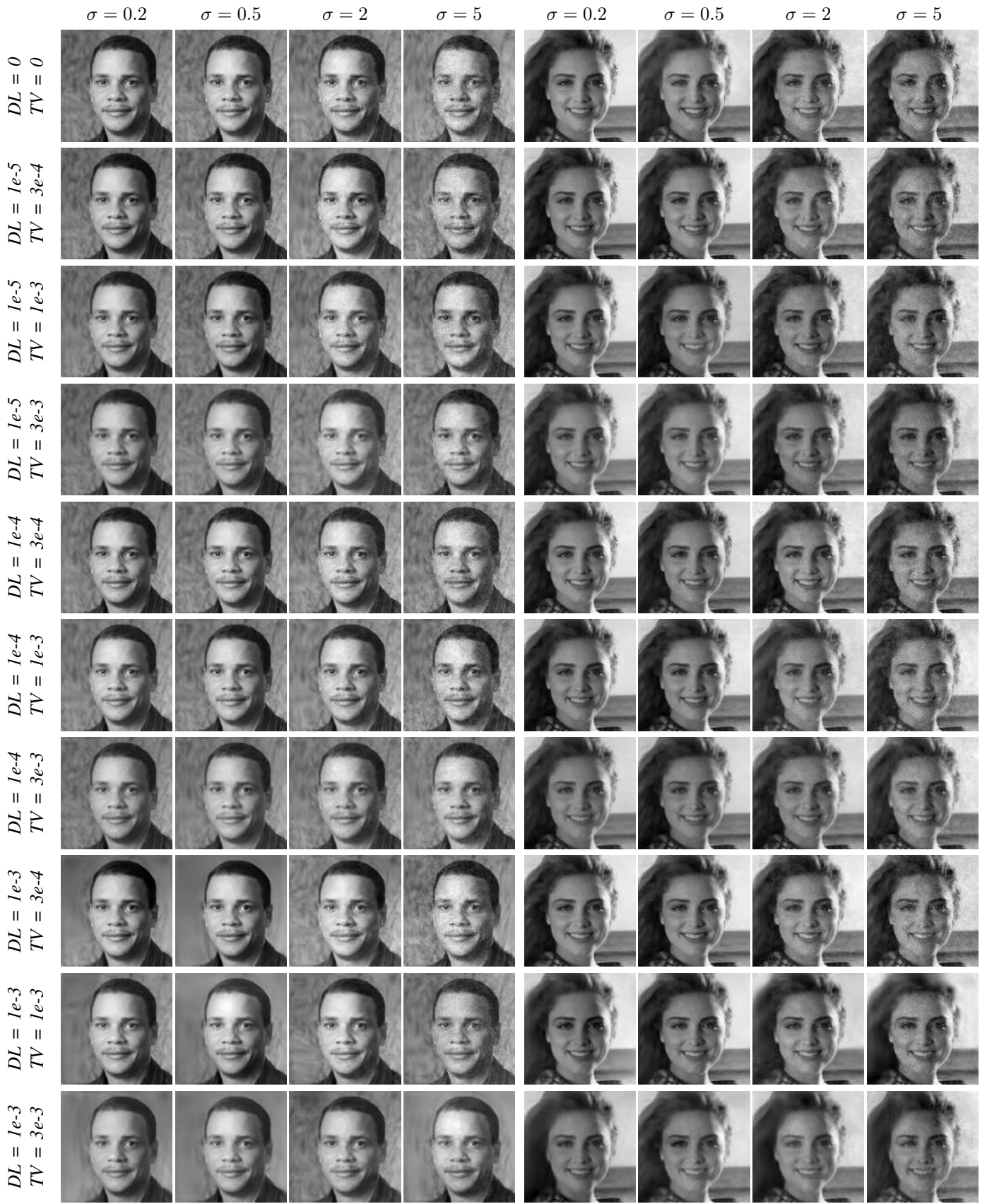


Fig. 5. Reconstructions from noisy data with varying noise levels and varying levels Discriminator Loss (DL) and Total Variation (TV) priors for 50% overlap cases

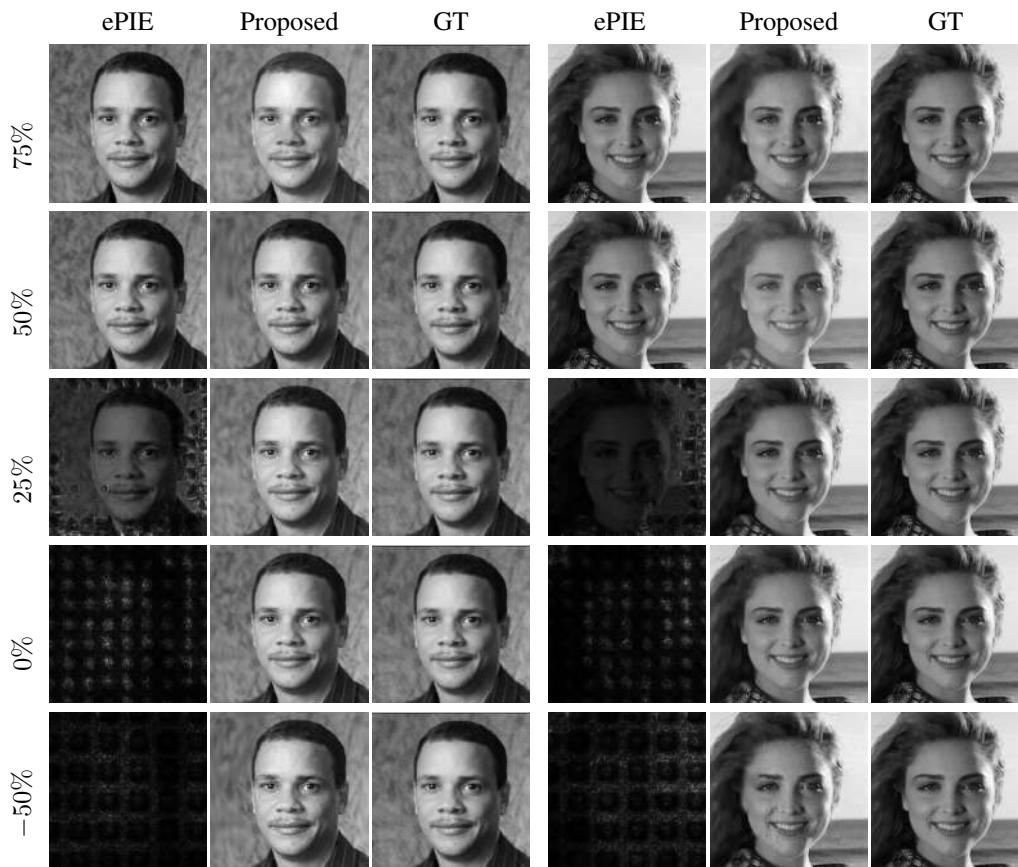


Fig. 6. Comparison of the reconstructions from the proposed method against ePIE and ground truth for noise-free case for different overlap percentages

divergence from the data itself, as can be seen from various cases in the bottom rows. In addition, discriminator loss forces the reconstruction to an “average face” from the learned domain, which means that it inherently forces smoothness. This effect can be seen to be helpful for high noise scenarios, resulting in higher SSIM scores with increased discriminator loss regularization effect. However, it also reduces the fit to the data itself, resulting in over-smoothed reconstructions in low noise scenarios.

Even though the goals of using total variation regularization and discriminator loss regularization are similar in terms of smoothing, the effects are significantly different from each other. Forcing the solution to be close to averaged out face prior does not necessarily enforce piece-wise smoothness as total variation regularization does. In Fig 5, it can be ob-

TABLE II
STRUCTURAL SIMILARITY INDEX (SSIM) COMPARISON FOR ePIE AND PROPOSED (PROP) METHODS AND DIFFERENT OVERLAP PERCENTAGES USING NOISE-FREE DATA.

| SSIM | Image 1 | | Image 2 | |
|----------------|-------------|-------------|-------------|-------------|
| | ePIE | Prop | ePIE | Prop |
| Overlap = 75% | 1.00 | 0.98 | 1.00 | 0.98 |
| Overlap = 50% | 0.99 | 0.98 | 0.99 | 0.97 |
| Overlap = 25% | 0.54 | 0.97 | 0.33 | 0.96 |
| Overlap = 0% | 0.16 | 0.94 | 0.14 | 0.95 |
| Overlap = -50% | 0.15 | 0.93 | 0.18 | 0.93 |

served that an increased multiplier of TV regularizer results in smoother areas in specific parts of the face such as the forehead or cheeks. However, since the face images are not piece-wise smooth in most cases, the effect of the regularization is lower than it could have been in other datasets. In addition, unlike the effect of discriminator loss regularization, the effect of total variation regularization is more linear and consistent, which results in a gradual increase or decrease in reconstruction quality. Similarly, the SSIM values in Table I are highest for the higher multiplier of TV regularizer for high noise levels and highest for the lower multiplier for low noise levels.

The most crucial advantage of our proposed method is the performance when there is a low overlap between the scanning positions in x-ray ptychography. The constraints that deep generative priors provide are crucial for having a good initial reconstruction, and the optimization that deep image priors drive the solution towards the ground truth. To test how effective these optimizations are, experiments with different overlap percentages are done with the standard Extended Ptychographic Iterative Engine (ePIE) method, and a comparison with the proposed method is made.

In Fig. 6, the comparison between the ePIE algorithm and the proposed algorithm can be seen for overlap percentages of 75%, 50%, 25%, 0%, and -50% are shown. When we look at the reconstructions with the highest overlap percentage, the reconstructions show similar properties for both algorithms. However, due to the increased computational complexity of the

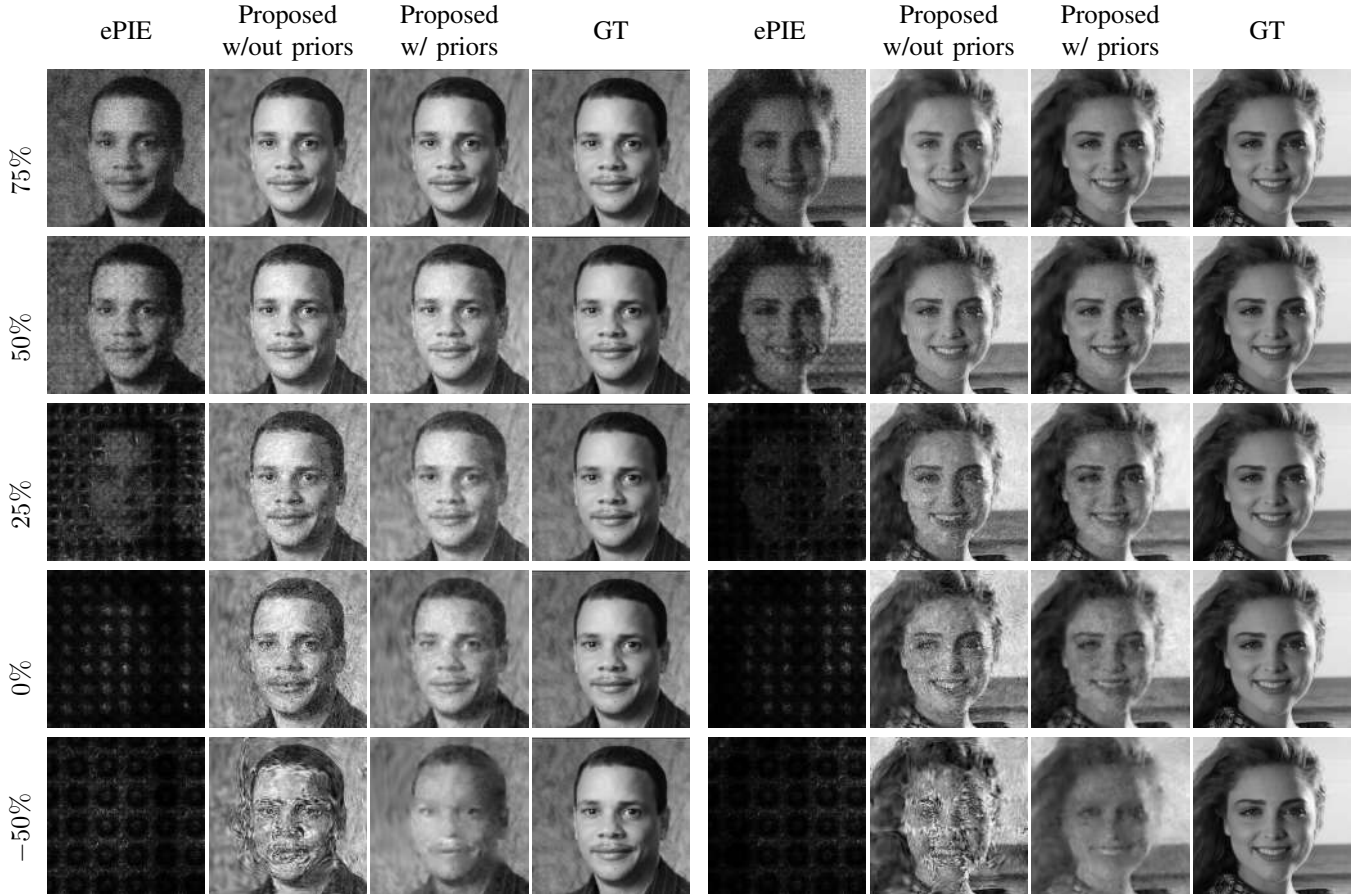


Fig. 7. Comparison of the the reconstructions from the proposed method against ePIE and ground truth for noisy case ($\sigma = 2$) for different overlap percentages

proposed algorithm and the disadvantage of the background for DGPs, the reconstruction from ePIE has slightly better reconstructions. This situation can also be observed in the Table II. A similar observations can be made for an overlap percentage of 50%. The reconstructions are good for both methods, where the analytical method edges the proposed method with a slight margin due to the imperfect optimization. However, the difference can be eliminated with the correct parameter selection in these cases.

The most easily observable and important result can be seen for the low overlap percentage case, that is 25% overlap. Although the reconstruction quality for the proposed method does not change much, the reconstructions from ePIE are significantly worsens, as can be seen in Fig. 6 and Table II. When we move on to no overlap cases, there is no possibility for ePIE to have a reconstruction. On the other hand, the proposed method results in a satisfactory reconstruction with small decreases in the SSIM scores. More importantly, the last three overlap percentages corresponds to approximately 33%, 50%, and 67% decrease in the scanning time for synchrotron experiments. The decrease in the scanning time results in higher quality scans with less noise in the diffraction patterns, which is crucial in many applications.

A similar situation can be observed for the noisy reconstructions in Fig. 7 and Table III. Here noise with relative value 2 ($\sigma = 2$) is used for all reconstructions, and the proposed

method is compared with ePIE, where the reconstructions with and without regularization are shown for fairness since ePIE does not have any regularization.

Reconstructions with ePIE shows observable artifacts in the case of noise. The advantage of solving the problem analytically for ePIE does not apply here since the proposed algorithm is able to generate good reconstructions even without regularization for a low high overlap percentage. Similar to the previous case, with the decreased overlap percentage, ePIE loses the ability to have a reconstruction at all. On the other hand, while the reconstruction quality for the proposed algorithm decreases gradually, the reconstructions without regularization remain good qualitatively. However, the reconstructions show significant artifacts for no overlap cases by the combination of lack of data and the noise, which can also be seen in the SSIM values in Table III.

The addition of regularization increases the reconstruction quality significantly quantitatively, as can be observed from the SSIM values. The qualitative evaluation shows that when there is no overlap, the reconstruction in high noise scenarios might not be acceptable. However, the reconstructions are significantly superior to the ones from ePIE in all cases.

IV. DISCUSSION AND OUTLOOK

In x-ray ptychography, reconstruction quality depends heavily on the overlap percentage between consecutive scanning

areas. Due to the limited time allocated for a sample to be scanned, the overlap amount cannot be kept high in all experiments. Similarly, with the increased overlap, the dwelling time for each scanning position is shortened, leading to noisier diffraction patterns recorded at the detector. The trade-off between these aspects leads to imperfect reconstructions in x-ray ptychography, which either do not have enough redundant data for good reconstructions or show artifacts related to noise. Existing methods in the literature approach the problem with different methods. While some try to increase the reconstruction quality by modifying the algorithm, others try increasing the output resolution via post-processing. However, these are not always good enough for some experiments.

Usage of deep generative priors (DGPs) gives an important advantage for achieving good reconstructions. Since the starting point for any ill-posed problem is vital for a good solution, having a good initialization is crucial for x-ray ptychography also. Usages of DGPs do not only enable us to start with a better initial reconstruction, but it also provides a network that is trained to give reconstructions in a constrained domain, which adds additional priors to the problem. Thus, the overlapping requirement is significantly reduced even just by the effect of DGPs.

In addition to DGPs, the reconstructions can be improved significantly with the usage of deep image priors (DIPs). DIPs allow a solution to be optimized using the structure of a network. Optimizing the weights and filters of a network with a single set of data specializes the network for a specific solution; in other words, the network over-fits the data. However, during this process, due to the abundance of network parameters, the initial points of the network weights and the network structure plays an important role in the reconstructions. Due to this advantage, the combination of DGPs with DIPs generates good reconstructions even though there is not enough data for an exact solution. The initial priors learned by DGPs, and the structural priors learned by DIPs constrain solutions to a smaller set of solutions with significantly lower error. The advantage can be best seen in Fig. 6 and Table II.

The combination of the DGPs and DIPs can be improved further based on the properties of the objects to be reconstructed. In case of a noisy measurements, other priors can be utilized as regularizers, such as total variation (TV) regularization and discriminator loss (DL) regularization. TV regularizer forces the solution to be piece-wise smooth, which

eliminates point-like artifacts caused by noisy data in the reconstructions as it is used in similar approaches. Moreover, the generator network is trained together with a discriminator network which specializes in classifying if an object belongs to a certain set of samples or not. This information can be used to punish reconstructions diverging too much from the desired range of outputs. Essentially, the regularizer pushes the solution towards a general understanding of a face image in our examples, which is an averaged out face concept. If the output is too noisy and has a lot of artifacts, it will be far from this range and will be punished by the regularizer. The effect of TV and DL regularizers on top of DIPs and DGPs can be best observed in the noisy reconstructions in Fig. 7 and Table III

Despite the improved quality of reconstructions, the most important drawback of the proposed method is the reconstruction time. The method is a self-optimizing algorithm, and the network is needed to be optimized for each reconstruction. That makes the proposed method considerably slower compared to the iterative reconstruction methods. In our simulations, the training is done on an Nvidia GeForce Titan X graphics card, where reconstruction with 50% overlap can take 23 minutes, and reconstruction with 25% overlap takes 15 minutes. The reconstruction time may increase or decrease based on the parameters chosen for optimization. Although that is an important drawback, the reconstruction time can be improved significantly by different methods. The largest improvement can be made by sharing the computational burden among multiple GPUs. Since the ptychographic process has multiple scanning points, scanning points can be scattered randomly among GPUs for improved speed. Also, all computations can be split among GPUs without high-level distribution, which can help overcome this bottleneck [50], [51].

An improvement in the reconstructions for low overlap scenarios can be made by a combination of an analytical method (such as ePIE) and the proposed algorithm for a simpler but more accurate reconstruction. In this method, reconstructions from insufficient overlap data can be made using the proposed algorithm, and new data can be acquired from the reconstruction with a higher overlap percentage using the forward model. If the low overlap data is combined with this generated high overlap data, an analytical method can be used for a more quality reconstruction.

V. CONCLUSION

In this paper, we have proposed an x-ray ptychography reconstruction method utilizing deep image priors, deep generative priors, and other data-based priors. We demonstrated that using deep image priors to optimize the solution for the highly ill-posed x-ray ptychography problem can generate improved reconstruction results. We also showed that using a pre-trained generator on a dataset that is similar to the target reconstruction can greatly improve the initial state for the deep image priors, drive the solution towards the desired domain, and provide an implicit prior for a quality reconstruction in a shorter time compared to a similar deep image priors method.

TABLE III
STRUCTURAL SIMILARITY INDEX (SSIM) COMPARISON FOR ePIE, PROPOSED METHOD WITHOUT REGULARIZATION (PROP), AND PROPOSED METHOD WITH REGULARIZATION (PROP W/ R) METHODS AND DIFFERENT OVERLAP PERCENTAGES USING NOISY DATA ($\sigma = 2$)

| SSIM | Image 1 | | | Image 2 | | |
|----------------|---------|------|-------------|---------|------|-------------|
| | ePIE | Prop | Prop w/ R | ePIE | Prop | Prop w/ R |
| Overlap = 75% | 0.50 | 0.93 | 0.95 | 0.51 | 0.87 | 0.93 |
| Overlap = 50% | 0.54 | 0.86 | 0.90 | 0.54 | 0.84 | 0.90 |
| Overlap = 25% | 0.26 | 0.73 | 0.85 | 0.22 | 0.75 | 0.87 |
| Overlap = 0% | 0.17 | 0.60 | 0.84 | 0.15 | 0.65 | 0.84 |
| Overlap = -50% | 0.16 | 0.52 | 0.80 | 0.15 | 0.52 | 0.75 |

Moreover, we have shown that combining the method total variation prior or a prior coming from the discriminator of the used generator, and using these priors as regularizers would lower the artifacts caused by noisy data. Finally, we demonstrated the ability of the proposed method, particularly in low-overlap data, to have significantly improved reconstruction quality compared to analytical methods.

ACKNOWLEDGMENTS

This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility at Argonne National Laboratory, and is based on research supported by the U.S. DOE Office of Science-Basic Energy Sciences under contract DE-AC02-06CH11357.

REFERENCES

- [1] D. Gürsoy, Y.-c. K. Chen-Wiegart, and C. Jacobsen, "Lensless x-ray nanoimaging: Revolutions and opportunities," *IEEE Signal Processing Magazine*, vol. 39, no. 1, pp. 44–54, 2021.
- [2] J. R. Fienup, "Reconstruction of an object from the modulus of its Fourier transform," *Optics Letters*, vol. 3, no. 1, pp. 27–29, 1978.
- [3] V. De Andrade, V. Nikitin, M. Wojcik, A. Deriy, S. Bean, D. Shu, T. Mooney, K. Peterson, P. Kc, K. Li *et al.*, "Fast x-ray nanotomography with sub-10 nm resolution as a powerful imaging tool for nanotechnology and energy storage applications," *Advanced Materials*, vol. 33, no. 21, p. 2008653, 2021.
- [4] M. Du, Z. Di, D. Gürsoy, R. P. Xian, Y. Kozorovitskiy, and C. Jacobsen, "Upscaling x-ray nanoimaging to macroscopic specimens," *Journal of applied crystallography*, vol. 54, no. 2, 2021.
- [5] P. Thibault, M. Dierolf, A. Menzel, O. Bunk, C. David, and F. Pfeiffer, "High-resolution scanning x-ray diffraction microscopy," *Science*, vol. 321, no. 5887, p. 379–382, Jul 2008. [Online]. Available: <https://www.science.org/doi/10.1126/science.1158573>
- [6] P. Thibault, M. Dierolf, O. Bunk, A. Menzel, and F. Pfeiffer, "Probe retrieval in ptychographic coherent diffractive imaging," *Ultramicroscopy*, vol. 109, no. 4, pp. 338–343, 2009.
- [7] M. Du, Y. S. Nashed, S. Kandel, D. Gürsoy, and C. Jacobsen, "Three dimensions, two microscopes, one code: Automatic differentiation for x-ray nanotomography beyond the depth of focus limit," *Science advances*, vol. 6, no. 13, p. eaay3700, 2020.
- [8] A. Maiden, D. Johnson, and P. Li, "Further improvements to the ptychographical iterative engine," *Optica*, vol. 4, no. 7, p. 736, Jul 2017. [Online]. Available: <https://opg.optica.org/abstract.cfm?URI=optica-4-7-736>
- [9] Q. Zhai, B. Wohlberg, G. T. Buzzard, and C. A. Bouman, "Projected multi-agent consensus equilibrium for ptychographic image reconstruction," 2021. [Online]. Available: <https://arxiv.org/abs/2111.14240>
- [10] P. M. Pelz, M. Guizar-Sicairos, P. Thibault, I. Johnson, M. Holler, and A. Menzel, "On-the-fly scans for x-ray ptychography," *Applied Physics Letters*, vol. 105, no. 25, p. 251101, 2014.
- [11] J. Deng, Y. S. G. Nashed, S. Chen, N. W. Phillips, T. Peterka, R. Ross, S. Vogt, C. Jacobsen, and D. J. Vine, "Continuous motion scan ptychography: characterization for increased speed in coherent x-ray imaging," *Optics Express*, vol. 23, no. 5, 2015.
- [12] J. N. Clark, X. Huang, R. J. Harder, and I. K. Robinson, "Continuous scanning mode for ptychography," *Optics Letters*, vol. 39, no. 20, p. 6066, 2014.
- [13] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [14] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [15] M. L. Moravec, J. K. Romberg, and R. G. Baraniuk, "Compressive phase retrieval," in *SPIE Proceedings*, D. V. D. Ville, V. K. Goyal, and M. Papadakis, Eds. SPIE, Sep. 2007. [Online]. Available: <https://doi.org/10.1117/12.736360>
- [16] H. Ohlsson, A. Yang, R. Dong, and S. Sastry, "Cprl – an extension of compressive sensing to the phase retrieval problem," in *Advances in Neural Information Processing Systems*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., vol. 25. Curran Associates, Inc., 2012. [Online]. Available: <https://proceedings.neurips.cc/paper/2012/file/3a066bda8c96b9478bb0512f0a43028c-Paper.pdf>
- [17] M. C. Newton, "Compressed sensing for phase retrieval," *Physical Review E*, vol. 85, no. 5, May 2012. [Online]. Available: <https://doi.org/10.1103/physreve.85.056706>
- [18] J. C. da Silva and A. Menzel, "Elementary signals in ptychography," *Optics Express*, vol. 23, no. 26, p. 33812, Dec. 2015. [Online]. Available: <https://doi.org/10.1364/oe.23.033812>
- [19] A. Stevens, H. Yang, W. Hao, L. Jones, C. Ophus, P. D. Nellist, and N. D. Browning, "Subsampled STEM-ptychography," *Applied Physics Letters*, vol. 113, no. 3, p. 033104, Jul. 2018. [Online]. Available: <https://doi.org/10.1063/1.5040496>
- [20] A. Sinha, J. Lee, S. Li, and G. Barbastathis, "Lensless computational imaging through deep learning," *Optica*, vol. 4, no. 9, p. 1117, Sep. 2017. [Online]. Available: <https://doi.org/10.1364/optica.4.001117>
- [21] M. J. Cherukara, Y. S. G. Nashed, and R. J. Harder, "Real-time coherent diffraction inversion using deep generative networks," *Scientific Reports*, vol. 8, no. 1, Nov. 2018. [Online]. Available: <https://doi.org/10.1038/s41598-018-34525-1>
- [22] Y. Xue, S. Cheng, Y. Li, and L. Tian, "Reliable deep-learning-based phase imaging with uncertainty quantification," *Optica*, vol. 6, no. 5, p. 618, May 2019. [Online]. Available: <https://doi.org/10.1364/optica.6.000618>
- [23] G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica*, vol. 6, no. 8, pp. 921–943, Aug 2019. [Online]. Available: <http://opg.optica.org/optica/abstract.cfm?URI=optica-6-8-921>
- [24] K. C. Zhou and R. Horstmeyer, "Diffraction tomography with a deep image prior," *Optics express*, vol. 28, no. 9, pp. 12872–12896, 2020.
- [25] A. Goy, K. Arthur, S. Li, and G. Barbastathis, "Low photon count phase retrieval using deep learning," *Physical Review Letters*, vol. 121, no. 24, Dec. 2018. [Online]. Available: <https://doi.org/10.1103/physrevlett.121.243902>
- [26] G. Barbastathis, A. Ozcan, and G. Situ, "On the use of deep learning for computational imaging," *Optica*, vol. 6, no. 8, p. 921, Jul. 2019. [Online]. Available: <https://doi.org/10.1364/optica.6.000921>
- [27] U. S. Kamilov, C. A. Bouman, G. T. Buzzard, and B. Wohlberg, "Plug-and-play methods for integrating physical and learned models in computational imaging," 2022. [Online]. Available: <https://arxiv.org/abs/2203.17061>
- [28] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," *International Journal of Computer Vision*, vol. 128, no. 7, p. 1867–1888, Jul 2020. [Online]. Available: <http://link.springer.com/10.1007/s11263-020-01303-4>
- [29] F. Shamshad and A. Ahmed, "Robust compressive phase retrieval via deep generative priors," *arXiv:1808.05854 [cs, stat]*, Aug 2018, arXiv: 1808.05854. [Online]. Available: <http://arxiv.org/abs/1808.05854>
- [30] R. VanRullen and L. Reddy, "Reconstructing faces from fmri patterns using deep generative neural networks," *Communications Biology*, vol. 2, no. 1, p. 193, Dec 2019. [Online]. Available: <http://www.nature.com/articles/s42003-019-0438-y>
- [31] F. Shamshad and A. Ahmed, "Compressed sensing-based robust phase retrieval via deep generative priors," *IEEE Sensors Journal*, vol. 21, no. 2, p. 2286–2298, Jan 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9174821/>
- [32] F. Wang, Y. Bian, H. Wang, M. Lyu, G. Pedrini, W. Osten, G. Barbastathis, and G. Situ, "Phase imaging with an untrained neural network," *Light: Science & Applications*, vol. 9, no. 1, May 2020. [Online]. Available: <https://doi.org/10.1038/s41377-020-0302-3>
- [33] X. Yang, M. Kahnt, D. Brückner, A. Schropp, Y. Fam, J. Becher, J.-D. Grunwaldt, T. L. Sheppard, and C. G. Schroer, "Tomographic reconstruction with a generative adversarial network," *Journal of Synchrotron Radiation*, vol. 27, no. 2, pp. 486–493, Feb. 2020. [Online]. Available: <https://doi.org/10.1107/s1600577520000831>
- [34] S. Aslan, Z. Liu, V. Nikitin, T. Bicer, S. Leyffer, and D. Gürsoy, "Joint ptycho-tomography with deep generative priors," *Machine Learning: Science and Technology*, vol. 2, no. 4, p. 045017, Dec 2021. [Online]. Available: <https://iopscience.iop.org/article/10.1088/2632-2153/ac1d35>
- [35] X. Pan, X. Zhan, B. Dai, D. Lin, C. C. Loy, and P. Luo, "Exploiting deep generative prior for versatile image restoration and manipulation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 1–1, 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9547753/>

- [36] F. Guzzi, G. Kourousias, A. Gianoncelli, F. Billè, and S. Carrato, "A parameter refinement method for ptychography based on deep learning concepts," *Condensed Matter*, vol. 6, no. 4, 2021. [Online]. Available: <https://www.mdpi.com/2410-3896/6/4/36>
- [37] T. Nguyen, Y. Xue, Y. Li, L. Tian, and G. Nehmetallah, "Deep learning approach for fourier ptychography microscopy," *Optics Express*, vol. 26, no. 20, p. 26470, Sep. 2018. [Online]. Available: <https://doi.org/10.1364/oe.26.026470>
- [38] Z. Guan, E. Tsai, X. Huang, K. Yager, and H. Qin, "PtychoNet: Fast and high quality phase retrieval for ptychography," Tech. Rep., Sep. 2019. [Online]. Available: <https://doi.org/10.2172/1599580>
- [39] R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.
- [40] J. R. Fienup, "Reconstruction of an object from the modulus of its fourier transform," *Optics letters*, vol. 3, no. 1, pp. 27–29, 1978.
- [41] —, "Phase retrieval algorithms: a comparison," *Applied optics*, vol. 21, no. 15, pp. 2758–2769, 1982.
- [42] —, "Reconstruction of a complex-valued object from the modulus of its fourier transform using a support constraint," *JOSA A*, vol. 4, no. 1, pp. 118–123, 1987.
- [43] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [44] A. Creswell, T. White, V. Dumoulin, K. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Processing Magazine*, vol. 35, no. 1, pp. 53–65, 2018.
- [45] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of stylegan," 2019. [Online]. Available: <https://arxiv.org/abs/1912.04958>
- [46] P. Wang, "Stylegan2 pytorch," <https://github.com/lucidrains/stylegan2-pytorch>, 2020.
- [47] T. Karras, S. Laine, and T. Aila, 2018. [Online]. Available: <https://arxiv.org/abs/1812.04948>
- [48] Y. S. Nashed, T. Peterka, J. Deng, and C. Jacobsen, "Distributed automatic differentiation for ptychography," *Procedia Computer Science*, vol. 108, p. 404–414, 2017. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S1877050917306531>
- [49] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, cite arxiv:1412.6980Comment: Published as a conference paper at the 3rd International Conference for Learning Representations, San Diego, 2015. [Online]. Available: <http://arxiv.org/abs/1412.6980>
- [50] M. Hidayetoğlu, T. Biçer, S. G. de Gonzalo, B. Ren, D. Gürsoy, R. Kettimuthu, I. T. Foster, and W. mei W. Hwu, "MemXCT," in *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*. ACM, Nov. 2019. [Online]. Available: <https://doi.org/10.1145/3295500.3356220>
- [51] M. Majchrowicz, P. Kapusta, L. Jackowska-Strumiłło, R. Banasiak, and D. Sankowski, "Multi-GPU, multi-node algorithms for acceleration of image reconstruction in 3d electrical capacitance tomography in heterogeneous distributed system," *Sensors*, vol. 20, no. 2, p. 391, Jan. 2020. [Online]. Available: <https://doi.org/10.3390/s20020391>



Aggelos K. Katsaggelos is Professor and Joseph Cummings Chair in the ECE Department at Northwestern University. He also runs the Image and Video Processing Laboratory (IVPL), whose objective is to generate cutting-edge research results in the fields of multimedia signal processing, multimedia communications, and computer vision. IVPL works in a variety of problems (e.g., recovery, compression, segmentation, and speech and speaker recognition) and applications areas (e.g., medical, multi-spectral, and astronomical image processing). Dr. Katsaggelos is a Fellow of the IEEE (1998) and SPIE (2009), the co-inventor of seventeen international patents, the recipient of the IEEE Third Millennium Medal (2000), the IEEE Signal Processing Society Meritorious Service Award (2001), the IEEE Signal Processing Society Technical Achievement Award (2010), and co-author of several award-winning papers.



Semih Barutcu received his B.Sc. degree in Electrical and Electronics Engineering from Bogazici University, Istanbul, Turkey, and his M.Sc degree in Electrical Engineering and Computer Science from Northwestern University, Evanston, IL, USA, in 2017 and 2018, respectively. He is currently working toward a Ph.D. degree at Northwestern University, Evanston, IL, USA. His research interests include computational imaging, deep learning, and inverse problems in computer vision.



Doğa Gürsoy is a computational scientist at the X-ray Science Division of Advanced Photon Source. He is also an adjunct associate professor in the Electrical Engineering and Computer Science Department at Northwestern University, and a fellow of the Northwestern Argonne Institute for Science and Engineering. Dr. Gürsoy's research focuses on computational imaging and inverse problems in imaging sciences. He is an associate editor for the IEEE Transactions on Computational Imaging, and a member of IEEE Signal Processing Society's