

# PROCEEDINGS OF SPIE

[SPIDigitalLibrary.org/conference-proceedings-of-spie](https://SPIDigitalLibrary.org/conference-proceedings-of-spie)

## Towards fast ptychography image reconstruction of EUV masks by deep neural networks

Paolo Ansuinelli, Benjamín Béjar Haro, Yasin Ekinici, Iacopo Mochi

Paolo Ansuinelli, Benjamín Béjar Haro, Yasin Ekinici, Iacopo Mochi, "Towards fast ptychography image reconstruction of EUV masks by deep neural networks," Proc. SPIE 12751, Photomask Technology 2023, 127510Q (21 November 2023); doi: 10.1117/12.2685227

**SPIE.**

Event: SPIE Photomask Technology + EUV Lithography, 2023, Monterey, California, United States

# Towards fast ptychography image reconstruction of EUV masks by deep neural networks

Paolo Ansuinelli<sup>a</sup>, Benjamín Béjar Haro<sup>a,b</sup>, Yasin Ekinici<sup>a</sup>, and Iacopo Mochi<sup>a</sup>

<sup>a</sup>Paul Scherrer Institut, 5232 Villigen PSI, Switzerland

<sup>b</sup>Swiss Data Science Center, 8006 Zurich, Switzerland

## ABSTRACT

Extreme ultraviolet (EUV) mask metrology and inspection is crucial to evaluate the quality of devices realized by EUV lithography and to obtain the required yield. Actinic (*i.e.*, at wavelength  $\lambda = 13.5$  nm) mask inspection is particularly essential, as this wavelength ensures an imaging resolution and overall imaging physics that matches the operative condition of the lithography scanner. In recent years, various groups have explored coherent diffractive imaging (CDI), and particularly ptychography, as a candidate method for actinic EUV mask inspection. The simplicity of the ptychography approach, the absence of expensive lenses, and the possibility to image both amplitude and phase structures make this method particularly appealing. Despite these advantages, ptychography suffers from throughput limitations dictated by both the long data acquisition process and the time-demanding phase retrieval step. While the former challenge can be mitigated by advancements on source brilliance and detector technology, the latter clearly demands improvements on the algorithmic front. In this paper, we present our recent results on the study of deep learning as a means to achieve fast, high quality, and phase-sensitive reconstruction of aerial images of EUV masks, given the acquired data and the abundant *a-priori* information on the geometrical layout and chemical composition of the samples. We show that, once trained, the selected Deep Neural Network (DNN) architecture achieves a much faster reconstruction of the sample compared to the standard ptychography approach, while retaining high quality in both magnitude and phase images of the object.

**Keywords:** Deep neural networks, Ptychography, Coherent diffractive imaging (CDI), EUV mask metrology, actinic mask inspection, image reconstruction.

## 1. INTRODUCTION

EUV lithography has reached the high volume manufacturing (HVM) threshold in 2019 and it is now used as a patterning method for leading-edge technology nodes in chip fabs worldwide.<sup>1</sup> The large-scale adoption of EUV lithography drives the need for technology advancements in EUV mask metrology, which aims to verify that the mask meets appropriate quality criteria and that eventual amplitude and phase defects do not print on wafer thereby causing a yield loss. The mask can be inspected by systems that deploy a different wavelength than the one used in the lithography step, *e.g.*, e-beam inspection or deep ultraviolet (DUV) tools.<sup>2</sup> However, these approaches fail to exactly replicate the imaging conditions in the scanner. Actinic mask inspection systems, on the other hand, mimic the operational conditions of the scanner by imaging at a wavelength of 13.5 nm.<sup>3–6</sup> Ptychography, a computational imaging method that images a sample by processing a series of diffraction patterns obtained by illuminating the object at overlapping probe positions, is a promising alternative to conventional actinic mask inspection methods, owing to its robustness, flexibility and relatively simple and lens-free optical design.<sup>7,8</sup> On the other hand, the need for a scanning illumination, the requirement for a large overlap among adjacent probe positions, the need to handle a significant volume of data, and the slow iterative phase retrieval process pose throughput challenges that need to be properly evaluated and addressed.

Particularly, the need for fast reconstruction methods demands substantial progress on the numerical algorithms side. The far-field ptychography methods are based on propagating the optical fields back and forth among real and Fourier space. This operation is usually performed via Fast Fourier Transforms (FFT) that

---

Further author information / E-mail: paolo.ansuinelli@psi.ch

favorably scale, for  $N$  unknowns, as  $\mathcal{O}(N \log N)$ . Nevertheless, ptychography methods typically require many iterations to converge, leading to an overall slow phase retrieval process.

In the last few years several works have been published on the topic of phase retrieval using deep neural networks (DNN), comprising both the single intensity phase retrieval problem and ptychography, *e.g.*<sup>9–12</sup> In these approaches, a DNN architecture is trained to learn the mapping from the diffraction data to the real space (complex-valued) object. Once trained, the network is used at inference time, to estimate an unseen object given a set of diffraction patterns generated from it. While training these networks is a long process, inference is usually very fast, hence making this approach suitable for fast and high-quality imaging of a given object from its diffraction data.

In this work, we demonstrate the use of a DNN architecture for fast ptychography imaging of a patterned EUV mask. Particularly, we show how the availability of *a-priori* geometrical and optical information about the sample can be included in the formulation of the imaging problem to improve its robustness.

## 2. MATERIALS AND METHODS

### 2.1 EUV mask pattern and materials.

The reticle used in this case consists of a random, logic-like pattern, inclusive of programmed defects, realized on top of a multilayer. A top view of the sample geometry is given in Figure 1.

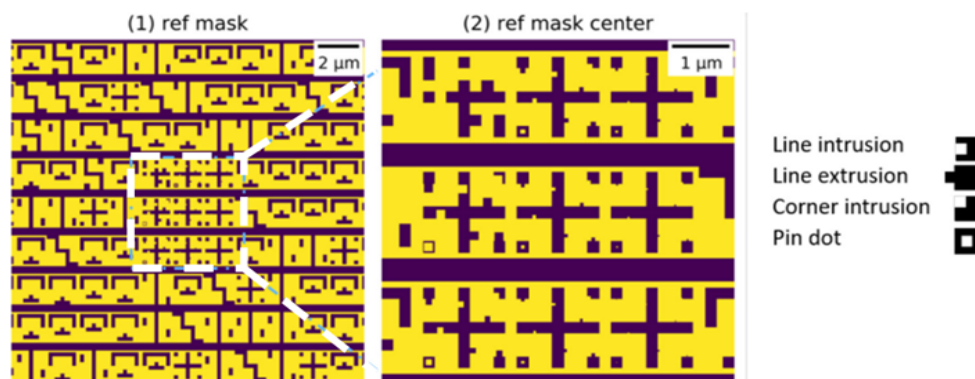


Figure 1: (1) The mask layout used in the present study (2) Zoom-in layout of the EUV mask at the central region and the types of programmed defects.

The absorber is a 140 nm thick layer of hydrogen silesquioxane (HSQ) on 40 bilayers of Mo/Si on top of a low thermal expansion glass substrate. The values of the complex refractive index,  $\underline{n} = n - ik$ , of each material and the thickness of the layers are reported in Table 1.<sup>13,14</sup> The programmed defects include phase structures,

Table 1: **Layers thicknesses and Materials at  $\lambda = 13.5$  nm**

layer	thickness [nm]	$n$	$k$
HSQ	140	0.98075	-0.0065
Si	2.506	0.99932	-0.00183
MoSi2	0.802	0.9693	-0.00433
Mo	1.904	0.92108	-0.00644
MoSi2	1.844	0.9693	-0.00433
Si	1e6	0.99932	-0.00183

3.5 nm carbon pillars buried under the multilayer and absorber defects. The effect of the buried defects is visible in Fig. 3, around the border of the main object. The programmed absorber defects are located in the center of the object in Fig. 3. A zoom on these defects is given in Fig. 4.

## 2.2 Computation of the mask *prior*.

The availability of the nominal mask design and the knowledge of its optical properties can be used to generate a *prior* for the image of the mask, which in ptychography terms is an estimation of the reflection function of the *a-priori* known object. The most robust and physically sound way to do so consists in using a rigorous Maxwell solver to evaluate the exit wave on top of the scattering geometry due to the probe-object interaction, and descramble from this exit wave the illumination function.<sup>8</sup> However this approach is cumbersome, slow, and can be computationally prohibitive when the object is non-periodic and the field of view (FOV) is much larger than the wavelength. In this work, we have adopted a different approach by approximating this reflection function by the transmission matrix method (TMM)<sup>15</sup> rather than estimating it from rigorous electromagnetic simulations. The TMM method, for a given multilayer, is an analytical procedure which is based on expressing the electric and magnetic fields in each slab of the multilayer, and on solving for the unknowns of the boundary value problem (*i.e.*, the coefficients of the electric and magnetic fields in each layer) by enforcing the continuity of the tangential components of the fields at every layer interface. It can be shown that for a system of  $N$  interfaces the tangential components of the field at the last slab and at the first one are interrelated by:

$$\begin{bmatrix} E_1 \\ H_1 \end{bmatrix} = \mathcal{M}_1 \mathcal{M}_2 \dots \mathcal{M}_{N-1} \begin{bmatrix} E_N \\ H_N \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} E_N \\ H_N \end{bmatrix}, \quad (1)$$

where the coefficients of the matrices are related to the thickness, the index of refraction, the angle of incidence, and the impedance of the layer.<sup>15</sup> The equivalent reflection function of the object is computed for a plane wave that illuminates the sample at  $6^\circ$  angle of incidence. It should be noted that this approach assumes layers that are infinite in the lateral direction. Because the pattern on top of the multilayer has critical dimensions that are much larger than the wavelength (about 20 times as big) and because of the low optical contrast, one can consider – as a first approximation – that diffraction effects are negligible during the generation of the real space and complex valued *prior*. Hence, in those parts of the mask where the absorber is not present, the reflection function is evaluated by a TMM method for the multilayer alone and where the absorber is present the reflection function is evaluated as given by the multilayer covered with a 140 nm thick HSQ absorber which extends infinitely in the horizontal direction. With the nominal mask design and the material properties as described in Table 1, we generate an *a-priori* complex-valued reflection function for the sample. This *prior* is subsequently padded to the same FOV of the reference ptychography reconstruction, rotated and aligned with the reference object by means of a phase correlation method.<sup>16</sup> Afterwards, the *prior* is propagated to the detector (NA=0.22) and inverse Fourier transformed so to have the same resolution as the reference. The amplitude and phase histograms of the *prior*, are adapted to the ones of the reference. The result of this procedure is illustrated in Figures 2 and 3, where amplitude and phase are normalized to have values between 0 and 1.

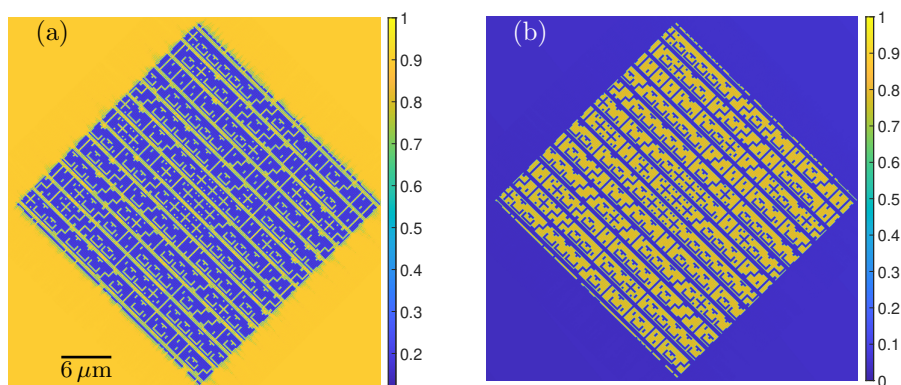


Figure 2: *Prior* of the object. a) Magnitude. b) Phase.

Note that the method used to generate the *prior* (TMM) is physically different from the one used to generate the reference object (ptychography). On top of this difference, the *prior* is generated from a version of the mask design layout without the programmed defects which are instead present in the fabricated EUV mask.



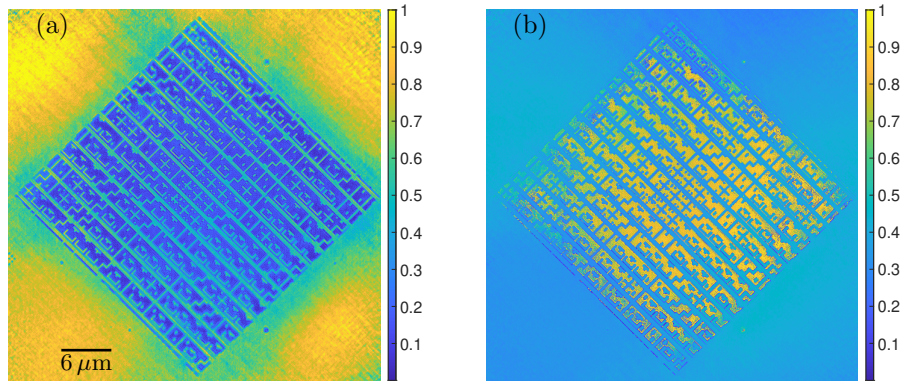


Figure 3: Reference ptychography reconstruction. a) Magnitude. b) Phase.

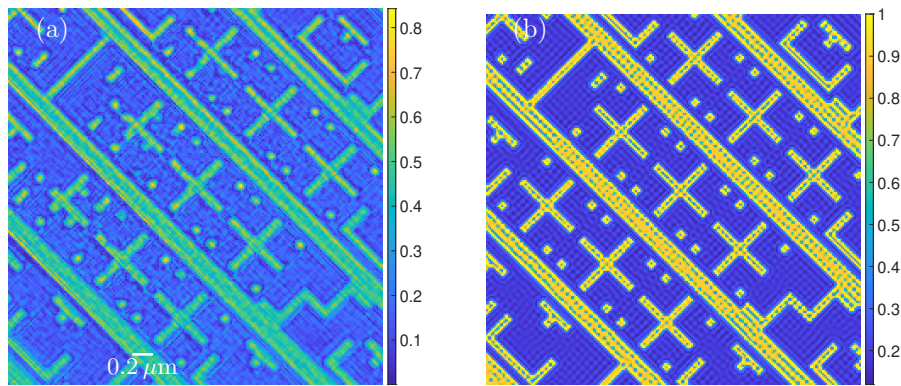


Figure 4: Zoom on the programmed absorber defects. a) Ptychography magnitude. b) *Prior* magnitude.

The DNN is tasked with learning how to efficiently reconstruct the object from the diffraction patterns and the ideal (without defects) mask prior. As ground-truth for the reconstruction we use the complex object estimated using ptychography.

### 2.3 DNN architecture.

The architecture chosen in this work is a U-net.<sup>17</sup> A U-net architecture consists of a contracting (encoder) and expanding (decoder) part. The encoder is a sequence of convolutional and max-pooling layers followed by a nonlinear activation unit. The convolutional layers extract features from their own input by a convolution with a learned filter, while the max-pooling layers downsample the input while preserving its most important characteristics. As the encoder approaches the inner layer the amount of features extracted, and their complexity, increases at the expense of lower spatial resolution. In the decoder part the information contained in the feature map generated by the encoder gets processed by additional convolutional layers and upsampled, and the final output gets reconstructed by learning its relation with the feature representation of the input. On top of the increased spatial resolution provided by the learned upsampling layer, the skip connections provide additional information to the decoder by “passing” the features map extracted by the encoder. The U-net is a widely adopted architecture, and its use in phase retrieval is documented *e.g.*, in references.<sup>9,18,19</sup>

### 2.4 Training dataset generation.

The measured data used in this work is acquired by illuminating the fabricated EUV mask in 106 probe positions according to a Fermat spiral trajectory. Because the training of DNN architectures requires diversity in the dataset and a high quantity of data we have increased the data-set by applying data augmentation (in this case rotation only) to the object and to the diffraction patterns. On top of the data augmentation we have enriched

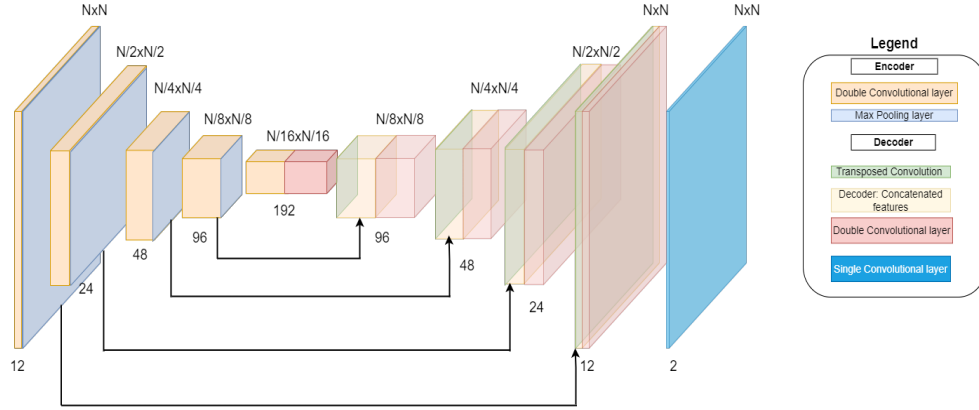


Figure 5: The DNN architecture used in this study.

the dataset by generating synthetic diffraction patterns. The synthetic dataset is obtained by simulating the far-field intensity generated by scanning the reconstructed probe through the reconstructed object in a raster scan. This process leads to overall about 5000 probe positions. 80% of both the experimental and synthetic data is used for training. The remaining 20% is used for testing. The input to the DNN, at each probe position, consists of two channels (matrices): one diffraction pattern and the phase of the *prior*. The magnitude of the *prior* is not used as an input channel in order to avoid over-feeding the DNN with information coming from the nominal mask design, which could result in excessive bias towards the *prior*. It should be noted that the acquired diffraction patterns, the respective *prior* and labels, have a high number of pixels (1800x1800). This translates in a long training time. To reduce the computational burden associated to the training we have binned the dataset by a factor 2 and correspondingly cropped the FOV of the object and of the *prior* of the same factor. The cropping is done, per probe position, by truncating the FOV around the center, while the object is shifted inside its matrix to account for the illumination shift in the ptychography method.

During the training phase we sample randomly – from the probe positions selected for training – a total of four inputs (*i.e.*, 4 diffraction patterns and 4 *priors*) to generate the mini-batch for training. Therefore the tensor in input to the DNN has a size of  $[4, 2, m, m]$ , where 4 is the number of probe positions, 2 is the number of channels and  $m$  is the number of pixels across one dimension of the input matrices (960 in this case). In this initial implementation we did not require for the probe positions to be overlapping, which is a requirement of conventional ptychography.

We introduce the following nomenclature:  $O_{j,DNN}$  is the complex object estimated by the DNN at the  $j^{th}$  probe position,  $O_{ptycho}$  is the reference ptychography object. The symbols  $|\cdot|$  and  $\angle$  indicate the magnitude and the phase of the complex object.

In this work we use a supervised training approach, where the weights of the DNN are adjusted by minimizing, per probe position, the following cost functional:

$$\min_{O_{j,DNN}} \left\| |\mathcal{F}(P(\mathbf{r} - \mathbf{R}_j)O_{j,DNN}(\mathbf{r}))| - \sqrt{I_j(\mathbf{k})} \right\|_2^2 + \left\| |O_{j,DNN}| - |O_{j,ptycho}| \right\|_2^2 + \left\| \angle O_{j,DNN} - \angle O_{j,ptycho} \right\|_2^2. \quad (2)$$

### 3. RESULTS.

In this section we show some of the results obtained by employing the DNN architecture described above. An example of the input channels to the DNN for the central probe position is reported in Fig. 6.

For the results presented in this Section we have trained the DNN for 50000 iterations. About 20% of the patterns acquired in the Fermat spiral have been used for testing. In what follows we highlight inference results obtained by using one of the probe positions in the Fermat spiral close to the object center, which we have kept out of the training dataset. The object predicted during inference by the DNN, in magnitude and phase, and the one given by ptychography are shown in Figs. 7, 8. The white boxes/circles in the images highlight those

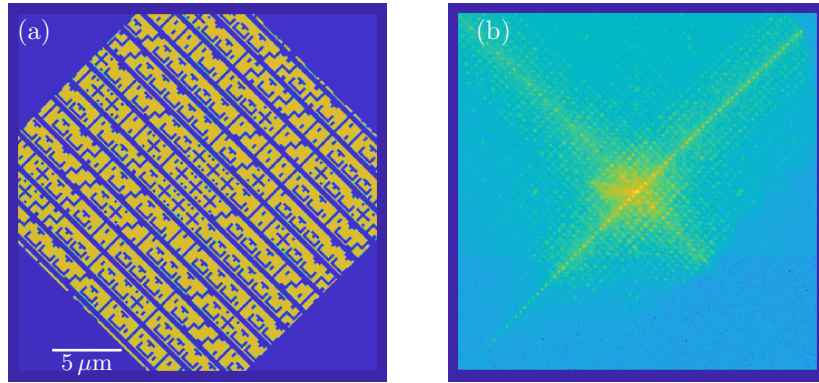


Figure 6: The two channels that constitute, per probe position, the input to the DNN. a) Phase of the *prior*. b) Diffraction pattern in log scale.

parts of the object where the *prior* and the reference ptychography object are dissimilar (*i.e.*, the white shapes are placed at the defective sites).

The presence of the defects in the DNN outputs in Figs. 7, 8 (c) shows that the DNN architecture is effectively learning about the defects' presence from the diffraction pattern and from the labels.

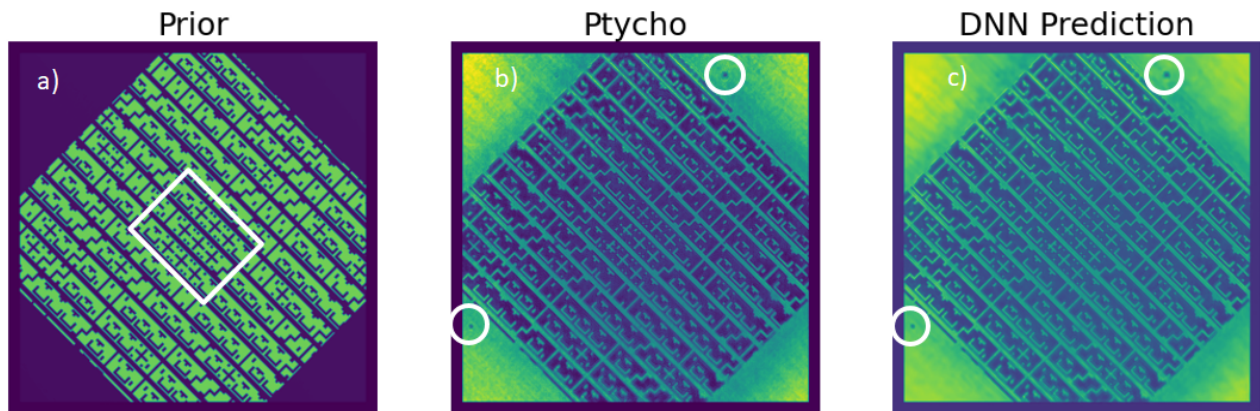


Figure 7: Magnitude of the complex object. a) *Prior*. b) Ptychography reconstruction. c) DNN prediction. The white circles and the white square highlight defective sites where the object and the *prior* are dissimilar.

Figures 9 and 10 show an inset of the two images within the white box in Figs. 7, 8.

Notice, in Figs. 9 and 10, that the absorber defects, present in the Ptychography and in the DNN reconstruction are absent in the *prior*. On top of this difference, we also notice that the overall image quality of these figures is improved, as the background appears overall more uniform and less impacted by the noise.

The images shown in Fig. 7 are relative to a ptychography dataset containing 106 scan points. The standard ptychography reconstruction, shown in Fig. 7b, was carried out with a GPU-accelerated Difference-Map algorithm<sup>20</sup> using a dedicated work station with 2× Intel Xeon CPU E5-2697 v4 2.30GHz processors (18 cores 36 threads) and 2× Tesla P100 GPU (3584 CUDA cores, 16GB memory). The reconstruction required a total time of 625 s and 2000 iterations of the ptychography algorithm. In comparison, the inference step using the trained DNN was able to obtain the image in Fig. 7c in 0.02 s on a similar machine from only 4 scan points. Although this is not a proper benchmark, these result gives an idea of the potential reconstruction speed improvement that can be achieved with the proposed approach.



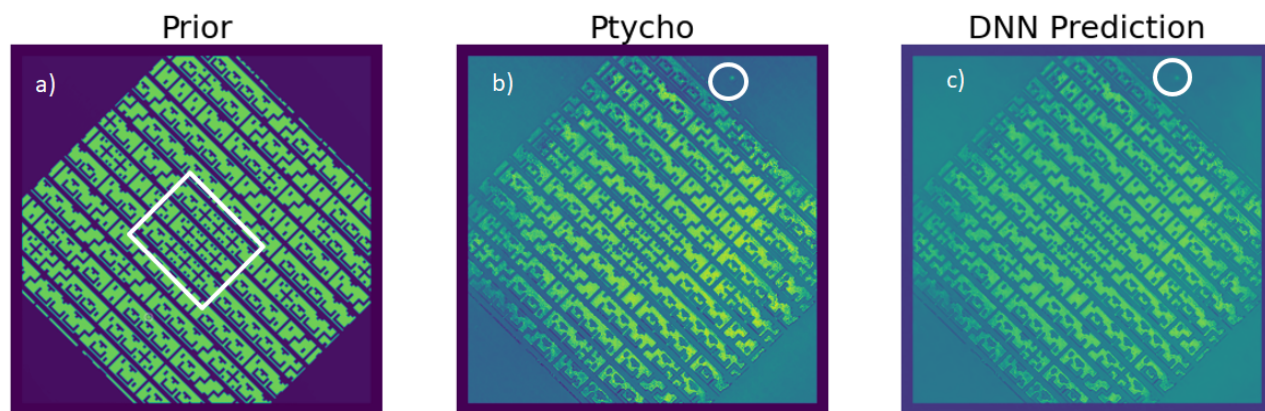


Figure 8: Phase of the complex object. a) *Prior*. b) Ptychography reconstruction. c) DNN prediction. The white circles and the white square highlight defective sites where the object and the *prior* are dissimilar.

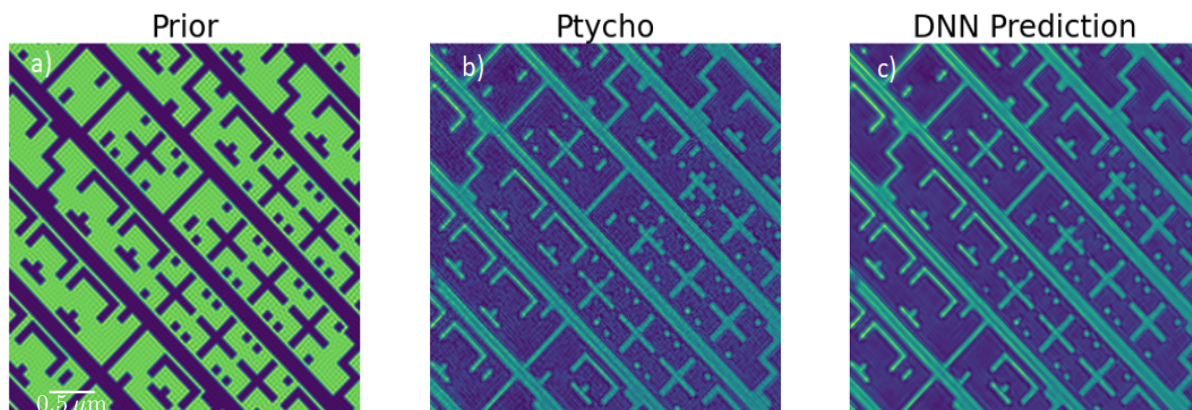


Figure 9: Magnitude of the complex object. a) *Prior*. b) Ptychography. c) DNN.

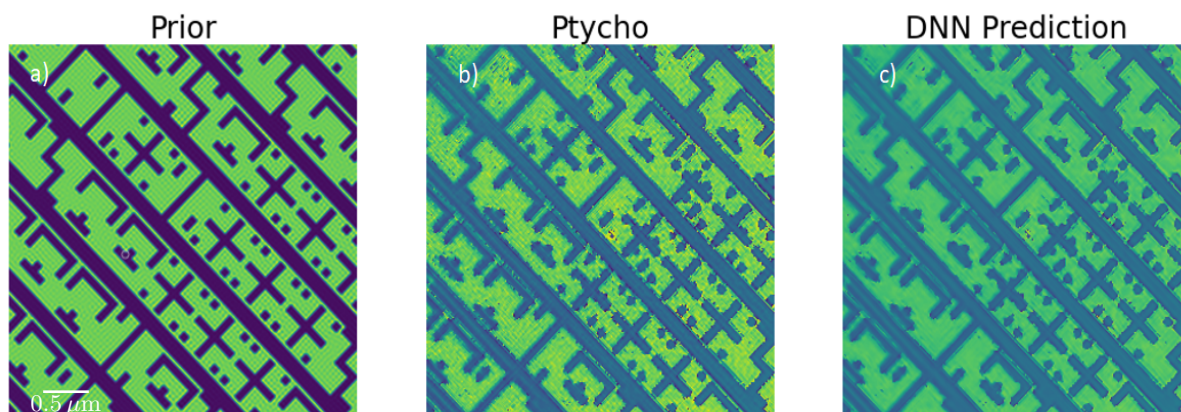


Figure 10: Phase of the complex object. a) *Prior*. b) Ptychography. c) DNN.

## 4. CONCLUSIONS

In this paper we have presented a study on the use of a neural network architecture for the imaging of a patterned EUV mask. Motivated by the need for fast ptychography reconstruction, we apply a U-net architecture to image an intricate random logic-like pattern. We use the nominal information of the mask to estimate a *prior* of the complex reflection function of the object. The phase of the *a-priori* estimated object reflection function is used – in addition to the diffraction pattern – as input to the DNN architecture. This paper proposes a supervised learning approach, where the weights of the DNN are adjusted by minimizing a loss functional which promotes, at the same time, adhesion of the estimated diffraction data with the one given by the ptychography model and adherence to the complex valued ptychography label (ground truth).

The initial reconstruction speed comparison between the proposed method and the conventional ptychography approach revealed a remarkable enhancement of four orders of magnitude. While it is imperative to validate this outcome through careful benchmarking across diverse datasets, it strongly implies that the use of a prior-primed Deep Neural Network (DNN) holds the potential to significantly augment the reconstruction speed and reduce the data collection requirements. This substantial leap in performance is the key to elevating CDI as a formidable tool for industrial semiconductor device metrology.

## 5. ACKNOWLEDGEMENTS.

The authors acknowledge Dr. Luis Felipe Barba Flores for his useful suggestions in the design of the learning problem.

Authors acknowledge financial support from the Swiss Data Science Center (SDSC) under grant no. C21-17L, the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No. 884104 (PSI-FELLOW-III-3i) and from the SAMSUNG Global Research Outreach program.

## REFERENCES

- [1] Mastenbroek, M., “Progress on 0.33 NA EUV systems for High-Volume Manufacturing,” in [*International Conference on Extreme Ultraviolet Lithography 2019*], Itani, T., Gargini, P. A., Naulleau, P. P., and Ronse, K. G., eds., **11147**, 1114703, International Society for Optics and Photonics, SPIE (2019).
- [2] Liang, T., Tejnil, E., and Stivers, A. R., “Pattern inspection of EUV masks using DUV light,” in [*22nd Annual BACUS Symposium on Photomask Technology*], Grenon, B. J. and Kimmel, K. R., eds., **4889**, 1065 – 1072, International Society for Optics and Photonics, SPIE (2002).
- [3] Goldberg, K., Benk, M., Wojdyla, A., Mochi, I., Rekawa, S., Allézy, A., Dickinson, M., Cork, C., Chao, W., Zehm, D., Macdougall, J., Naulleau, P., and Rudack, A., “Actinic mask imaging: Recent results and future directions from the sharp euv microscope,” **9048**, 90480Y (04 2014).
- [4] Mochi, I., Kim, H.-S., Dejkameh, A., Nebling, R., Kazazis, D., Locans, U., Shen, T., and Ekinici, Y., “Lensless EUV mask inspection for anamorphic patterns,” in [*Extreme Ultraviolet (EUV) Lithography XII*], Felix, N. M. and Lio, A., eds., **11609**, 116090M, International Society for Optics and Photonics, SPIE (2021).
- [5] Ansuinelli, P., Coene, W. M., and Urbach, H. P., “Refinement strategies for optimal inclusion of prior information in ptychography,” in [*Extreme Ultraviolet Lithography 2020*], Naulleau, P. P., Gargini, P. A., Itani, T., and Ronse, K. G., eds., **11517**, 1151711, International Society for Optics and Photonics, SPIE (2020).
- [6] Bahrenberg, L., Herbert, S., Tempeler, J., Maryasov, A., Hofmann, O., Danylyuk, S., Lebert, R., Loosen, P., and Juschkin, L., “Analysis of distinct scattering of extreme ultraviolet phase and amplitude multilayer defects with an actinic dark-field microscope,” in [*Extreme Ultraviolet (EUV) Lithography VI*], II, O. R. W. and Panning, E. M., eds., **9422**, 942229, International Society for Optics and Photonics, SPIE (2015).
- [7] Rajendran, R., Mochi, I., Helfenstein, P., Mohacsi, I., Redford, S., Mozzanica, A., Schmitt, B., Yoshitake, S., and Ekinici, Y., “Towards a stand-alone high-throughput EUV actinic photomask inspection tool: RESCAN,” in [*Metrology, Inspection, and Process Control for Microlithography XXXI*], Sanchez, M. I., ed., **10145**, 101450N, International Society for Optics and Photonics, SPIE (2017).

- [8] Ansuinelli, P., Coene, W. M. J., and Urbach, H. P., “Improved ptychographic inspection of euv reticles via inclusion of prior information,” *Appl. Opt.* **59**, 5937–5947 (Jul 2020).
- [9] Yang, D., Zhang, J., Tao, Y., Lv, W., Lu, S., Chen, H., Xu, W., and Shi, Y., “Dynamic coherent diffractive imaging with a physics-driven untrained learning method,” *Opt. Express* **29**, 31426–31442 (Sep 2021).
- [10] Kim, J. W., Messerschmidt, M., and Graves, W. S., “Performance Evaluation of Deep Neural Network Model for Coherent X-ray Imaging,” *Ai* **3**(2), 318–330 (2022).
- [11] Cherukara, M. J., Zhou, T., Nashed, Y., Enfedaque, P., Hexemer, A., Harder, R. J., and Holt, M. V., “AI-enabled high-resolution scanning coherent diffraction imaging,” *Applied Physics Letters* **117**(4) (2020).
- [12] Yao, Y., Chan, H., Sankaranarayanan, S., Balaprakash, P., Harder, R. J., and Cherukara, M. J., “AutoPhaseNN: unsupervised physics-aware deep learning of 3D nanoscale Bragg coherent diffraction imaging,” *npj Computational Materials* **8**(1), 124 (2022).
- [13] Fernandez, S., Kazazis, D., Rajendran, R., Mochi, I., Helfenstein, P., Yoshitake, S., and Ekinici, Y., “Comparative study of extreme ultraviolet absorber materials using lensless actinic imaging,” *Journal of Micro/Nanolithography, MEMS, and MOEMS* **18**(1), 013506 (2019).
- [14] Mochi, I., Fernandez, S., Nebling, R., Locans, U., Rajeev, R., Dejkameh, A., Kazazis, D., Tseng, L.-T., Danylyuk, S., Juschkin, L., and Ekinici, Y., “Quantitative characterization of absorber and phase defects on EUV reticles using coherent diffraction imaging,” *Journal of Micro/Nanolithography, MEMS, and MOEMS* **19**(1), 014002 (2020).
- [15] Hecht, E., [*Optics*], Pearson (2012).
- [16] Reddy, B. and Chatterji, B., “An fft-based technique for translation, rotation, and scale-invariant image registration,” *IEEE Transactions on Image Processing* **5**(8), 1266–1271 (1996).
- [17] Ronneberger, O., Fischer, P., and Brox, T., “U-net: Convolutional networks for biomedical image segmentation,” (2015).
- [18] Zhang, F., Liu, X., Guo, C., Lin, S., Jiang, J., and Ji, X., “Physics-based iterative projection complex neural network for phase retrieval in lensless microscopy imaging,” in [*2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*], 10518–10526 (2021).
- [19] Nagahama, Y., “Phase retrieval using hologram transformation with u-net in digital holography,” *Opt. Continuum* **1**, 1506–1515 (Jul 2022).
- [20] Thibault, P., Dierolf, M., Bunk, O., Menzel, A., and Pfeiffer, F., “Probe retrieval in ptychographic coherent diffractive imaging,” *Ultramicroscopy* **109**(4), 338–343 (2009).