

Data augmentation in extreme ultraviolet lithography simulation using convolutional neural network

Hiroyoshi Tanabe* and Atsushi Takahashi

Tokyo Institute of Technology, Tokyo, Japan

Abstract

Background: In the previous work, we developed a convolutional neural network (CNN), which reproduces the results of the rigorous electromagnetic (EM) simulations in a small mask area. The prediction time of CNN was 5000 times faster than the calculation time of EM simulation. We trained the CNN using 200,000 data, which were the results of EM simulation. Although the prediction time of CNN was very short, it took a long time to build a huge amount of the training data. Especially when we enlarge the mask area, the calculation time to prepare the training data becomes unacceptably long.

Aim: Reducing the calculation time to prepare the training data.

Approach: We apply data augmentation technique to increase the number of training data using limited original data. The training data of our CNN are the diffraction amplitudes of mask patterns. Assuming a periodic boundary condition, the diffraction amplitudes of the shifted or flipped mask pattern can be easily calculated using the diffraction amplitudes of the original mask pattern.

Results: The number of training data after the data augmentation is multiplied by 200 from 2500 to 500,000. Using a large amount of training data, the validation loss of CNN was reduced. The accuracy of CNN with augmented data is verified by comparing the CNN predictions with the results of EM simulation.

Conclusions: Data augmentation technique is applied to the diffraction amplitude of the mask pattern. The data preparation time is reduced by a factor of 200. Our CNN almost reproduces the results of EM simulation. In this work, the mask patterns are restricted to line and space patterns. It is a challenge to build several CNNs for specific mask patterns or ultimately a single CNN for arbitrary mask patterns.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 International License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JMM.21.4.041602](https://doi.org/10.1117/1.JMM.21.4.041602)]

Keywords: lithography simulation; neural network; extremely ultraviolet mask.

Paper 22021SS received May 12, 2022; accepted for publication Jul. 29, 2022; published online Oct. 14, 2022.

1 Introduction

High aspect absorbers used in extremely ultraviolet (EUV) masks induce several mask three-dimensional (3D) effects, such as critical dimension (CD) and image placement errors.^{1,2} It is necessary to include the mask 3D effects in EUV lithography simulation. Mask 3D effects can be calculated rigorously using electromagnetic (EM) simulators.³⁻⁷ However, these simulators are highly time-consuming for full-chip applications.

Recently, many attempts have been made to simulate the mask 3D effects using deep neural networks (DNNs). They are classified into three models depending on the targets of DNNs. Three possible targets are, from the mask plane to the wafer plane, the near-field amplitude on the mask, the far-field amplitude (diffraction spectrum) at the pupil of the projection optics, and the image intensity on the wafer. In the first model, the target is the near-field amplitude on the mask calculated by EM simulation.⁸⁻¹¹ This model requires many DNNs to reproduce

*Address all correspondence to Hiroyoshi Tanabe, tanabe.h.af@m.titech.ac.jp

different near-field amplitudes depending on the source position. In the second model, which is our model,¹² the target of DNN is the far-field amplitude at the pupil of the projection optics. Because the far-field amplitudes are described in momentum (wave vector) space and the source position corresponds to the incident momentum in Koehler illumination, our model naturally parametrizes the source position dependence of the amplitude. The third model^{13,14} uses the image intensity on the wafer as the target of DNN. This model is much straightforward than other models because the image intensity is used in the following resist simulation. However, the phase information is lost when the diffraction amplitude is converted to the image intensity. The phase of the amplitude is not included in the targets of this model. The phase indirectly influences the focus dependence of the intensity. Therefore, the model needs many intensity targets at different focus positions for each mask pattern.

In our previous work,¹² we developed a convolutional neural network (CNN), which reproduces the results of the rigorous EM simulation in a small mask area. The prediction time of CNN was 5000 times faster than the calculation time of EM simulation. We trained the CNN using 200,000 data, which were the results of EM simulation. Such a large amount of the data was necessary to reduce the validation loss during the training. Although the prediction time of CNN was very short, it took a long time (~1 week) to build the training data. Creating the training data in the work was possible because the mask area was small. However, when we enlarge the mask area for optical proximity correction (OPC) in large area, the calculation time to prepare the training data becomes unacceptably long. In the large area OPC process, the large mask area is clipped into many small mask areas. The size of the clipped areas needs to be large enough to avoid the influence from the surrounding mask pattern, at least near the center of the clipped area.

In this work, we apply data augmentation to our CNN, which is a standard technique in DNN. The technique allows us to increase the number of the training data without performing EM calculation, which significantly reduces the time to prepare the training data. In Sec. 2, we explain the detail of our data augmentation technique. In this work, we focus on the application to metal layers and assume that the mask patterns are simple line and space patterns. In Sec. 3, we study the accuracy of our CNN prediction on CDs and edge placement errors (EPEs). Section 4 is the summary.

2 Data Augmentation for Large Mask Patterns

In the previous work,¹² we assumed a periodic mask pattern with 720 nm × 720 nm mask area. When we clip out a small mask area from the mask data, we should not use the edges of the mask area to avoid the influence of the neighboring mask pattern. According to Ref. 15, the optical interaction range R_{opt} is calculated as

$$R_{\text{opt}} = \frac{1.12\lambda}{\sigma\text{NA}}, \quad (1)$$

where λ , σ , and NA represent the wavelength, coherence factor, and numerical aperture of the scanner, respectively. The wavelength of EUV light is 13.5 nm, and the numerical aperture of the current EUV scanner is 0.33. The coherence factor depends on the illumination setting, and the typical value is 0.5. Inserting these values in Eq. (1), the optical interaction range $R_{\text{opt}} = 90$ nm. This value is the length on the wafer, and the number is multiplied by four on the mask. Therefore, the optical interaction range on the mask is $4 \times R_{\text{opt}} = 360$ nm. Figure 1 shows the usable mask area excluding the area influenced by the neighboring mask pattern. The mask size L should be larger than 720 nm to get the usable mask area. Therefore, there was no usable area for large area OPC in our previous work.

In this work, we choose 1024 nm × 1024 nm mask area. The usable mask area is 300 nm². The usable area is not large, but the EM simulation time highly depends on the size of the mask area. The calculation time of 1024 nm × 1024 nm mask area takes 162 s using Core i9-9900K CPU. In the simulation, we use 3D waveguide model,⁵⁻⁷ which solves coupled wave equations in momentum space. The calculation time highly depends on the cut-off momentum. In this work, we include the momentum (k_x, k_y) , which satisfies

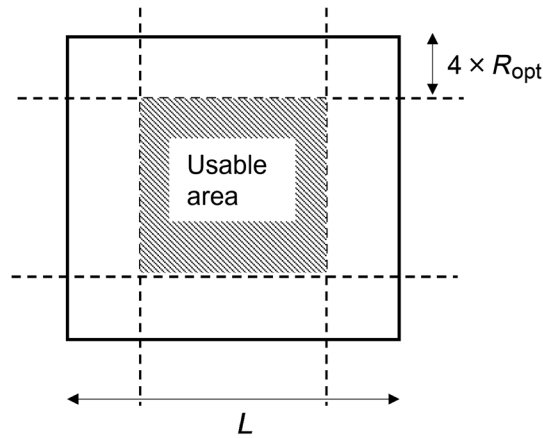


Fig. 1 Usable mask area.

$$\left(\frac{|k_x|}{k_x^{\max}} + 1\right) \left(\frac{|k_y|}{k_y^{\max}} + 1\right) \leq 2, \tag{2}$$

where $k_x^{\max} = k_y^{\max} = 6 \cdot \frac{NA}{4} \frac{2\pi}{\lambda}$. This number is six times larger than the size of the pupil $\frac{NA}{4} \frac{2\pi}{\lambda}$. Discretizing the momentum by $2\pi/L$, there are 2121 (k_x, k_y) pairs, which satisfy Eq. (2). The size of the matrix solving the coupled wave equations is 4242×4242 because there are two polarizations. The region in Eq. (2) is quasi-hyperbola, which resembles the diffraction spectrum of mask patterns consisting of vertical and horizontal lines or holes. Mask patterns are conventionally designed using X - Y coordinates. The minimum pattern pitch in X or Y direction is small compared to the minimum pattern pitch in the diagonal direction. Therefore, the diffraction amplitude in the diagonal direction decreases rapidly compared to the amplitude in X or Y direction in momentum space.

DNNs require a large amount of the training data. In the previous work, we used 200,000 training data. It will take a year if we calculate the same number of the data with the mask area in this work. Data augmentation is a powerful technique in deep learning to increase the number of the training data with limited original data. In our CNN, the input is the mask pattern, and the outputs are the far-field diffraction amplitude $A(l, m; l_s, m_s)$, where (l, m) is the diffraction order, and (l_s, m_s) is the source position (Fig. 2). In 3D waveguide model,⁵⁻⁷ not only the diffraction momentum but also the source position (or incident momentum) is discretized by $2\pi/L$. As discussed in Ref. 12, assuming the largest σ value to be 1, the diffraction order and the source position are restricted by the pupil shape and the source shape as follows:

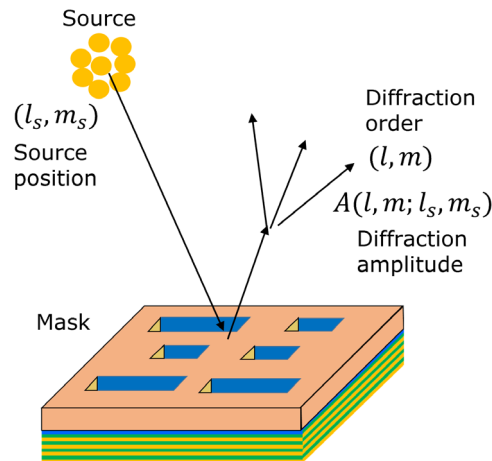


Fig. 2 Schematic view of light diffraction by an EUV mask.

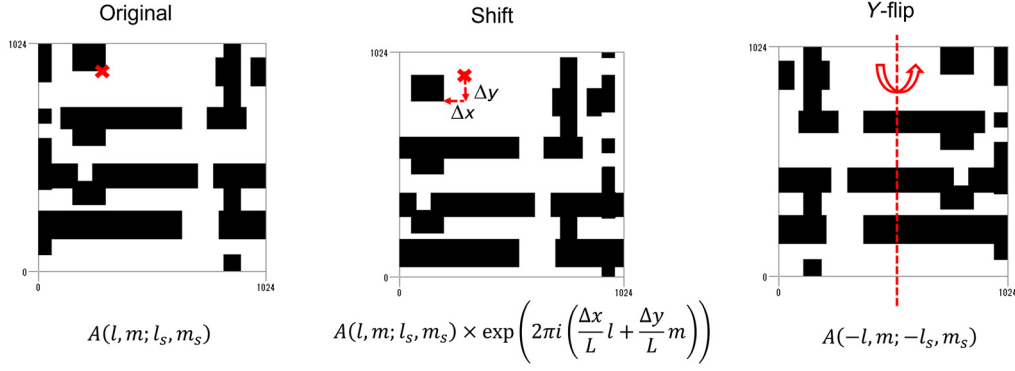


Fig. 3 Original, shifted, and Y-flipped mask patterns.

$$\sqrt{(l + l_s)^2 + (m + m_s)^2} \leq \frac{\text{NA} L}{4 \lambda}, \quad (3)$$

$$\sqrt{l_s^2 + m_s^2} \leq \frac{\text{NA} L}{4 \lambda}. \quad (4)$$

When $L = 1024$ nm, the number of the possible combination of (l, m) is 457.

When the mask pattern is shifted or Y-flipped as shown in Fig. 3, the diffraction amplitudes of these patterns can be easily calculated from that of the original pattern. Note that in EUV reflective optics, X-flip is not symmetrical to the chief ray because it is tilted 6 deg in Y direction.

Following Ref. 12, the far-field diffraction amplitude $\mathbf{A}(l, m; l_s, m_s)$ is divided into the thin mask amplitude (Fourier transform of the mask pattern) $\mathbf{A}^{\text{FT}}(l, m)$, which does not depend on the source position (l_s, m_s) , and the mask 3D amplitude $\mathbf{A}^{\text{3D}}(l, m; l_s, m_s)$

$$\mathbf{A}(l, m; l_s, m_s) = \mathbf{A}^{\text{FT}}(l, m) + \mathbf{A}^{\text{3D}}(l, m; l_s, m_s). \quad (5)$$

Figure 4 shows the source position dependence of the mask 3D amplitude. The source position where the amplitude contributes to the image intensity is limited by the source shape and the pupil shape. Only the overlapping area in Fig. 4 contributes to the image intensity. We approximate the mask 3D amplitude in this area by a linear function of the source position (l_s, m_s) as follows:

$$A_x^{\text{3D}}(l, m; l_s, m_s) \cong a_0(l, m) + a_x(l, m)(l_s + l/2) + a_y(l, m)(m_s + m/2), \quad (6)$$

where a_0 is the mask 3D amplitude at the center of the overlapping area: $(l_s, m_s) = (-l/2, -m/2)$, and a_x and a_y are the slopes of the amplitude in the X and Y directions on the source plane, respectively. We call these three numbers as mask 3D parameters.

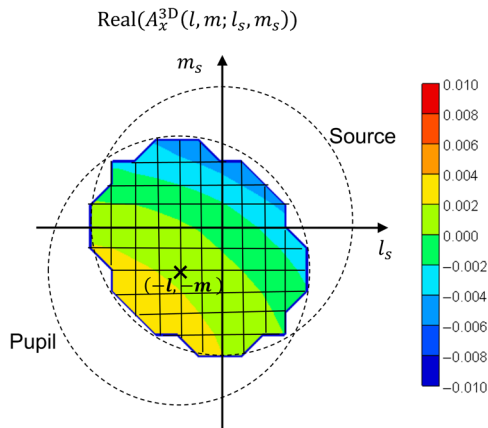


Fig. 4 Source position dependence of mask 3D amplitude.

Equation (6) is slightly different from equation (11) in Ref. 12. We modify the equation for the following reason. 3D waveguide model calculates the diffraction amplitudes at the grid points in Fig. 4. The model solves coupled wave equations, so all the amplitudes are calculated simultaneously. Mask 3D parameters are derived by least square fitting to the amplitudes in the overlapping area. The larger (l, m) , the fewer grid points inside the overlapping area. If the number of the grid points is too small, the overlapping area becomes a line or just a point. In such case, we approximate the amplitude in the area using only $a_0(l, m)$ as the average of the amplitudes and do not use $a_x(l, m)$ and $a_y(l, m)$. Therefore, the number of $a_0(l, m)$ is 457, whereas the number of $a_x(l, m)$ or $a_y(l, m)$ is 349.

We use two different methods to derive $a_0(l, m)$ for small (l, m) and large (l, m) . Equation (6) is used to ensure that $a_0(l, m)$ is always the average of the amplitudes. When we define the overlapping area as O and the number of the grid points in the area as N , the result of the least square fitting to Eq. (6) gives

$$a_0(l, m) = \frac{1}{N} \sum_{(l_s, m_s) \in O} A_x^{3D}(l, m; l_s, m_s), \tag{7}$$

because

$$\sum_{(l_s, m_s) \in O} (l_s + l/2) = \sum_{(l_s, m_s) \in O} (m_s + m/2) = 0. \tag{8}$$

Equation (6) contributes to reducing the loss of CNN training in a_0 compared to using equation (11) in Ref. 12.

In Eq. (6), we consider only the x polarized amplitude. The difference between the diffraction amplitudes of x and y polarizations is very small, and the polarization change by EUV mask diffraction is negligible as shown in Ref. 12.

The values of mask 3D parameters are determined by the mask pattern and the absorber. In this work, we use Ta absorber with 50-nm thickness. We construct a CNN model, which predicts the mask 3D parameters from the mask pattern. Figure 5 shows the architecture of our CNN. Six independent CNNs are used for the real part and the imaginary part of three mask 3D parameters. Six CNNs are merged into one model after the training. The input is a random line and space pattern with a mask area of $1024 \text{ nm} \times 1024 \text{ nm}$. The pattern size is randomly selected from 60 to 160 nm (15 to 40 nm on wafer). Half of the training data are bright field (BF) masks, and the rest are dark field (DF) masks. 1024×1024 binary data are averaged to 256×256 float data before inputting to the CNNs. Circular padding¹⁶ is used because we assume periodic boundary conditions for input mask patterns.

Figure 6 shows the loss functions of training and validation data for $\text{Real}(a_0)$ with/without data augmentation. The number of the original data for training is 2500, and the number of the data for validation is 1000. With data augmentation, the original data are shifted by 103 nm

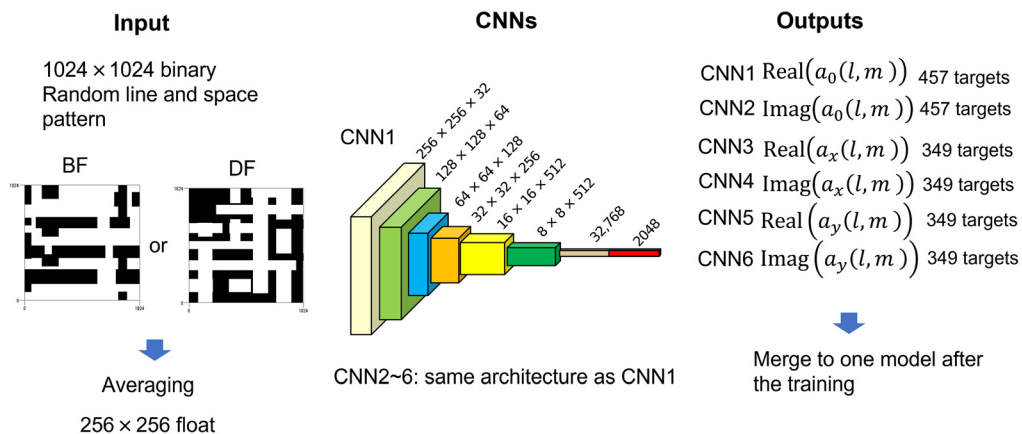


Fig. 5 Architecture of our CNN.

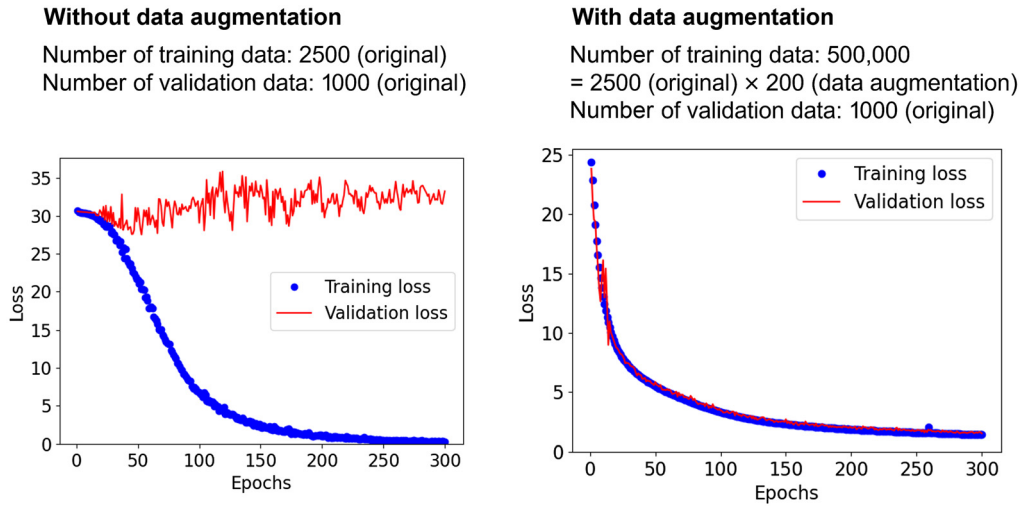


Fig. 6 Training and validation loss for $\text{Real}(a_0)$ with/without data augmentation.

increments in both X and Y directions and flipped along the Y axis. Therefore, the number of the training data after the data augmentation is multiplied by 200 to 500,000.

Without data augmentation, the training loss decreases during the training, whereas the validation loss does not. This is a typical overfitting phenomenon. With data augmentation, both the training loss and the validation loss decrease during the training. The validation loss after the training is small. The mean square error per target after the training is $1.52/457 = 0.0033$. For every 457 targets, the maximum value of the input data is normalized to 1 in the training.

Figure 7 compares the mask 3D parameters at several diffraction orders for 100 test data. The correlation between the parameters by EM simulation and CNN predictions is generally good. There are some exceptions such as $\text{Real}(a_x(0,0))$ where the correlation is poor. However, the value is very small compared to the values of other parameters.

3 CNN Prediction Accuracy

The accuracy of our CNN is verified by calculating the image intensities of test mask patterns. Training data in this work are random line and space patterns. Standard line and space test mask

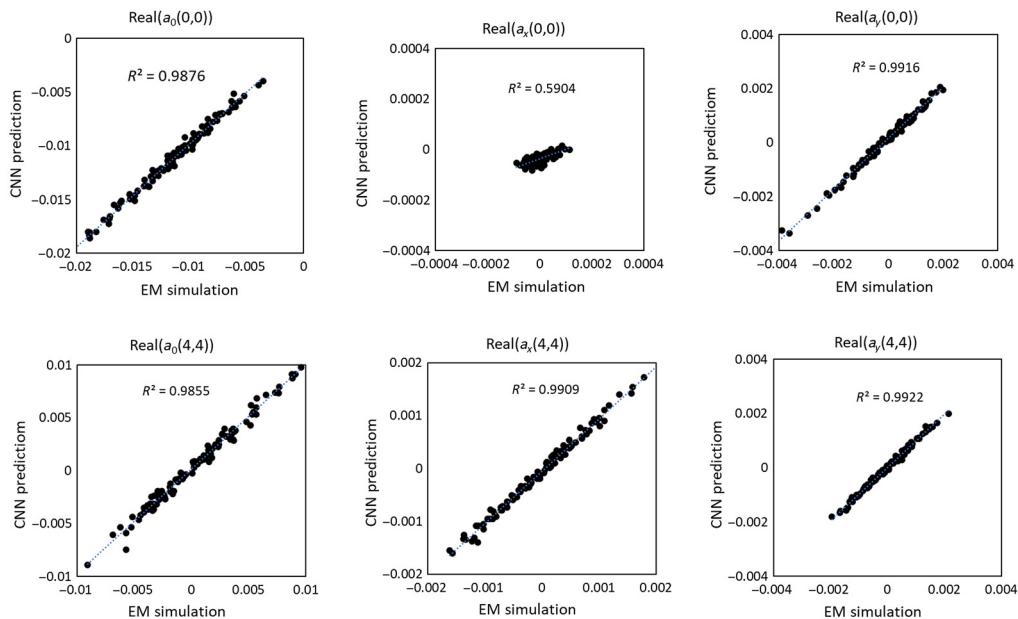


Fig. 7 Mask 3D parameters calculated by EM simulation and parameters predicted by CNN.

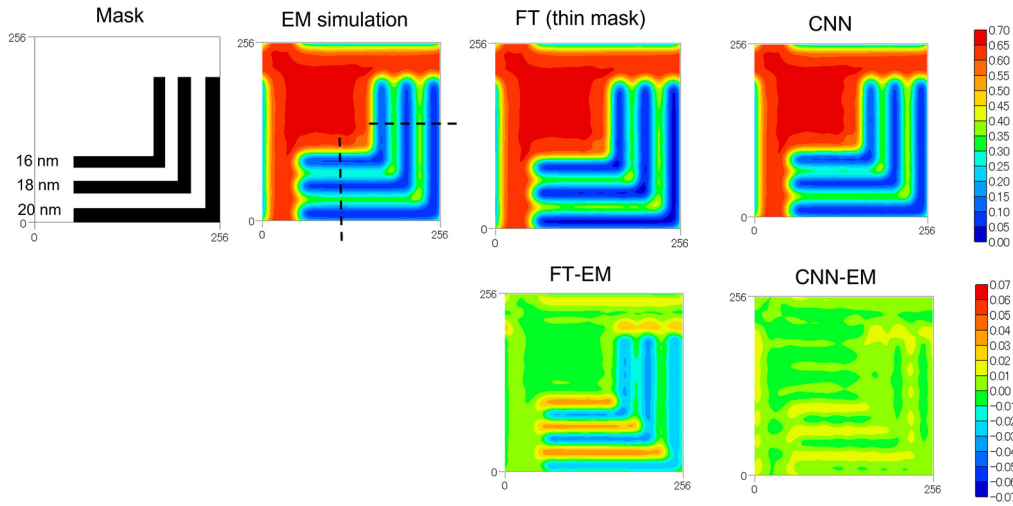


Fig. 8 Line mask pattern and its image intensities by EM simulation, FT, and CNN prediction.

patterns are used to confirm the accuracy of CNN. Figure 8 compares the image intensities of a line mask pattern by EM simulation, Fourier transformation (FT), and CNN prediction. In the calculations, we assume $\lambda = 13.5$ nm, $NA = 0.33$, and annular illumination with $\sigma_{in}/\sigma_{out} = 0.3/0.8$. The bottom figures show the difference of the intensities between EM simulation and FT or CNN prediction. The difference between EM and CNN is much smaller than the difference between FT and EM.

Figure 9 compares the CDs and the EPEs of vertical (V) lines with several line widths. They are measured at the cut line across the V lines in Fig. 8. In addition to EM, CNN, and FT, we plot the result of the simulation using the linear (LIN) approximation of the diffraction amplitude in Eq. (6) (LIN in Fig. 9). The difference between EM and LIN indicates the accuracy of the linear approximation in Eq. (6). Also, the difference between LIN and CNN indicates the accuracy of CNN prediction. The agreement among EM simulation, the linear approximation, and CNN prediction is good. Figure 10 shows the results for horizontal (H) lines. The agreement between LIN and CNN is good but we see small difference between EM and LIN. Adding higher-order terms to Eq. (6) may help reduce this error.

Figure 11 compares the image intensities of a space mask pattern, and Figs. 12 and 13 show the CDs and EPEs of V and H spaces with several space widths. Similar results can be seen with space patterns as with line patterns.

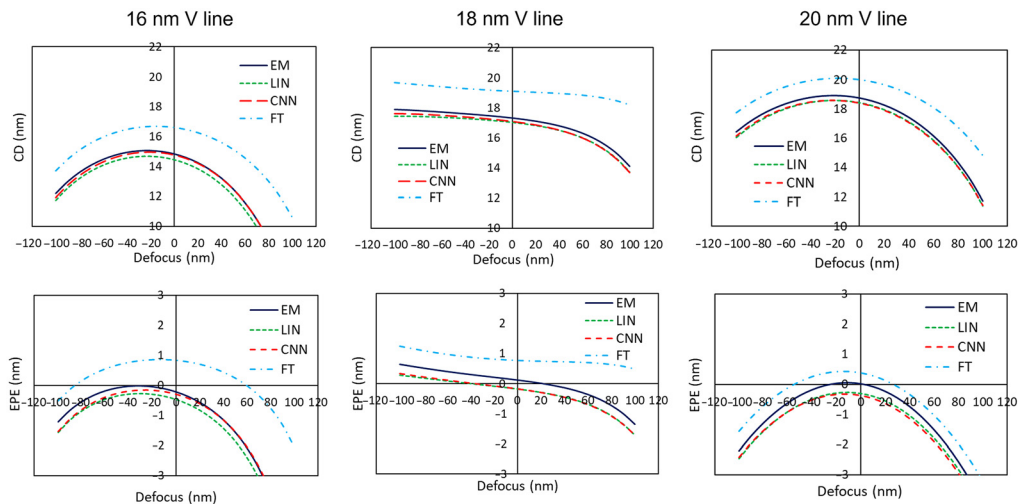


Fig. 9 CDs and EPEs of vertical lines.

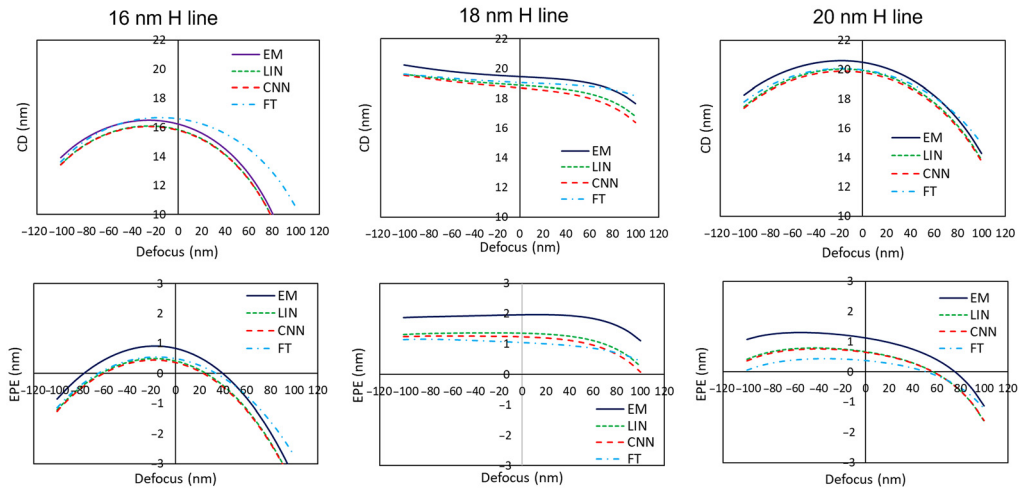


Fig. 10 CDs and EPEs of horizontal lines.

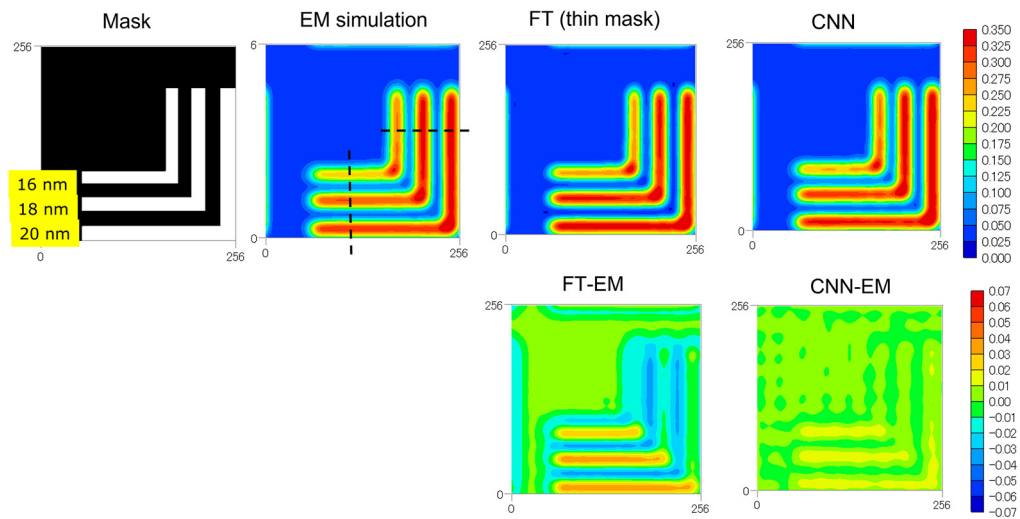


Fig. 11 Space mask pattern and its image intensities by EM simulation, FT, and CNN prediction.

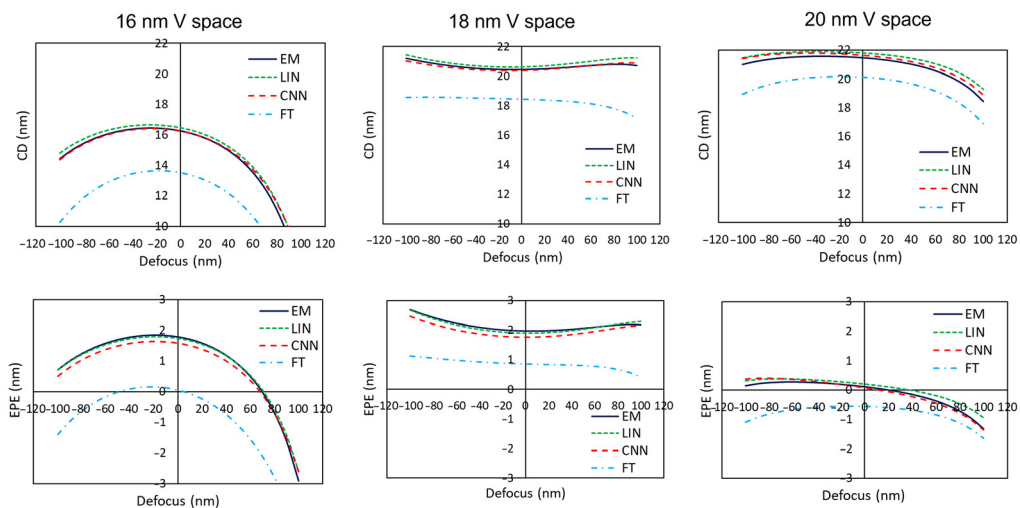


Fig. 12 CDs and EPEs of vertical spaces.

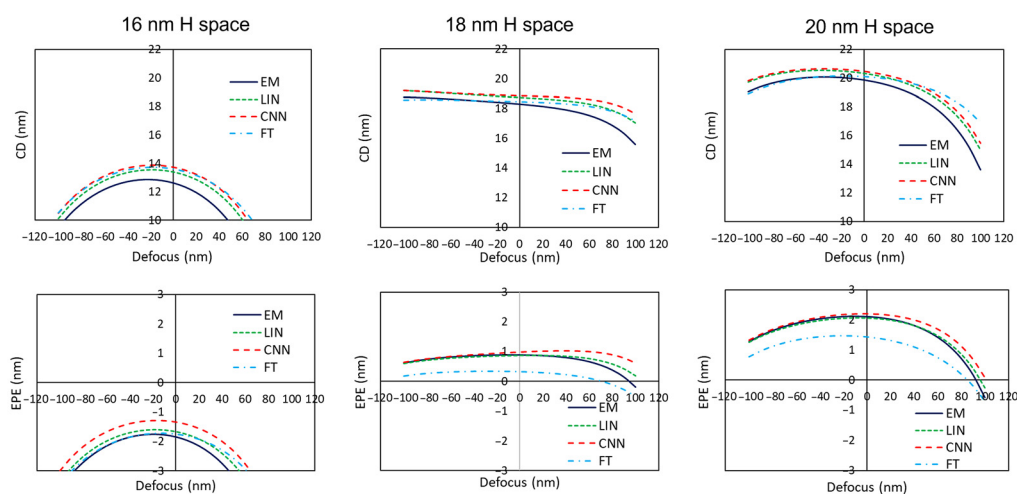


Fig. 13 CDs and EPEs of horizontal spaces.

4 Summary

Data augmentation technique was applied to the diffraction amplitude of the mask pattern. Diffraction amplitudes of shifted or Y -flipped mask patterns were calculated using the diffraction amplitude of the original mask pattern. The number of the training data after the data augmentation is multiplied by 200 from 2500 to 500,000. Using a large amount of training data, the validation loss of CNN was significantly reduced compared to the validation loss without augmentation.

We verified the accuracy of our CNN by comparing the results of EM simulation with CNN predictions. Our CNN almost reproduced the CDs and EPEs of line and space patterns.

In this work, the mask patterns are restricted to line and space patterns. We did not include hole patterns, patterns with serifs and assist bars, or curvilinear patterns in the training data. We do not expect our CNN to correctly predict images for these patterns. Neural network is only as good as the data we feed it. It is a challenge to build several CNNs for specific mask patterns or ultimately a single CNN for arbitrary mask patterns.

This work is based on the prior SPIE proceedings paper.¹⁷

References

1. V. Philipsen, "Mask is key to unlock full EUV potential," *Proc. SPIE* **11609**, 1160904 (2021).
2. A. Erdmann et al., "3D mask effects in high NA EUV imaging," *Proc. SPIE* **10957**, 109570Z (2019).
3. A. Wong, "TEMPEST users' guide," UCB/ERL M94/64 (1994).
4. M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction," *J. Opt. Soc. Am.* **71**(7), 811 (1981).
5. H. Tanabe, "Modeling of optical images in resist by vector potentials," *Proc. SPIE* **1674**, 637 (1992).
6. K. D. Lucas, H. Tanabe, and A. J. Strojwas, "Efficient and rigorous three-dimensional model for optical lithography simulation," *J. Opt. Soc. Am. A* **13**, 2187 (1996).
7. P. Evanschitzky and A. Erdmann, "Fast near field simulation of optical and EUV masks using the waveguide method," *Proc. SPIE* **6533**, 65330Y (2007).
8. S. Lan et al., "Deep learning assisted fast mask optimization," *Proc. SPIE* **10587**, 105870H (2018).
9. P. Liu, "Mask synthesis using machine learning software and hardware platforms," *Proc. SPIE* **11327**, 1132707 (2020).
10. R. Pearman et al., "Fast all-angle mask 3D ILT patterning," *Proc. SPIE* **11327**, 113270F (2020).

11. J. Lin et al., "Fast mask near-field calculation using fully convolution network," in *Int. Workshop Adv. Pattern. Solutions* (2020).
12. H. Tanabe, S. Sato, and A. Takahashi, "Fast EUV lithography simulation using convolutional neural network," *J. Micro/Nanopattern. Mater. Metrol.* **20**(4), 041202 (2021).
13. W. Ye et al., "TEMPO: fast mask topography effect modeling with deep learning," in *Int. Symp. Phys. Design*, Vol. 20, pp. 127–134 (2020).
14. A. Awad et al., "Accurate prediction of EUV lithographic images and 3D mask effects using generative networks," *J. Micro/Nanopattern. Mater. Metrol.* **20**(4), 043201 (2021).
15. A. Wong, *Resolution Enhancement Techniques in Optical Lithography*, SPIE Press, Bellingham, WA (2001).
16. S. Schubert et al., "Circular convolutional neural networks for panoramic images and laser data," in *Proc. IEEE Intell. Veh. Symp. (IV)* (2019).
17. H. Tanabe and A. Takahashi, "Data augmentation in EUV lithography simulation based on convolutional neural network," *Proc. SPIE* **12052**, 120520T (2022).

Hiroyoshi Tanabe is a researcher at Tokyo Institute of Technology. He received his PhD in physics from the University of Tokyo in 1986. He has more than 30 years of experience in optical and EUV lithography. He is the author of more than 30 papers. He was the program committee chair of Photomask Japan in 2003 and 2004. His current research interests include EUV masks and lithography simulation. He is a member of SPIE.

Atsushi Takahashi received his BE, ME, and DE degrees in electrical and electronic engineering from Tokyo Institute of Technology, Tokyo, Japan, in 1989, 1991, and 1996, respectively. He is currently a professor in the Department of Information and Communications Engineering, School of Engineering, Tokyo Institute of Technology. His research interests are in VLSI layout design and combinational algorithms. He is a fellow of IEICE, a senior member of IEEE and IPSJ, and a member of ACM.