IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 4 December 2024, accepted 12 December 2024, date of publication 16 December 2024,
date of current version 24 December 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3517875

RESEARCH ARTICLE

Machine Learning-Enhanced Model-Based Optical Proximity Correction by Using Convolutional Neural Network-Based Variable Threshold Method

JINHAO ZHU[1], ZHIWEI REN[1], YING LI[2], XIANHE LIU[1,2], QIANG WU[1,2], YANLI LI[1,2], AND QI WANG[1,2]

[1]School of Micro-Electronics, Fudan University, Shanghai 200433, China
[2]National Integrated Circuit Innovation Center, Shanghai 201203, China

Corresponding authors: Yanli Li (li_yanli@fudan.edu.cn) and Qi Wang (wangqi_fd@fudan.edu.cn)

ABSTRACT As the lithography process continues to become more rigorous in advanced technology nodes, the model-based optical proximity correction (MBOPC), as a core component within computational lithography, necessitates the development of highly precise techniques. In this paper, we propose an approach to enhance MBOPC through the integration of machine learning (ML), utilizing convolutional neural network (CNN)-based variable threshold method. This OPC framework is characterized by retaining the physical lithography model while integrating the mapping capability of the neural network model to rectify errors encountered in MBOPC. We validate the CNN-based MBOPC model's feasibility at implant layers in advanced nodes. The results demonstrate an improvement in the accuracy of threshold value regression compared to conventional variable threshold methods, and confirm the positive impact of ML integration in simulation accuracy across various patterns. The enhanced MBOPC model effectively compensates for lithography differences in both one-dimension (1D) and two-dimension (2D) regions. This research aims to enhance the simulation precision of MBOPC, thereby ultimately contributing to the ongoing advancement of computational lithography technology.

INDEX TERMS Computational lithography, optical proximity correction (OPC), convolutional neural network, variable threshold.

I. INTRODUCTION

A. PROBLEM DESCRIPTION AND CONTEXT

The miniaturization of semiconductor manufacturing is limited by the imaging performance of optical lithography systems [1], [2]. Computational lithography further pushes the boundaries of achievable resolution and enables more precise and efficient semiconductor manufacturing [3], [4], [5]. Optical proximity correction (OPC) is one of the key resolution enhancement technologies (RET) used within computational lithography to compensate for aerial image distortions that occur during the lithography process [6], [7], [8], [9], [10]. This technology can facilitate the enhancement of lithographic process window near the diffraction limit. As the critical dimension (CD) continues to shrink, the diffraction-dominated proximity effects become more pronounced, leading to distortions in the patterns transferred onto the wafer. To address these issues, OPC determines the aerial images and modifies the original mask patterns by comparing the predicted images with the design target. Thus, precise calculation of the aerial image generated by a specific mask design is necessary to achieve accurate OPC.

In practical applications, there are primarily two methods for calculating aerial image using OPC: rule-based OPC (RBOPC) [11], [12], and model-based OPC (MBOPC) [13], [14]. RBOPC primarily obtains predicted aerial image

The associate editor coordinating the review of this manuscript and approving it for publication was Sukhdev Roy.

through referencing a lookup table, with the correction amounts varying based on the geometry of the patterns. RBOPC achieves high accuracy for one-dimensional (1D) patterns, but the complexity significantly increases with two-dimensional (2D) graphics due to their non-periodic nature. In contrast, MBOPC, which encompasses both edge-based and pixel-based OPC techniques, meticulously models the entire lithography process. The optical imaging of spatial images relies on calculating the matrix of transmission cross-coefficients (TCC) of the designed patterns in the spatial domain [15], [16]. Nowadays, MBOPC is the mainstream technology of OPC in the semiconductor industry, as it can effectively enhance the image fidelity and process robustness in full chip layouts, replacing RBOPC at advanced nodes.

For MBOPC, the models often show discrepancies when implemented in actual lithography processes. Actual lithography systems are highly complex, incorporating factors like nonlinear photoacid diffusion in photoresists [17], [18], [19], mask corner rounding [20], mask 3D effects [21], [22]. Therefore, to address the challenges posed by MBOPC modeling, it is necessary to properly introduce additional mathematical terms and tools to describe the lithography process more accurately. If the model's simulation lacks representativeness and applicability, it will result in numerous weak points within the OPC model results. Consequently, a series of compensatory patches and subsequent refinements will become necessary to address these deficiencies and facilitate the practical implementation of MBOPC. This, in turn, leads to an increase in cycle time and cost associated with layout tape-out process.

B. PRIOR STUDIES

In early 0.18 μm technology nodes, a polynomial-based variable threshold method has been proposed as a mathematical approach to improving the applicability and accuracy of MBOPC simulations [23], [24], [25], [26], [27]. This method primarily consists of sampling optical intensity signals along the cutline direction of polygon edges, commonly referred to as 'slices' in lithography. By extracting key optical parameters from these slices, including minimum intensity, maximum intensity, and intensity slope at the edge, a polynomial fit is employed to determine the imaging threshold for the edge, thereby defining the simulated aerial image of the pattern. This fitting process based on partial optical parameters offers a representation of lithography characteristics. However, polynomial-based variable threshold method also has limitations. This approach solely relies on optical signals from slices, using a selected set of optical characteristics along straight lines as parameters. Consequently, for intricate mask patterns and other features that cannot be easily fitted with polynomials functions using partial optical parameters, this variable threshold method may encounter the applicability risks discussed earlier for MBOPC, restricting its applicability to real lithography processes.

Recently, machine learning-based OPC (MLOPC) has gained attention as a novel OPC model, including several traditional machine learning (ML) models [28], [29] and deep learning models [30], [31], [32], [33]. It is designed to further mitigate the runtime challenges of MBOPC for large-scale layouts by eliminating the reliance on repeated iterations of lithographic calculations for aerial image during the mask correction process, while preserving accuracy in layout correction through the utilization of the robust learning capabilities inherent in ML technology. Nevertheless, the current limitation of MLOPC models is that they always need a training dataset generated from other mask generation tools based on lithographic simulator for training [34], and the complete substitution of lithographic models with ML models poses a risk of loss in interpretability [35], [36]. Thus, the performance of MLOPC models can hardly surpass the physical MBOPC output. Hence, considering MBOPC's enduring status in the industry, there exists considerable potential in integrating MBOPC with ML technology, rather than replacing the existing MBOPC framework altogether.

C. RESEARCH HYPOTHESIS, OBJECTIVES, AND OVERVIEW OF RESULTS

In this study, ML technology is utilized to facilitate the rectification of deviations within the MBOPC model, leading to an innovative implementation of an ML-enhanced MBOPC that exhibits enhanced performance capabilities for calculating aerial image. We propose a convolutional neural network (CNN)-based variable threshold method to improve the simulation performance of the MBOPC model. CNN is a pivotal model in ML, known for its excellence in feature extraction and recognition [37].

The selection of CNN models for enhancing MBOPC can be attributed to their suitability from two perspectives. From an engineering application perspective, CNNs offer lightweight structures with fewer parameters, enabling efficient learning and inference speeds, especially valuable in resource-constrained and time-sensitive industrial applications. Their ability to extract and reduce the dimensionality of local features enhances generalization, mitigating the risk of overfitting with limited data. From a physical essence perspective, considering the local nature of optical proximity effects (OPE), CNNs, with their close-range feature extraction capabilities, exhibit compatibility with the underlying physical phenomenon.

The main contributions of this paper are summarized as follows:

- We propose a novel CNN-based variable threshold method can extract mask features around edge, subsequently regressing these optical features from another input of slices to determine the MBOPC imaging threshold.
- This threshold regression process, based on CNN for different patterns and optical conditions, could accurately reflect the characteristics of various lithography

conditions and rectify errors inherent in the MBOPC approach.

- We compare the performance of two methods: the traditional polynomial variable threshold approach based on least squares fitting, and the CNN-based variable threshold approach which considers surrounding pattern features. The results show that the threshold regression accuracy of the CNN method proposed in this paper is superior to the traditional polynomial-based method both in the training set and the test set.
- This work further validates the effectiveness of the CNN variable threshold method in enhancing the simulation accuracy of the OPC model. This ML-enhanced MBOPC, which integrates both data-driven and knowledge-driven modeling, achieves a more accurate and versatile OPC framework.

This research endeavor aims to advance the application of OPC, improve the simulation precision of MBOPC, and eliminate the need for cumbersome compensation patches, consequently optimizing the layout tape-out process.

D. PAPER ORGANIZATION

The remainder of this paper is organized as follows: In Section II, data sets, polynomial-based variable threshold method, and CNN-based variable threshold method are introduced briefly. In Section III, we present experiments for evaluating the performance of the proposed methods and analyze the enhanced effect of the CNN model on OPC model simulation. Finally, this paper is summarized in Section IV.

II. PRINCIPLE

A. DATA SETS, OPTICAL PARAMETERS, AND THRESHOLD VALUE

In this paper, the MBOPC model consists of an optical model based on the principles of partially coherent imaging, as well as a resist model that incorporates Gaussian convolution and other higher-order terms. The dataset used in our study is based on actual measurements of KrF lithography performed with conventional partial coherence illumination, where partial coherence factor σ_{out} and σ_{in} are 0.8 and 0, respectively. This dataset in our research utilizes actual mask data and silicon wafer measurement data. Consequently, the challenges addressed in our paper are rooted in typical issues encountered in actual lithography processes. This ensures that our findings are relevant and applicable to real industrial scenarios. Additionally, as 248 nm KrF lithography with more than 100 nm pitch could neglect the influence of mask 3D effect and simply analyze the aerial imaging simulation, we evaluate the performance of CNN-based variable threshold method based on KrF lithography. In addition, KrF lithography is still used in advanced nodes for implant layers [38], [39], and precise OPC is still required in these implant layers.

To extract the optical parameters and establish a threshold value, we perform sampling of intensity along the direction perpendicular to the pattern edge, as shown in Fig. 1. This procedure generates intensity profiles of slice curves, which are expressed as a function of 1D location. From these slice curves, we identify the maximum intensity and minimum intensity values. Additionally, slopes are defined as the absolute values of the gradient of intensity at the mask CD positions, and threshold values are determined based on the intensity at the wafer measured line CD corresponding to the slices. The threshold is a normalized value that reflects the critical exposure intensity required for imaging process in MBOPC model. Therefore, each specific pattern corresponds to the three optical parameters (I_{max}, I_{min}, slope) mentioned above and one practical threshold.

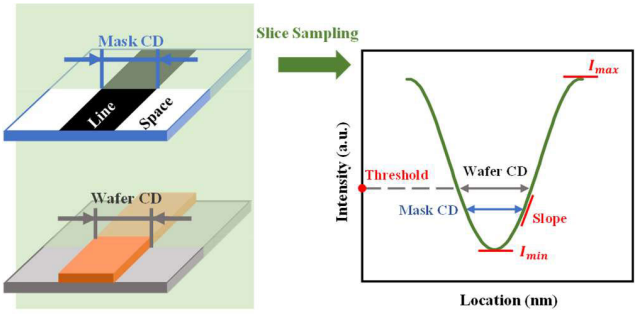

FIGURE 1. Schematic diagram of optical parameters and threshold.

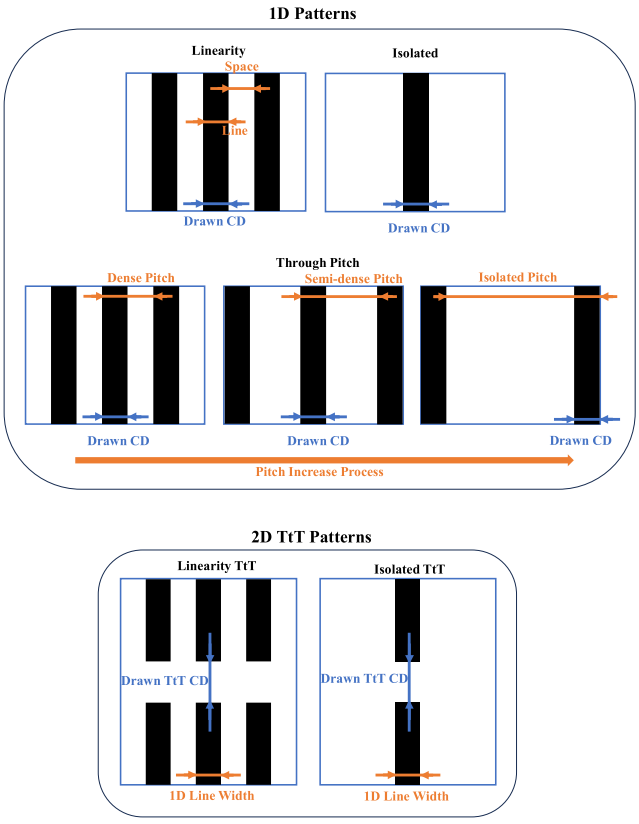

FIGURE 2. Schematic diagram of five types of patterns in dataset.

The mask patterns in dataset include 1D patterns and 2D tip-to-tip (TtT) patterns as shown in Fig. 2, specifically classified into five categories: linearity pattern, isolated pattern,

through pitch pattern, linearity TtT pattern, and isolated TtT pattern. Herein, the term 'linearity' refers to a type of patterns where the line width is equal to space width, 'isolated' represents a category of patterns with only one isolated line, and 'through pitch' reflects a series of patterns with a specified line width from relatively dense pitch to relatively isolated pitch. In all patterns, the term 'drawn CD' refers to the mask design, which constitutes the layout feature that OPC models are required to simulate. In 1D patterns, the pattern line width is labeled as the drawn CD, while in 2D TtT patterns, the pattern TtT space width is labeled as the drawn CD. This means that for 1D patterns, we sample the intensity along the slice of their opaque line area, shown in Fig. 1. In contrast, for 2D TtT patterns, we determine the optical parameters and threshold value by sampling the intensity along the slice of the TtT region, which represents the non-blocked space area.

The data are allocated to training sets and test sets with a ratio of approximately 4:1 for each category. Given that our OPC simulation operates on an edge-based framework, we complete the simulation of drawn CD through the implementation of two separate edge threshold regressions. Finally, the dataset consists of 460 training samples and 120 test samples, each sample features a pattern centered on a specific edge. Every individual pattern exhibits a dimension of 200 pixels × 200 pixels, corresponding to spatial coordinates of 2 μm × 2 μm.

B. VARIABLE THRESHOLD METHOD BASED ON POLYNOMIAL

For the illustration presented in Fig. 1, three optical features can be extracted and employed as input parameters for a dataset. In the variable threshold method that relies on polynomial fitting, these three optical parameters serve as the basis for determining the threshold. An n-th order polynomial with three variables can be expanded in the following equation:

$$y = k_{00} + k_{11}x_1 + k_{12}x_2 + k_{13}x_3 + k_{21}x_1^2 + k_{22}x_1x_2 + \ldots + k_{n,c}x_3^n \quad (1)$$

where x_i is an optical parameter in this work, and $c = C_{n+2}^2$. This symbol c represents the count of distinct ways to distribute the degree n among the three variables. The undetermined coefficients k in (1), which are used in the variable threshold method based on polynomial fitting, can be determined using the method of least squares [40].

Considering an overdetermined matrix equation $A\boldsymbol{x} = \boldsymbol{b}$, where $\boldsymbol{b} \in \mathrm{R}^{m\times 1}$, $A \in \mathrm{R}^{m\times n}$, and $m > n$. Using the method of least squares, we can find an estimate $\hat{\boldsymbol{x}}$ as:

$$\left|\left|\boldsymbol{b} - A\hat{\boldsymbol{x}}\right|\right| \leq ||\boldsymbol{b} - A\boldsymbol{x}|| \quad (2)$$

for all $\boldsymbol{x} \in \mathrm{R}^{n\times 1}$.

To solve the ordinary least squares problem, follow these steps: Given A and $\boldsymbol{b}$, we define $\hat{\boldsymbol{b}}$ as the projection of $\boldsymbol{b}$ onto the column space of A, expressed mathematically as:

$$\hat{\boldsymbol{b}} = \mathrm{Proj}_{\mathrm{col}A}\boldsymbol{b} \quad (3)$$

Since $\hat{\boldsymbol{b}}$ belongs to the column space of A, the equation $A\boldsymbol{x} = \hat{\boldsymbol{b}}$ is compatible, and there exists a vector $\hat{\boldsymbol{x}} \in \mathrm{R}^{n\times 1}$ such that:

$$A\hat{\boldsymbol{x}} = \hat{\boldsymbol{b}} \quad (4)$$

As $\hat{\boldsymbol{b}}$ is the point in the column space of A that is closest to $\boldsymbol{b}$, the $\hat{\boldsymbol{x}}$ that satisfies this above equation is therefore a least squares solution to the equation $A\boldsymbol{x} = \boldsymbol{b}$.

Based on the orthogonal decomposition theorem, $(\boldsymbol{b} - \hat{\boldsymbol{b}})$ is orthogonal to Col A, which implies that $(\boldsymbol{b} - A\hat{\boldsymbol{x}})$ is orthogonal to each column of A. If $\boldsymbol{a}_j$ represents any column of A, then $\boldsymbol{a}_j \cdot (\boldsymbol{b} - A\hat{\boldsymbol{x}}) = 0$ and $\boldsymbol{a}_j^{\mathrm{T}} \cdot (\boldsymbol{b} - A\hat{\boldsymbol{x}}) = 0$. Since each $\boldsymbol{a}_j^{\mathrm{T}}$ corresponds to a row of A^{T}, we have:

$$A^{\mathrm{T}}(\boldsymbol{b} - A\hat{\boldsymbol{x}}) = 0 \quad (5)$$

Therefore, the solution to (2) is:

$$\hat{\boldsymbol{x}} = \left(A^{\mathrm{T}}A\right)^{-1}A^{\mathrm{T}}\boldsymbol{b} \quad (6)$$

C. VARIABLE THRESHOLD METHOD BASED ON CNN

Fig. 3 depicts the schematic representation of the CNN-based variable threshold method. The architecture of this network comprises two types of inputs: the pattern image and optical parameters. The pattern image undergoes a series of interleaved convolutional and pooling layers for feature extraction. Subsequently, these extracted features are concatenated with the features derived from the optical parameters within subsequent fully connected layers. This fusion process facilitates a joint regression to determine a threshold value.

Specifically, the input mask image possesses a resolution of 200 × 200 pixels. Employing 3 × 3 convolutional kernels, we generate an initial set of four (200 × 200) feature maps within the first convolutional layer. Subsequently, these maps undergo downsampling through a 4 × 4 max-pooling layer, resulting in four (50 × 50) feature maps. This sequence of convolutional (using 3 × 3 kernels) and max-pooling (using 2 × 2 kernels) operations is repeated four additional times. The resultant mask features are then flattened and forwarded to a fully connected layer comprising 100 nodes. Concurrently, the optical parameters input is directly input into a fully connected layer with 40 nodes. In subsequent fully connected layers, the extracted mask features and optical parameters are concatenated, enabling the network to jointly regress towards a final threshold value.

In both convolutional and fully connected layers, a commonly employed activation function known as the rectified linear unit (ReLU) is used. For network training, the Adam Optimizer, which has garnered widespread adoption in the realm of deep learning, is deployed with an initial learning rate set at 5 × 10^4. The network undergoes training utilizing a batch size of 8, with the training process spanning 200 iterations. When training the CNN, the mean square error (MSE) between the predicted threshold values and the actual threshold values serves as the loss function. MSE is a commonly used loss function in regression problems, measuring the average squared error between predicted values and actual

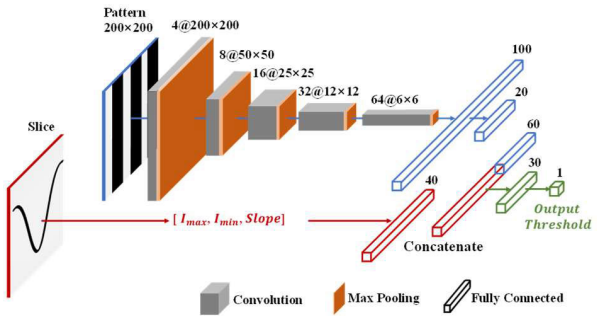

FIGURE 3. Schematic diagram of the CNN-based variable threshold method.

values. Its definition is as follows:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(a_i - p_i)^2 \tag{7}$$

where N represents the number of samples, a_i and p_i are the actual and predicted threshold value respectively.

The PyTorch framework and Python language are used to design, train, and test the custom CNN on a computer. The relevant environment parameters are shown in Table 1.

TABLE 1. Hardware and software environment.

Configuration Item	Value
CPU	Inter(R) Xeon(R) Platinum 8367B
GPU	NVIDIA GeForce RTX 4060
Memory	1 TB
Hard disk	8 TB
Operating system	Windows 10 64bits

III. RESULTS AND DISCUSSION

A. COMPARISON FOR THRESHOLD PREDICTION ACCURACY

Fig. 4(a) compares the accuracy of predicted thresholds between two variable threshold methods based on polynomial and CNN. To verify the accuracy of methods, the indicator of root mean square error (RMSE) is adopted:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(a_i - p_i)^2} \tag{8}$$

where N represents the number of samples, a_i and p_i are the actual and predicted threshold values, respectively. The unit of the RMSE value coincides with that of the source data, facilitating an intuitive assessment of error, thereby enhancing the comprehensibility of the results.

For the variable threshold method of polynomials, it is observed that the overall trend of RMSE in the training set decreases as the degree of fit increases from 1st order to 5th order. This indicates that the ability of polynomial fitting to the normalized threshold value in training set improves with higher degrees. However, it is also noteworthy that the RMSE of the 4th order polynomial method in the training set is 0.00874, higher than that of the 3rd order method. This may be due to different fitting characteristics for odd and even-order polynomial terms. Typically, as the degree of fit increases, polynomials tend to transition from underfitting to overfitting. In our test dataset, this behavior reflects in the initial decrease and subsequent gradual increase of RMSE with the increases in polynomial degree. Therefore, from Fig. 4(a), we observe that in this work, the performance of the variable threshold method based on a 3rd order polynomial is better than methods of other degrees. The RMSE values of 3rd order polynomial in the training set and test set are 0.00564 and 0.00829, respectively.

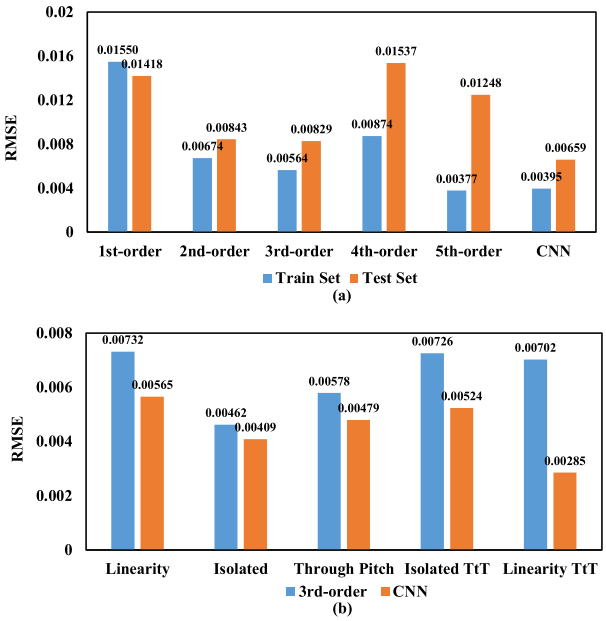

FIGURE 4. (a) The comparison of two variable threshold methods based on 1st order to 5th order polynomial and CNN in the training set and the test set. (b) The comparison of two variable threshold methods based on 3rd order polynomial and CNN in five distinct pattern types.

For the variable threshold method based on CNN, it has an RMSE of 0.00395 on the training set and 0.00659 on the test set. Compared to the 3rd order polynomial method, this represents a reduction in RMSE by 29.9% in the training set and 20.5% in the test set. Therefore, with the same input o optical parameters, the CNN method, which additionally takes into account nearby pattern features, demonstrates superior performance compared to the traditional polynomial variable threshold method, both in the training set and the test set.

Furthermore, we compare the specific performance of the 3rd order polynomial method and the CNN method in various patterns within the dataset. From Fig. 4(b), the CNN method consistently exhibits lower RMSE values than the 3rd method through all graphics categories, including linearity pattern, isolated pattern, through pitch pattern, linearity TtT pattern, and isolated TtT pattern. In summary, these results demonstrate that the CNN variable threshold method possesses a more accurate regression threshold capability.

B. CNN FEATURE MAPS ANALYSIS

In the mentioned discussion, we validated the high threshold regression accuracy of the variable threshold method based on CNN proposed in this paper. However, further attention needs to be paid to the interpretability of the model. Introducing feature map visualization, which enables the visualization of the output features from specific layers within the network, is commonly employed in practical neural network research to aid in enhancing the model's interpretability [41], [42]. In this part, we leveraged the visualization of convolutional layer feature maps to demonstrate how the CNN model processes mask data and extracts mask features.

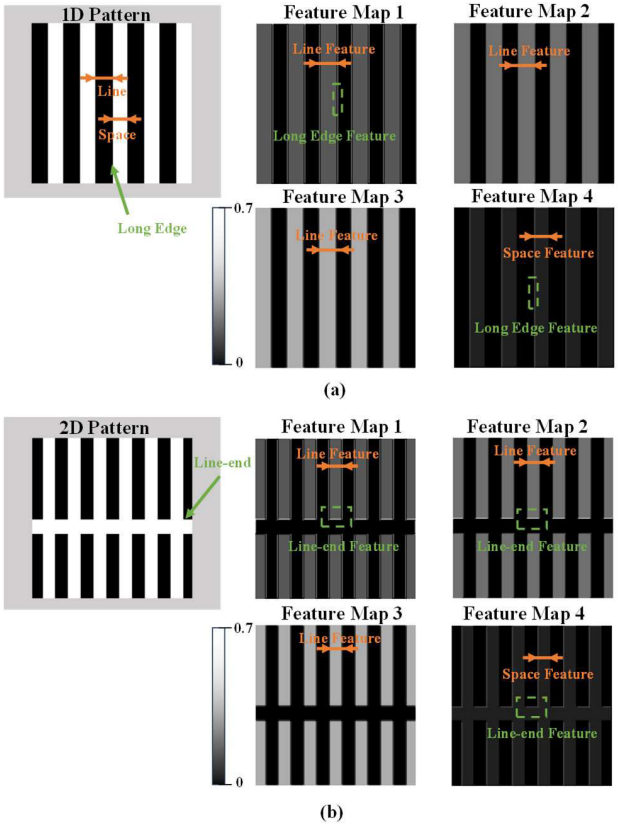

FIGURE 5. The CNN feature maps results corresponding to raw mask data and the abstract features extracted from the first convolutional layer for the typical (a) 1D patterns, and (b) 2D patterns.

Fig. 5(a) and Fig. 5(b) show the feature maps results corresponding to raw mask data and the abstract features extracted from the first convolutional layer for the typical 1D and 2D patterns, respectively. After undergoing normalization through the first convolutional layer and the ReLU activation function, the resulting feature maps are then converted into grayscale images, with the maximum intensity value mapped to 0.7 for improved contrast visualization. Within these feature maps, maps 1-3 primarily capture features associated with lines, demonstrating varying intensities in their depiction of line, space, and edge features. Meanwhile, map 4, which specifically focuses on space features, exhibits a relatively weaker intensity compared to the other maps. This observation suggests a tendency of the CNN model to emphasize line features, namely the spatial relationships within the patterns of the mask. Furthermore, some feature maps exhibit enhanced extraction of long edges and line-ends, which exhibit higher intensities in some maps and will significantly influence OPE. This visualization analysis, based on specific mask data, indicates that the CNN model in this study effectively captures crucial lithographic features in the mask, which will contribute to its excellent regression performance potentially.

C. CNN VARIABLE THRESHOLD MODEL ENHANCES OPC IN 1D PATTERNS

Next, we further examine the enhancement impact of CNN variable threshold methods on the original constant threshold MBOPC model across different patterns. Through the intensity simulation of the MBOPC model and the determination of thresholds by the CNN model, the finer simulation results of the drawn CD of specific patterns can be obtained. The verification of CD accuracy and error trends for simulating images of the different patterns is facilitated by the simulation of both the original OPC model and the enhanced OPC+CNN model.

In typical OPC simulation results figures, model- predicted results are shown as lines, while actual measurement data are represented by data points. However, in this section, we reverse the typical figures plotting to discuss the CD accuracy of the results in two models, where the measurement data are represented by lines, while simulating results of the two models are represented by data points. This reversal is intended to highlight the deviations between the simulated results and the actual measurements, enabling a more direct comparison of CD accuracy. Moreover, the simulation results of the OPC model are consistent with the actual sampling step of the measurement data in this work.

Fig. 6 shows the simulation results of drawn CD in linearity patterns and isolated patterns with original OPC model and the CNN variable threshold OPC model. This figure encompasses three components: schematic diagrams of patterns, comparative analysis of the simulation results of two models, and quantitative assessment of the absolute error incurred by each model in relation to the measurements. By comparing the CDs of the two models with actual wafer measurement results, the simulation accuracy of the model can be analyzed. Fig. 6(a) and Fig. 6(d) are the diagrams of linearity pattern and isolated pattern respectively. As shown in Fig. 6(b) and Fig. 6(e), the original OPC model already provides generally accurate wafer simulation results for mask drawn CD in both pattern types, matching actual wafer measurement results. The CNN variable threshold OPC model maintains the precision relative to the original model for these patterns, and demonstrates slight improvements in specific graphics, such as large drawn CD conditions in the isolated pattern in Fig. 6(f).

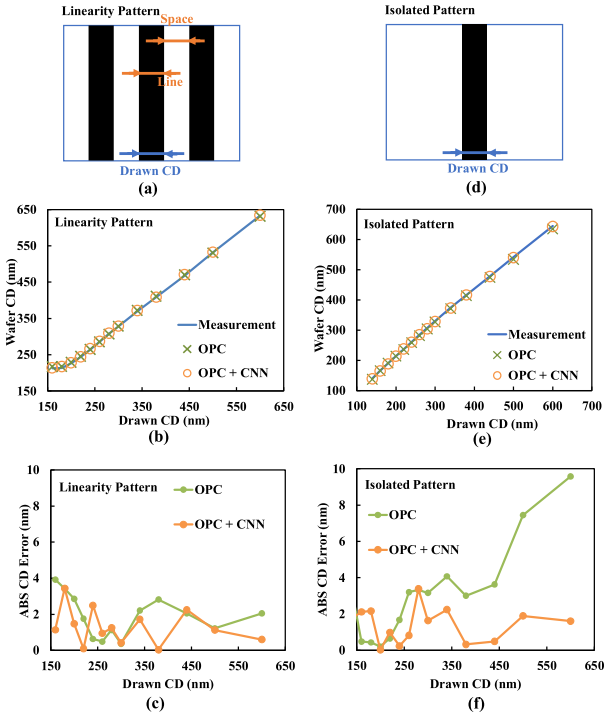

FIGURE 6. The simulation results of (a-c) linearity patterns and (d-f) isolated patterns with original OPC model and the CNN variable threshold OPC model.

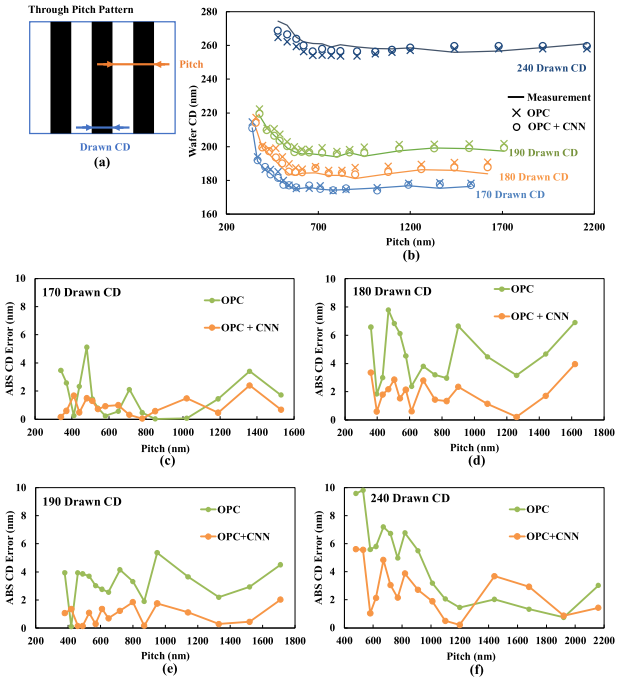

FIGURE 7. (a) The schematic diagram, (b) simulation results of the original OPC model and the CNN variable threshold OPC model for drawn CD in through pitch pattern. (c-f) The quantitative assessment of the absolute error incurred by each model in relation to the measurements with mask drawn CD from 170 nm to 240 nm.

Fig. 7 illustrates the simulation results of mask drawn CDs at 170 nm, 180 nm, 190 nm, and 240 nm with increasing pitch for both the original OPC model and CNN variable threshold OPC model. The results of through pitch patterns often receive significant attention, providing intuitive insight into optical proximity effects within a lithography system. Notably, as pitch increases in Fig. 7(b), the measurement of wafer CD initially decreases rapidly in small pitch range, but then stabilizes in large pitch range. This behavior in small pitch range arises from the imaging process near the diffraction limit, where only a few orders of diffraction light are collected by the pupil contributing to the image formation. Increasing the pitch allows additional diffraction orders to be included, significantly improving imaging quality. Under large pitch conditions, sufficient diffraction orders are already included, and higher diffraction orders have less impact on the current imaging, resulting in stable wafer CD values. Consequently, as pitch continues to increase, only minor nanometer scale oscillations are observed.

For the original OPC model, within the small pitch range, the simulation results for the 170 nm drawn CD pattern closely align with the measured results. However, for the 180 nm and 190 nm drawn CD patterns, the simulation results slightly exceed the measured values, while the 240 nm drawn CD pattern results are underestimated slightly. In contrast, after enhancing the model with CNN-based variable thresholds, we observe an improvement in reducing this deviation in Fig. 7(c-f). Thus, the enhanced simulation results exhibit a closer correspondence with the measured line depicted in Fig. 7(b). Additionally, under larger pitch conditions, both the original OPC model and the CNN-enhanced OPC model tend to stabilize. This result indicates that these models still effectively capture optical characteristics and avoid overfitting to measurement errors.

As depicted in Fig. 6 and Fig. 7, the original OPC model performs well in these 1D patterns, with an average absolute CD error of less than 4 nm and a maximum absolute CD error of less than 10 nm. In general, the original model demonstrates satisfactory accuracy. In contrast, the ML-enhanced MBOPC model (OPC+CNN) achieves an even higher level of accuracy, with an average absolute CD error of less than 2 nm and a maximum absolute CD error of less than 6 nm.

D. CNN VARIABLE THRESHOLD MODEL ENHANCES OPC IN 2D PATTERNS

In the preceding discourse, our primary focus has been on the simulation outcomes pertaining to two models in the context of 1D patterns. Subsequently, we will further discuss their performances on 2D patterns.

Fig. 8 and Fig. 9 show the simulation results of linearity TtT pattern and isolated TtT pattern with drawn TtT CD of 200 nm and 250 nm. Both TtT patterns pose significant challenges for original OPC models due to 2D line-end shortening effect. At this endpoint, the simulation conditions encompass both bright field and 2D regions, where the photoresist diffusion behavior differs significantly from that of polygons with longer edge. Fig. 8(b) and Fig. 9(b) illustrate that the simulated TtT CDs at small linewidths are significantly smaller than the actual measured results for original

OPC model. The maximum errors for linearity and isolated TtT are approximately 20 nm and 50 nm, respectively. For these weak points, it is usually necessary to patch the OPC recipe, which leads to the extension of the layout tape-out cycle and may increase the tape out times. As the 1D line width increases, simulation conditions of line-end graphics gradually approach those of long edge of 1D graphics. Thus, the TtT simulation results of original OPC model tend to be closer to the actual measurements.

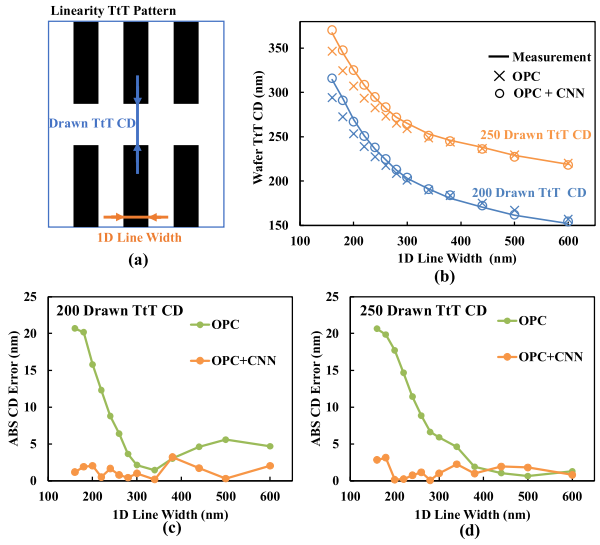

FIGURE 8. The simulation results of the original OPC model and the CNN variable threshold OPC model for drawn CD in linearity TtT patterns.

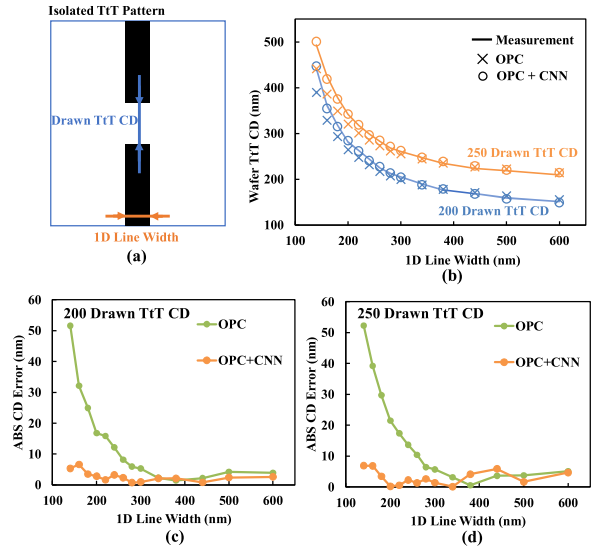

FIGURE 9. The simulation results of the original OPC model and the CNN variable threshold OPC model for drawn CD in linearity isolated TtT patterns.

On contrast, CNN variable threshold OPC model significantly improves the simulation accuracy for both types of TtT patterns through the entire 1D line width. Especially in the narrow line width region, after enhancing the OPC model using CNN-based variable threshold method, the max TtT CD error in the linearity TtT pattern is reduced from 21 nm to 2 nm, and the max TtT CD error in the isolated TtT pattern is reduced from 52 nm to 7 nm. This means that the MBOPC model enhanced by CNN significantly reduces errors on these weak points, enabling the simulation performance of 2D patterns to achieve the accuracy of 1D patterns. Thus, the calculation accuracy of the enhanced CNN-based variable threshold OPC can meet the demands of industrial mass production. These results demonstrate optical simulation capability of ML-enhanced MBOPC under different patterns conditions.

IV. CONCLUSION

This study proposes a novel ML-enhanced MBOPC by CNN-based variable threshold method. This OPC framework retains the inherent physical property of lithography simulation model while integrating the robust mapping capabilities of neural network model to rectify errors encountered in MBOPC. The CNN-based variable threshold MBOPC model finely regresses the threshold value by inputting pattern information and optical parameters, thereby improving the simulation precision of the OPC model. This work demonstrates that CNN variable threshold method outperforms the traditional polynomial variable threshold method in terms of threshold accuracy, achieving a reduction in RMSE by 29.9% in the training set and 20.5% in the test set. Notably, it achieves substantial enhancement in the simulation accuracy of 2D TtT weak points, and modestly improving accuracy for 1D patterns as well. It reveals that the variable threshold model based on CNN can well compensate for the different lithography characteristics for both 1D and 2D regions. Looking ahead, we plan to investigate the performance of ML-enhanced OPC framework in more complex lithography scenarios and consider the applicability of more complex neural network structures in this context. Using the data-knowledge dual-driven OPC modeling approach, we aim to explore its adaptability in various lithography scenarios and actively contribute to the ongoing advancement of computational lithography technology.

V. DATA AVAILABILITY

The data that support the findings of this study are available from the National Integrated Circuit Innovation Center. Restrictions apply to the availability of these data, which were used under license for this study. Data are available from the corresponding authors with the permission of National Integrated Circuit Innovation Center.

REFERENCES

[1] Q. Wu, Y. Li, and X. Liu, "Development and future outlook of modern lithography machines," *Micro/nano Electron. Intell. Manuf.*, vol. 4, no. 1, pp. 30–56, 2022.

[2] Y. Li, X. Liu, and Q. Wu, "Evolution and updates of advanced photolithography technology," *Laser Optoelectron. Prog.*, vol. 59, no. 9, Sep. 2022.

[3] M. Yuan, Y. Y. Sun, and Y. Q. Li, "Advanced computational lithography," (in Chinese), *Laser Optoelectron. Prog.*, vol. 59, no. 9, p. 16, May 2022.

[4] M. Kuramochi, Y. Kohira, H. Tanabe, T. Matsunawa, and C. Kodama, "Development of a lithography simulation tool set in various optical conditions for source mask optimization," *IEEE Access*, vol. 12, pp. 58490–58501, 2024.

[5] J. Liu, Q. Y. Xu, M. Su, N. Fang, and B. F. Wang, "A hybrid eigenmode restoration algorithm for computational lithography problems based on mode matching principle," *IEEE Access*, vol. 10, pp. 45799–45810, 2022.

[6] C. Spence, "Full-chip lithography simulation and design analysis: How OPC is changing IC design," *Proc. SPIE*, vol. 5751, pp. 1–14, May 2005.

[7] X. Ma and G. R. Arce, "Pixel-based simultaneous source and mask optimization for resolution enhancement in optical lithography," *Opt. Exp.*, vol. 17, no. 7, p. 5783, 2009.

[8] J. F. Chen, T. Laidig, K. Wampler, and R. Caldwell, "Practical method for full-chip optical proximity correction," *Proc. SPIE*, vol. 3051, pp. 790–803, 1997, doi: 10.1117/12.276060.

[9] Y. J. Shen, N. Wong, and E. Y. Lam, "Aberration-aware robust mask design with level-set-based inverse lithography," *Proc. SPIE*, vol. 7748, pp. 567–574, May 2010.

[10] Y. J. Shen, "Efficient level-set based mask optimization with a vector imaging model," *Proc. SPIE*, vol. 10587, pp. 345–355, Mar. 2018.

[11] M. E. Kling, K. D. Lucas, A. J. Reich, B. J. Roman, H. Chuang, P. V. Gilbert, W. D. Grobman, E. O. Travis, P. Tsui, T. Vuong, and J. P. West, "0.25μm logic manufacturability using practical 2-D optical proximity correction," *Proc. SPIE*, vol. 3334, pp. 204–214, Jun. 1998.

[12] O. W. Otto, J. G. Garofalo, K. K. Low, C.-M. Yuan, R. C. Henderson, C. Pierrat, R. L. Kostelak, S. Vaidya, and P. K. Vasudev, "Automated optical proximity correction: A rules-based approach," *Proc. SPIE*, vol. 2197, pp. 278–293, May 1994.

[13] C. K. Kim, J. S. Choi, B. H. Nam, and D. Yim, "Considerations of model based OPC verification for sub-70 nm memory device," *Proc. SPIE*, vol. 6156, pp. 438–443, Mar. 2006.

[14] B. Lin, X. L. Yan, Z. Shi, and Y. W. Yang, "A sparse matrix model-based optical proximity correction algorithm with model-based mapping between segments and control sites," *J. Zhejiang Univ.-Sci. C-Comput. Electron.*, vol. 12, no. 5, pp. 436–442, 2011.

[15] H. H. Hopkins, "On the diffraction theory of optical images," *Proc. Royal Soc. London. A, Math. Phys. Sci.*, vol. 217, no. 1130, pp. 408–432, 1953.

[16] N. Cobb, A. Zakhor, and E. Miloslavsky, "Mathematical and CAD framework for proximity correction," *Proc. SPIE*, vol. 2726, pp. 208–222, Jun. 1996, doi: 10.1117/12.240907.

[17] Q. Wu, "An optical proximity model for negative toned developing photoresists," in *Proc. China Semiconductor Technol. Int. Conf. (CSTIC)*, Shanghai, China, Mar. 2018, pp. 1–4.

[18] T. Kozawa, S. Tagawa, J. J. Santillan, and T. Itani, "Impact of nonconstant diffusion coefficient on latent image quality in 22 nm fabrication using extreme ultraviolet lithography," *J. Photopolymer Sci. Technol.*, vol. 21, no. 3, pp. 421–427, 2008.

[19] F. Tsai, T.-H. Yeh, C. C. Yang, E. Yang, T. H. Yang, and K. C. Chen, "CD uniformity improvement of dense contact array in negative tone development process," *Proc. SPIE*, vol. 9427, pp. 807–816, Mar. 2015.

[20] H. B. Kim, W. K. Ma, C. N. Ahn, and K. S. Shin, "Optical lithography simulation considering impact of mask errors," *Proc. SPIE*, vol. 2002, pp. 1278–1286, Jul. 2002.

[21] A. Erdmann, P. Evanschitzky, J. T. Neumann, and P. Gräupner, "Mask-induced best-focus-shifts in DUV and EUV lithography," *Proc. SPIE*, vol. 9426, pp. 112–122, Mar. 2015.

[22] J. Finders and T. Hollink, "Mask 3D effects: Impact on imaging and placement," *Proc. SPIE*, vol. 7985, pp. 156–178, 2011.

[23] H. M. Liao, S. Palmer, and K. Sadra, "Variable threshold optical proximity correction (OPC) models for high performance 0.18 μm process," *Proc. SPIE*, vol. 4000, pp. 1033–1040, 2000.

[24] N. Cobb and A. Zakhor, "Experimental results on optical proximity correction with variable threshold resist model," *Proc. SPIE*, vol. 3051, pp. 458–468, Jul. 1997.

[25] C. H. Park, S. U. Rhie, J. H. Choi, J. S. Park, H. W. Seo, Y. H. Kim, Y. K. Park, W. S. Han, W. S. Lee, and W. S. Kong, "Optical proximity correction considering mask manufacturability and its application to 0.25μm DRAM for enhanced device performance," *Proc. SPIE*, vol. 2000, pp. 1041–1046, Jul. 2000.

[26] T. H. Coskun, "Uncertainty aware site selection method for OPC model calibration," *Proc. SPIE*, vol. 9426, pp. 217–227, Mar. 2015.

[27] B. S. Ward, "Improving hyper-NA OPC using targeted measurements for model parameter extraction," *Proc. SPIE*, vol. 6730, pp. 898–908, Oct. 2007.

[28] X. Ma, S. Jiang, J. Wang, B. Wu, Z. Song, and Y. Li, "A fast and manufacture-friendly optical proximity correction based on machine learning," *Microelectronic Eng.*, vol. 168, pp. 15–26, Jan. 2017.

[29] J. Cho, G. Cho, and Y. Shin, "Optimization of machine learning guided optical proximity correction," *Proc. SPIE*, vol. 2018, pp. 921–924, Aug. 2018.

[30] S. Choi, S. Shim, and Y. Shin, "Machine learning (ML)-guided OPC using basis functions of polar Fourier transform," *Proc. SPIE*, vol. 9780, pp. 63–70, Mar. 2016.

[31] S. Choi, S. Shim, and Y. Shin, "Neural network classifier-based OPC with imbalanced training data," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 38, no. 5, pp. 938–948, May 2019.

[32] Y. Kwon and Y. Shin, "Optical proximity correction using bidirectional recurrent neural network with attention mechanism," *IEEE Trans. Semiconductor Manuf.*, vol. 34, no. 2, pp. 168–176, May 2021.

[33] P. Yuan, P. Xu, L. Ma, and Y. Wei, "Optical proximity correction by using unsupervised learning and the patch loss function," *Appl. Opt.*, vol. 61, no. 14, p. 3924, 2022.

[34] A. T. Ngo, B. Dey, S. Halder, S. De Gendt, and C. Wang, "Machine learning-based edge placement error analysis and optimization: A systematic review," *IEEE Trans. Semiconductor Manuf.*, vol. 36, no. 1, pp. 1–13, Feb. 2023.

[35] D. V. Carvalho, E. M. Pereira, and J. S. Cardoso, "Machine learning interpretability: A survey on methods and metrics," *Electronics*, vol. 8, no. 8, p. 832, Jul. 2019.

[36] Q.-S. Zhang and S.-C. Zhu, "Visual interpretability for deep learning: A survey," *Frontiers Inf. Technol. Electron. Eng.*, vol. 19, no. 1, pp. 27–39, Jan. 2018.

[37] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A survey of convolutional neural networks: Analysis, applications, and prospects," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022.

[38] T. You, T. Lee, G. Yoo, Y. Park, C. Kim, and D. Yima, "Simple method for decreasing wafer topography effect for implant mask," *Proc. SPIE*, vol. 9780, pp. 177–184, Mar. 2016.

[39] S. W. Kim, P. Zheng, K. Kato, L. Rubin, and T. J. K. Liu, "Enhanced patterning by tilted ion implantation," *Proc. SPIE*, vol. 9777, pp. 273–278, Mar. 2016.

[40] D. C. Lay, S. R. Lay, and J. J. McDonald, *Linear Algebra and Its Applications*, 5th ed., London, U.K.: Pearson, 2014.

[41] Y. Yu, J. Huang, L. Wang, and S. Liang, "A 1D-inception-ResNet based global detection model for thin-skinned multifruit spectral quantitative analysis," *Food Control*, vol. 167, Jan. 2025, Art. no. 110823.

[42] Y. Yu and M. Yao, "When convolutional neural networks meet laser-induced breakdown spectroscopy: End-to-end quantitative analysis modeling of ChemCam spectral data for major elements based on ensemble convolutional neural networks," (in English), *Remote Sens.*, vol. 15, no. 13, p. 3422, Jul. 2023.

JINHAO ZHU received the M.S. degree from the University of Chinese Academy of Sciences, Beijing, China, in 2021. He is currently pursuing the Ph.D. degree in integrated circuit science and engineering with Fudan University, Shanghai, China. His research interests include optical proximity correction and computational lithography simulation.

ZHIWEI REN received the B.E. degree from Hangzhou Dianzi University, in 2023. He is currently pursuing the master's degree in microelectronics with Fudan University. His current research interest includes mask pattern spilt algorithms.

YING LI received the M.S. degree in measurement technology and instruments from Xi'an University of Technology, in 2016. She is currently pursuing the Ph.D. degree in integrated circuit science and engineering with Fudan University. Between 2016 and 2019, she was a Key Contributor with the Lithography Department of SMIC, where she spearheaded research and development efforts in lithography processes for nodes at and below 14 nm. Her extensive research accomplishments have been recognized through publications in prestigious channels like IEEE. She has showcased her work by delivering presentations at esteemed international conferences, such as IWAPS and CSTIC.

XIANHE LIU received the joint Graduate degree from the Institut National des Sciences Appliquées (INSA) and Rouen University, France, in 2015, the master's degree in chemistry and in material, and the Ph.D. degree in physical-chemistry from the University of Strasbourg, France, in 2018. She is currently an Assistant Professor with Fudan University. After the Ph.D. degree, she was with the French National Centre for Scientific Research (CNRS), Charles Sadron Institute, for two years. Then, she joined the School of Microelectronics, Fudan University, in 2021. She has authored or co-authored 30 papers and has presented invited oral in CSTIC and ASICON international conferences. Her research interests include advanced lithography process and material development.

QIANG WU received the Ph.D. degree from Yale University, in 1999. He is currently a Professor with Fudan University. Then, he worked for several companies as the Photo Engineer and the Manager. He has published 83 articles and 112 patents. His research interests include lithography process, equipment/material development, and computational lithography.

YANLI LI received the Graduate degree from Fudan University, in 2015. She worked for SMIC and ICRD and then Fudan University. She has experience in the developing of logic photolithographic process for the 28 nm, 14 nm, 10 nm, and 5 nm technological nodes. She has authored or co-authored 22 technical articles in photolithography and 42 patents. Among these patents, two have been granted with one U.S. patent. She won the second prize of Outstanding Young Engineer of CSTIC 2020 and Outstanding Young Scholar Paper of ICSICT 2022. A professional book title *Advanced Lithography Process R&D Methodology and Procedures in Modern Integrated Circuit Factories* (2024).

QI WANG received the Ph.D. degree from Fudan University, in 2019. He is currently the Young Associate Researcher with the School of Micro-Electronics, Fudan University. His research interests include lithography process, computational lithography, and optical metrology.