Diffraction-Based Overlay Metrology With Optical Convolution Layer

Jinyang Li and Hung-Fei Kuo, *Member, IEEE*

***Abstract*—Overlay is a crucial indicator of manufacturing processing between layers. Currently, diffraction-based overlay (DBO) is widely adopted in overlay metrology. In response to existing challenges in DBO metrology, this study applied the concept of optical computing to establish a DBO metrology with an optical convolution layer that can predict the overlay of a DBO target, thereby providing a promising measurement method that increases the intuitiveness and efficiency of overlay metrology. In this study, pupil images from DBO metrology were directly output to the optical convolution layer and integrated with subsequent CNN layers in electronic hardware. A total of 630 pupil images with 21 types of designed overlay were generated for network training and testing. The experimental testing results revealed that the root mean square error of the predicted values of 90 pieces of testing pupil image data was 0.17 nm. The completed optical convolution layer is the fundamental unit for establishing an optical interconnection with the functions of weighting and matrix summation, which will contribute to further exploration of the DBO technique.**

***Index Terms*—Convolution neural network (CNN), diffraction-based overlay (DBO), optical convolution, optical neural network.**

I. Introduction

OVERLAY metrology can be employed to detect unwanted positional shifts as well as process variation between layers. Overlay is a crucial measurement for processes such as semiconductor manufacturing, advanced packaging, and stage alignment; this is because it can affect a product's performance, yield, and reliability. To increase the quality of process control, precise measurement of the overlay between layers is required [1], [2], [3]. According to a technology trend report published by the International Roadmap for Devices and Systems, the continually increasing complexity of layers is reducing overlay budgets [4]. For example, the single-exposure overlay for dynamic random-access memory had to be $\leq$ 3.4 nm in 2021; by 2027, this number will be 2.2 nm, and the uncertainty of overlay metrology will be reduced from 0.7 to 0.3 nm.

Advanced processes often involve the application of overlay metrology technologies such as image-based overlay, diffraction-based overlay (DBO), and scanning electron microscopy [5]. Among these technologies, DBO is the overlay metrology technique that has the highest potential and is receiving the most attention [6], [7]. However, DBO also has several disadvantages; for example, it is easily affected by noise, produces complex and unintuitive measurement results, and exhibits poor goodness-of-fit for in-line measurements [8], [9], [10], [11], [12]. Given the excellent performance and widespread application of DBO, researchers have been testing various means to optimize DBO. Regarding light sources, multiwavelength measurement methods have been applied to considerably increase the precision and consistency of overlay metrology [13]. Various DBO targets, such as μDBO and continuous bias DBO, were designed in response to process development for high-quality measurement. Prediction methods incorporating neural networks have been widely employed [14], [15], [16], [17]. Compared with the weighted moving average method, a machine-learning algorithm can increase measurement precision by 25%–29% [18].

An advantage of optical computing over electronic computing is that it can process a large amount of data at high speed while using only a small amount of computing resources; thus, optical computing is often integrated with neural networks to establish optical neural networks (ONNs) [19]. Researchers have developed optics using microlens array, spatial light modulator (SLM), and holographic elements to realize ONN interconnections in free space [20]. A hybrid optical–electronic convolutional neural network (CNN) that uses an optimized diffractive optical element was developed for image classification tasks. Based on the 4f optical system, the network achieved a classification accuracy of 77.9% in simulation tests [21]. For CNNs, extensive optical interconnections are required to construct a network with numerous parameters [20]. To address this challenge, an optical computing loop must be generated to reduce the complexity of ONNs significantly, that is, a basic unit of ONNs, which is an optical convolution layer in this case [22].

To address the problems of complex and unintuitive measurement results and poor goodness-of-fit of in-line measurements, the present study established a CNN model for overlay prediction, and an optical convolution layer was developed to provide a new connection between a metrology tool and a prediction model. Section II of this paper describes the application of

Manuscript received 20 July 2023; revised 20 October 2023; accepted 14 November 2023. Date of publication 20 November 2023; date of current version 1 December 2023. This work was supported by the National Science and Technology Council of Taiwan through Project NSTC under Grant 110-2221-E-011-085-MY2. *(Corresponding author: Hung-Fei Kuo.)*

Jinyang Li is with the Graduate Institute of Automation and Control, National Taiwan University of Science and Technology, Taipei City 106335, Taiwan.

Hung-Fei Kuo is with the Graduate Institute of Automation and Control, National Taiwan University of Science and Technology, Taipei City 106335, Taiwan, and also with the Advanced Manufacturing Research Center, National Taiwan University of Science and Technology, Taipei City 106335, Taiwan (e-mail: hfkuo@mail.ntust.edu.tw).

Digital Object Identifier 10.1109/JPHOT.2023.3334263

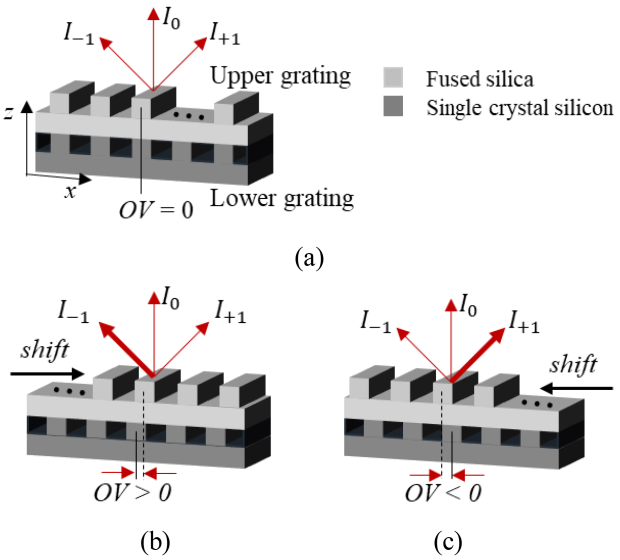

Fig. 1. Offsets of the grating target cause changes in the intensities I_{+1} and I_{-1}. (a) When the two gratings are not offset, $OV = 0$, and I_{+1} and I_{-1} are the same. (b) When a positive offset exists between the two gratings, $OV > 0$. (c) When a negative offset exists between the two gratings, $OV < 0$.

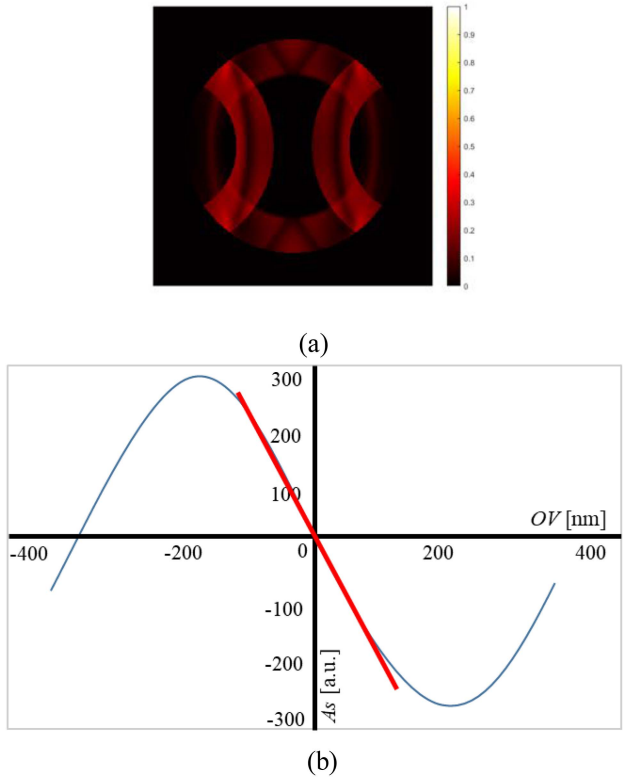

Fig. 2. *As* and *OV* are proportional in a small *OV* range (indicated by the red line). (a) The diffraction map corresponding to an overlay of +10.0 nm generated through rigorous coupled-wave analysis. (b) Relationship between *OV* and *As*.

the DBO technique and our pupil image model as a database for machine learning. Section III introduces the CNN model established for overlay prediction and an optics design based on the first convolution layer in the CNN model. Section IV compares the performance of established models to verify the appropriateness of integration of optical convolution in DBO metrology. Finally, Section V presents the conclusion.

II. Modeling of DBO Pupil Images

The DBO metrology technique incorporates angle-resolved scattering. Fig. 1 presents the stack grating used as the DBO target in the present study. The stack grating structure comprised an upper and lower grating layer. The periods of the upper and lower gratings were both 800 nm. The upper grating was made of fused silica, and its linewidth and thickness were 235 and 100 nm, respectively. The lower grating was made of single-crystal silicon, and its linewidth and thickness were 425 and 200 nm, respectively. Light is shone on the overlay target, and the beams diffracted by gratings are captured [11]. I_0 denotes a zero-order diffracted beam, whereas I_{+1} and I_{-1} represent positive and negative first-order diffracted beams, respectively. In Fig. 1(a), the upper and lower gratings do not have relative displacement; $OV = 0$, and I_{+1} and I_{-1} are equal. When the upper grating is shifted positively or negatively along the x direction relative to $OV = 0$, the overlay of the two gratings is not zero [Fig. 1(b) and (c)], I_{+1} and I_{-1} exhibit asymmetrical intensity.

A beam diffracted by DBO targets can be represented as a diffraction image map. Rigorous coupled-wave analysis (RCWA) was employed to calculate I_0, I_{+1}, and I_{-1} to obtain a diffraction image map. The RCWA tool (Gsolver) is used to analyze diffraction efficiency by the grating targets [23], [24]. This study treated the DBO target as 1D planar dielectric grating divided into 25 slices. Then apply the RCWA algorithm to generate the diffraction image map using a TE polarized plane wave incident on the grating. Fig. 2(a) presents a calculated diffraction image map with a corresponding overlay of +10.0 nm; the annular part indicates the zero-order diffracted beam, whereas the half rings on the left and right sections indicate the positive and negative first-order diffracted beams. After the intensity difference between the positive and negative first-order beams diffracted by the DBO targets with different overlays had been calculated, a scatter plot indicating the relationship between the overlay and intensity was obtained [Fig. 2(b)]; this plot revealed that, for an overlay within a small range, the differences between I_{+1} and I_{-1} (represented by *As*) and their corresponding overlays exhibited a linear relationship. Thus, the relative displacement of the two gratings, which caused differences in the intensity of first-order diffracted beams, directly reflected the alignment of the different layers within the overlay target, thereby enabling the overlay to be analyzed through its diffracted beams.

When overlay information is embedded in a diffraction image map, using a neural network as a prediction model for DBO metrology is a promising method. Furthermore, when using a neural network prediction model, additional bias does not need to be introduced into the overlay target to assist with calculations. A group of diffraction image maps was generated through RCWA to obtain a data set for the neural network;

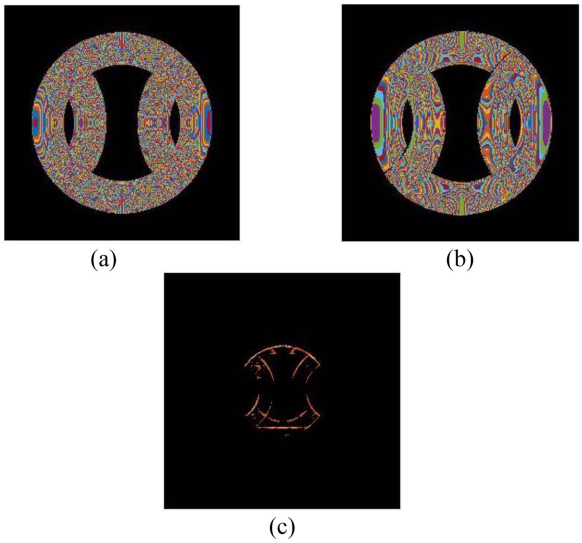

Fig. 3. Reduced complexity and image size after convolution layer operation results in a smaller network burden. (a) Diffraction image map; (b) image map after one convolution layer was used; (c) image after two convolution layers were used and max pooling was performed.

for this process, an annular laser beam with a wavelength at 638 nm, an incidence angle of between 30° and 50°, and an azimuth of between 0° and 180° was used. Specifically, the lower grating was set to have a randomly selected, asymmetric sidewall angle between 85° and 90°, whereas the upper grating was set to have no sidewall angle. Diffraction image maps with 21 types of overlay were generated; the overlays ranged from −10.0 to +10.0 nm in intervals of 1.0 nm, the condition of no overlay (i.e., $OV = 0$) was included, and 30 diffraction image maps were obtained for each type of overlay. A total of 630 designed overlay images were thus generated. From the 630 images, 450, 90, and 90 were randomly selected to form the training, validation, and testing data sets. The subsequent section presents the CNN-based overlay calculation model and optical design implemented in the present study.

III. Design of Optical Convolution Layer

An ONN is faster than a conventional neural network in electronic hardware and uses fewer computing resources. Thus, developing an ONN is crucial for the large-scale application of artificial intelligence [20]. In conventional neural networks, calculating the network weights of layers (e.g., convolution layers) is time-consuming. When the propagation weights are replaced with optical modules, the computational burden on a processor is substantially reduced because of the benefits of optical computing. Fig. 3(a) displays a generated diffraction image map, and Fig. 3(b) presents the image of that map after the use of one convolution layer; complexity was effectively reduced because of the feature extraction function of the convolution layer. Fig. 3(c) reveals that the size and complexity of the image

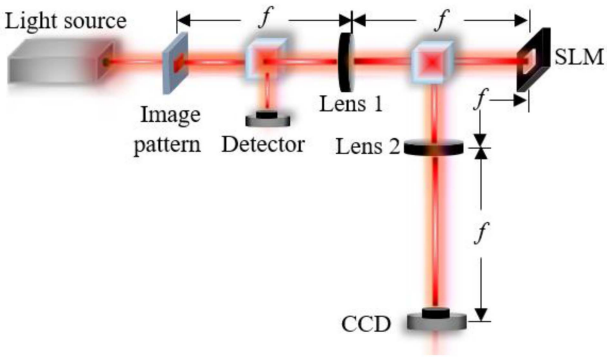

Fig. 4. Design of an optical convolution layer based on a 4f system.

were further reduced by using two convolution layers and performing max pooling, which simplified the calculation process. Thus, the development of an optical replacement of convolution layers to simplify networks is promising research. The present study designed an optical interconnection that enables optical convolution [25], [26], [27]. In addition, an SLM was used to modulate the spatial distribution of an optical field in real-time.

In the convolution layer, without the addition of deviations, each output channel j could be expressed as:

$$I_{out_j} = \sum_{i=1}^{C_{\text{in}}} [Corr(I_{in_i}, W_{i,j})] \tag{1}$$

where $Corr(\cdot)$ represents a two-dimensional correlation and $W_{i,j}$ represents the jth kernel in the ith channel. The summation operation was not required for the first layer of convolution because $C_{\text{in}} = 1$. Therefore, the present study used a 4f optical system to implement the optical convolution layer, in which the output image I_{out} was formed by conducting a convolution operation with the incident light field and system point spread function (PSF):

$$I_{\text{out}}(x, y) = I_{\text{in}}(x, y) \otimes PSF(x, y). \tag{2}$$

where "⊗" denotes a two-dimensional convolution operation. The PSF on the Fourier plane was regarded as a kernel in the frequency domain. Component parameters were edited on the Fourier plane to perform an expected convolution. At this point, the convolution could be expressed as:

$$I_{\text{in}} \otimes PSF = \mathscr{F}^{-1}[\mathscr{F}(I_{in}) \cdot \mathscr{F}(PSF)]. \tag{3}$$

where "·" denotes the pixel-wise multiplication of matrix elements, and $\mathscr{F}$ and $\mathscr{F}^{-1}$ are the two-dimensional Fourier transform and two-dimensional inverse Fourier transform, respectively.

To test the designed optical convolution layer, the optic simulation software VirtualLab (LightTrans GmbH) was employed to model the optics shown in Fig. 4. The light source, a 638-nm laser beam, was shone on the image pattern to shape a beam as a pupil image, which served as the input for the subsequent convolution layer. A detector was added to verify the input image. Lens 1 and Lens 2 were identical biconvex lenses with

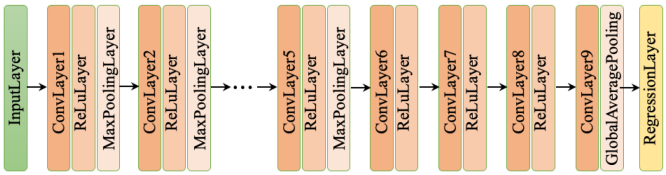

Fig. 5. Structure of the CNN model.

front and rear focal lengths of 75 mm. The SLM was employed to display the kernel in the frequency domain and placed on the focal plane between the two lenses. The charge-coupled device (CCD) was positioned on the rear focal plane of Lens 2 and read the output obtained after performing the convolution. The shaped beam passed through Lens 1, the SLM, and Lens 2, finally arriving at the CCD; the distance between any two neighboring components was 75 mm, the focal length. Thus, a 4f system was established.

A CNN model was trained for overlay prediction using images from the training data set. Fig. 5 shows the structure of the regression CNN model. There are nine convolution layers in total. The input layer has a size of $256 \times 256 \times 1$. The first five convolution layers have a kernel size of $3 \times 3 \times 32$, following a max pooling layer with a pool size of 2. A gradient descent algorithm was employed for the training process; the minibatch size was 15, and 900 iterations were performed. The training results revealed that both the root mean square error (RMSE) and loss decreased rapidly as the number of iterations increased. The training process stabilized after approximately the 200th iteration; the final RMSE was 0.14 nm. This pre-trained CNN is used as the reference for the optics design in the study.

In the designed optics, input images were filtered on the Fourier plane on which the SLM was located. The size of the filters had to correspond to that of the input images to allow for the two to be multiplied. The pre-trained CNN model comprised multiple convolution layers, of which the first employed a kernel that had 32 channels and for which $h_k \times w_k = 3 \times 3$. The output of this layer was set to have equal size of the input and zero padding. In other words, when the input was a single-channel grayscale image for which $h_i \times w_i \times C_{\text{in}} = 256 \times 256 \times 1$, the output of the convolution layer became stacked image data in the form of $h_o \times w_o \times C_{\text{out}} = 256 \times 256 \times 32$. To implement this convolution layer through optics, after the CNN model was confirmed, the relative parameters of the trained network had to be converted appropriately for the optical calculations. Because the kernel size in the convolution layer was much smaller than that of the input image, dilation of the kernels was required. A kernel in the pre-trained CNN model is used as an example; a 3×3 kernel is presented in Fig. 6(a), and the right side of the figure illustrates the dilation method. The kernel is divided into four parts and placed in the four corners of a matrix that is the same size as the input image, and the center is filled with zeros. Because of the periodicity of a fast Fourier transform, the kernels had to be arranged correspondingly to support periodic

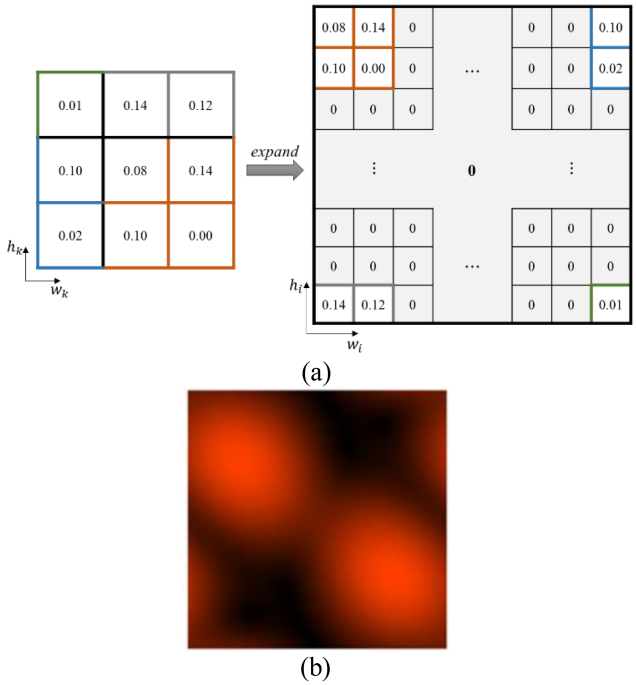

Fig. 6. Processing of a kernel. (a) Kernel dilation; (b) corresponding kernel image in frequency domain installed in the spatial light modulator.

calculations. When the kernels were divided into various sizes, the appearance of specific numbers during periodic tiling, which could cause erroneous periods and affect the calculation process, was prevented. After the 32 kernels in the first convolution layer were dilated, they were converted into the frequency domain through the two-dimensional Fourier transform. Subsequently, they could be loaded onto the SLM to participate in optical convolution. Fig. 6(b) displays the kernel image in the frequency domain corresponding to the kernel presented in Fig. 6(a) and was installed in the SLM.

The output image of the designed optical convolution layer was compared with the image from the theoretical results obtained from MATLAB to verify the layer's performance (Fig. 7). Fig. 7(a) displays the output image generated in theoretical two-dimensional convolution performed using MATLAB, whereas Fig. 7(b) displays the image obtained from the designed optical convolution layer that used the input and kernel used in Fig. 7(a). Fig. 7(c) presents the discrepancy between Fig. 7(a) and (b). The maximum difference was 0.01 per pixel. The designed optical convolution layer thus successfully outputs a high-quality image.

For the output form of the optical convolution layer, single output images were collected one by one. That is, in the developed optical setup, the output images obtained by the CCD were two-dimensional images that corresponded to every single kernel. Because the selected convolution layer consisted of 32 kernels, the SLM installed 32 kernels turn by turn for each input. The SLM could be controlled to automatically alter the installed images in accordance with the set image sequence. It

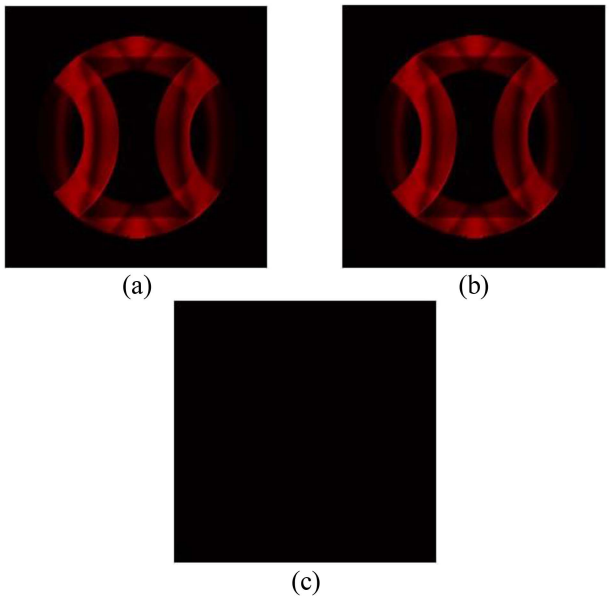

Fig. 7. Comparison of modeling output of the designed optical convolution layer from VirtualLab and MATLAB. (a) Output obtained from MATLAB; (b) output obtained from VirtualLab; and (c) discrepancy image.

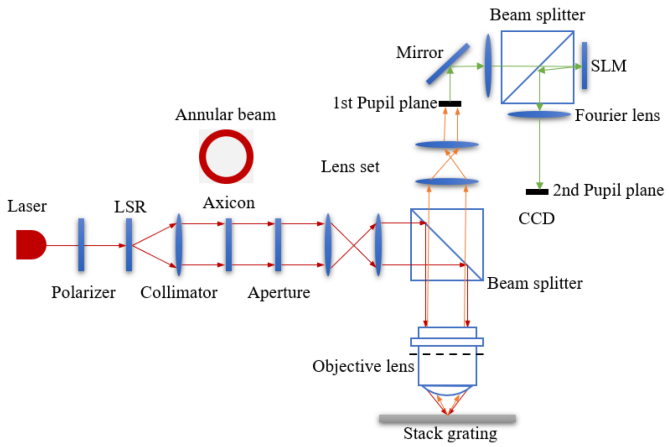

Fig. 8. Design of diffraction-based overlay metrology with an optical convolution layer.

had a usual maximum input frame rate of 60 Hz. Subsequently, the outputs from these kernels, captured by the CCD, were uploaded to the hardware and integrated to generate the 256 $\times$ 256 $\times$ 32 stacking data required for the electronic neural network that was subsequently established.

IV. Testing and Verification

Based on the aforementioned theories, the present study established an integrated optical module that incorporated DBO metrology and a convolution layer to verify the feasibility of the optical interconnection between them and to test their function for optical calculations (Fig. 8). The laser from the source in the module was diffused by a laser speckle reducer (LSR) and converted into a collimated annular beam by using a collimator

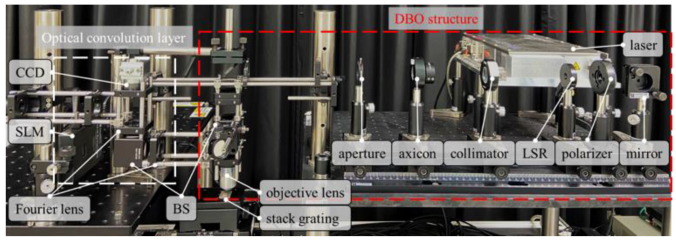

Fig. 9. Optical setup with integrated DBO metrology and an optical convolution layer.

and an axicon, following a lens set adjusting the beam into desired size. An annular beam passed through an objective lens and shone on the stack grating target. After the 1st pupil plane, where the output pupil image was located, the 4f optical system with the SLM was connected to perform the designed optical convolution layer. The CCD was employed to capture the output images on the 2nd pupil plane.

The developed optical setup is presented in Fig. 9, where the red frame outlines the DBO structure. A laser diode with a wavelength of 638 nm was used as the light source. It was converted using the collimator and polarizer into an annular beam form that was incident on the two-layer grating of the DBO target. The grating parameters were described in Section II. The objective lens (SIGMAKOKI, EPLE-50) collected the diffracted pupil images that were formed to enable further transmission to subsequent structures. The upper grating was located on the focal plane of the objective lens, whereas the lower grating was fixed onto the positioning system. The positioning system consists of two stages. The X-axis piezo stage (MotionSolutions, NPX60-223) with a resolution of 1 nm can introduce overlay from -10 nm to 10 nm with a step of 1 nm. The other is the 3-axis & rotation motorized stage (UNICE) under the piezo stage for positioning the above objects. The diffracted pupil images were first captured and input to the pre-trained CNN model to calculate the overlay of the stack grating; subsequently, the position of the lower grating was adjusted until the status of $OV = 0$ was able to be determined. Based on this reference, we could introduce an overlay by 1.0 nm into the stack grating.

In Fig. 9, the white frame indicates the optical convolution layer positioned after the 1st pupil plane behind the DBO structure. The optical convolution layer corresponded to the designed optics incorporated a 4f optical system that used two biconvex lenses with focal lengths of 75 mm. The front focal plane of the first lens was fixed on the 1st pupil plane as the input for the convolution layer. A reflective SLM (ARK JD7555, Jasper Display) with 1080 $\times$ 1920-pixel resolution, 6.4-μm pixel pitch, an effective active area of 12.29mm $\times$ 6.91 mm, and 93% fill factor was installed. The dilated kernels were displayed on the screen center with a resolution of 256 $\times$ 256. While the 32 dilated kernels were being sequentially loaded by the SLM, the CCD (Altair, U500C, resolution 2592 $\times$ 1944, pixel size 2.2 μm) captured each kernel's corresponding convolution layer output and uploaded it to the hardware for storage. The established optical setup (Fig. 9) generated 21 pupil images

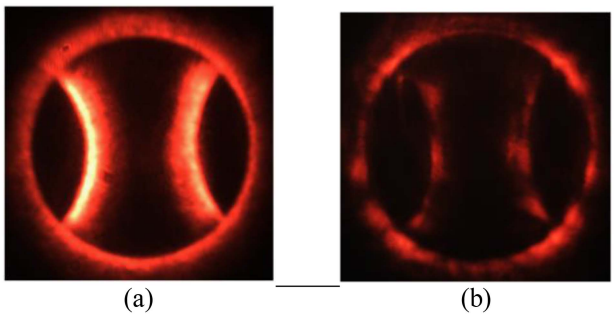

Fig. 10. Captured pupil images on 1st and 2nd pupil planes. (a) A pupil image with OV = +1 nm on the 1st pupil plane; (b) a pupil image on the 2nd pupil plane corresponding to OV = +1 nm.

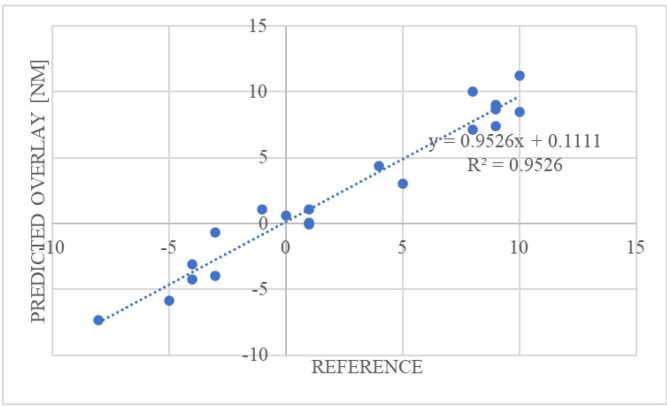

Fig. 11. Experimental results of 20 tests by the proposed model.

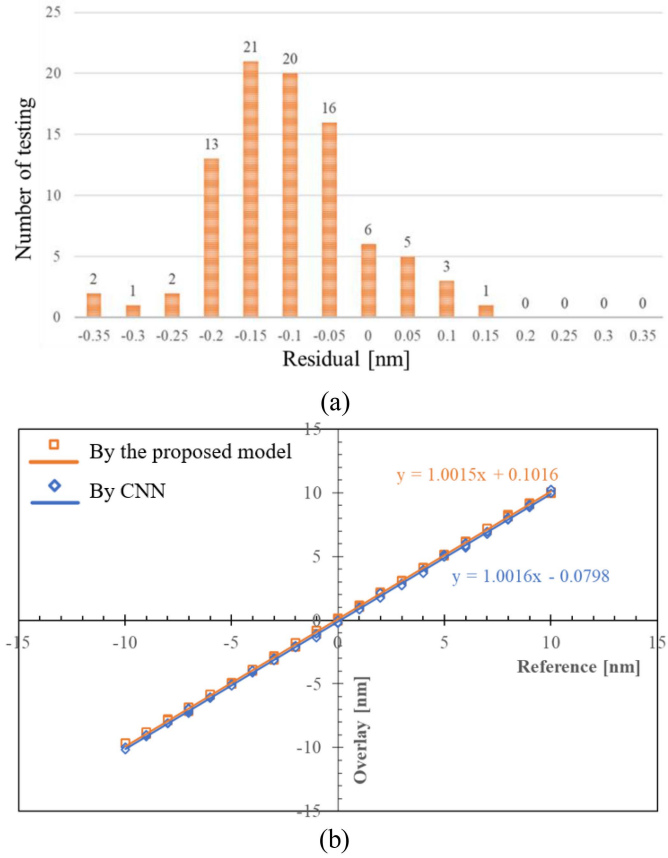

Fig. 12. (a) Residual for the proposed model. (b) Comparison of the proposed model with conventional convolution neural network.

corresponding to the introduced overlay from −10.0 to +10.0 nm. Fig. 10(a) presents an example of a pupil image that corresponded to an overlay of +1 nm and was captured on the 1st pupil plane. Fig. 10(b) presents the same image on the 2nd pupil plane after optical convolution was performed with the designed kernels. These captured images were then adjusted to produce 256 × 256 × 32 stacked image data, which were input to the subsequent network for overlay prediction. The numerical results revealed a coefficient of determination of 0.95 between the predicted and introduced overlays, implying high fitting linearity, as shown in Fig. 11. Many aspects could cause the experimental variation, including the motorized stage for rotation control having a minimum step of 0.1 degrees, the piezo stage having a position noise of 0.2 nm, and alignment error when stacking the gratings. In addition, the SLM can show 8-bit images, where the maximum numbers of colors are 256, and the CCD noise and the optical alignment could introduce extra errors in the measurement. Further research should focus on removing image noise and improving environmental conditions.

The present study employed MATLAB to verify the proposed model with an optical convolution layer. The modeling process involved the employment of a fast Fourier transform to perform image conversion at the front and rear focal planes of Lens 1 and Lens 2. The encoded kernels were implemented in the SLM for the convolution operation in the frequency domain. Ninety

TABLE I
TESTING RESULTS FROM THE PROPOSED MODEL AND CONVOLUTION NEURAL NETWORK

-	Number of Overlay Testing	RMSE [nm]	Max Error [nm]	3σ [nm]
Proposed model	90	0.17	0.38	0.43
CNN model	90	0.14	0.36	0.42

images were randomly selected from the testing data set for the proposed model. Fig. 12(a) shows the residual results for the ninety testings in the proposed model. Fig. 12(b) presents the correlation and residual analysis results from the two models used to predict the 90 testing images on the 1st pupil plane. The blue line indicates the results from the conventional CNN model, whereas the orange line indicates the results from the proposed model with an optical convolution layer. The reference axis indicates the designed overlay. Fig. 12(b) reveals strong correlations for both models, with the coefficients of the determination being up to 0.99. Table I reveals that the proposed model with an optical convolution layer exhibited performance comparable to that of the conventional CNN model with respect to their RMSEs (proposed vs. conventional CNN models, 0.17 vs. 0.14 nm) and maximum errors (proposed vs. conventional CNN models, 0.38 vs. 0.36 nm). The 3σ is calculated

based on the residual data and regards a residual of zero as the mean position to estimate the uncertainty of the prediction model. The 3σ values of the proposed and conventional models were 0.42 and 0.43, respectively.

V. Conclusion

The present study successfully integrated DBO metrology with an optical convolution layer. Through experimental verification, the overlay prediction performance of the completed hybrid CNN model was comparable to that of the conventional CNN model. Insertion of the optical convolution layer into the overlay prediction model suggested that in-line real-time measurements can be made in semiconductor manufacturing. With respect to the application of the developed optical convolution layer as a fundamental unit in ONNs, future studies should explore the inclusion of DBO metrology with an optical CNN that fully implements the ONN concept.

References

[1] H. Zhang, T. Feng, and D. Djurdjanovic, "Dynamic down-selection of measurement markers for optimized robust control of overlay errors in photolithography processes," *IEEE Trans. Semicond. Manuf.*, vol. 35, no. 2, pp. 241–255, May 2022.

[2] Y.-L. Liu et al., "Imaging-based overlay metrology optimized by HV-SEM in 3D-NAND process," *Proc. SPIE*, vol. 11611, pp. 659–666, 2021.

[3] T. C. Tin et al., "A realizable overlay virtual metrology system in semiconductor manufacturing: Proposal, challenges and future perspective," *IEEE Access*, vol. 9, pp. 65418–65439, 2021.

[4] B. Hoefflinger, "IRDS—International roadmap for devices and systems, rebooting computing, S3S," in *Nano-Chips 2030*. Berlin. Germany: Springer, 2020, pp. 9–17.

[5] T. Tamulevičius, I. Gražulevičiūtė, A. Jurkevičiūtė, and S. Tamulevičius, "The calculation, fabrication and verification of diffraction grating based on laser beam splitters employing a white light scatterometry technique," *Opt. Lasers Eng.*, vol. 51, no. 10, pp. 1185–1191, 2013.

[6] B. Bringoltz et al., "Accuracy in optical overlay metrology," *Proc. SPIE*, vol. 9778, pp. 483–501, 2016.

[7] M. Adel, D. Kandel, V. Levinski, J. Seligson, and A. Kuniavsky, "Diffraction order control in overlay metrology: A review of the roadmap options," *Proc. SPIE*, vol. 6922, pp. 23–41, 2008.

[8] B. Bunday, A. Bello, E. Solecky, and A. Vaid, "7/5nm logic manufacturing capabilities and requirements of metrology," *Proc. SPIE*, vol. 10585, 2018, Art. no. 105850I.

[9] X. Ma, X. Zhao, Z. Wang, Y. Li, S. Zhao, and L. Zhang, "Fast lithography aerial image calculation method based on machine learning," *Appl. Opt.*, vol. 56, no. 23, pp. 6485–6495, 2017.

[10] Y.-S. Kim et al., "Improving full-wafer on-product overlay using computationally designed process-robust and device-like metrology targets," *Proc. SPIE*, vol. 9424, pp. 375–385, 2015.

[11] S. Lozenko et al., "Matching between simulations and measurements as a key driver for reliable overlay target design," *Proc. SPIE*, vol. 10585, 2018, Art. no. 105851E.

[12] H.-J. H. Smilde et al., "Target design optimization for overlay scatterometry to improve on-product overlay," *Proc. SPIE*, vol. 9424, 2015, Art. no. 942412.

[13] K. Bhattacharyya et al., "Multi-wavelength approach towards on-product overlay accuracy and robustness," *Proc. SPIE*, vol. 10585, 2018, Art. no. 105851F.

[14] J. Zhu, S. Liu, C. Zhang, X. Chen, and Z. Dong, "Identification and reconstruction of diffraction structures in optical scatterometry using support vector machine method," *Proc. SPIE*, vol. 12, 2013, Art. no. 013004.

[15] W. Jin, J. Bao, and S. Li, "Optical metrology using support vector machine with profile parameter inputs," U.S. Patent US7483809B2, 2009.

[16] E. Schmitt-Weaver et al., "Overlay improvements using a real time machine learning algorithm," *Proc. SPIE*, vol. 9050, pp. 496–502, 2014.

[17] C. Hwang et al., "Smart overlay metrology pairing adaptive deep learning with the physics-based models used by a lithographic apparatus," *Proc. SPIE*, vol. 10587, 2018, Art. no. 105870B.

[18] C.-H. Su, Z.-H. Lin, Y.-S. Lin, and H.-F. Kuo, "Enhancement of diffraction-based overlay model for overlay target with asymmetric sidewall," *IEEE Trans. Semicond. Manuf.*, vol. 33, no. 3, pp. 373–382, Aug. 2020.

[19] W. Hu, X. Li, J. Yang, and D. Kong, "Crosstalk analysis of aligned and misaligned free-space optical interconnect systems," *J. Opt. Soc. Amer. A*, vol. 27, no. 2, pp. 200–205, 2010.

[20] X. Sui, Q. Wu, J. Liu, Q. Chen, and G. Gu, "A review of optical neural networks," *IEEE Access*, vol. 8, pp. 70773–70783, 2020.

[21] J. Chang, V. Sitzmann, X. Dun, W. Heidrich, and G. Wetzstein, "Hybrid optical-electronic convolutional neural networks with optimized diffractive optics for image classification," *Sci. Rep.*, vol. 8, no. 1, pp. 1–10, 2018.

[22] S. Xu, J. Wang, R. Wang, J. Chen, and W. Zou, "High-accuracy optical convolution unit architecture for convolutional neural networks by cascaded acousto-optical modulator arrays," *Opt. Exp.*, vol. 27, no. 14, pp. 19778–19787, 2019.

[23] M. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of metallic surface-relief gratings," *J. Opt. Soc. Amer. A*, vol. 3, no. 11, pp. 1780–1787, 1986.

[24] M. Moharam and T. K. Gaylord, "Diffraction analysis of dielectric surface-relief gratings," *J. Opt. Soc. Amer.*, vol. 72, no. 10, pp. 1385–1392, 1982.

[25] X. Lin et al., "All-optical machine learning using diffractive deep neural networks," *Science*, vol. 361, no. 6406, pp. 1004–1008, 2018.

[26] B. Wang, J. Yang, X. Liu, P. Gu, J. Tang, and Y. Zhang, "Realization of parallel subtraction using LCLV in optical neural network," *Proc. SPIE*, vol. 2897, pp. 89–94, 1996.

[27] J. Bueno et al., "Reinforcement learning in a large-scale photonic recurrent neural network," *Optica*, vol. 5, no. 6, pp. 756–760, 2018.