

Data-driven approaches to optical patterned defect detection

MARK-ALEXANDER HENN,* HUI ZHOU, AND BRYAN M. BARNES

Nanoscale Device Characterization Division, Physical Measurement Laboratory, National Institute of Standards and Technology, 100 Bureau Drive, Gaithersburg, MD 20899, USA

*mark.henn@nist.gov

Abstract: Computer vision and classification methods have become increasingly wide-spread in recent years due to ever-increasing access to computation power. Advances in semiconductor devices are the basis for this growth, but few publications have probed the benefits of data-driven methods for improving a critical component of semiconductor manufacturing, the detection and inspection of defects for such devices. As defects become smaller, intensity threshold-based approaches eventually fail to adequately discern differences between faulty and non-faulty structures. To overcome these challenges we present machine learning methods including convolutional neural networks (CNN) applied to image-based defect detection. These images are formed from the simulated scattering of realistic geometries with and without key defects while also taking into account line edge roughness (LER). LER is a known and challenging problem in fabrication as it yields additional scattering that further complicates defect inspection. Simulating images of an intentional defect array, a CNN approach is applied to extend detectability and enhance classification to these defects, even those that are more than 20 times smaller than the inspection wavelength.

© 2019 Optical Society of America under the terms of the [OSA Open Access Publishing Agreement](#)

1. Introduction

Compared to the Apollo 11's onboard guidance computer, a modern cellphone is about 1,400 times faster and has 4,000,000 times more memory [1]. These dramatic increases illustrate the substantial impact, observed by Gordon E. Moore, that with the decrease in production costs, the number of transistors in a dense integrated circuit will double about every two years [2]. Transistor count is the most common measure of integrated circuit complexity and is closely related to computational performance [3] - the main force driving the feasibility and wide-spread availability of the different techniques of data-driven methods. The manufacturing of these integrated circuits as of 2017 has become a \$ 400 billion industry [4], and even as the semiconductor industry struggles to perpetuate Moore's law [5], crucial challenges exist in monitoring the production process for decreasing feature sizes [6].

One of the most pressing challenges is the detection of so-called "killer defects" i.e., deviations that would lead to device failure due to shorted or broken electrical connections within the layers that are printed using photolithography [7,8]. Not detecting such defects can lead to systematic imperfections within other devices on subsequent wafers, resulting in multiple failed devices that are only observed after fabrication is complete. Examples of such observations include electrical testing and the X-ray based detection of die warpage after packaging [9]. The latter however is limited by long scan times for laboratory-based X-ray sources.

Defect metrology concentrates on locating and identifying these defects during manufacturing to increase yield. Optical tools, such as scatterometry [10] or imaging techniques [11], are the only way to successfully inspect these defects non-destructively at high speeds over the area of the typical 300 mm diameter wafer. As killer defects decrease in size with shrinking device dimensions the scattered intensity from these defects becomes harder to detect, thus for either approach a large amount of data need to be processed. Converting these low-intensity data into

meaningful results requires exploiting the very increase in computation power that results from successfully producing more powerful devices. In this work, this virtuous cycle is illustrated by adding several key aspects from machine learning to image-based defect detection, comparing a contemporary deep ultraviolet (DUV) inspection wavelength against proposed and potential vacuum- and extreme-ultraviolet (VUV, EUV) wavelengths.

While machine learning (ML) has successfully been reported in the analysis of patterns of poor device yield across such wafers after electrical testing [12–14] with some even using convolutional neural networks (CNN) [15–17], only recently has the imaging of defects been treated in a ML setting, more specifically by using principal component analysis [18]. This work broadens the application of ML to improve localized, image-based defect metrology by comparing linear classifiers and CNNs. Note, image-based defect detection with machine learning has been realized in other industries e.g., textiles [19–22], steel [23], and wood [24], but a key difference is that due to the decreased dimensions in semiconductors these defects must be detected even as they are often unresolved.

2. Simulation details

Shown schematically in Fig. 1 are two types of bridging defects and two types of line extensions which in general are harder to detect. These layouts are based upon public information about recent manufacturing processes [25] of the fins for 3-D field effect transistors (finFETs) and also upon an intentional defect array defined by SEMATECH [26]. The latter provides the naming convention for the defectuous wafers, see panels (b) to (e) in Fig. 1.

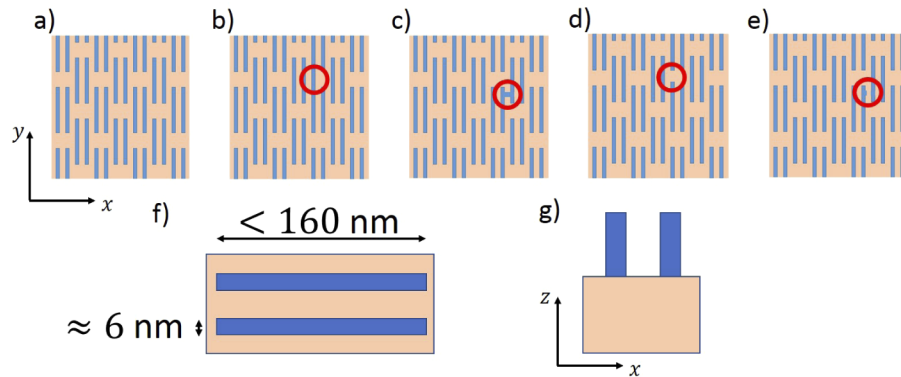


Fig. 1. Schematic representation of a) ideal layout, b) Bx and c) By bridging defects, d) Cx and e) Cy line extension defects, and f) and g) key dimensions of the unit cell. The lighter color is simulated as crystalline silicon and the blue as amorphous silicon. For clarity two 2 nm thick conformal layers that coat the amorphous silicon are not shown. Geometry and materials details are available at [27].

The simulations were performed using a well-verified [28–30] in-house implementation of the finite-difference time-domain [31] (FDTD) method to model the electromagnetic field scattered from the patterned layout and its defects. The incident angle of the illumination is chosen to be normal to the substrate for clarity; prior simulation results indicate that the defect detection often varies when using oblique illumination [29,32]. The linear polarization basis within the simulation is defined with respect to the long axis of the nominal pattern. (Note, the x and y directions do not correspond to the defect naming convention.) The scattered and reflected fields are converted to images through an idealized modeling of the Fourier optics assuming a collection numerical aperture of 0.95 for simplicity. To account for the measurement noise, Poisson (shot) noise [33] is applied to the raw images. Throughout the remainder of this paper

the pixel size at the sample is assumed to be $10 \text{ nm} \times 10 \text{ nm}$. This number has been proven to be a good compromise between the changes in the noise model due to an increased photon count for larger pixels and aliasing effects.

We have previously utilized intensity and area thresholding to extract a signal-to-noise ratio (SNR) from differential images. In [30,34,35] the SNR was strongest for bridging defects when applying a wavelength of $\lambda = 47 \text{ nm}$ for the simulated structures, while the SNR varied notably across the remaining wavelengths. While SNR is straightforward and convenient it does require an informed choice of these intensity and area thresholds to exclude the shot noise, and for sufficiently low photon densities a SNR may not be reportable.

In this work three wavelengths are employed, 193 nm as a common inspection wavelength, 47 nm that performed best for the SNR metric, and 13 nm which was proven to be a challenging wavelength in our prior reports. The photon densities that are applied here are based on current estimates for the intensity of the three wavelengths used in this study, see Refs. [36,37], and can be found in Tab. 1 below.

Table 1. Benchmark photon densities ρ_{ph} from the literature [36,37] for the wavelengths used in this study.

$\lambda \text{ (nm)}$	13	47	193
$\rho_{\text{ph}} \text{ (nm}^{-2}\text{)}$	10	10	10^5

In addition to measurement noise, “wafer noise” due to process variations is included [38–42]. Line edge roughness (LER) is known to be present in every lithographically manufactured device, either reducing the signal or increasing the noise [43]. The geometries in these simulations include LER [44,45] that is based on the current state-of-the-art with $3 \cdot \sigma_{\text{LER}} = 0.6 \text{ nm}$, i.e. 10 % of the line width and a correlation length of $\xi = 10 \text{ nm}$ [46]. Both types of noise are applied separately to defect and no-defect images that are used to form the absolute value differential images (AVDI). These AVDIs form the basis of the investigations, and are realized by subtracting these two images and taking the absolute value of the resulting pixel values, see Fig. 2 for examples. Here, the intensities in the defect and no-defect images are given relative to the intensity of the incident light, the AVDIs are converted to 8-bit integers and normed to their respective maximum values for use in the machine learning algorithms. This subtraction almost completely removes the background at a cost of combining two realizations of shot noise.

Separating the shot noise from the noise due to LER for experimental data is a very challenging task for optical tools, more easily achieved in scatterometry-based critical dimension metrology [47,48] than for imaging defect inspection tools. For this simulation study, by definition we have perfect knowledge of the distribution of the shot noise and removing this would lead to unrealistically good results. Noise filtering in this work is therefore limited only to wavelet-based compression techniques.

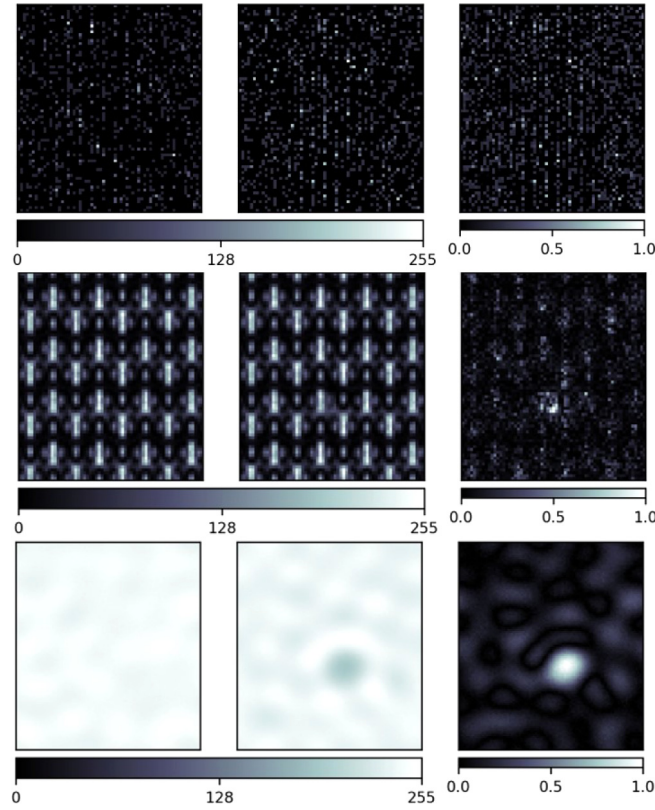


Fig. 2. Example images with Poisson noise generated using photon densities from Tab. 1, (left) no defect, (center) defect, (right) AVDI. (top row) Bx defect, Y polarization, $\lambda = 13$ nm, (middle row) Cx defect, Y polarization, $\lambda = 47$ nm, (bottom row) By defect, X polarization, $\lambda = 193$ nm. While the longer wavelengths are able to identify the defect, it is almost indistinguishable from noise at $\lambda = 13$ nm.

3. Implementation of ML algorithms

Two types of ML algorithms are applied to this classification problem. While the first type, linear classifiers (LC) [49] can in some sense be seen as extensions of our previously used SNR metric, the second one, convolutional neural networks (CNNs) [50] are a class of algorithms that are widely used across a vast number of image recognition tasks. Each algorithm requires a set of features to operate on, and the selection of these features is an integral part in any ML setup. Limited computation resources, especially memory, have guided the selection of these methods. The linear classifier uses histograms, while the CNN processes wavelet-compressed AVDIs.

The histogram of its pixel intensities is an easily obtained image feature. Even though one discards the spatial information of the image, several applications in such diverse fields as wood [24] and fabric inspection [51] have proven that histograms can be a very valuable feature for classification. The intensities for each image are normed relative to its respective maximum value and a total of 100 bins to create the histogram has been used (although not shown, setting the number of bins to less than 50 in this work has a negative effect on the performance of the ML algorithms used). The histograms are normalized to show the relative pixel frequencies. A training set is created that consists of $n_t = 10000$ histograms $x^{(i)} \in \mathbb{R}^{100}$, $i = 1, \dots, n_t$, one half of which are images from simulations without a defect and labeled by $y^{(i)} = [0, 1]^T$, $i = 1, \dots, 5000$, and the other half are images from simulations with a defect $y^{(i)} = [1, 0]^T$, $i = 5001, \dots, 10000$.

Figure 3 presents the histograms for two critical wavelength/defect combinations. Note the clear difference in the histograms for $\lambda = 193$ nm while the histograms for $\lambda = 13$ nm are virtually indistinguishable.

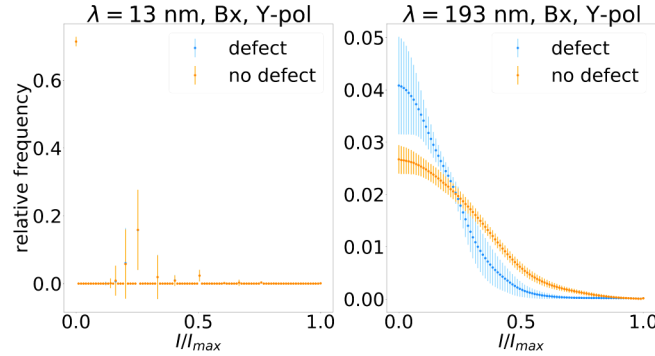


Fig. 3. Histograms of pixel intensities as used as features in the linear classifier, $\rho_{ph}(193 \text{ nm}) = 10^5 \text{ nm}^{-2}$, $\rho_{ph}(13 \text{ nm}) = 100 \text{ nm}^{-2}$.

The classifier that has been applied to these histograms is a simple linear classifier (LC). Figure 4 has the classification success rates (CSRs) for the above algorithm if applied to histogram data that has been generated using realistic photon densities as given in Tab. 1. Ten optimizations have been performed to determine the mean CSRs. The corresponding standard deviations in all cases were below 0.01 and are therefore not plotted. The LC performs quite well for $\lambda = 193$ nm yielding a CSR of approximately 0.98, i.e., on average it successfully classifies an image in 98% of all cases. For the $\lambda = 13$ nm data with a low photon count $\rho_{ph} = 1 \text{ nm}^{-2}$ to 10 nm^{-2} , however, the LC performs poorly as the CSR is just above 0.5. With increasing ρ_{ph} , the CSR increases to 0.82 for the Bx defect and 0.98 for the By defect at $\rho_{ph} = 1000 \text{ nm}^{-2}$. It has been reported that the SNR is a good metric for defect detection at $\lambda = 47$ nm, so it is not surprising to see that the LC also performs very well for a reference photon density of 10 nm^{-2} . Even for photon counts that are a magnitude less, a CSR of around 0.9 is achieved here. However, for real-world process control this value is not satisfying. The same can be said for the LC at $\lambda = 13$ nm, even for an increased photon count. Therefore the CNN is to be applied which will need a different feature that ideally contains more information without requiring too much memory.

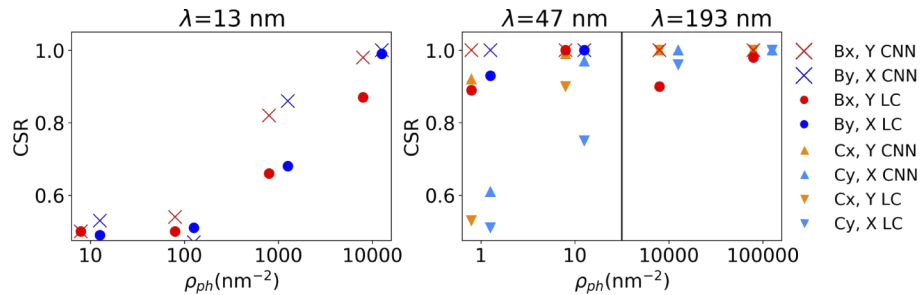


Fig. 4. CSRs as a function of photon density, capital letter after defect type denotes polarization of incident light, CNN denotes convolutional neural network, LC denotes linear classifier.

Memory constraints using the full images with a pixel size of $10 \text{ nm} \times 10 \text{ nm}$ arise as images at $\lambda = 13 \text{ nm}$ and $\lambda = 47 \text{ nm}$ consist of 71×63 pixels and 107×95 for $\lambda = 193 \text{ nm}$. Building a sufficiently large library of training data with these images is not possible given our resources,

hence the information provided by the images is condensed by applying a two-step wavelet-based image compression using a 'db1' wavelet [52], by high-pass filtering the original images. The resulting images are then low-pass filtered and downsampled, yielding approximated subimages for which the procedure is repeated. This approach leads to a reduction of the image sizes to 18×16 pixels for $\lambda = 13$ nm and $\lambda = 47$ nm and 27×24 pixels for $\lambda = 193$ nm, and hence an increase of the pixel sizes at the sample from 10 nm to 39.4 nm. Even with these larger pixel sizes, the wavelet-based compressed images preserve the details of the original defect images while they tend to disappear if one simply rebins, see Fig. 5 for an example.

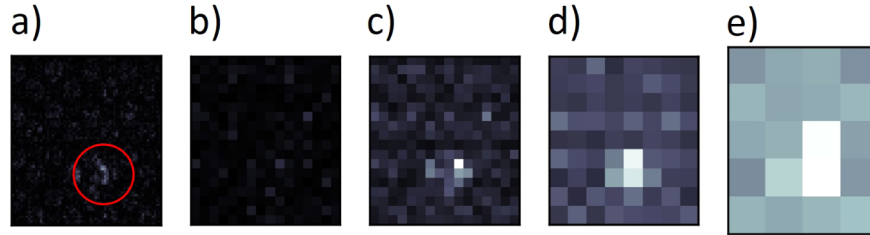


Fig. 5. Effect of compression and binning for the Cx defect (circled), $\lambda = 47$ nm, Y-polarization, $\rho_{ph} = 10 \text{ nm}^{-2}$, a) original AVDI, b) rebinned image, c)-e) wavelet-based compression for two, three, and four steps.

With the 16-fold data reduction due to the compression, it is now possible to train a convolutional neural network that uses the spatial information contained in the compressed images as features to classify the defect/no-defect AVDIs. A known, fundamental architecture that has been proven to successfully detect defects in a slightly different field [23] is used, given in Fig. 6, and implemented using the TensorFlow toolbox [54]. Just as with in the histogram case, a training set for each different λ - ρ_{ph} -defect-polarization combination is created.

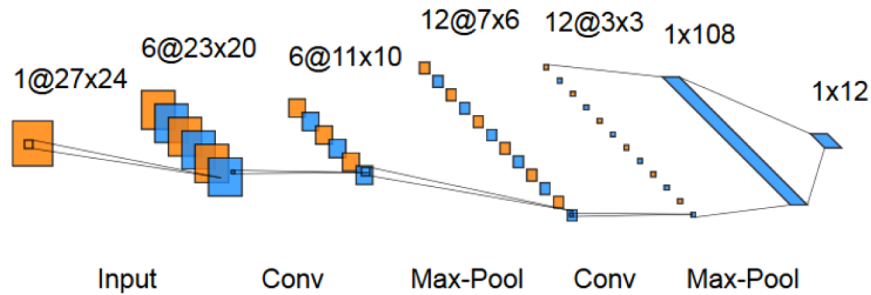


Fig. 6. Schematic representation of used CNN, the filter size for the convolution layers was set to 5×5 pixels. For supplementary information see [53].

Initially the LC and the CNN are both used for binary classification, i.e., defect no-defect. The capabilities of the CNN will be evaluated further by using an order of magnitude less light at $\lambda = 193$ nm and also attempting defect classification among the Bx, By, Cx, and Cy defects and the no-defect case.

4. CNN results

Starting with the $\lambda = 47$ nm case and the reduced photon density of $\rho_{ph} = 1 \text{ nm}^{-2}$, recall that the histogram approach resulted in CSRs of approximately 0.9 for both the Bx and By defects. Using the CNN approach yields CSRs of almost 1, cf. the red and blue 'x's in Fig. 4. The same improvement in performance is observed for the Bx defect and $\lambda = 193$ nm case, with

the CSR increasing from 0.88 for the LC to basically 1 for the CNN. For $\lambda = 13$ nm with current photon densities, even the CNN approach cannot detect these defects at this wavelength. Therefore the change in the CSRs is to be determined for increased ρ_{ph} . With a photon density of $\rho_{ph} = 1000 \text{ nm}^{-2}$ values for the CSRs are close enough to 1 if the CNN is used. That is about an order of magnitude less in photon density than would be needed for the SNR metric to successfully separate defects from no-defect images [35].

Finally the presented approaches are applied to smaller defects, excluding $\lambda = 13$ nm due to the difficulties this wavelength presented for larger defects. Figure 1 c) and d) have a schematic representation of the non-bridging defects that, following the SEMATECH convention, will be denoted by Cx and Cy, and that shall be investigated here using $\lambda = 47$ nm and $\lambda = 193$ nm. Although not shown, neither defect could be classified adequately using the SNR metric for any wavelength, further motivating the use of machine learning based methods. The results for applying the LC and CNN approaches are presented in Fig. 4 as light blue and orange triangles, respectively. The shorter wavelength does not have any problems detecting the Cx and Cy defects at the current ρ_{ph} if the CNN is used, but only reaches a maximum CSR of 0.92 for the Cx defect for lower photon densities. As expected the LC is not sufficient to detect either defect at $\lambda = 47$ nm for any given photon density due to the very small scattering volume. On the other hand, $\lambda = 193$ nm performs very well on those small defects, especially given their size, that is approximately less than $\frac{1}{20}$ -th the length of the inspection wavelength. The Cy defect at this wavelength does not even require the CNN approach and yields a CSR of 0.995 using the LC.

For processing images for high-volume manufacturing however, the image's size in memory may need to be decreased beyond the 16-fold reduction from the two-step wavelet compression. Figure 7 shows the effect that further compressing of the images has on the obtained CSRs for $\lambda = 13$ nm. Increased compression does indeed have a negative impact on CSRs > 0.6 , decreasing the values for both defect types at the two larger investigated photon densities. For the smaller photon densities, better CSRs might occur for higher compression, however with CSR < 0.6 this is of negligible impact on defect detection. It is however surprising to see that the drop in CSRs is not as dramatic as expected, for example the CSR decreased from 0.985 to 0.938 for the Bx defect at $\rho_{ph} = 10000 \text{ nm}^{-2}$. While the CSRs from the highly compressed data are insufficient for practical application, the size of the image in memory is one variable of many that must be optimized for data-driven defect inspection.

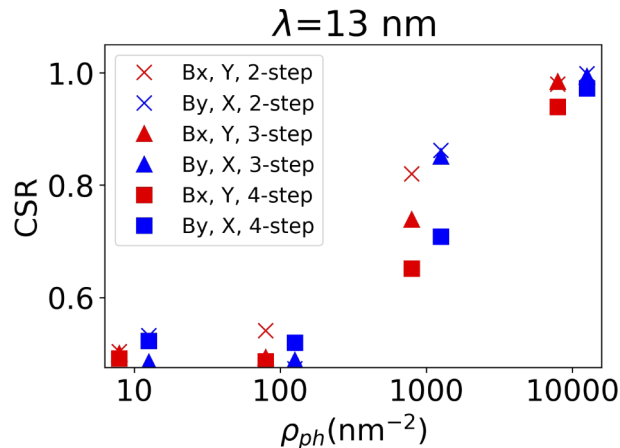


Fig. 7. Effect of compression on CSRs for $\lambda = 13$ nm.

One advantage of simulating an intentional defect array is the perfect prior knowledge of each image's defect type, and this enables further testing of the CNN beyond binary classification

as shown in Fig. 4 using training and test data unique to each defect type. In the following we therefore train the same CNN architecture using a five-fold classification to distinguish among the no-defect case and the four types of defects as shown in Fig. 2. Specifically we use AVDIs that were generated for a wavelength of $\lambda = 193$ nm and an order of magnitude less photon density than the benchmark value to represent possible inefficiencies in source strengths and faster data acquisition with examples shown in Fig. 8. From each of these five classes 4000 images have been generated leading to a total of $n_t = 20000$ images that are separated into 16000 training and 4000 test images.

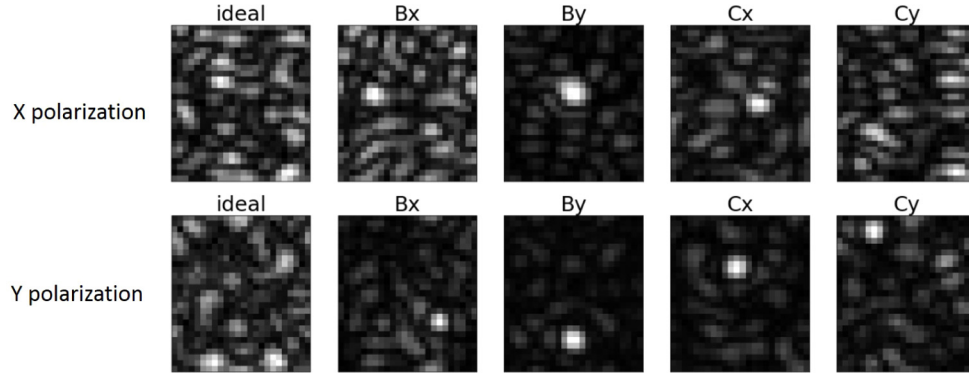


Fig. 8. Defect AVDIs at $\lambda = 193$ nm and $\rho_{ph} = 10000$ nm⁻². While many defects are easily identified by eye, some defect and polarization combinations yield difference images that are visually similar.

The classification works quite well with the confusion matrix presented in Tab. 2 showing the accurate classification of three of the four defect types and of the no-defect case. For X polarization only 14 % of the By images were misclassified as Cx defects; for Y polarization (not shown) the trends are the same with 16 % of By images misclassified similarly and an overall CSR of 0.968. Note, that despite the small error rate, all defects were accurately flagged as a defect for both linear polarizations. Another key result from these data is that the success of the CNN does not depend upon polarization optimization. Contrast this with Fig. 4 where these highly directional defects are illuminated using each defect's optimal linear polarization axis, e.g., Bx with Y polarization. While defects encountered in nanoelectronics fabrication often defy such straightforward classification, the presented results demonstrate the versatility of a CNN approach to addressing the ever-pressing challenge of detecting killer defects.

Table 2. Confusion matrix for $\lambda = 193$ nm, $\rho_{ph} = 10000$ nm⁻², X-polarization, CSR = 0.974 for multiple defect classification. As is apparent the random draw of 4000 test images did not precisely select 800 of each class.

Ground Truth \ Inferred	no defect	Bx	By	Cx	Cy
no defect	802	0	0	0	0
Bx	0	848	0	0	0
By	0	0	646	103	0
Cx	0	0	0	780	0
Cy	0	0	0	0	821

5. Conclusion

We have applied two data-driven approaches to defect detection, namely linear classifiers and convolutional neural networks to simulated images computed using normally incident illumination using three wavelengths. As expected, CNN outperforms both the linear classifier and the SNR, due to the conservation of spatial information of the images. A very straightforward CNN approach can be used to extend the defect detectability to smaller defects, even as some are more than 20 times smaller in one dimension than the inspection wavelength of $\lambda = 193$ nm. Successful classification from an intentional defect array has been demonstrated for these data at this longer wavelength.

However, the prospects for defect metrology remain challenging at $\lambda = 13$ nm despite the implementation of ML algorithms, partly due to the very low photon density realistically expected at this wavelength. It has been demonstrated that an increase in photon density can help to improve the detectability significantly using longer wavelengths for “killer defects” and further improvements in defect detectability through optimal combinations of illumination angle, polarization, photon density, defect type, and wavelength can be expected.

Acknowledgments

The authors would like to thank Richard M. Silver of NIST and Petru Manescu of University College London for useful discussions and comments. Certain commercial materials are identified in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials are necessarily the best available for the purpose.

References

1. D. Silverman, “Your smartphone is light years ahead of NASA computers that guided Apollo moon landings,” *Houston Chronicle* (11 April 2019).
2. G. E. Moore, “Cramming more components onto integrated circuits,” *Proc. IEEE* **86**(1), 82–85 (1998).
3. J. Koomey, S. Berard, M. Sanchez, and H. Wong, “Implications of historical trends in the electrical efficiency of computing,” *IEEE Annals Hist. Comput.* **33**(3), 46–54 (2011).
4. Semiconductor Industry Association, “Annual semiconductor sales,” (2017).
5. M. M. Waldrop, “The chips are down for Moore’s law,” *Nature (London, U. K.)* **530**(7589), 144–147 (2016).
6. N. G. Orji, M. Badaroglu, B. M. Barnes, C. Beitia, B. D. Bunday, U. Celano, R. J. Kline, M. Neisser, Y. Obeng, and A. Vladar, “Metrology for the next generation of semiconductor devices,” *Nat. Electron.* **1**(10), 532–547 (2018).
7. J. E. Bjorkholm, “EUV lithography - the successor to optical lithography,” *Intel Technol. J.* **2**, 1–8 (1998).
8. T. F. Crimmins, “Defect metrology challenges at the 11-nm node and beyond,” in *Metrology, Inspection, and Process Control for Microlithography XXIV*, vol. 7638 (International Society for Optics and Photonics, 2010), 76380H.
9. N. Gorji, B. Tanner, R. Vijayaraghavan, A. Danilewsky, and P. McNally, “Nondestructive, in situ mapping of die surface displacements in encapsulated ic chip packages using x-ray diffraction imaging techniques,” in *2017 IEEE 67th Electronic Components and Technology Conference (ECTC)*, (IEEE, 2017), pp. 520–525.
10. T. Harada, M. Nakasuji, A. Tokimasa, T. Watanabe, Y. Usui, and H. Kinoshita, “Defect characterization of an extreme-ultraviolet mask using a coherent extreme-ultraviolet scatterometry microscope,” *Jpn. J. Appl. Phys.* **51**, 06FB08 (2012).
11. R. M. Silver, B. M. Barnes, Y. Sohn, R. Quintanilha, H. Zhou, C. Deeb, M. Johnson, M. Goodwin, and D. Patel, “The limits and extensibility of optical patterned defect inspection,” *Proc. SPIE* **7638**, 76380J (2010).
12. F. Adly, O. Alhussein, P. D. Yoo, Y. Al-Hammadi, K. Taha, S. Muhaidat, Y.-S. Jeong, U. Lee, and M. Ismail, “Simplified Subspaced Regression Network for Identification of Defect Patterns in Semiconductor Wafer Maps,” *IEEE Trans. Ind. Inf.* **11**(6), 1267–1276 (2015).
13. F. Adly, P. D. Yoo, S. Muhaidat, Y. Al-Hammadi, U. Lee, and M. Ismail, “Randomized General Regression Network for Identification of Defect Patterns in Semiconductor Wafer Maps,” *IEEE Trans. Semicond. Manufact.* **28**(2), 145–152 (2015).
14. K. Nakata, R. Orihara, Y. Mizuoka, and K. Takagi, “A Comprehensive Big-Data-Based Monitoring System for Yield Enhancement in Semiconductor Manufacturing,” *IEEE Trans. Semicond. Manufact.* **30**(4), 339–344 (2017).
15. T. Nakazawa and D. V. Kulkarni, “Wafer Map Defect Pattern Classification and Image Retrieval Using Convolutional Neural Network,” *IEEE Trans. Semicond. Manufact.* **31**(2), 309–314 (2018).

16. G. Tello, O. Y. Al-Jarrah, P. D. Yoo, Y. Al-Hammadi, S. Muhaidat, and U. Lee, "Deep-Structured Machine Learning Model for the Recognition of Mixed-Defect Patterns in Semiconductor Fabrication Processes," *IEEE Trans. Semicond. Manuf.* **31**(2), 315–322 (2018).
17. K. Kyeong and H. Kim, "Classification of Mixed-Type Defect Patterns in Wafer Bin Maps Using Convolutional Neural Networks," *IEEE Trans. Semicond. Manuf.* **31**(3), 395–402 (2018).
18. S. Purandare, J. Zhu, R. Zhou, G. Popescu, A. Schwing, and L. L. Goddard, "Optical inspection of nanoscale structures using a novel machine learning based synthetic image generation algorithm," *Opt. Express* **27**(13), 17743–17762 (2019).
19. C. Li, G. Gao, Z. Liu, M. Yu, and D. Huang, "Fabric Defect Detection Based on Biological Vision Modeling," *IEEE Access* **6**, 27659–27670 (2018).
20. K. Zhang, Y. Yan, P. Li, J. Jing, X. Liu, and Z. Wang, "Fabric Defect Detection Using Saliency Metric for Color Dissimilarity and Positional Aggregation," *IEEE Access* **6**, 49170–49181 (2018).
21. R. A. Lizarraga-Morales, F. E. Correa-Tome, R. E. Sanchez-Yanez, and J. Cepeda-Negrete, "On the Use of Binary Features in a Rule-Based Approach for Defect on Patterned Textiles," *IEEE Access* **7**, 18042–18049 (2019).
22. L. Tong, W. K. Wong, and C. K. Kwong, "Fabric Defect Detection for Apparel Industry: A Nonlocal Sparse Representation Approach," *IEEE Access* **5**, 5947–5964 (2017).
23. D. Soukup and R. Huber-Mörk, "Convolutional neural networks for steel surface defect detection from photometric stereo images," *Int. Symp. on Vis. Comput.* (2014).
24. M. Niskanen, O. Silvén, and H. Kauppinen, "Color and texture based wood inspection with non-supervised clustering," *Proc. scandinavian Conf. on image analysis* (2001).
25. S. Natarajan, M. Agostinelli, S. Akbar, M. Bost, A. Bowonder, V. Chikarmane, S. Chouksey, A. Dasgupta, K. Fischer, Q. Fu, T. Ghani, M. Giles, S. Govindaraju, R. Grover, W. Han, D. Hanken, E. Haralson, M. Haran, M. Heckscher, R. Heussner, P. Jain, R. James, R. Jhaveri, I. Jin, H. Kam, E. Karl, C. Kenyon, M. Liu, Y. Luo, and R. M. Khandru, "A 14nm logic technology featuring 2nd-generation FinFET, air-gapped interconnects, self-aligned double patterning and a 0.0588 μm^2 SRAM cell size," *Electron Devices Meeting (IEDM)* (2014).
26. A. Raghunathan, S. Bennett, H. O. Stamper, J. G. Hartley, A. Arceo, M. Johnson, C. Deeb, D. Patel, and J. Nadeau, "13nm gate Intentional Defect Array (IDA) wafer patterning by e-beam lithography for defect metrology evaluation," *Microelectron. Eng.* **88**(8), 2729–2731 (2011).
27. B. Barnes, "Geometries and material properties for simulating semiconductor patterned bridge defects using the finite-difference time-domain (FDTD) method," (2018), <https://doi.org/10.18434/T4/1500937>.
28. B. M. Barnes, M. Y. Sohn, F. Goasmat, H. Zhou, A. E. Vladár, R. M. Silver, and A. Arceo, "Three-dimensional deep sub-wavelength defect detection using $\lambda = 193$ nm optical microscopy," *Opt. Express* **21**(22), 26219–26226 (2013).
29. B. M. Barnes, F. Goasmat, M. Y. Sohn, H. Zhou, A. E. Vladár, and R. M. Silver, "Effects of wafer noise on the detection of 20-nm defects using optical volumetric inspection," *J. Micro/Nanolithography, MEMS, MOEMS* **14**(1), 014001 (2015).
30. B. M. Barnes, H. Zhou, M.-A. Henn, M. Y. Sohn, and R. M. Silver, "Optimizing image-based patterned defect inspection through FDTD simulations at multiple ultraviolet wavelengths," *Proc. SPIE* **10330**, 103300W (2017).
31. A. Taflov, "Application of the finite-difference time-domain method to sinusoidal steady-state electromagnetic-penetration problems," *IEEE Trans. Electromagn. Compat.* **EMC-22**(3), 191–202 (1980).
32. B. M. Barnes, R. Quinhanilha, Y.-J. Sohn, H. Zhou, and R. M. Silver, "Optical illumination optimization for patterned defect inspection," in *Metrology, Inspection, and Process Control for Microlithography XXV*, vol. 7971 (International Society for Optics and Photonics, 2011), p. 79710D.
33. R. Durrett, *Probability: Theory and Examples* (Cambridge Univ., 2010).
34. B. M. Barnes, H. Zhou, M.-A. Henn, M. Y. Sohn, and R. M. Silver, "Assessing the wavelength extensibility of optical patterned defect inspection," *Proc. SPIE* **10145**, 1014516 (2017).
35. B. M. Barnes, M.-A. Henn, M. Y. Sohn, H. Zhou, and R. M. Silver, "Assessing form-dependent optical scattering at vacuum- and extreme-ultraviolet wavelengths off nanostructures with two-dimensional periodicity," *Phys. Rev. Appl.* **11**(6), 064056 (2019).
36. M. Y. Sohn, B. M. Barnes, and R. M. Silver, "Design of angle-resolved illumination optics using nonimaging bi-telescopicity for 193 nm scatterfield microscopy," *Optik (Munich, Ger.)* **156**, 635–645 (2018).
37. A. Wojdyła, A. Donoghue, M. P. Benk, P. P. Naulleau, and K. A. Goldberg, "Aerial imaging study of the mask-induced line-width roughness of euv lithography masks," *Proc. SPIE* **9776**, 97760H (2016).
38. B. M. Barnes, F. Goasmat, M. Y. Sohn, H. Zhou, A. E. Vladár, and R. M. Silver, "Effects of wafer noise on the detection of 20-nm defects using optical volumetric inspection," *J. Micro/Nanolith. MEMS MOEMS* **14**(1), 014001 (2015).
39. M. Yoshizawa and S. Moriya, "Study of the acid-diffusion effect on line edge roughness using the edge roughness evaluation method," *J. Vac. Sci. Technol. B* **20**(4), 1342 (2002).
40. N. Kubota, T. Hayashi, T. Iwai, H. Komano, and A. Kawai, "Advanced resist design using AFM analysis for ArF lithography," *J. Photopolym. Sci. Technol.* **16**(3), 467–474 (2003).
41. S. C. Palmateer, S. G. Cann, J. E. Curtin, S. P. Doran, L. M. Eriksen, A. R. Forte, R. R. Kunz, T. M. Lyszczyk, M. B. Bern, and C. M. Nelson-Thomas, "Line-edge roughness in sub-0.18- μm resist patterns," *Proc. SPIE* **3333**, 634 (1998).

42. A. Yamaguchi and O. Komuro, "Characterization of line edge roughness in resist patterns by using fourier analysis and auto-correlation function," *Jpn. J. Appl. Phys.* **42**(Part 1), 3763–3770 (2003).
43. J. Croon, G. Storms, S. Winkelmeier, I. Pollentier, M. Ercken, S. Decoutere, W. Sansen, and H. Maes, "Line edge roughness: characterization, modeling and impact on device behavior," in *Electron Devices Meeting, 2002. IEDM'02. International*, (IEEE, 2002).
44. C. A. Mack, "Analytic form for the power spectral density in one, two, and three dimensions," *J. Micro/Nanolith. MEMS MOEMS* **10**(4), 040501 (2011).
45. C. A. Mack, "Generating random rough edges, surfaces, and volumes," *Appl. Opt.* **52**(7), 1472 (2013).
46. P. Oldiges, Q. Lin, K. Petrillo, M. Sanchez, M. Jeong, and M. Hargrove, "Modeling line edge roughness effects in sub 100 nanometer gate length devices," *Simulation of Semiconductor Processes and Devices* (2000).
47. H. Gross, S. Heidenreich, M.-A. Henn, G. Dai, F. Scholze, and M. Bär, "Modelling line edge roughness in periodic line-space structures by fourier optics to improve scatterometry," *J. Eur. Opt. Soc. Publications* **9**, 14003 (2014).
48. M.-A. Henn, S. Heidenreich, H. Gross, A. Rathsfeld, F. Scholze, and M. Bär, "Improved grating reconstruction by determination of line roughness in extreme ultraviolet scatterometry," *Opt. Lett.* **37**(24), 5229–5231 (2012).
49. A. Agresti, *Categorical Data Analysis* (John Wiley & Sons, 2003).
50. I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio, *Deep Learning* (MIT Press, 2016).
51. I.-S. Tsai, C.-H. Lin, and J.-J. Lin, "Applying an artificial neural network to pattern recognition in fabric defects," *Text. Res. J.* **65**(3), 123–130 (1995).
52. M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Trans. on Image Process.* **1**(2), 205–220 (1992).
53. M.-A. Henn, Z. Hui, R. M. Silver, and B. M. Barnes, "Applications of machine learning at the limits of form-dependent scattering for defect metrology," *Proc. SPIE* **10959**, 109590Z (2019).
54. Google Brain, "TensorFlow," (2015).