Machine Learning Hotspot Prediction Significantly Improve Capture Rate on Wafer

Wei Yuan
TD OPC
ICRD
Shanghai, China
yuanwei@icrd.com.cn

Yifei Lu
TD OPC
ICRD
Shanghai, China
luyifei@icrd.com.cn

Ming Li
Process Technologies Dept
ICRD
Shanghai, China
panbingyang@icrd.com.cn

Bingyang Pan
TD OPC
ICRD
Shanghai, China
panbingyang@icrd.com.cn

Ying Gao
TD OPC
ICRD
Shanghai, China
gaoying.@icrd.com.cn

Yu Tian
TD OPC
ICRD
Shanghai, China
tianyu@icrd.com.cn

Zhi-qin Li
Product Engineering Group
ASML
ShenZhen, China
zhi-qin.li@asml.com

Liang Ji
Product Engineering Group
ASML
ShenZhen, China
liang.ji@asml.com

Ying Huang
Product Engineering Group
ASML
ShenZhen, China
ying.huang@asml.com

Hao Chen
Product Engineering Group
ASML
ShenZhen, China
hao.chen@asml.com

Yueliang Yao
Product Engineering Group
ASML
San Jose, USA
yue-liang.yao@asml.com

Sean Park
Marketing
ASML
San Jose, USA
seung-hoon.park@asml.com

In a real mask tape-out (MTO) process, an end user would typically use simulation tools to capture hotspot candidates which are at risk of appearing on wafers. The tight turn-around-time(TAT) in a fab requires an efficient method to categorize these candidates and sampling before measurement. Traditionally, in order to capture hotspots, verification tools mainly focus on limited parameters such as contours, local image contrast and parameters extracted from the full aerial and resist information. This approach makes it difficult to quickly pinpoint high risk hotspots, especially when the hotspot count is large. In contrast, by using advanced machine learning techniques, Newron hotspot prediction is an innovative method that makes full use of whole simulated images to generate accurate prediction information for every hotspot candidate.

Newron hotspot prediction is able to significantly reduce the amount of required input information and improve the hotspot capture rate.

Keywords—Machine learning, hotspot prediction, image based

I. INTRODUCTION

In the mask tape-out flow, before handing over the post-OPC layout to the mask shop for manufacturing, verification tools are used to check the layout through different process conditions, e.g. through focus, dose, mask bias, and etc. Ideally this flow can capture most of the yield limiting hotspots with an accurate OPC model in mature nodes. As technology nodes continuous to shrink, stricter thresholds are needed to avoid missing killer hotspots. These strict thresholds results in many hotspot candidates including a large fraction of false alarms. Different ways are implemented to identify the real hotspots and reduce the false alarm during OPC stage: 1) Using different methods to construct KPI to judge whether a hotspot candidate is real or not; however traditional detectors contain intuitive physics that require the end user to empirically manipulate the criteria to highlight desired high risky hotspots. 2) Introducing more advanced and accurate models into OPC and its verification flow [1][2][3][4]. Despite such improvements, complicated interactions between design, OPC and lithography result in ever more hotspots are being inspected in a fab.

In this paper, an innovative method of hotspot prediction based on Tachyon Newron engine is introduced that aims to drastically increase the capture rate of real defects while decreasing the false alarms. Tachyon Newron hotspot prediction with an image based defect detection flow is constructed on Deep Convolutional Neural Network which is capable of fitting the majority physical and non-physical effects with sufficient data.

Newron hotspot is part of ASML Tachyon Newron product family which provide deep learning enhanced OPC solutions, including ADI/AEI OPC model calibration [1][2], SRAF generation [5], and OPC [6].

II. MACHINE LEARNING HOTSPOT PREDICTION

Supervised machine learning algorithms in TensorFlow are employed in Newron hotspot prediction. The training flow is illustrated in three steps below.

1) Generate and crop resist images of the hotspots using Tachyon model; and convert the images to latency space vectors through a CNN networks.

2) Prepare training data with wide coverage in litho and design spaces; and label hotspots ('risky' or 'safe') based on the SEM-verified ground truth.

978-1-7281-7577-5/20/$31.00 ©20XX IEEE

3) Feed training data into desitn and behavior space prediction models of the Newron engine.

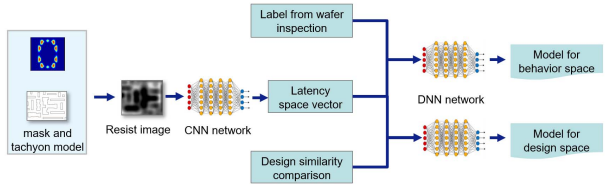

Fig. 1. Training flow of Newron hotspot models

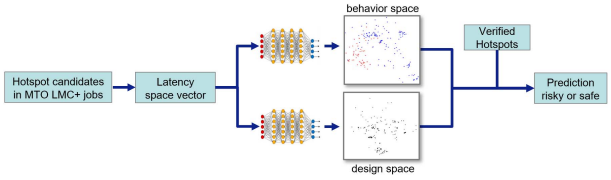

Fig. 2. The application flow of Newron hotspot prediction

When Newron hotspot prediction models are applied to new hotspot candidates, resist images are generated around hotspots and then converted to latency space vectors in the same fashion as described the in the training flow above. Once the latency space vectors are ready, those vectors are projected to new feature spaces by prediction models.
In design space, design similarities are gauged based on, the Euclidean distance between input each hotspot candidates and the hotspots in the a training set while the severity judgement, 'risky' or 'safe', is made in behavior space. indicates In respect to the relative distribution of these two feature spaces, new hotspot candidates are labeled in are labeled in the final prediction.

III. Prediction Result Analysis

In this study, Newron hotspot prediction was explored in two aspects:

1) Predict lithography defects with simulation data.
2) Predict lithography model error with simulation data.

A. Ground truth labeling based on real wafer data

ADI wafer data collected from ICRD 14nm M1 NTD lithography process was employed in this paper. 375 unique hotspot candidates are selected for the CD-SEM data collection on 7 through-focus conditions. Each hotspot was measured with 10 repetitions at different locations. It resulted in 26,250 measurements (7x10x375) to build the ground truth labels based on two KPIs.

Firstly, CD variation in PW conditions is used for KPI #1. It indicates that the hotspots with larger CD variation are riskier. 125 candidates from the top cd variation are marked as "Risky" while the 250 hotspot candidates with smallest CD range are categorized to "Safe". This is the ground truth label (GT #1) to be assessed with Newron hotspot.

Secondly, KPI #2 is rooted from lithography model errors. It is calculated based on the delta between average wafer CDs and simulation CDs of each at nominal condition. The bottom line is that the hotspots with bigger model error is riskier. In the similar manner with KPI #1, the top 125 and bottom 250 of the biggest model errors are marked as "Risky" and "Safe", respectively. This is added to the ground truth label as GT #2.

B. Baseline performance

Computational lithography verification tools for instance LMC+, have been widely used for addressing wafer hotspot detection in the early phase of the technology node development. It inevitably requires high model accuracy. Tachyon FEM+ has been well received in the major chip makers when it comes to accuracy and stability. As shown in Table 1, the model accuracy is meeting the specifications for 14nm technology node by some margin. Based on the results simulated with this ADI model, several legacy detectors are examined to predict GT #1 and GT #2.

TABLE I. Conventional Model Accuracy

Type	Spec	In Spec ratio	Overall RMS	Anchor error
1D	3nm	96.7%	1.05	
2D	5nm	99.8%	1.4	
Total			1.24	-0.09

For GT #1, 6 LMC+ detectors are used to capture the worst 125 hotspots.

1) *Process variation band(PVB)*
2) *Nominal contour CD*
3) *Image log slope(ILS)*
4) *Normalized image log slope(NILS)*
5) *Aerial image intensity(AI)*
6) *Mask error enhancement factor(MEEF)*

Take a PVB detector as an example, the top 125 hotspots with bigger process variation band are predicted as "Risky" by PVB detector while the rest 250 hotspots are put in the "Safe" bucket. As key performance indicators of prediction, the nuisance and capture rates can be drawn as shown in Table 2 to evaluate the accuracy of the baseline detectors. Also F1 score is constructed to evaluate the overall prediction accuracy.

$$Nuisance\ Rate(NR) = \frac{FP}{TP + FP}$$

$$Capture\ Rate(CR) = \frac{TP}{TP + FN}$$

$$F1 = \frac{2}{(1 - NR)^{-1} + CR^{-1}}$$

TABLE II. Confusion Matrix Definition

		Prediction	
		Risky	Safe
Ground Truth	Risky	True Positive(TP)	False Negative(FN)
	Safe	False Positive(FP)	True Negative(TN)

Among the detectors in the experiment, PVB is the best performing one to predict GT #1 (Fig 3). Ii is somehow expected since the PVB detector measures the process variation which is in line with the physical meaning of KPI #1. From the correlation between PVB defect size and wafer CD variation of KPI #1 (Fig 4), some points were drifted away from a certain linear correlation due to the model error.

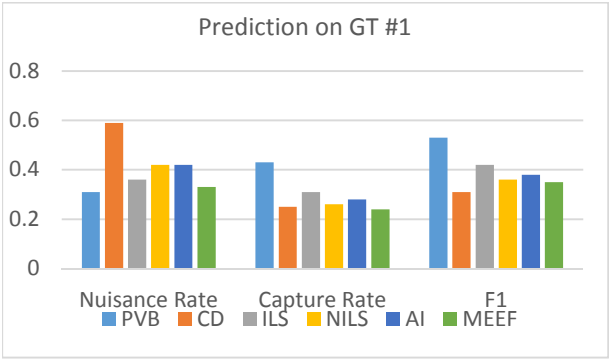

Fig. 3. The results of predicting GT #1 with baseline detectors

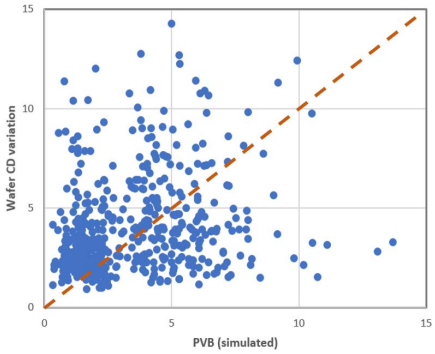

Fig. 4. Correlation between PVB defect size and wafer CD Variation

For GT #2, traditional detectors do not have the capability to predict the model error because all the detectors use the information simulated by the model itself.

C. *Newron Hotspot Prediction performance*

A huge amount of simulation data are fed into the design space model for the training. The purpose of this training is to acquire a distribution for quantitatively evaluating the design similarity between any two hotspot candidates. The training data is manipulated in a way that the central area of each resist image contributes more to the design similarity than the margin area.

The design space model training do not require wafer verification data input, thus the overfitting issue can be largely suppressed by increasing a training data amount. By using transfer learning skills, the design space model result can be used to reduce over-fitting issues in the behavior space model training. Fig 5.f and Fig 5.a-5.e show the PCA(Principle Component Analysis) visualization of design space of full 375 hotspot candidates and resist image example of 5 HS, respectively. Design space is constructed with consideration of different weighting in the resist images. Center region (orange box) has higher weight than outer region. In these 5 hotspot examples, Euclidean distance in design space between (a) and other 4 hotspot candidates are calculated and ranked from small to large as (b), (c), (d), and (e). These examples show the smaller the Euclidean distance, the higher the similarity.

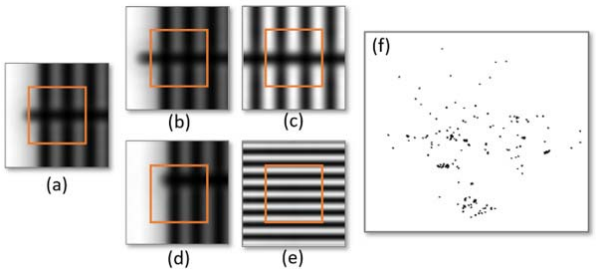

Fig. 5. (a)-(e) Resist images of five hotspot candidates. (f) PCA visualization of design space generated from full data set of 375 hotspot candidates

In the behavior space model training, the 375 hotspot candidates were randomly separated, 50% to training set and the rest 50% to the test set. Fig 6 shows the behavior space distribution of the training and test sets. The separations between the risky and safe hotspots are clear in the training set. Although some safe hotspots are mixed with risky hotspots in the test set, the overall distribution is matched well to the training set and exhibits good generalization ability.

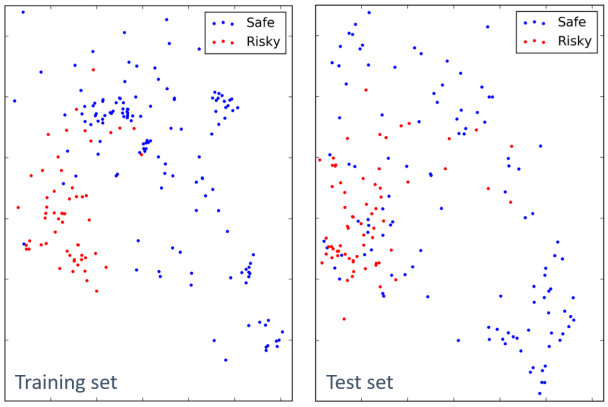

Fig. 6. PCA visualization of hotspot distribution in behavior space

D. *Comparison between Newron Hotspot Prediction and baseline*

Fig. 7. shows the comparison between traditional detectors and Newron hotspot prediction results. For KPI #1, PVband-focus has the best prediction accuracy among the traditional detectors due to its physical modeling; however, Newron hotspot prediction shows even better prediction accuracy with 47% higher in the F1 score. The capture rate is above 90%, which matches with the behavior space distribution.

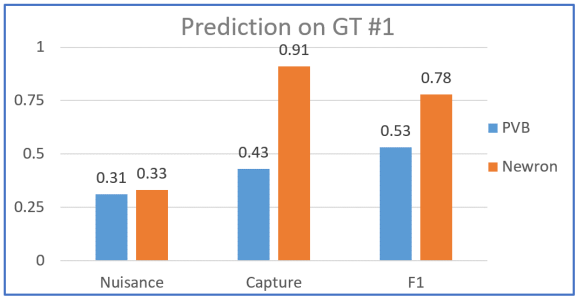

Fig. 7. Histogram of prediction accuracy of baseline detector and Newron Hotspot for KPI #1

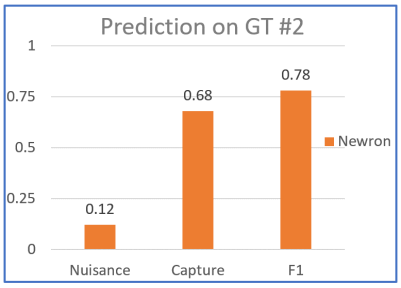

Fig. 8. Histogram of prediction accuracy of Newron Hotspot for KPI #2

Fig 8 shows the Newron hotspot prediction accuracy on KPI #2. The high F1 score indicates the Newron hotspot solution can also effectively distinguish hotspots marked by non-physical effects. This is a big breakthrough to decouple the prediction power from the lithography model error.

IV. Conclusion

The conventional hotspot prediction capability is highly bound to the model accuracy; and the new model calibration cadence is very long and expensive to embrace both OPC on-target and hotspot detection at the same time. To address this difficulties, Newron hotspot flow was demonstrated its superior capability distinguishing risky and safe hotspot candidates by complementing the traditional computational lithography verification deficiency. The trained model shows good generalization ability despite the limitation on the wafer verification data; however it was clearly shown that the design space training and transfer learning technique can benefit hotspot prediction on both physical and non-physical effects.

References

[1] Wei Yuan, Yifei Lu, Yuhang Zhao, et. al. "Metrology and deep learning integrated solution to drive OPC model accuracy improvement", Proc. SPIE 10961, Optical Microlithography XXXII, 109610N (17 May 2019)

[2] Yifei Lu, Yuhang Zhao, Ming Li, Wei Yuan, et. al. "Accurate etch modeling with massive metrology and deep-learning technology" , Proc. SPIE 11327, Optical Microlithography XXXIII, 113270B (23 March 2020)

[3] Woojoo Sim, Kibok Lee, Dingdong Yang, et. al. “Automatic correction of lithography hotspots with a deep generative model”, Proc. SPIE 10961, Optical Microlithography XXXII, 1096105 (20 Mar 2019)

[4] Bei Yu, Jhih-Rong Gao, Duo Ding et. al. “Accurate lithography hotspot detection based on principal component analysis-support vector machine classifier with hierarchical data clustering”, J. Micro/Nanolith. MEMS MOEMS 14(1) 011003 (4 Nov 2014)

[5] Hu Hongmei, Wei Yuan, Yifei Lu, et. al. "Machine Learning SRAF improves OPC Performance at 1X node and below" , SEMICON China, Symposium II (2019)

[6] Xuelong Shi, YuHang Zhao, Shoumian Cheng, Ming Li, Wei Yuan, et. al. "Optimal feature vector design for computational lithography", Proc. SPIE 10961, Optical Microlithography XXXII, 109610O (20 March 2019)