Enhanced Defect Detection in After Develop Inspection with Machine Learning Disposition

Matthew P. McLaughlin, Andrew Stamper, Gabriel Barber, Janice Paduano, Petra Mennell, Emerson Benn, Michael Linnane, Justin Zwick, Chetan Khatumria, Robert L. Isaacson, Nathan Hoffman, Clayton Menser

GLOBALFOUNDRIES
Hopewell Junction, NY
12533

Matthew.McLaughlin@Globalfoundries.com

2021 32nd Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC) | 978-1-7281-8645-0/20/$31.00 ©2021 IEEE | DOI: 10.1109/ASMC51741.2021.9435721

***Abstract*— A complementary Machine Learning disposition method was generated and tested for after develop inspections in lithography. For lithography coating defects, this new method showed twice the sensitivity and five times the specificity in a controlled experiment versus the baseline system. Applying the detection method along with process improvements, preventative measures and rework for splatter defects, reduced yield loss from splatters by over 30x. Herein we describe learnings on the use of image enhancement for training and disposition, an Explainable AI system to support understanding, and a process flow to train augmentation based on performance.**

Keywords— Machine Learning, Semiconductor, Smart Manufacturing, ADI, Lithography

I. Introduction

After Develop Inspection (ADI) is used in semiconductor manufacturing as an inline control for the lithography process.[1] The control allows the detection of both coating and incoming defects, where most coating defects caused during the lithography process can be removed through rework to improve yield. At the cost of resolution, most macro ADI systems can image each wafer without limiting tool speed. Where the majority of macro ADI systems use visible light for bright and dark field imaging with resolutions of 30 microns and above. Existing methods for macro ADI defect detection are known to have low sensitivity for certain coating defects.[2] This is due in part because nuisance defects from normal process variation can cause a similar amount of total color variation to real defects (e.g. incoming CMP, test sites, ADI image stitching, edge topography, etc.). And so the recipe settings for color difference and affected area are insufficient for many processes. And despite progress in methods to distinguish critical and nuisance defects, including the use of Global defect libraries, sensitivity and selectivity for color streaks and resist splatters remain a noted issue (Fig. 1 and 2). Therefore the ADI area there is an interest in improved detection methods for faint resist splatters and other global coating defects with large yield impacts. One method for improving ADI performance normalizes image color by tool intensity to create a more representative image set for model training and dispositioning.[3] For the defect model itself, commercial approaches also allow the model to be trained and adjusted to optimize sensitivity and selectivity. But even with a commercial trained model, detection of faint color defects requires recipes to be sensitized to a very high false fail rate (e.g. worst recipes >50%). Where improving the sensitivity and selectivity for the subtlest defects requires a system capable of distinguishing normal process variation and defects.

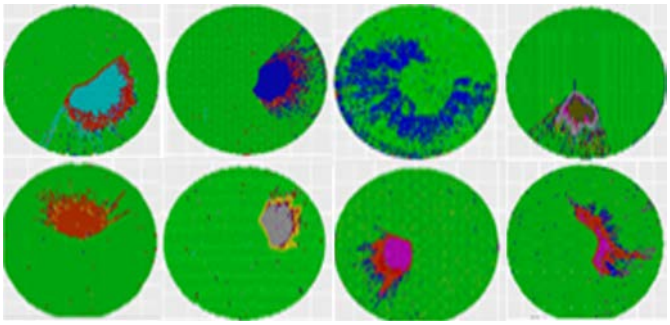

Fig. 1. Wafer test maps showing resist splatter yield impact when undetected at ADI: green color = pass, other colors are fails. The overall shape of these fails is common for splatters.at ADI.

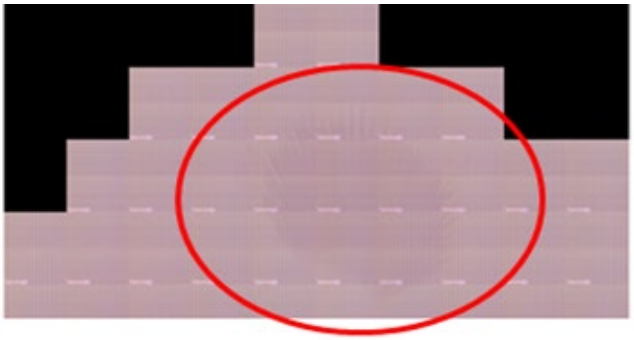

Fig. 2. Representative faint resist splatter at ADI circled in red, which correlates with yield loss. These defects are challenging to see and detect with ADI systems (see Fig. 3 for defect image enhancement).

Herein we describe the evaluation of a complementary Machine Learning (ML) method for detecting splatter and color defects in ADI images. We detail the use of image enhancement through equalization to support ML model training, and dispositioning; the use of an explainable artificial intelligence (AI) approach to understand our model better; and a process flow ML training approach to systematically improve our models. Taken together, the complimentary ML system showed substantially higher sensitivity and selectivity for coating defects, which supported work to improve yield.

978-1-7281-8645-0/20/$31.00 ©2021 IEEE ASMC 2021

II. Discussion and Methods

To improve sensitivity and selectivity of ADI for certain coating defects, we pursued the evaluation of complementary ML models. We explored models in order to evaluate the ADI images for defects in parallel to the ADI module. From this work, a complementary automated disposition method was generated and tested for ADI. During this work, methods in image enhancement, explainable AI, and systematic ML model augmentation were useful supplementary tools for training.

To support ML model creation and evaluation of its predictions, a command line batch image processor was used to automate the enhancement of images. In particular, a process

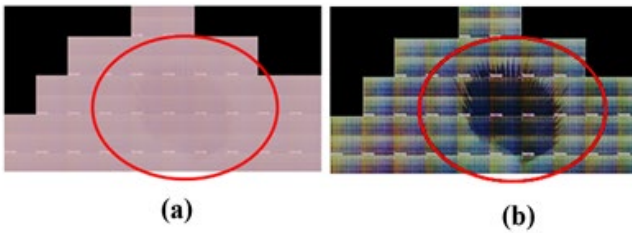

Fig. 3. Shows the faint splatter defect from introduction (a) that is more easily seen with the equalized images used to enhance dispositioning (b).

was evaluated to equalize images, which provided a consistently superior visualization of defects for the human eye (Fig. 3). The enhanced images showed direct benefits, improving the accuracy and speed of dispositioning but also have anticipated indirect benefits by reducing mistakes from eye fatigue.[4]

These image equalizations provided superior ML training on defects that were hardest to see, limiting defect misclassification, which is known to undermine both the sensitivity and selectivity of models. The enhanced images also supported rapid and confident dispositioning. For instance, the creation of the model described in the abstract was created after 6 generations of ML model refinement, which required manual review of hundreds of thousands of images.

While refining each generation of the model, at times it was hard to understand why some defects were readily detected with an ML model and some were detected at a lower rate (see Fig. 4 and Table I).

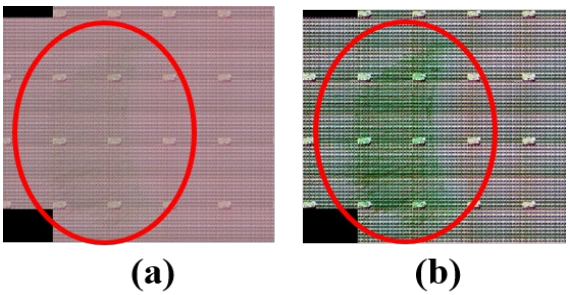

Fig. 4. Shows the faint splatter not detected with early ML models (a) that is more easily seen with the equalized images used to enhance dispositioning (b). The failure to detect this defect with ML contrasts with a similar defect in Fig. 3, which was detected with ML with very high confidence.

To enhance our understanding of the early ML models, which only provide a label prediction without explanation, we

Table I. Confusion matrix for a preliminary ML model with the 4 labels used. Where BAD = coating defect; GOOD = no defect; NI = no image available; WP = ADI alignment error. The ML predicted label is shown in red and the true label assigned by an engineer is shown in green. The correct assignments by ML are filled with blue. ML false positives and false negatives are filled with grey, highlighting weak areas for the model.

True/Predicted	BAD	GOOD	NOIMAGE	WP
BAD	78%	22%	-	-
GOOD	1%	99%	-	-
NI	-	-	100%	-
WP	-	-	-	100%

explored the use of an Explainable AI model to try to understand the label predictions. The AI model provided maps of where in each image was most correlated with a possible defect (Fig. 5). Using the model heat maps for defectivity, we found that many of the predicted defects, were located near the edge of our wafer map, even on the black border were no defects were located. By using the AI model to visualize the label predictions to try to identify ML model weaknesses, our attention was drawn to the dimension of our ADI images and how the ML system managed those dimensions.

Investigating the image dimensions revealed the ADI delivered a wide range of image dimensions (with many as high

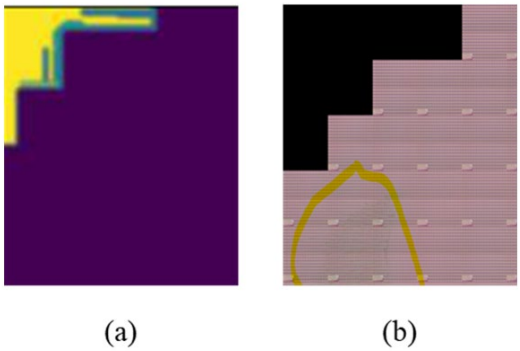

Fig. 5. (a) Explainable AI "heat map" showing the top left as having a defect (yellow) rather than in the middle left where the defect is actually located, circled in yellow (b). In other words, the preliminary models showed confusion on defect location from training set.

Table II. Representative metrics for preliminary ML model shown in Table I.

Total Images	7,787
Test Items	869
Precision	96.1%
Recall	93.3%

as 4,256 but most below 2,240 pixels); and that the ML system used crops images to a max x or y dimension of 2,240 pixels. This cropping was a problem because images in the ML training set, or sent to the model for a prediction, could have the defect removed from the image if the dimension was too large and/or the defect was close to the edge. To limit undesirable cropping of edge defects, which results in mislabeling and model

confusion, we used a command line batch image processor to proportionally scale the dimension of all images to max x/y dimensions of 2,240 pixels. This one change substantially

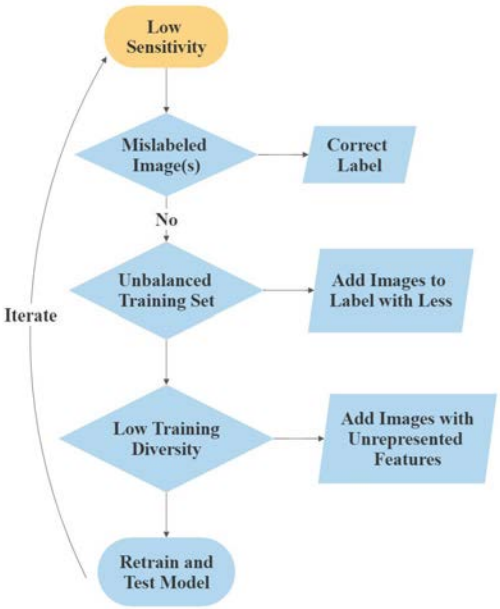

Fig. 6. Process flow diagram summarizing iterative training method to improve sensitivity for label features.

improved the sensitivity and selectivity of the model, while enhancing the models ability to detect edge defects, which were being cropped in early models.

To ensure the continuous improvement of our model sensitivity and specificity for defects with each of the 6 generations of the model, we compared each model's performance against hundreds of thousands of images (Table II). If any of the ML metrics were lower than expected, the model could be redone using different training or test images.

Through an iterative process of model refinement, a number of practices were applied to optimize our ML model sensitivity and selectivity for defects, which are summarized in a process flow diagram (Fig. 6). For this application, sensitivity and selectivity optimizations focused on using performance to guide image augmentation, identification of inaccurate labels, and identifying the best labels to use. The best practices used have been well documented elsewhere, and so are only described briefly in the next 2 paragraphs.[5]

In order to improve defect detection in each of the 6 generations of the ML model we sought to optimize sensitivity for the coating defects of interest. The practices to optimize ML model sensitivity for defects outlined in Fig. 6 focused on balancing the training set (i.e. using a similar quantity of images for each label) and identifying mislabeling in the training set. Since defect examples were limiting, the sensitivity of the ML model was optimized by using as many defect examples as could be identified in the training set. Because we could not identify enough defect examples for a balanced training set, the model was augmented with additional images at designs showing low sensitivity for defects. These augmented images were made with image processing software, where small variations in the image were made. Additionally, if low sensitivity for a defect was

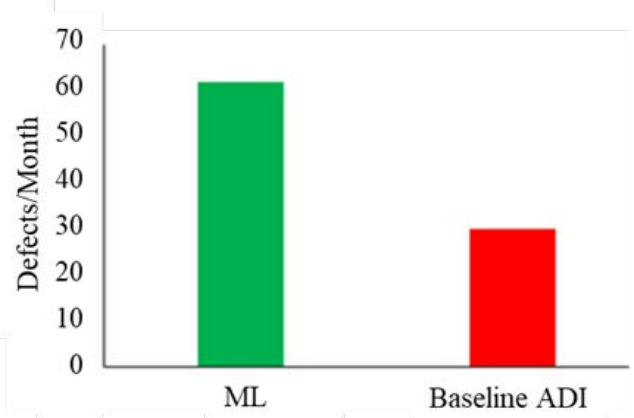

Fig. 7. Wafers detected comparing baseline and ML enhanced method in a controlled experiment. Shows the coating defects impacting 3 or more fields per month.

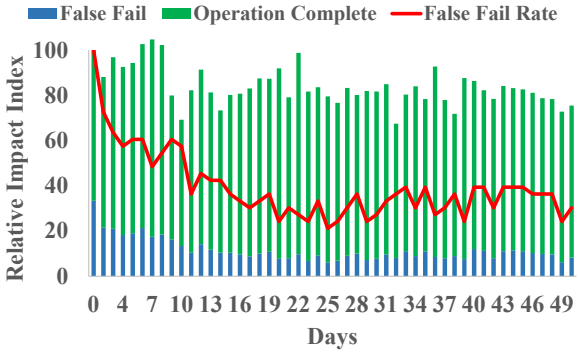

Fig. 8. Trend of ADI related false fails expressed as a Relative Impact Index, shown in arbitrary relative units. A 77% reduction in false fails was achieved as the system defect selectivity improved.

identified, enhanced images of wafers labeled "good" in the training set were reviewed for any mislabeling. For example, Fig. 3 above showed a subtle defect that could be mislabeled without careful review of enhanced images. Also of note, on two occasions an ML model validation and test set weres chosen that did not represent all of the important defect features adequately. This caused the models to converge with low sensitivity for the underrepresented defect. To improve these models, the test examples were reselected with one per type of defect represented.

We also sought to improve selectivity for the defects to reduce false fails. This required the model to have good sensitivity for design images without any defects. Through the iterative process of model refinement, the method used to optimize sensitivity for defects were also applied to improve sensitivity for wafers without defects (Fig. 6). Some more specific guiding principles were identified to maximize defect selectivity. For instance, selectivity was improved substantially by including images with the GOOD label (i.e. without any defects) for each design type. A clear pattern emerged, where unrepresented designs for the GOOD label received a higher false fails rate than designs with representation. Some designs

had more variation in the GOOD label, which resulted in more false fails. In cases where a represented design had persistently high false fails, additional images from the false fails were added to the training set to diversify the model. And in instances where chronic false fails were observed for a design, defect training sets with similar designs were inspected to identify if any wafers were mislabeled BAD (i.e. good wafers misclassified as having a defect).

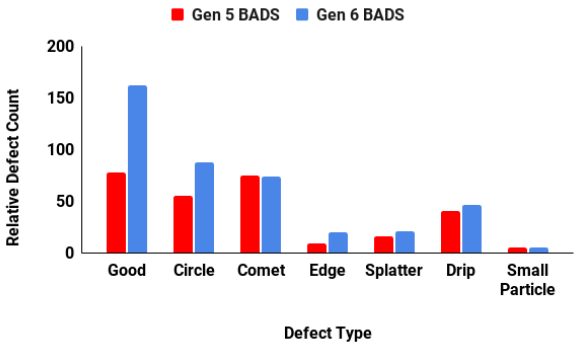

Fig. 9. BAD label assignments of ML Generation (Gen) 5 and 6 models compared for the same image set in an experiment. The higher defect sensitivity of Gen 6 is apparent from greater number of true defects detected but there were also more false fails (i.e. GOOD wafers mislabeled BAD).

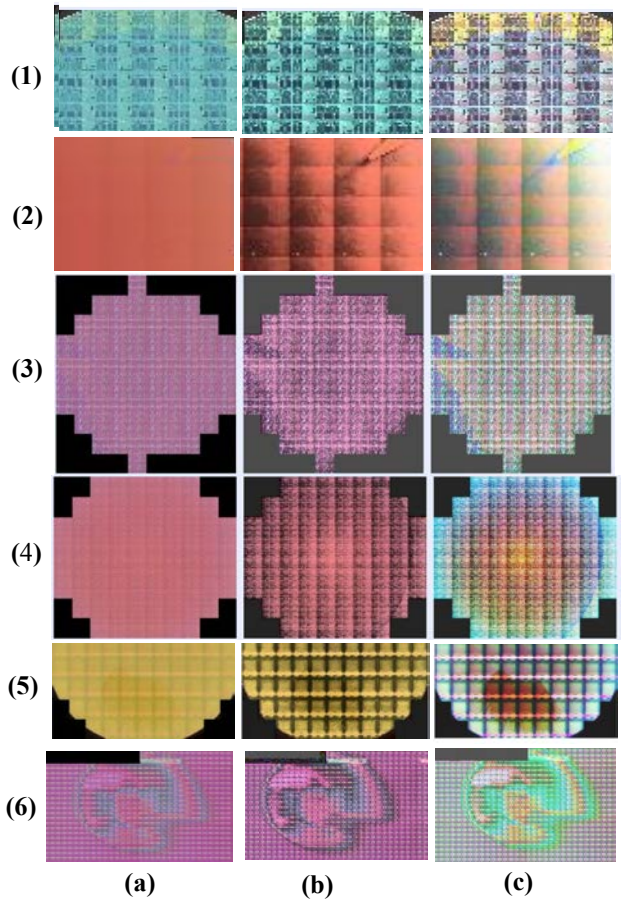

Fig. 10. Representative images of coating defects shown as original image (a), HSB equalized (b), and RGB equalized (c). The defects shown include short shot (1), comet (2), edge (3), circle (4), splatter (5), and drip (6). Enhancing the image with HSB and RGB equalizations help to see the defect.

To summarize, the iterative training of a complementary ML model for coating defects was successful. Recognized ML best practices,[5] large training sets, and past learnings from automated use of ADI for defect detection all supported our success.[6-8] And although there are other examples of ML used to identify defects in the lithography area,[8-10] this work uniquely used image enhancement, Explainable AI, and an ADI defect library for defect model training.

III. Results and Conclusions

After six generations of ML model refinement, the complimentary detection was compared against a commercial ADI system, showing over twice the sensitivity and five times the selectivity for global coating defects (Fig. 7 and 8). Each generation of the model was compared to the previous generation for a number of defect types, where sensitivity and selectivity were both evaluated (Fig. 9). As described in the methods, image enhancement was used to support better visualization of defects. Hue-Saturation-Brightness (HSB) and Red-Green-Blue (RGB) equalized images most consistently improved defect detection by a manual inspection (Fig. 10).

During the course of this study, increasingly sensitive ML models were used to detect defects for rework. The improving ML model supported and coincided with an increase in preventative maintenance and process improvements, which over time reduced the incidence of coating defects (Fig. 11). Depending on the fail, the maintenance and process activities were varied and numerous. Including targeted routine actions, like the cleaning and replacement of failing nozzles. And also process changes, including a switch to a walk-off rinse process. Consistent with the reduction in defects, an approximate 30x reduction in yield loss from coating defects was observed over the course of this work (Fig. 12).

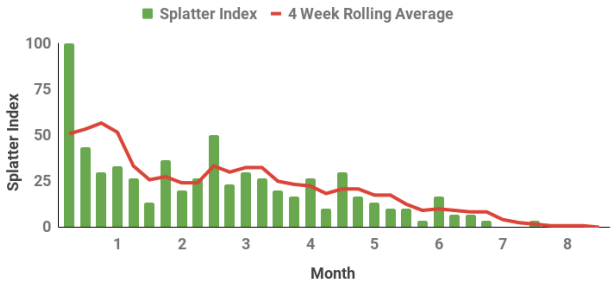

Fig. 11. Chart showing improvements in an inline Splatter Index over time. The Splatter Index is a description of the number of wafers dispositioned to have a splatter at ADI (arbitrary units). The time is shown as the week and calendar year. Note, the largest spikes correlated with the increase in defect sensitivity provided by the ML system, not an increase in splatters.

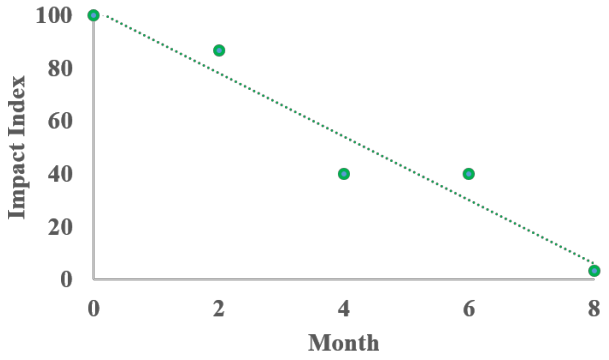

Fig. 12. Chart showing improvements in an Impact Index over time. The Impact Index is a description of the number of wafers determined to have yield loss from splatter incidents (arbitrary units). The yield impact is shown versus months of project actions.

REFERENCES

[1] T. L. Taylor, P. Shirley, D. Dixon, S. Yanagi, E. Makimura, "Product and tool control using integrated auto macro defect inspection in the photolithography cluster," SPIE, vol. 7638, pp. 763821-763829, Apr. 2010. DOI. 10.1117/12.

[2] M. Darwin, P. Kinikoglu, Y. Liu, K. Darwin, J. Clerico, "Automatic classification of microlithography macrodefects using a knowledge-based system," Proc. SPIE 5752, Metrology, Inspection, and Process Control for Microlithography XIX, (10 May 2005); doi: 10.1117/12.599660

[3] D. Dixon, L. Lee, "Method and system to compensate for lamp intensity differences in a photolithographic inspection," United States Patent US7359545B2, Dec. 31, 2003.

[4] M. J. Wang and C. L. Huang, "Evaluating the eye fatigue problem in wafer inspection," IEEE Trans Semi Mfg, vol. 17, no. 3, pp. 444-447, Aug. 2004, DOI. 10.1109/TSM.2004.831943.

[5] M. Zinkevich, M. (2017). "Rules of Machine Learning: Best Practices for ML Engineering. http://martin.zinkevich.org/rules of ml/rules of ml.pdf.

[6] V. C. Menon, R. L. Isaacson, M. C. Nicholls, S. J. Lickteig, T. Forstner, A. R. Barnett, J. Mulhall, "Product and tool control using integrated auto-macro defect inspection in the photolithography cluster," Proc. SPIE, vol. 6152, pp. 61521R1-61521R9, Mar. 2006, DOI. 10.1117/12.656394

[7] A. Berezin, D. Schaffer and A. Villarreal, "Automating after develop inspection," IEEE/SEMI 1996 Advanced Semiconductor Manufacturing Conference and Workshop. Theme-Innovative Approaches to Growth in the Semiconductor Industry. ASMC 96 Proceedings, Cambridge, MA, USA, 1996, pp. 103-106, DOI: 10.1109/ASMC.1996.557981.

[8] N. Rao, P. Kinney, A. Gupta, "High throughput wafer defect monitor for integrated metrology applications in photolithography", Proc. SPIE 6922, Metrology, Inspection, and Process Control for Microlithography XXII, 69223B (25 March 2008); DOI: 10.1117/12.772384

[9] S. Purandare, J. Zhu, R. Zhou, G. Popescu, A. Schwing, L. L. Goddard, "Optical inspection of nanoscale structures using a novel machine learning based synthetic image generation algorithm," Opt. Express, vol. 27, no. 13, pp. 17743-17760, Jun 2019, DOI. 10.1364/OE.27.017743

[10] S.H. Lee, A. Jain, M. Plihal, S. Paramasivam, T.-K. Ng, E. Soltanmohammadi, I. Tolle, D. Salvador "Machine Learning Approaches for Nuisance filtering in Inline Defect Inspection," 2019 30th Annual SEMI Advanced Semiconductor Manufacturing Conference (ASMC), Saratoga Springs, NY, USA, 2019, pp. 1-4, doi: 10.1109/ASMC.2019.8791805.