A Generative Approach for SEM Images Towards Advanced Node Defect Inspection in Semiconductor Manufacturing

Vic De Ridder[1,2,*], Bappaditya Dey[2,*], Sandip Halder[2]
, Victor Blanco[2], Bartel Van Waeyenberge[1]
[1] Ghent University, 9000 Ghent, Belgium
[2] Interuniversity Microelectronics Centre (imec), 3001 Leuven, Belgium
[*] Equal Contribution

I. Introduction

As Moore's Law drives the semiconductor industry towards achieving ever-smaller feature sizes ($\leq$10nm) and increased transistor density, the traditional methods of patterning face challenges. New approaches, including emerging lithography technologies like Extreme-Ultra-Violet-Lithography (EUVL) ($\leq$7nm), high-NA EUVL ($\leq$2nm), and other alternatives, are (being) developed to keep pace with Moore's Law and maintain the relentless pursuit of smaller feature sizes. The escalating complexity of semiconductor devices necessitates a corresponding elevation in process control. This entails the integration of precise metrology, sophisticated data analysis, and cutting-edge defect inspection methodologies. The prevailing state-of-the-art (SOTA) defect detection tools, whether optical or e-beam based, exhibit specific limitations. These tools rely on rule-based techniques for defect classification and detection, which introduces constraints in their adaptability and effectiveness. The use of rule-based approaches implies that these tools are programmed with predefined criteria to identify and classify defects. While this methodology is effective for well-understood and predictable defect patterns, it becomes increasingly challenging when dealing with complex, evolving, or stochastic defects [1], specifically in the presence of reduced signal-to-noise ratio (SNR) and image contrast.

Due to the inadequacy of rule-based methods at advanced nodes [2], DL-based object detectors have emerged as the state-of-the-art for stochastic defect inspection [3]. However, the acquisition of a relevant stochastic defect dataset for training ML models faces considerable challenges within the semiconductor manufacturing domain. Not only is such a dataset rare and inherently noisy, but its acquisition is also a costly endeavor. The rarity of stochastic defect instances makes it challenging to compile a comprehensive dataset that accurately represents the diverse range of stochastic defects encountered in real-world semiconductor manufacturing processes. Additionally, two significant bottlenecks further complicate the use of stochastic defect datasets in semiconductor manufacturing defect detection: (a) class imbalance, which arises when certain defect types are underrepresented or occur infrequently in the dataset, leading to a skewed distribution. This imbalance can compromise the model's ability to generalize and accurately detect defects across all classes. (b) insufficient dataset size, as a limited amount of data may not adequately capture the variability and complexity of stochastic defects. The inherent diversity of semiconductor manufacturing processes demands large and representative datasets to ensure the robust training of machine learning models. Addressing these challenges requires innovative approaches to dataset acquisition, including strategic data augmentation techniques to enhance dataset diversity. Collaboration within the industry and the development of shared datasets could also contribute to mitigating the issues associated with rare, noisy, and expensive stochastic defect datasets. Overcoming these challenges is pivotal for advancing the capabilities of machine learning models in semiconductor manufacturing defect detection. In this research work, we use Denoising Diffusion Probabilistic Models (DDPM) to generate realistic semiconductor wafer SEM images, thereby increasing defect inspection training data and improving defect inspection performance. Our main contributions are:

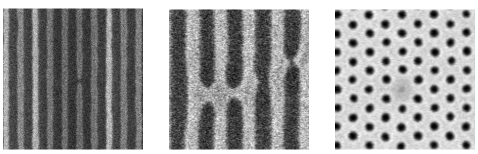


Fig. 1: Representative sample SEM images illustrating example defect types in the datasets used in this study

i) we propose a patch-based generative framework utilizing DDPM to generate SEM images that include intended defect classes with randomly variable instances, aiming to address class-imbalance and data insufficiency bottlenecks. This approach leads to an enhancement in defect detection performance, particularly in terms of precision and recall.

ii) our proposed approach generates synthetic images that closely resemble real ones, preserving actual characteristics without the need for prior knowledge of imaging settings (Best-Known-Methods).

iii) we demonstrated that a defect detector trained on a generated defect dataset, either independently or in combination with a limited real dataset, can achieve a similar or improved mAP on real wafer SEM images during validation/testing compared to when trained exclusively on a real defect dataset. This trend was consistent across three different SEM datasets, validating the capability of DDPM to generate images with characteristics identical to real SEM images.

Finally, iv) our proposed approach demonstrates the capability to transfer defect types, critical dimensions, and imaging conditions from one specified CD/Pitch and metrology specifications to another CD/Pitch and metrology specifications.

II. Methodology

A. Proposed Diffusion-based Approach

Due to their success in numerous other applications, and flexible usage, we investigate the potential of DDPM as a generative tool to solve the problem of low data availability in semiconductor defect inspection application. Our proposed approach does not train the diffusion model on the real SEM wafer images directly. Instead, various small patches are extracted from the original image. Each patch has a class label as either the defect type present inside the patch, or background (defect-free). The DDPM model is then trained in a class conditional manner on these patches. Fig.2 depicts our proposed framework towards generating synthetic realistic SEM image dataset containing multiple defect types.

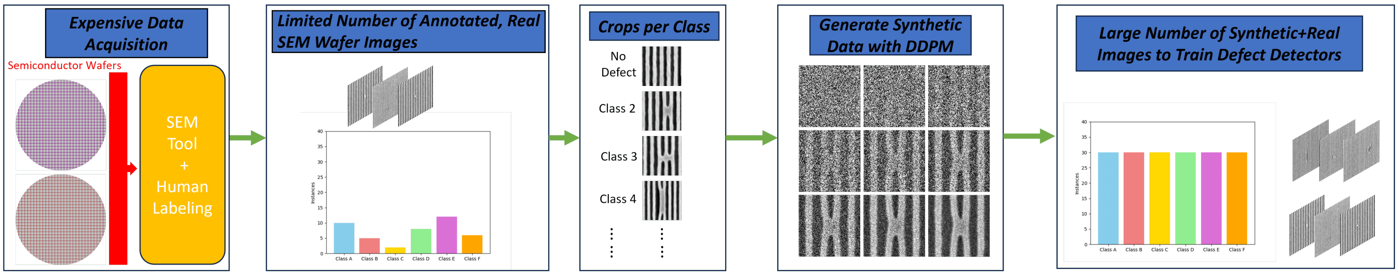


Fig. 2: High level overview of the proposed approach towards synthetic SEM image dataset generation containing multi defect types.

After training, synthetic images are generated using an inpainting procedure. First, the method displayed in Fig.3 is used to generate full-size, defect-free, synthetic images. Afterwards, crops of these full-size images are inpainted to simulate intended defect types, resulting in the final synthetic images containing defects.

This patch-based approach offers three advantages over training directly on full-size SEM images: i) significantly reduced training time. ii) control over the number of defects in the generated images, enabling the generation of full-scale images with defect numbers not present in the real SEM dataset. iii) Training on patches results in larger datasets, thereby enhancing the learning process.

B. Datasets

In this research work, we validate our proposed approach on three semiconductor SEM datasets: Hexagonal Contact-Hole arrays (HEXCH-DSA), Line-Space After Develop Inspection (LS-ADI), and Line-Space After Etch Inspection (LS-AEI). Each of these contains only real SEM wafer images, and no defects are synthetic or intentionally placed. The different defect types for each dataset are: i) partially closed hole, missing hole, and closed patch for HEXCH-DSA, ii) gap, probable gap, bridge, microbridge, linecollapse for LS-ADI, and finally iii) thin bridge, single bridge, line collapse, multi brige horizontal, and multi bridge non-horizontal for LS-AEI.

C. Diffusion Model: Implementation and Training

The diffusion model used in this research work is as implemented by Nichol et al. [4], with cosine noise schedule and 1000 sample steps. On each real SEM dataset, the model is trained using a learning rate of 0.0001 until convergence. We have added the inpainting functionality to the existing code, with implementation inspired by [5]. We have trained the models on a default image size of 128 pixels. However, some defect types such as linecollapse or closed patch exceed this limit. Thus, we have separately trained a model instance on larger image sizes to generate these defect types.

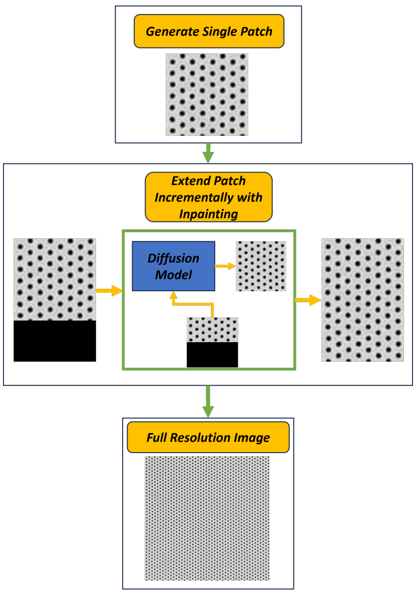


Fig. 3: Proposed approach to generate full-size, defect-free SEM image (archetype) using patch based method

D. Training: Object Detection

In section III-B, defect detector is trained on three datasets (LS-ADI, HEXCH-DSA, LS-AEI) under different configurations (real, synthetic, combined). Due to its fast training time, YOLOv5n has been selected as defect detector [6] to validate the use of generated synthetic images in training object detectors. Each model is initialised from COCO pretrained weights, trained for 200 epochs with batch size of 32, and with early stopping criteria enabled. The weights with best performance on validation dataset are selected for use in hereafter mentioned experimentations. Code and all other hyperparameters of YOLOv5n are used as implemented by Ultralytics [7].

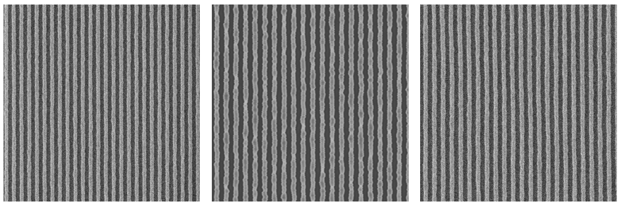
E. Labeling of Synthetic Images

To utilize synthetic images in training supervised defect detectors, such as YOLO, they must first be annotated/labelled. We propose labeling synthetic images by applying a defect detector already trained on real data. However, training defect detectors on synthetic data poses an additional challenge, as it may yield worse predictions compared to human annotation. This challenge arises due to two reasons: i) Synthetic images lack sufficient resemblance to the original data, and ii) Labeling errors in synthetic training data result in suboptimal learning signals, affecting performance on real test data. Not only have we tackled above mentioned challenges with our proposed approach, but we have also demonstrated in the next section how generated synthetic images and associated labeling quality improved or performed as per on defect detection task.

III. Results

A. Qualitative Evaluation of Synthetic Images

Synthetic images generated by the proposed DDPM-based approach are evaluated qualitatively. First, visual comparison does not yield any differentiating characteristics between synthetic defects (Fig.6) and those obtained from SEM tools (validated with several anonymous SEM image experts). Beyond visual comparison, line-edge-roughness and critical dimension (CD) are crucial metrology parameters in semiconductor patterning, towards validating device electrical characteristics and performance. To generate synthetic or artificial SEM images, with or without defects, using conventional software such as ARTIMAGEN [8], [9], it is essential to be aware of industry Best-Known-Methods (BKM) settings to comply with tool



(a) Real (b) Simulation (c) Proposed

Fig. 4: Comparison between (a) real SEM image, (b) image generated with software simulation, and (c) image generated with our proposed method.

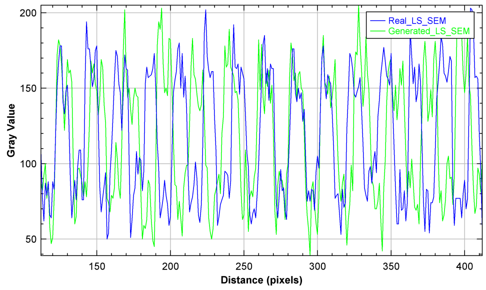


Fig. 5: Linescan plot for generated (using our proposed approach) and real SEM image

imaging conditions. Without appropriate values for parameters like pixel size, number of frames in acquisition, accelerating voltage, probe current, etc., it becomes quite challenging to generate synthetic images that closely resemble real images. Additionally, using incorrect parameter values can introduce digital artifacts into the synthetic images, rendering them unsuitable for ML model training and, consequently, compromising the preservation of original device characteristics. Contrary to this, our proposed approach generates synthetic images that closely resemble real images and preserve actual characteristics without requiring prior knowledge of BKM settings, as shown in Fig.4. Fig.5 shows the line-scan plots of generated (by our proposed approach) and real SEM image (for Line-Space feature), which can be used to compare CD and roughness parameters between these two. Lastly, Fig.7 presents inference results on generated synthetic and real images (test set) from a defect detector trained solely on real data. Both the inference confidence and classification accuracy on semantic contexts (such as probable gap and gap), as well as the line-scan plots of original and synthetic data (for line-space feature), appear nearly identical. This strongly indicates that generated synthetic images are sufficiently similar to the original data, supporting their use in expanding the size of the semiconductor defect inspection dataset.

B. Training Defect Detectors with Synthetic Images

Second, we demonstrate that synthetic images generated with the proposed approach can be succesfully used in training defect

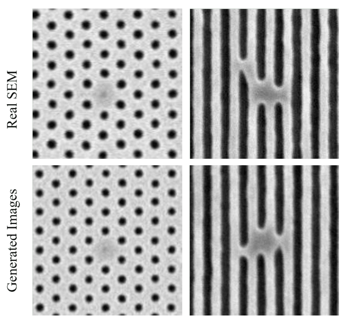
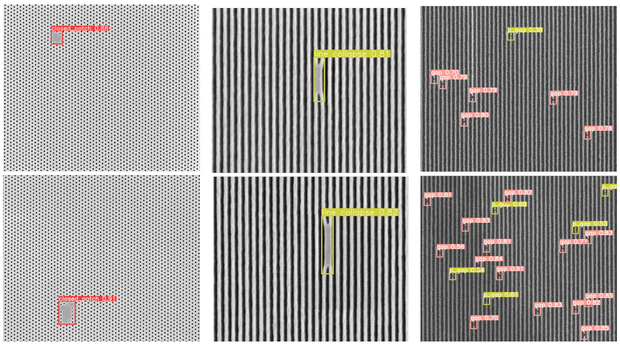


Fig. 6: Examples of generated defect instances against real SEM image of same defect type



(a) Closed Patch (b) Linecollapse (c) (p)gap

Fig. 7: Inference results for real (top), and synthetic (down) images

TABLE I: Statistics of real and synthetic training datasets

			Real	*Synthetic*	*Real+ Synthetic*
LS-ADI	**Instances**	*pgap*	315	1375	1690
		microbridge	380	1477	1857
		linecollapse	550	501	1051
		bridge	238	406	644
		gap	1046	2155	3201
	Total Images		1053	1199	2252
HEXCH	**Instances**	*cp*	74	210	284
		pch	94	434	528
		mh	30	591	621
	Total Images		174	420	594
LS-AEI	**Instances**	*multi bridge nh*	160	120	280
		multi bridge h	80	143	223
		linecollapse	202	248	450
		single bridge	240	245	485
		thin bridge	241	354	595
	Total Images		920	932	1864

detectors. We have trained defect detectors with generated synthetic counterparts of each investigated dataset in two configurations as: (1) with synthetic dataset only and, (2) combined with real dataset. Statistics of real and synthetic datasets are shown in table I.

Fig.8 shows the AP and AR scores per defect class on LS-ADI real test dataset, achieved by YOLOv5n model trained on either real, synthetic, or combined datasets. While some deviations and deficits are present, no major performance drops are observed when training on synthetic data. Thus, the proposed DDPM approach generates images based on the LS-ADI dataset, which can be properly utilized for training defect detectors, as no major performance deficits are encountered when switching from training on real data, to training only on synthetic data. Training only on the combination of the real and synthetic dataset does not yield any performance gains, despite the larger size of the combined dataset.

Fig.9 shows the AP and AR scores per defect class on LS-AEI real test dataset by model trained on real, synthetic or combined datasets. A scenario similar to the LS-ADI is observed, where training on the synthetic dataset does not lead to major performance deficits.

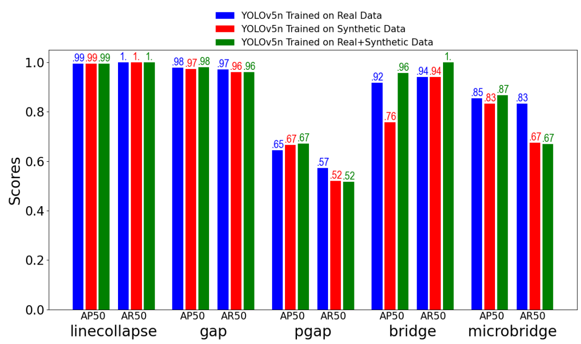


Fig. 8: AP and AR scores achieved on real ADI test dataset

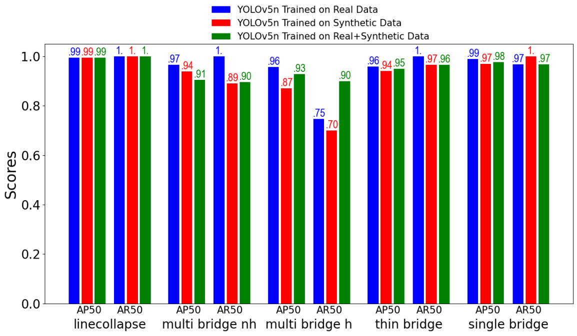


Fig. 9: AP and AR scores achieved on real AEI test dataset

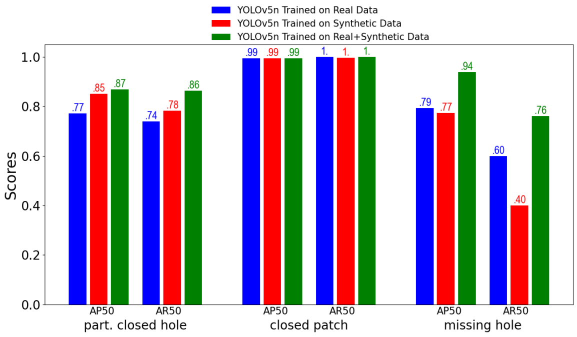


Fig. 10: AP and AR scores achieved on real HEXCH-DSA test dataset

Finally, Fig.10 shows the AP and AR scores on HEXCH-DSA real test dataset by model trained on the different dataset configurations. On HEXCH test data, a definite performance improvement has been observed when model is trained on the combined synthetic+real dataset. HEXCH dataset is significantly smaller in size compared to the other two datasets (table.I). This may explain that, while model did not benefit from training on synthetic+real data for LS datasets, the model significantly benefited for HEXCH-DSA dataset, as combining both synthetic and real data increased the dataset size (without altering real characteristics of the image/defect features) to properly learn the required defect features.

Table II shows the mAP and mAR scores achieved on the different real SEM test datasets, for different training dataset configurations. This table summarises that, while training only on synthetic data does not provide clear benefits in all scenarios, it never causes significant performance drops, against training only on real dataset. This validates that the synthetic images generated by proposed method can have valid usage in training defect detection models, by replacing or combining with real SEM images datasets for different process steps.

C. Defect Transfer

SEM wafer images can vary significantly depending on factors such as dose/focus used, design geometrical patterns, critical dimension, resist profiles, or underlayers. Consequently, significant numbers of SEM images have to be acquired and labeled for each set/combination of process parameters to train a defect detection model. Furthermore, as defect types are stochastic in nature, two challenging scenarios may occur

TABLE II: mAP and mAR scores achieved on real test data

Dataset (test split)	*Metric*	*YOLOv5n with Training Data*		
		Real	*Synthetic*	*Real+Synthetic*
Real ADI-LS	mAP	**0.878**	0.845	0.839
	mAR	**0.864**	0.819	0.839
Real HEXCH-DSA	mAP	0.853	0.873	**0.933**
	mAR	0.78	0.85	**0.874**
Real AEI-LS	mAP	**0.973**	0.943	0.951
	mAR	0.943	0.911	**0.946**

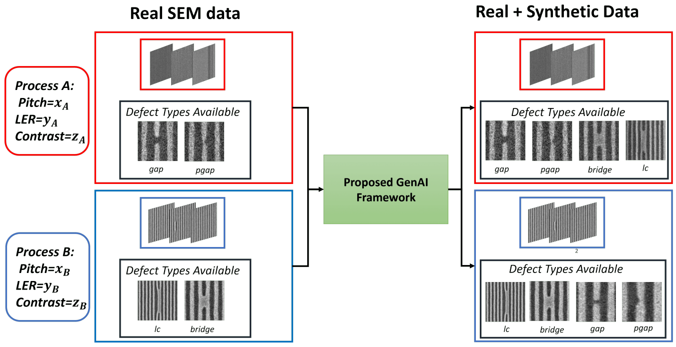


Fig. 11: Proposed framework can generate defect types outside its typical process parameters, to prepare defect detectors for unexpected scenarios.

quite frequently, as (i) a given defect type/class has a very small probability to occur in that process while training defect detection models (class imbalance), or (ii) relevant defect SEM images dataset to train a model is not just rare and noisy, but also very expensive to get (limited training dataset size).

In these cases, without sufficient images of certain defect types available at a certain process step, deploying an industry-compliant ML-based defect detection framework may be problematic, as overall model convergence can not be guaranteed towards generalizability and robustness due to model's underfitting for those defect type's features.

To mitigate this, we have examined whether the proposed approach can generate instances outside of the extent of corresponding defect type's typical process context. In this way, the proposed generative model can be trained on different processes and their associated defect types concurrently. Afterwards, defect instances can be generated for a process where the given type has not been encountered yet (or encountered in limited numbers). This generated dataset can then be used to train ML-based defect detectors towards detecting the given defect types in the new environment. Our proposed approach is demonstrated in Fig.11

The proposed approach manages to successfully generate defect instances outside of the process parameters they were encountered in during training. However, as of now, no dataset which allows extensive investigation of this proposed "defect transfer" approach, is available to the authors. Therefore, quantitative experimentation and validation with this approach is left for future research directions.

REFERENCES

[1] P. De Bisschop, "Stochastic effects in euv lithography: random, local cd variability, and printing failures," *Journal of Micro/Nanolithography, MEMS, and MOEMS*, vol. 16, no. 4, pp. 041013–041013, 2017.
[2] G. F. Lorusso *et al.*, "Metrology of thin resist for high na euvl," SPIE Advanced Lithography+Patterning, 2022.
[3] T. Kondo *et al.*, "Massive metrology and inspection solution for euv by area inspection sem with machine learning technology," in *Metrology, Inspection, and Process Control for Semiconductor Manufacturing XXXV*, vol. 11611, pp. 210–219, SPIE, 2021.
[4] A. Q. Nichol and P. Dhariwal, "Improved denoising diffusion probabilistic models," in *International Conference on Machine Learning*, pp. 8162–8171, PMLR, 2021.
[5] A. Lugmayr *et al.*, "Repaint: Inpainting using denoising diffusion probabilistic models," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11461–11471, 2022.
[6] B. Dey *et al.*, "Towards improving challenging stochastic defect detection in sem images based on improved yolov5," in *Photomask Technology 2022*, vol. 12293, pp. 28–37, SPIE, 2022.
[7] G. Jocher *et al.*, "Yolo by ultralytics. 2023," *URL: https://github.com/ultralytics/ultralytics*, 2023.
[8] P. Cizmar *et al.*, "Simulated sem images for resolution measurement," *Scanning*, vol. 30, no. 5, pp. 381–391, 2008.
[9] P. Cizmar *et al.*, "Optimization of accurate sem imaging by use of artificial images," in *Scanning Microscopy 2009*, pp. 208–213, SPIE, 2009.